

BANGLA CLASSIFIERS: MEDIATING BETWEEN KINDS AND OBJECTS¹

Veneeta Dayal
Rutgers University

To appear in the Italian Journal of Linguistics

ABSTRACT

This paper is concerned with the expression of kind, definite, and indefinite readings in Bangla. Bangla is a classifier language and allows bare nominal arguments to have kind readings. The definite-indefinite contrast, however, requires the presence of classifiers and is sensitive to the relative position of the noun and the classifier. The paper first presents diagnostics to establish the semantic import of the variation in word order. It then provides an explicit syntax and semantics for each variant. Finally, it draws out the implications of the Bangla facts for theories of mapping from morpho-syntax to meaning in the nominal domain. It establishes that Bangla adds a significant new dimension to our current understanding of the range of possible variation, while acknowledging the challenges it presents for an empirically adequate theory of cross-linguistic variation.

1: BACKGROUND

1.1. Bangla as a Classifier Language

Masica (1976) in his book *Defining a Linguistic Area: South Asia*, notes the following:

“THE USE OF NUMERAL CLASSIFIERS OR ‘COUNTER WORDS’ ...LINKS CERTAIN LANGUAGES MAINLY ON THE EASTERN SIDE OF INDIA WITH THE LANGUAGES OF EAST AND SOUTHEAST ASIA. IN THIS CASE, THE FEATURES IN THE INDIAN LANGUAGES CONCERNED ARE MARGINAL INSTANCES OF A PHENOMENON THAT SEEMS CLEARLY TO HAVE ITS CENTER IN SOUTHEAST ASIA.”

¹ Thanks are due to Probal Dasgupta for extensive comments on a first draft, only some of which I could address here. Thanks also to Mark Baker, Gennaro Chierchia, Jane Grimshaw, Peter Jenks, Li Jiang, Ayesha Kidwai, Xu-Ping Li, Mingming Liu, Beibei Xu and audiences at the Conference on Bantu, Chinese and Romance DPs (University of Leiden) in October 2010 and the International Workshop on Syntax-Semantics Interface (Academia Sinica) in June 2011, where earlier versions of this work were presented. I gratefully acknowledge the financial support from a Fulbright Senior Research Award in 2004-05 for the initial research on classifiers in South Asian languages and to Joyshree Sutradhar whose research assistance in that project was invaluable.

Masica's statement is accurate in the sense that classifiers do not constitute an areal feature and to that extent they are not definitive of South Asia as a linguistic area. The statement is not to be understood as claiming either that classifier languages of the area are few in number or that they are a marginal phenomenon in the languages that have them.

Sutradhar (2006) lists a number of such languages, from different language families. From Indo-Aryan: Bangla (aka Bengali), Asamiya (aka Assamese), Oriya, Bihari, Nepali, Sinhala. From Dravidian: Telugu (listed by Masica, not in Sutradhar), Malto, Kolami, Parji, Kui-Kuwi, Kurux. From Tibeto Burman: Ahom, Apatani, Bodo, Dimasa, Garo, Kokborok, Rabha, Mishmi, Karbi, Rangboli, Chtiya, Mikir, Empeo/Kachcha Naga, Lotha Naga. From Austro-Asiatic: Khasi, Pnar, Korowa, Kharia. And finally, from the contact language group: Nagamese, Bishnupuriya. This list may or may not be exhaustive but it is enough to establish that classifiers are well attested in South Asia.

In this paper I will be mainly concerned with the relation between morpho-syntax and interpretation in the nominal system of Bangla. Much of what I have to say also applies to Asamiya, though I do not provide examples or arguments from it. It is quite possible that the crucial features of the Bangla-Asamiya system may be manifested in other South Asian classifier languages as well.

A quick examination of the Bangla nominal system shows the characteristic properties of classifier languages:

- 1a.** kal ek *(-tʰa) / du *(-tʰo) tʃʰatro efe tʃʰilo
yesterday one CL two CL student came
“Yesterday a student/two students came.”
- b.** anu ek *(-tʰa) / du *(-tʰo) boi kinetʃʰilo
Anu one CL two CL book bought
“Anu bought a book/two books.”
- c.** anu ek / du *(peala) tʃʰa kheyetʃʰe
Anu one two cup tea drank
“Anu drank a cup of / two cups of tea.”

Bangla treats discrete entities like students/books and liquid substances like tea alike, in that neither can be counted directly. The noun does not change form whether the numeral is singular or plural. That is, Bangla does not show the typical mass-count distinction characteristic of number marking languages.

Furthermore, the classifiers themselves may be restricted to specific noun classes. Against the general classifier *-ta*, and its morphological variants *-to* and *-te*, are the following:²

- 2a. *-ta/to/te* general classifier for count nouns
- b. *-jon* classifier restricted to humans
- c. *-k^hana* classifier restricted to inanimate count nouns
- d. *-ra* number-neutral classifier restricted to animate nouns
- e. *-gulo* plural classifier applicable to all count and mass nouns
- f. *-k^hani* classifier restricted to mass nouns

The discussion draws on earlier work on Bangla classifiers, most notably by Bhattacharya, by Dasgupta and by R. Ghosh (see references). It also draws on fieldwork conducted at Jawaharlal Nehru University in a course on *(In)definiteness and Genericity* taught by Ayesha Kidwai and myself in Spring 2010, as well as on follow-up fieldwork by me.³ The aspect of the Bangla nominal system that I am interested in here is the expression of definite, indefinite and kind-related readings. Apart from showing how these readings are derived, the discussion aims to draw out the implications for an understanding of classifier systems across languages as well as to nominal systems more generally. While the ultimate goal is to develop a theory of cross-linguistic variation, the goal of the present paper is more modest. I use the Bangla facts to show the challenges inherent in such an enterprise, using a small sample of classifier languages to illustrate the theoretical issues involved in fitting Bangla into the typology.

1.2. *Kinds and Objects*

The readings we are interested in draw on a fundamental ontological distinction that is due to Carlson (1977). His study of English bare plurals established that the grammar of natural language is sensitive to the difference between kinds and objects. That is, we need to distinguish between reference to kinds vs. reference to objects in order to

² Even though the general classifier *-ta* is typically used for canonical count nouns, those that correspond to discrete entities, it can also be used for mass nouns with certain specific interpretations. This is discussed in Dayal (in prep).

³ Participants in the course: Diti Bhadra, Shiladitya Bhattacharya, Atanu Saha, Anu Beshears, Raj Lakshmi Singh, Aravind Kumar, Janani Kandhadai, Hima S., Jyoti Iyer, Bipasha Patgiri, Akansha Bansal, Thoibi Oinam, Manji Bhadra, Madhav Gopal. The facts in section 2.2 are from the fieldwork done in class. The class was divided into groups that met with and elicited data from Diti Bhadra, Shiladitya Bhattacharya, Atanu Saha, who worked as consultants for the groups they were not part of. Follow-up fieldwork included Anannya Dasgupta in addition to the three named above.

differentiate between the acceptability and interpretations of English sentences with bare plurals and definite noun phrases:

- 3a.** Dinosaurs are extinct.
- b.** The dinosaurs are extinct.
- c.** The dinosaur is extinct.

An important starting point for Carlson is the fact that the predicate *extinct* can only hold of species, not of members of the species. (3a) shows that the bare plural can denote the kind *dinosaurs* and serve as the argument of the kind-level predicate *extinct*. (3b) does not have this reading, showing that the plural definite does not denote the kind *dinosaurs*. It does have a reading, however, in which extinction is predicated of the set of dinosaur sub-kinds. Replacing *the dinosaurs* with *the dodos* would lead to unacceptability since dodos are not known to have sub-kinds. (3c), on the other hand, is acceptable as referring to the kind again, and is therefore replaceable by *the dodo*. The generalization that emerges is that in English, bare plurals and definite singulars can denote kinds while definite plurals cannot. Both definite singulars and definite plurals can, of course, denote at the object level. (4a-b) predicate the activity of barking at a particular time and place to a unique dog/the plurality of dogs in the context:

- 4a.** The dog is barking.
- b.** The dogs are barking.

Carlson also argued that kind denoting terms can be used in contexts where object level reference is required:

- 5a.** Dogs are barking.
- b.** Dogs are not barking.

The bare plurals in (5a-b) refer not to kinds but rather to (some) instantiations of the kind. A crucial feature of Carlson's analysis of bare plurals as kind terms was to deliver the obligatory narrow scope reading observed in sentences like (5b). The sentence can only mean that there are no dogs barking, not that some are and some are not.

There are, as is well known, two broad developments with respect to the phenomena studied by Carlson. While reference to kinds in the case of English bare plurals is accepted by all for kind-level predication, there is a difference of opinion in the case of object-level predication. The proponents of the view that bare plurals are simply indefinites in object-level statements, I believe, have not met the bar for explaining the special properties of bare plurals noted by Carlson (see Dayal 2004, 2011 for discussion). There is also no specific theory for predicting cross-linguistic variation in the form and interpretation of noun phrases, which we will discuss in the next subsection. I will therefore follow the Neo-Carlsonian approach of Chierchia (1998), with the specific modification of ranking from Dayal (2004). The key ingredients of this account are given below:

6. *Set of type shifts*: nom ($\langle s, \langle e, t \rangle \rangle \rightarrow \langle s, e \rangle$)
 Iota ($\langle e, t \rangle \rightarrow e$)
 Pred ($\langle s, e \rangle \rightarrow \langle e, t \rangle$)
 \exists ($\langle e, t \rangle \rightarrow \langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$)

7a. *Ranking of Type Shifts* (from $\langle e, t \rangle \rightarrow \langle e \rangle / \langle \langle e, t \rangle, t \rangle$): $\{\text{nom, iota}\} > \exists$

b. *Blocking Principle* (Type Shifting as Last Resort): For any type shifting operation τ and any X: * $\tau(X)$ if there is a determiner D s.t. for any set X in its domain, $D(X) = \tau(X)$.

c. *Derived Kind Predication* (DKP): If P applies to objects (ie. ordinary individuals) and k denotes a kind, then $P(k) = \exists x [\cup k(x) \wedge P(x)]$, where $\text{PRED}(\cup): \lambda k_{\langle s, e \rangle} \lambda x [x \leq k_s]$

Chierchia takes bare plurals to be predicative terms that can shift to argumental types through the three basic type shifts available in language: *nom*, *iota*, \exists . In English *nom* is a covert type shift while *iota* is lexicalized in the definite determiner *the*. The substantive difference between the sort adjusting rule of *DKP* vs. \exists is that the former yields obligatory narrow scope for kind denoting bare plurals. The latter is associated with scopal flexibility. We will refer to indefinites of the first type as ‘weak indefinites’ and those of the latter type ‘strong indefinites’. In Chierchia’s system, \exists only applies when *nom* and *iota* are unavailable, either due to conceptual blocking (of *nom*) or lexical blocking (of *iota*). These concepts are perhaps best explained by looking at some example sentences.

Consider the following, where the first sentence has a kind-level predicate, the second is an episodic statement and the last includes a scope bearing expression:

8a. $\llbracket \text{Dogs have evolved from wolves} \rrbracket = \text{evolve-from}(\cap \text{dogs}, \cap \text{wolves})$

b. $\llbracket \text{Dogs are barking} \rrbracket = \text{barking}(\cap \text{dogs}) \Rightarrow_{\text{DKP}}$

$$\exists x [\cup \text{dogs}_s(x) \wedge \text{barking}_s(x)]$$

c. $\llbracket \text{Dogs are not barking} \rrbracket = \neg \text{barking}(\cap \text{dogs}) \Rightarrow_{\text{DKP}}$

$$\neg \exists x [\cup \text{dogs}_s(x) \wedge \text{barking}_s(x)]$$

In (8a) the predicative expressions, *dogs* and *wolves*, are type-shifted by *nom* to make them into argumental type $\langle e^k \rangle$.⁴ Interpretation proceeds smoothly since the sort of the argument and the predicate match, they are both kind-level. The sentence will be true if and only if it is a fact that the dog-species evolved from the wolf-species. Let us now turn to (8b), which has an object-level predicate. The property of barking at a given time and

⁴ I subscript $e^{o/k}$ to indicate whether reference is to object level or kind level individuals.

place cannot hold of the species as a whole but only to some instantiations of it. Feeding in an argument of type $\langle e^k \rangle$ results in a sort mismatch and requires the mediation of the rule of *DKP*, given in (7c). There are two aspects of the rule worth noting. Instead of reference to kinds, we now have reference to instantiations of the kind, via *pred*. Furthermore, we have existential quantification over such instances. The sentence will be true if and only if there are some instances of dogs that are barking. But the introduction of an existential quantifier raises the issue of scopal interaction with other scope bearing expressions. (8c) shows that *DKP* yields obligatory narrow scope since it is a local sort adjusting operation that takes place when a kind-level argument combines with an object-level predicate. To see this, it may be worth comparing (8c) to cases where scopal flexibility is possible.

Consider (9a) which has a bare plural, but one that does not conceptually denote a kind. It is too grounded in a particular machine for its parts to count as a kind.

According to Chierchia, such a bare plural will be shifted by the lower ranked \exists type-shift, *nom* being conceptually blocked and *iota* being lexically blocked by *the*. Crucially, this bare plural can take scope over negation, as witnessed by the acceptability of a follow-up like *but parts of it are*. (9b) represents another case of an indefinite taking scope over negation. The sentence can be true even if there are some barking dogs, as long as three are not barking:

- 9a.** Parts of this machine are not new.
b. Three dogs are not barking.

While the ranking argued for by Chierchia seems reasonable enough for English, a closer study of languages without articles, such as Hindi/Russian/Mandarin, prompted the revision in (7a). As discussed in Dayal (2004), bare plurals in such languages do not show variable scope readings. Contrary to the view that the absence of articles leaves open the possibility of bare nominals being definite or indefinite, it is established that in fact they can only have kind or definite readings. Their indefinite readings are of the type associated with kind terms that shift through *DKP*, rather than the variable scope interpretation expected of \exists -type shifted expressions. This shows that *nom* and *iota* do not compete with each other and they jointly outrank \exists . In these languages, neither *nom* nor *iota* is lexically blocked, so both are available and the lower ranked \exists is never tapped. This is significantly different from English where bare plurals do not have definite readings, because of lexical blocking. Whether they have strong indefinite readings or not, therefore, turns on whether they are defined for *nom*.⁵

⁵ It is worth noting that the translation of (9a) into Hindi results in a definite interpretation, as predicted by the revision. The translation of (9b), however, has the same interpretation as in English. This will be of relevance in later sections.

1.3. A Theory of Cross-linguistic Variation

In this section we will consider the theory of cross-linguistic variation forwarded in Chierchia (1998). Although much has changed since then in our understanding of the empirical domain of variation, partly as a result of reactions to that proposal, it remains to date the most explicit theory of variation. In section 4 we will entertain a different theoretical tack but for now let us see how Chierchia seeks to capture the phenomena. The two key features of this system are as follows:

10a. The Nominal Mapping Parameter (NMP): $N \Rightarrow [+/- \text{ pred}, +/- \text{ arg}]$

Languages without Mass-Count Distinction

- i. [- pred, + arg] every lexical noun is mass. *Chinese*

Languages with Mass-Count Distinction

- ii. [+pred, +arg] bare arguments are allowed. *With articles: Germanic*
Without articles: Slavic
- iii. [+pred, -arg] bare arguments disallowed. *With $\delta_{\text{null-det}}$: Italian*
Without $\delta_{\text{null-det}}$: French

b. Avoid Structure: Apply SHIFT at the earliest level.

Consider [+arg, +pred] languages like English and Hindi, one with a definite determiner, the other without. NPs that start out as $\langle e^o, t \rangle$ and can shift covertly via *nom* to $\langle e^k \rangle$, can do so at the level of NP without violating the *NMP*. Bare plurals therefore can refer to kinds and have weak indefinite readings at the object level. This was demonstrated for English in the previous section. The two languages part company when it comes to *iota*, because of lexical blocking. This accounts for the fact that Hindi bare NPs can also have definite readings. In the case of French, which is classified as [-arg, +pred], no type shift from $\langle e^o, t \rangle \rightarrow e$ can take place at NP. DP is obligatorily projected and must be filled with either a null or an overt D. This rules out bare arguments altogether if the language does not have null determiners. The presence of D in languages with null determiners is detected through licensing requirements that restrict bare arguments to specific syntactic positions.

Our current concern is with classifier languages, which instantiate the [+arg, -pred] setting of the parameter. Chierchia takes NPs in classifier languages to obligatorily denote type $\langle e^k \rangle$. The idea that bare plurals in number neutral languages denote kinds and that the role of classifiers is to mediate between the kind-level meaning and a predicate of objects goes back to Krifka (1995) who analyzed Mandarin in these terms. For Chierchia, the projection of classifiers is a way for the language to shift to type $\langle e^o, t \rangle$ without violating the *NMP*. As for definite readings, consider the fact that kinds can be shifted to properties by *pred*. Thus the following schematic possibilities are available in principle. Although Chierchia's system does not have expletive determiners, I include the option here for completeness. I should clarify though that I do not personally endorse the possibility of expletive determiners:

- 11 a. [DP D_{expl} [NP N_{<e>}]]
 b. [DP D_{<<e,t>, e>} [NP PRED(N_{<e>})_{<e,t>}]]
 c. [NP N_{<e>}]

Chinese being [+arg, -pred], (11b) is not an option: by hypothesis NPs cannot denote properties. *Avoid Structure* rules out an expletive structure like (11a) since there is a smaller structure with the same meaning, namely (11c). This predicts that classifier languages should generally allow bare arguments and not have definite determiners, a prediction that was believed at the time to be correct.⁶

There are three other properties of [+arg, -pred] languages that Chierchia seeks to capture: the extension of all nouns as mass, no pluralization and a generalized classifier system. It is worth noting that the explanations for these do not follow directly from the NMP but rather from other aspects of his theory, specifically his view of mass nouns and kinds as having number neutral denotations. Contrary to the view of mass nouns as mereological sums, Chierchia takes them to have the same atomic structure as count nouns. Thus an individual chair or table would be identified as atomic parts of the denotation of mass nouns like *furniture*. Similarly, the denotation of mass nouns like *water* also includes atoms, even though those atomic entities may not be ordinarily identifiable. The real difference, he claims, is that count nouns denote a set of atomic entities, with plural entities entering the denotation as a result of pluralization, whereas mass nouns come out of the lexicon with both atomic and plural entities in their denotations.⁷ The way in which he defines the pluralization operation ensures that it would be undefined for such terms. Furthermore, counting requires a salient level of individuation. In the case of count nouns, the set of atomic entities provides this level but in the case of mass nouns there is no distinguished level of individuation. As such, a measure phrase is required. This is how mass nouns work in a [-arg, +pred] language. In a [+arg] language, mass nouns can denote kinds but when shifted to type $\langle e^o, t \rangle$ (via *pred*), a number neutral property with no distinguished level of individuation is obtained. Once again, the mediation of measure phrases is required for counting.⁸ From this one can plausibly conclude that languages in which all nouns denote kinds, the [+arg, -pred] languages, properties obtained via *pred* will not be of the appropriate sort for counting. Classifiers provide the same function as measure phrases in these languages by making

⁶ See section 4 for an explicit statement of how the definite readings can be derived from kind readings.

⁷ See also Chierchia 2010 and references cited there for discussion and further developments.

⁸ An anonymous reviewer asks for clarification on how the theory applies to [-arg, +pred] languages like French. For Chierchia, mass nouns are lexically pluralized. This means that even though they have an atomic structure, the atomic level is not salient enough to be visible for counting. As such, measure phrases will always be needed. In a [+arg, +pred] language like English, Chierchia holds that mass nouns denote kinds and the measure phrase serves to mediate between kinds and objects, much like classifiers in Chinese type languages. For present purposes, we can set aside how mass nouns are treated in languages with number (see Chierchia 2010 for further developments) and focus on classifier languages.

available a set of atoms for purposes of counting. Pluralization fails for the same reason, the properties denoted by all kind denoting nouns are number neutral and thus are already pluralized, so to say. It is in this sense that Chierchia's claim that the extension of all nouns in [+arg, -pred] languages is mass is to be understood.

I have been at some pains to separate out which properties are explained through the *NMP* and which properties are explained through the view of mass nouns and kind terms as denoting number neutral properties. The *NMP* yields the ubiquity of bare arguments in classifier languages, the view of mass/kind terms the absence of plural marking and the need for classifiers in counting, even for nouns referring to discrete entities. It will be important to keep these distinctions in mind when we return to the issue of cross-linguistic variation. For now, we turn to Bangla and take note of how the nominal system expresses reference at the level of kinds and reference at the level of objects.

2. (IN)DEFINITENESS IN BANGLA

In this section we will look at core cases involving reference at the object level and probe the relationship between word order and specific, definite and indefinite readings.

2.1. NP Raising and Specificity

Bhattacharya (1999a, 1999b) pays considerable attention to variations in word order and their accompanying semantic consequences. Briefly put, there are two orders that are relevant. One is the order in which the numeral+classifier occurs before the noun: (*Possessive*) (*Demonstrative*) *Num CL NP*, the other is the order in which the noun comes before the numeral+classifier: (*Possessor*) (*Demonstrative*) *NP Num CL*. Some relevant cases are shown below:⁹

- 12a.** du ʈo lal boi
 two CL red book
 “Two red books.”
- a’.** lal boi du ʈo
 red book two CL
- 12b.** ei du ʈo lal boi
 this two CL red book
 “These two red books.”

⁹ Probal Dasgupta (p.c.) points out that there is another version of the demonstrative *e-* (and *o-* for the distal *oi*) which is not acceptable in dem+numeral+CL+N combinations but is acceptable in dem+N+CL, dem+CL and dem+num+CL. I refer the reader to related discussion in Dasgupta (1992). Unfortunately, I must leave the analysis of this contrast for another occasion.

b' ei lal boi du ʈo
 this red book two CL

12c. amar ei du ʈo lal boi
 my this two CL red book
 “These two red books of mine.”

c' amar ei lal boi du ʈo
 my this red book two CL

Bhattacharya explains the paradigm in the following way. The base order of the Bangla noun phrase is as given in (13a), where the numeral and the classifier form a complex head of the Classifier Phrase. The classifier optionally carries a specificity feature that must be checked by movement of the post-classifier NP into the Spec of the Classifier Phrase. The fact that the adjective moves with the noun shows that this is an instance of NP movement, rather than the kind of N to D movement seen in Italian (Longobardi 1994):¹⁰

13a. [_{DP} amar [_{FP} ei [_{Clap} du-ʈo [_{NP} lal boi]]]]
 my this two-CL red book

b. [_{DP} amar [_{FP} ei [_{Clap} [_{NP} lal boi] du- ʈo t_{NP}]]]
 my this red book two-CL

It is worth noting here that it is possible for a classifier to occur without a numeral but in that case NP raising is obligatory, giving rise to the following:

14a. boi ʈa
 book CL
 “The book.”

a'. * ʈa boi
 CL book

b. ek ʈa boi
 one CL book
 “A/One book.”

b'. * boi ek ʈa
 book one CL

¹⁰ See, in particular, Bhattacharya (1999a) pp. 96-102 for details.

As we can see, if there is no numeral, the only acceptable order is NP+CL. This is presumably because the classifier needs to cliticize to an expression to its left. It can be argued that there is a null numeral *ek* ‘one’ in such cases, given their strictly singular interpretation. Because a null numeral does not allow cliticization, NP raising is forced (cf. 14a vs. 14a’). (14b) is the only grammatical option for a non-raised structure. One could take (14b’) to be blocked by the presence of (14a), on par with non-emphatic **the one book* vs. *the two books* in English.¹¹ I will attempt a better explanation for this gap after I have presented my analysis of the core facts.

To return to Bhattacharya’s account of NP raising, he notes that the versions with NP movement are more specific than the ones without movement. So, for example, (12a) *du fo lal boi* ‘two CL red book’ would be translated as the indefinite term ‘two red books’ while (12a’) *lal boi du fo* ‘red book two CL’ would be translated as the definite term ‘the two red books’. Similarly, (14a) would be translated as ‘the book’ while (14b) would be translated as ‘a/one book’. He does not elaborate much further on the notion of specificity that is relevant for NP raising.

2.2. NP Raising and Definiteness

While I agree with Bhattacharya’s basic claim, I believe the difference between the raised and non-raised versions is better characterized in terms of definiteness rather than specificity. This modification of Bhattacharya’s claim is based on field work conducted in 2010 but I should note that the idea that it is definiteness that is at issue has precedents in the literature.¹² Here I establish the claim on the basis of diagnostics that distinguish between definiteness and specificity. As we will see, the word order differences turn on presuppositions of uniqueness/maximality, a characteristic property of definites.

Consider first the examples in (15). The entities in the first sentence establish the existence of a set of students from which the set of students in the second sentence is to be drawn. That is, the context is one which supports partitive specificity (see Enç 1991, Diesing 1992):

¹¹ English allows the numeral *one* for emphasis: *The one book I wanted I couldn’t find*. In cases where a noun modified by a relative clause has a definite reading, Bangla adopts the correlative strategy typical of Indo-Aryan languages:

i. *jei ek-ta boi ami kinetʃʰilam ʃei boi-ta...*
 wh one-CL book I bought that book-CL

¹² For a good overview of the literature see R. Ghosh (2001). According to Dasgupta (p.c.) Azad (1983) claimed that definiteness was at issue in NP raising and more recently, Hildegunn Dirdal has argued against Bhattacharya’s claim of specificity in favor of definiteness. I do not have access to these works, so I do not know if the diagnostics presented here were also discussed by them. I mention them to acknowledge possible predecessors of the ideas presented here.

- 15a.** tin tɛ tʃʰatro eʃe tʃʰilo. du tɔ tʃʰatro boʃlo
 three CL student came two CL student sat
- b.** tin tɛ tʃʰatro eʃe tʃʰilo. #tʃʰatro du tɔ boʃlo
 three CL student came student two CL sat
 “Three students came. Two (of the) students sat down.”

We see here that partitive specificity can only be expressed by the base order with the NP following the classifier. The unacceptability of (15b) is to be expected if we correlate NP raising with the maximality typically associated with definite descriptions. The first sentence establishes a context in which there is no unique maximal individual made up of two atomic students.

We see a similar divergence in the case of referential specificity, discussed by Fodor and Sag (1982), among others:

- 16a.** jodi du tɔ tʃʰatro aʃe, ami parabo
 if two CL student come I will teach
 “If two students come, I will teach.”
- b.** jodi tʃʰatro du tɔ aʃe, ami parabo
 if student two CL come I will teach
 “If the two students come, I will teach.”

(16a) is ambiguous between a regular indefinite reading and a specific indefinite reading. Under the first, I will teach as long as I have more than one student in class, regardless of who those students are. Under the second, I may have many students but I care only about the presence of two particular students. This is the specific indefinite reading, where the identity of the individuals is not presumed to be known to the hearer. (16b), on the other hand, presupposes that I have exactly two students, familiar to speaker and hearer both, and the sentence is about their attendance. If there are more than two students salient in the discourse, the sentence will be infelicitous. In other words, (16a) can have a specific indefinite reading, while (16b) has only a definite reading.¹³

The distinction between specificity and definiteness can also be established by examining the possibility of intermediate readings in sentences like the following (Farkas 1981, among others):

- 17a.** prottek tʃʰatro du tɔ biʃoy-e ʃɔb pepar porlo
 every student two CL topic all paper read
 “Every student read all the papers on two topics.”

¹³ Another way to test this is to use singular terms *ɛk-ta tʃʰatro* vs. *tʃʰatro-ta*. The first implies the existence of a plurality of students, with reference being made to one of them; the second presupposes that there is exactly one student in the class. This might make it easier to explore intuitions but I do not include them in the text since, strictly speaking, they do not form a minimal morpho-syntactic pair.

- b.** prottek- ʈi tʃʰatro biʃoy-e du ʈo- te ʃob pepar porlo
 Every-CL Student topic two CL all paper read
 “Every student read all the papers on the two topics.”

Only (17a) has a reading in which the choice of topics varies with each student. (17b) presupposes that every student had to study the same two topics.

Finally, the specific/non-specific distinction is sometimes correlated with the referential/attributive distinction of Donnellan (1966). We see below that the NP preposed versions are used in contexts where a definite in English has an attributive reading. That is, when the description is expected to apply to a unique person, but the identity of that person has not been settled – the attributive reading. In (18a) the outcome of the contest is yet to be determined; and (18b) can be uttered based on circumstantial evidence at the scene of the crime, rather than on observations about an individual’s behavior:

- 18a.** je lok-ʈa jitbe, ʃe εk ʃo ʈaka pabe
 which person-CL wins that one hundred rupees will get
 “Whichever person wins will get a hundred rupees.”
- b.** k^huni: niʃtʃoi pagol
 murderer-CL certainly mad
 “The murderer must be insane.”¹⁴

What is clear in both cases is that there is a unique individual at issue, showing that this is what NP raising is sensitive to.

We have seen then, on the basis of familiar diagnostics, that the two versions of classifier phrases correspond to indefinite and definite readings of the noun phrase, where maximality is taken to be a key indicator of definiteness. Let us now consider the two versions with the demonstrative. It is not immediately obvious how specificity or definiteness can distinguish between them. Bhattacharya explains such cases in terms of a difference between deictic (19a) and specific readings (19b):

- 19a.** ei du ʈo boi
 this two CL book
- b.** ei boi du ʈo
 this book two CL

¹⁴ *principal-ke b^halo skolar hote hɔbe*, ‘The principal has to be a good scholar’ under the attributive reading, allows the bare NP without a classifier. However, this is a dative subject construction. Also, possibly, in this case ‘principal’ may have the status of a proper name. There are other such cases: *montri* ‘minister’ for example. A relevant article suggested by Probal Dasgupta (p.c.) is S. Ghosh (2006) but I have not had the opportunity to look at it yet. I mention it here, however, for completeness.

Demonstrative noun phrases, then, in the base order are compatible with a maximal or a non-maximal interpretation, though perhaps pragmatically favoring the latter. Demonstrative phrases in which the NP is preposed is only compatible with a maximal interpretation.

To sum up, we agree with Bhattacharya that there is a semantic correlate of DP-internal NP raising but differ somewhat in our characterization of it. We take such raising to correlate with a presupposition that there exist a unique maximal entity of the relevant kind, just as is the case with definites of the familiar sort. This is transparently reflected in the translations when there is no demonstrative present, but can also be established in cases where the presence of a demonstrative does not allow for a generalization in terms of an indefinite-definite opposition.

3. DERIVING (IN)DEFINITENESS IN BANGLA

In this section I will provide an explicit account of the facts given in section 2. In doing so, I will first expand the discussion to include bare NPs and reference to kinds. Specifically, I will try to account for the following facts. One, the Bangla nominal obligatorily requires the mediation of classifiers for counting. Two, cardinality phrases are indefinites with specific and non-specific readings. Three, a bare NP can have kind, generic or weak indefinite readings only, not strong indefinite/specific or definite readings. Four, the projection of a classifier with obligatory preposing of NP is required for the definite reading.

3.1. Bare NPs and Weak Indefiniteness

I will start by making the assumption, standard for classifier languages, that the Bangla NP denotes kinds, and may optionally be shifted to predicative meanings via the application of *pred* and to narrow scope existential readings via DKP. This is the default assumption for number neutral languages since Krifka (1995) and Chierchia (1998). Though we will probe this assumption in section 4, it makes sense to take this as our starting point. In (23) we see that Bangla bare NPs can be arguments to kind-level predicates:

- 23a. jatayat-er sahodʒ mad^hyom holo gaʒi
 transport-GEN convenient means be car
 “Cars are a convenient mode of transport.”
- b. pã tʃ^hodʒʃ ʃʌtabdi-te boi tʃ^hapa ʃuru hoi tʃ^hilo
 fifteenth century-LOC book printed start happened
 “Books started to be printed in the 15th century.”

As expected, these kind terms are fully acceptable in generic statements:

- 24a.** gaṛi petrol e chole
 car petrol on run
 “Cars run on petrol.”
- b.** ami roj rate boi pori
 I every night book read
 “I read books every night.”

Finally, again as expected, Bangla bare NPs have existential import in episodic statements:

- 25a.** ami ei rastay kal rate gaṛi dek^hetj^hilam
 I this road-LOC yesterday night car saw
 “I saw a car/cars on this road last night.”
- b.** ami gato botj^hor ei dokan t^heke boi kinetj^hilam
 I last year this shop at book bought
 “I bought a book/books from this shop last year.”

Bangla bare NPs, as we see, appear to be fully tractable under the standard approach. As predicted, the indefinite reading in (25) is a weak indefinite reading and the existential associated with the bare nominal would take narrowest scope if there were other operators in the sentence.

Applying the neo-Carlsonian approach discussed in section 1 to the data in (23-25) captures the intuitions correctly. We derive the following logical representations for the (a) sentences:

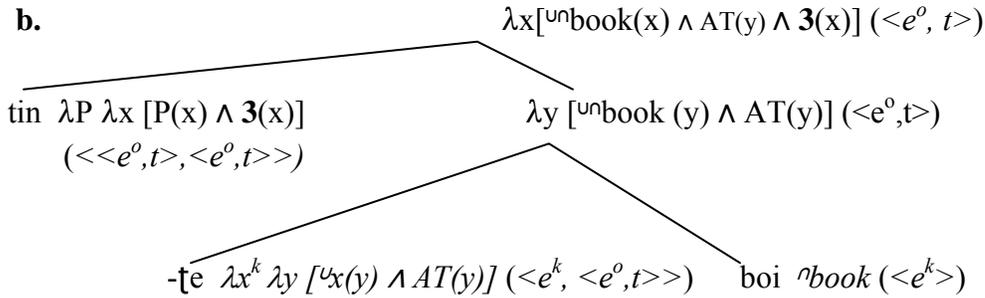
- 26a.** convenient-mode-of-transport(ⁿcar)
- b.** Gen s, x [ⁿcar(s)(x)] [run-on-petrol(s)(x)]
- c.** $\exists s \exists x$ [ⁿcar (s)(x) \wedge saw-yesterday(s)(I, x)]

So far, there is nothing special about Bangla, as compared to any language with kind denoting bare NPs. Its special properties, as we have already observed in the preceding section, are evident with object level readings that cannot be anchored to kinds. Here the mediation of classifiers is needed. In accounting for these readings, we will therefore start with cardinal phrases which involve the obligatory presence of classifiers.

3.2. Classifier/Cardinality Phrases – Predicative Meanings

In analyzing cardinal phrases, I will assume a basic layered nominal structure, in line with much current work. This division can roughly be demarcated as follows, where the Cardinality Phrase is treated as optional: [*DP* ([*Cardinality-P*]) *CL-P* [*NP*]]. In (27) I present a derivation for a phrase with a numeral, a classifier and a noun. The logical translations are given in italics and the type in parentheses:

27a. [_{Cardinality-P} tin [_{CL} -t̄e [_{NP} boi]]]
 3 -CL book



As shown above, I adopt the view that a classifier is a function from kinds to sets of object-level individuals. The classifier $-\text{t̄a}$ is a total function, but other quantifiers may impose restrictions on their domain in terms of shape, size etc. The set of individuals denoted by the classifier phrase is a set of atomic entities: AT in the logical translations above is a predicate denoting the set of those individuals that do not have proper parts. I further adopt the view of numerals in Ionin & Matushansky (2006) where a numeral is a predicate modifier:

28a. [[numeral]] =

$$\lambda P_{\langle e, t \rangle} \lambda x_e \exists Y_{\langle e, t \rangle} [\prod(Y)(x) \wedge |Y| = \text{numeral} \wedge \forall y \in Y P(y)]$$

b. [[tin-t̄e boi]] =

$$\lambda P \lambda x \exists Y [\prod(Y)(x) \wedge |Y| = 3 \wedge \forall y \in Y [\text{un}^n \text{book}(y) \wedge \text{AT}(y)]]$$

A numeral takes a set of atoms and yields a set of individuals, each of which can be partitioned into sets of the specified cardinality. For ease of exposition, I simply represent numerals as a predicate n , but it should be read as a short-hand for (28a). Under this view, numerals always combine with atomic predicates. Plural morphology in languages like English is considered a case of agreement rather than true semantic plurality. The role of classifiers fits in nicely with this account because the classifier is a function that takes a kind or its number neutral object level property and extracts the atoms from it.

To complete the picture, let us consider some of the other classifiers listed in (2), and repeated below:

29a. $-\text{j}^n$ classifier restricted to humans

b. $-\text{k}^n \text{ana}$ classifier restricted to inanimate count nouns

c. $\lambda x^k: \forall x [\text{un}^k z(x) \rightarrow \text{human/inanimate}(x)]. \lambda y [\text{un}^k x(y) \wedge \text{AT}(y)]$

The restriction to humans/inanimates is a presupposition, the part between the colon and the period, that restricts the domain of the function but the semantic import of the classifier is the same as in the case of the general classifier *-fa*.

The two plural classifiers in Bangla show variations on these options:

30a. *-ra* plural classifier restricted to animate nouns

b. $\lambda x^k: \forall x [\cup z(x) \rightarrow \text{animate}(x)]. \lambda y [\cup x(y)]$

31a *-gulo* plural classifier applicable to all count and mass nouns

b. $\lambda x^k \lambda y [\cup x(y) \wedge \neg \text{AT}(y)]$

I take *-ra* to have a presupposition restricting it to animate kinds, but lacking the atomization function. It thus yields the set of entities that instantiates the kind, a number neutral property that includes both atoms and pluralities. *-gulo*, on the other hand, lacks the presuppositional dimension but delivers a set of entities that are strictly plural. In this sense, it is the complement of *-fa*.

So, we have established that a classifier phrase has a predicative meaning at the object-level. *-jɔn* and *-kʰana* can both be arguments of cardinal expressions, but not *-gulo* and *-ra*. I look in more depth at these plural classifiers in Dayal (to appear). However, we will briefly mention a reason why the two plural classifiers do not combine with numerals. As pointed out by Ionin & Matushansky, counting requires elements of the set to be of standard sizes. Neither *-gulo* nor *-ra* delivers the right type of meaning. *-ra* marked phrases include atoms and pluralities, and *-gulo* marked expressions deliver pure pluralities, which can be of varying cardinalities.

3.3 Argumental Meanings – Strong Indefinite Readings

As indicated above, the cardinality/classifier phrase has a predicative meaning. In order to function as an argument its type has to be shifted, either to $\lambda Q \exists x [\cup^{n} \text{book}(x) \wedge 3(x) \wedge Q(x)]$ or to $\iota x [\cup^{n} \text{book}(x) \wedge 3(x) \wedge Q(x)]$. The indefinite meaning is often derived by appealing to the default type \exists shift (Landman 2004, Ionin & Matushansky 2006, Chierchia 2010, among others) and the definite meaning by the predicative meaning feeding into the lexical meaning of a definite determiner. The question of interest is how the predicative meaning turns into argumental type in Bangla.

Let us assume that NP and DP are the two projections that can appear in argument position. This implies that once a classifier phrase is projected, there is necessarily a DP projection above it. I will furthermore assume, departing from Bhattacharya in specifics but keeping to the spirit of his analysis of Bangla, that once a DP is projected, its head D has to be specified for the feature +/- definite. I will claim that numerals are, in fact, lexically ambiguous between predicative meanings of the type discussed above and generalized quantifiers of the kind shown in (32a):

- 32a.** $\llbracket \text{numeral}_2 \rrbracket = \lambda P \lambda Q \exists x [P(x) \wedge \text{numeral}(x) \wedge Q(x)]$
b. $[\text{NUM}_{\text{GQ}} [\text{NUM}_{\text{PRED-MODIFIER}} [\text{NP}]]]$
c. * $[\text{NUM}_{\text{PRED-MODIFIER}} \text{NUM}_{\text{GQ}} [\text{NP}]]]$

Ionin & Matushansky (2006) argue against a generalized quantifier meaning for numerals because such a meaning does not lend itself to the kind of complex cardinal expressions they are interested in: *twenty two N*, *two hundred N* etc. However, under the present proposal such expressions can be handled straightforwardly. The generalized quantifier meaning of a numeral can only work if it is the highest numeral expression in the phrase, as in (32b). A derivation like (32c) will simply lead to an irresolvable type mismatch since the higher numeral looks for something of type $\langle e, t \rangle$ and the lower numeral is of type $\langle \langle e, t \rangle, t \rangle$. So, Ionin & Matushansky's objection does not hold against this particular version of treating numerals as generalized quantifiers.

In (33a), I give the structural analysis of a phrase with the following order: numeral, classifier, NP. Here, the numeral is in D and has the generalized quantifier meaning, given in (32a) and the classifier phrase denotes a set of atoms, as given in (27b). The meanings combine by functional application to yield (33b), the set of properties of some set of three books, a strong indefinite meaning:

- 33a.** $[\text{DP } \text{tin} [\text{CL-P } \text{-te} [\text{NP } \text{boi}]]]$
b. $\llbracket \text{tin} \rrbracket (\llbracket \text{-te } \text{boi} \rrbracket)$
 three CL book
 = $\lambda Q \exists x [\text{book}(x) \wedge \mathbf{3}(x) \wedge Q(x)]$
c. * \exists -type shift $[\text{NUM}_{\text{PRED-MODIFIER}} [\text{NUM}_{\text{PRED-MODIFIER}} [\text{NP}]]]$

In positing a lexical ambiguity for numerals, I am specifically rejecting the option of a default type shift of the kind shown in (33c). I should note that the idea of having cardinal expressions encode an \exists quantifier is not to be taken too literally. \exists quantification is intended to capture the fact that cardinal phrases are not weak indefinites. We know, of course, that it is a matter of some debate what the appropriate way to capture the indefinite readings identified in section 2 might be: indefinites as generalized quantifiers (Montague 1974, Barwise & Cooper 1981), indefinites as set denoting terms that shift to argumental type via \exists -type shift (Partee 1986, Landman 2004), indefinites as choice functions, free or existentially closed (Reinhart 1997, Winter 1997, Kratzer 1998 among others), or quantifiers with a singleton domain (Schwarzschild 2002). For present purposes, however, I want to stay away from that discussion because there seems to be no significant cross-linguistic variation in terms of the possible readings for cardinal phrases – cross-linguistically they always allow strong indefinite readings.¹⁵ We can take the \exists -type shift, then, as a cover term for appropriate ways of fleshing out a strong indefinite meaning for a phrase headed by a numeral.

¹⁵ But see Trinh (2011) on Vietnamese. See also the discussion in section 4.

3.4 Argumental Meanings – Definite Readings

Turning to how the +def feature on D can be satisfied, recall that there are three different possibilities for Bangla cardinal phrases to have definite readings. We take each of these in turn. The first option is shown below, where there is an overt demonstrative in D:

34a. $[_{DP} [_{D+def} \text{ Demonstrative}] [(NUM) [CL [NP]]]]$

b. Demonstrative: $\lambda P \iota x [P(x) \wedge x_i = x]$

35a. ei tin-ṭe boi
this 3-CL book

b. $[[34b]]$ ($[[27b]]$)

$\Rightarrow \iota x [\text{book}(x) \wedge \mathbf{3}(x) \wedge x_i = x]$

Crucial in this derivation is the view that a demonstrative includes an indexical in its meaning (Kaplan 1989, Wolter 2006, Robinson 2006). (35b) denotes the unique entity that has two properties: it is a plurality of three books and it is the intended referent. Thus we get through simple compositionality, what Bhattacharya calls the “deictic” reading of demonstrative phrases. Phrases like (35a) can be felicitously used if the intended referent is a set of three books. There is no presupposition ruling out the possibility of other books in the context.

Next consider the possibility of satisfying the +def feature through NP raising:

36a. $[_{DP} NP_i D^0_{+Def} [_{Cardinality-P} \mathbf{3} [CL-P \text{ t}_e [t_i]]]]$

b. $[[3 CL t_{NP-i}]] = \lambda x [\iota x_i (x) \wedge \mathbf{3}(x)]$

c. $[[NP_i \mathbf{3} CL t_{NP-i}]] = \lambda x_i [[3 CL t_{NP-i}]]$ ($[[NP]]$)
 $= \lambda x_i \lambda x [\iota x_i (x) \wedge \mathbf{3}(x)]$ (${}^n\text{book}$)
 $= \lambda x [\iota x [\iota x_i (x) \wedge \mathbf{3}(x)]]$
 $= \iota x [\iota x [\iota x_i (x) \wedge \mathbf{3}(x)]]$

Here I take the moved NP to leave behind a trace of type $\langle e^k \rangle$, which is interpreted in the standard way as an indexed variable. At the level of DP, this variable is abstracted over and the NP meaning substituted for it via lambda conversion. The final step assumes type-shift via *iota*, conforming to the ranked system of type shifts discussed in section 1. This ensures that there be exactly three books in the context, otherwise *iota* will be undefined.

Finally, we look at the third option, which involves a demonstrative as well as NP raising. The derivation for this structure, I suggest, is as follows:

$$\begin{aligned}
\text{d. } * & \llbracket [\text{DP NP}_i \text{D}_{+\text{def}} [\text{Cardinality-P } 3 \text{ CL } t_{\text{NP-}i}]] \rrbracket \\
& = \llbracket \lambda x_i \llbracket [3\text{-CL } t_i] \rrbracket (\llbracket \text{NP} \rrbracket) \rrbracket \\
& = \lambda x_i \lambda Q \exists y [\text{u}_{x_i}(y) \wedge \mathbf{3}(y) \wedge Q(y)] (\text{book}) \\
& = \lambda Q \exists y [\text{u}^{\text{book}}(y) \wedge \mathbf{3}(y) \wedge Q(y)]
\end{aligned}$$

We can rule out (38a) by assuming that determiners and nouns are the only lexical categories that are specified +/-def. This means that DPs and NPs are the two phrasal categories with these features, so a numeral cannot raise to Spec of DP with $D_{+\text{def}}$ for feature checking. If such feature checking by a numeral were possible, we would get a predicative meaning that would then have to undergo covert type shift. Given the ranking in (7a), we would get a definite meaning via the *iota* type-shift, an incorrect result.

The possibilities in (38b-d) require some elaboration. In (38b) we have an $\text{NP}_{-\text{def}}$ in Spec of DP with $D_{-\text{def}}$ with a numeral interpreted as a predicate modifier. In (38c) we have an $\text{NP}_{-\text{def}}$ in Spec of DP with $D_{-\text{def}}$ with a numeral interpreted as a generalized quantifier. In (38d) we have an $\text{NP}_{+\text{def}}$ in Spec of DP with $D_{+\text{def}}$ with a numeral interpreted as a generalized quantifier. The meanings that would be derived by the compositional semantics in all three cases do not accord with intuitions: a structure with NP raising can never have an indefinite interpretation. Let us see how we can block these possibilities.

A simple way to rule out (38b-c), suggested by an anonymous reviewer, is to take nouns to be positively specified only for +def, so that NP raising correlates with checking the +def feature on D. However, (38d) still needs to be ruled out since NP raising here correlates with +def feature. Since the numeral is interpreted as a generalized quantifier, however, we end up with an indefinite reading after lambda conversion. This derivation can be ruled out by stipulating that only NP and DP can have argumental meanings and cardinals in intermediate functional structure must be interpreted as predicate modifiers or by appealing to the mismatch between the +def feature specification at the DP level and the indefinite semantics that the generalized quantifier produces.

Note that blocking generalized quantifier meanings in intermediate positions would also block the derivation in (38c), even if nouns could be specified –def and NP raising could check –def on D. Similarly, (38b) would be blocked in spite of NP raising with –def specification, because the \exists type shift would be outranked by *iota*. The result would be a definite meaning which would be incompatible with the –def feature of the DP. Thus we see that there is a principled way to derive the correct mapping from morpho-syntax to meaning.

Turning to a different point, recall the mysterious behavior of the cardinal *one* alluded to in section 2. While *one book* and *ek-ta boi* are both acceptable indefinite terms in English and Bangla respectively, the cardinal *one* is not allowed in the definite counterpart: **the one book* and **boi-ek-ta*. This, in fact, now has a straightforward explanation. Since English *book* and Bangla *ta boi* each denote a set of atoms, adding the numeral *one* as a predicate modifier inside the +def DP has no semantic impact

whatsoever. By *Avoid Structure* the extra syntactic projection within the DP is therefore ruled out.

Finally, note that the path of an NP raised to DP is made transparent by the presence of the classifier in Bangla. I should note, however, that there are some contexts in which a Bangla bare NP seems to have definite readings. Locative phrases, for example, seem to yield a definite reading relatively easily:

- 39a.** *tebal-e* *boi* *pore* *atʃe*
 table-on book lying is
 “Books are lying on the table.”
- b.** *pata* *maʃi-te* *portʃe*
 leaf ground-LOC be-falling
 “Leaves are falling on the ground.”

One might plausibly analyze *maʃi-te* ‘ground-LOC’ in (39b) as a contraction of *maʃi- ta-te* ‘ground-CL-LOC’ which would adhere to the generalization discussed above. One might also posit *tebal-ta-te* as a possible source for *tebal-e* ‘table-LOC’, though it is unclear why the full forms do not sound as natural. I leave the status of locative NPs open here, suggesting provisionally that they may involve NP → DP movement and that such movement is somehow obscured by the presence of the locative marker. One piece of evidence that supports this is the fact that the definite interpretation of the noun phrases is strictly singular, as is standard for structures with raised NPs and the classifier *-ta*. If the context is one in which reference is made to several equally salient tables, the plural classifier *tebal-gulo-e* ‘table-PL-Loc’ would need to be used, *tebal-e* would be undefined.¹⁶

Taking stock, we have successfully accounted for the fact that Bangla bare NPs denote kinds, with their associated weak indefinite readings at the object-level. We have also accounted for the fact that counting requires the mediation of classifiers. Assuming that the projection of classifiers entails the projection of the higher DP layer and the view that D comes with a +/- definite feature we have delivered the strong indefinite reading by positing an ambiguity in numeral expressions such that they can either function as predicate modifiers within cardinality phrases or as generalized quantifiers in D. We have also accounted for definite readings, which can arise in three different ways. One, through the lexical meaning of the demonstrative taking the predicative cardinality/classifier phrase as its argument, or through NP raising to spec of DP to value the +def feature on D and undergoing *iota* type shift, or through a combination of the two. We have also ruled out the possibility of definite readings for structures without NP movement, except when there is a demonstrative in D, and indefinite readings for

¹⁶ Anannya Dasgupta (p.c.) informs me that in colloquial Bangla it may be possible to use bare NPs with a definite reading, but there specific restrictions which bear further investigation.

structures with NP movement. In the next section, we will probe further into these issues, providing cross-linguistic arguments to either confirm or revise the particulars of the account given here.

4. BANGLA FROM A CROSS-LINGUISTIC PERSPECTIVE

It is typically the case that evidence from the morpho-syntax and semantics of a language under-determines the analysis. The facts of Bangla given in section 2, while consistent with the account in section 3, are also amenable to alternative analyses. In this section we will look at the Bangla facts through a cross-linguistic lens. Krifka (1995) provided an account of classifier languages like Mandarin Chinese in which the NP denotes a kind and the job of the classifier is to yield a predicate of objects. This was formalized in the *NMP* of Chierchia (1998), where all classifier languages, not just Mandarin, were classified as [+arg, -pred] languages. The key prediction is that bare nominals are expected to have kind and definite readings across the board in such languages. This is in contrast to [+/- arg, +pred] languages, in which the ability of a bare nominal to have definite readings depends on the presence or absence of lexical determiners. It is this particular aspect of NMP that we will focus our attention on below.

4.1. *Kind and Definite Readings in Classifier Languages*

Let us begin by reminding ourselves that in Chierchia's system classifying a language as [+arg, -pred] explains the absence of definite determiners in a language, while number neutrality in the nominal system explains the absence of plural morphology and the need for classifiers (cf. section 1.2). Implicit in the classification of a language as +arg is the view that NPs can be kind terms and that they can get definite readings through covert type shifts. One of the ways of repairing the sort mismatch between an object-level predicate and a kind-level argument, is to apply *pred* to the kind term, derive a number neutral property, and then apply *iota* covertly. Chierchia envisions the process in (40a), as does Yang (2001) in her analysis of Mandarin within Chierchia's system. Dayal (2011) and Trinh (2011), however, propose that a simpler way to derive the definite reading is to take the extension of the kind at the evaluation index, as shown in (40b):

40a. $\text{barking}(s)(^n\text{dogs}) \Rightarrow \text{barking}(\iota(^n\text{dogs}(s)))$

b. $\text{barking}(s)(^n\text{dogs}) \Rightarrow \text{barking}(^n\text{dogs}(s))$

The problem, of course, is that Bangla bare nominals do not behave as either version of (40) would predict. In fact, the Bangla facts are the same as was pointed out by Cheng and Sybesma (1999) for Cantonese (see also Li 2011). The further fact that Bangla puts on the table is the need for the definite reading to involve NP raising. Schematically, the following variation is attested:

41.		<i>Kind</i>	<i>Definite</i>
	<i>Mandarin</i>	NP	NP
	<i>Cantonese</i>	NP	CL NP
	<i>Bangla</i>	NP	[_{DP} NP _i [CL [t _i]]]
	<i>Yi</i>	NP	NP <u>or</u> [_{DP} Det [CL [NP]]]

In (41), I use the category NP descriptively, to indicate that there is no overt evidence of a higher structure when a bare nominal can denote a kind or have definite readings.

In this mix, the Tibeto-Burman language Yi adds an interesting twist. According to Jiang and Hu (2010) and Jiang (2011), there are two ways to derive a definite reading in Yi:

- 42a.** mu
Horse
“The horse/the horses”
- b.** mu ma su
horse CL the
“The horse”
- c.** mu sɔ ma su
horse three CL the
“The three horses”

As shown in (42a), the bare nominal can have a definite reading and when it does, it can refer to a singular or a plural individual, just as in Mandarin. Additionally, Yi has a lexical definite determiner *su*, which occurs above the Classifier Phrase and yields unambiguously singular (42b) or unambiguously plural (42c) definite readings. The obvious question this raises is the following: why does the lexical determiner in Yi not block the definite reading of the bare kind term, on analogy with English? According to Jiang and Hu, blocking does not apply in Yi because the two forms, with and without the classifier, have distinct definite interpretations. Once the classifier is projected, the resulting definite is either strictly singular or strictly plural. When there is no classifier, the definite is number-neutral. In English, the plural form is number neutral (see Zweig 2009, among others, for arguments) so the potential definite reading of a bare plural is blocked by the lexical option with an equivalent reading.

There is, however, a problem in extending this line of argumentation cross-linguistically. It would predict that Bangla too should allow the bare NP to have a definite reading. The bare NP would be ambiguous between singular and plural reference, while the NP raised structure would have the functionally distinct strictly singular/strictly plural reference. That this is not so, suggests that we need to continue looking for a theory for variation.

4.2. Number Marking vs Neutrality; Determiners vs Type-shift

Let us try a different tack to the problem. Let us assume that languages differ in whether number marking is encoded in NP. Languages like English, Hindi or Italian would be languages with number marking in NP, languages like Mandarin, Bangla, Cantonese and Yi would be languages where the NP is number neutral. Let us further assume that languages may have overt or null determiners for referring to kinds and for referring to unique/maximal object level individuals. This would be what is recognized as the definite determiner in languages like English or Italian. As discussed in Dayal (2004), if a language has a lexical determiner for *nom* it must have it also for *iota* but not vice versa. We now extend this idea to null determiners as well, so that a language may have a null determiner with the meaning of *iota* or *nom*. If it has it for *nom*, it must also have it for *iota* but not vice versa. This gives us the following typology:

43.	#-marking	<i>iota</i>	<i>nom</i>	
	√	D _{LEX}	D _{LEX}	<i>Italian</i>
	√	D _{LEX}	covert type-shift	<i>English</i>
	√	covert type-shift	covert type-shift	<i>Hindi</i>
	X	D _{NULL}	D _{NULL}	<i>Bangla</i>
	X	D _{LEX}	covert type-shift	<i>Yi</i>
	X	covert type-shift	covert type-shift	<i>Mandarin</i>

Let us go over (43) and see what this predicts for the structures listed. I use capitals in the following as a cover term for the particular language-specific lexical items:

44.	<i>Kind</i>	<i>Definite</i>	
	[DP THE [NP DOG]]	[DP THE [NP DOGS]]	<i>Italian</i>
	∩ _[NP DOGS]	[DP THE [NP DOGS]]	<i>English</i>
	∩ _[NP DOGS]	ι _[NP DOGS]	<i>Hindi</i>
	[DP [NP DOG] ∅ _{nom} t _i]	[DP [NP DOG] ∅ _{iota} [CL-P CL t _i]]	<i>Bangla</i>
	∩ _[NP DOG]	∩ _{[NP DOG](s)}	<i>Yi</i>
		[DP THE [CL-P CL DOG]]	<i>Yi</i>
	∩ _[NP DOG]	∩ _{[NP DOG](s)}	<i>Mandarin</i>

What we see in the above is that some languages are effectively [-arg, +pred], namely those with determiners for both *nom* and *iota*. But these now include not only Italian type number marking languages but also Bangla type number neutral languages. The only difference between them is that Bangla is treated as having null determiners. The

DP layer is projected in both, but for Bangla it is observationally indistinct from an NP in the case of kind reference. The presence of the classifier in the case of object level reference, however, tracks that movement. The facts of Bangla (and Cantonese) force us to reconsider the tight connection that Chierchia had posited between number neutrality and the ability of bare nominals to have definite readings.

There are, however, two questions that the present typology raises. One, why does the definite reading require the mediation of classifiers in Bangla? What rules out the possibility in (45a)? The second question is whether there could be a language like Bangla with an overt determiner?

- 45a.** $[_{DP} [DOG]_i \emptyset_{iota} [t_i]] = \iota[DOG]$
b. $[_{DP} THE_{nom/iota} [_{CL-P} CL [DOG]]]$

(45a), if it were possible, would have the *iota* operator apply to a number neutral property and yield a definite reading that would be ambiguous between singular and plural. (45b), if it were possible, would be the overt determiner counterpart of Bangla. It would be an obligatory determiner language like Italian, but for the fact that the bare nominal would denote a number neutral property so and would require classifiers. Such languages have not so far been attested.

I have barely scratched the surface of attested variation among languages with respect to kind vs. definite readings but the challenges inherent in the enterprise are pretty clear. It is relatively easy to come up with a plausible analysis for a language or even a group of languages but much harder to develop a theory that can hold up cross-linguistically, accounting for attested patterns as well as making explicit predictions about possible unattested patterns. Chierchia's *NMP*, clearly, did not make the right cut in this regard. However, it marked an important threshold in our understanding of the mapping between structure and meaning in the nominal domain. In responding to it, researchers have enriched the empirical landscape, thereby raising the bar for theoretical explanation. The current analysis of Bangla is to be seen as part of the ongoing search for an empirically adequate theory of cross-linguistic variation in the nominal domain.

BIOGRAPHICAL REFERENCES

- BARWISE, Jon & Robin COOPER 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4.
- BHATTACHARYA, Tanmoy 1999a. *Structure of the Bangla DP*. University College, London. PhD dissertation.
- 1999b. DP internal specificity in Bangla. *Yearbook of South Asian Languages and Linguistics* 2. 71-99.
- CARLSON, Gregory 1977. *Reference to Kinds in English*. University of Massachusetts, Amherst. PhD dissertation.

- CHENG, Lisa Lai-Shen, and Rint SYBESMA 1999. Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry*, 30.
- CHIERCHIA, Gennaro 1998. Reference to kinds across languages. *Natural Language Semantics* 6. 339–405.
- 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149.
- DASGUPTA, Probal 1983. The Bangla classifier /Ta/, its penumbra and definiteness. *Indian Linguistics* 44. 11-26.
- 1985. On Bangla nouns. *Indian Linguistics* 46. 37-65.
- 1992. Pronominality and deixis in Bangla. *Linguistic Analysis* 22. 61-77.
- DASGUPTA, Probal and Tanmoy BHATTACHARYA 1993. Classifiers and the Bangla DP. In Alice Davison and F.M. Smith (eds.). *Papers from the SALA Round Table Conference XV*. University of Iowa: Iowa. 59-69.
- DAYAL, Veneeta 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27.
- 2011. Bare noun phrases. In Klaus von Stechow, Claudia Maienborn, and Paul Portner (eds.). *Semantics: An International Handbook of Natural Language Meaning*, 33.2.
- to appear. Bangla plural classifiers. *Language and Linguistics*.
- in prep. *Bare Noun Phrases, Genericity and (In)definiteness: A Cross-Linguistic Perspective*. Wiley-Blackwell.
- DIESING, Molly 1992. *Indefinites*. Cambridge: MIT Press.
- DONNELLAN, Keith 1966. Reference and definite descriptions. *Philosophical Review* 75. 281-304.
- ENÇ, Mervet 1986. Tense Without Scope: an Analysis of Nouns as Indexicals. University of Southern California. PhD dissertation.
- FARKAS, Donka 1981. Quantifier scope and syntactic islands. In: *Papers from the Seventeenth Regional Meeting of the Chicago Linguistic Society*. Roberta Hendrick, Carrie Masek, and Mary F. Miller (eds.). 59–66. Chicago: Chicago Linguistic Society.

- FODOR, Janet Dean and Ivan SAG 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5. 355–398.
- GHOSH, Rajat. 2001. Some Aspects of Determiner Phrase in Bangla and Asamiya. Tezpur University. PhD dissertation.
- GHOSH, Sanjukta 2006. Honorificity-marking words of Bangla and Hindi: classifiers or not? *Bhashacintana* 1. 21-27.
- IONIN, Tanya & Ora MATUSHANSKY 2006. The composition of complex cardinals. *Journal of Semantics* 23. 315–360.
- JIANG, Julie L. 2011. Marking definite in classifier languages. Unpublished Harvard University ms.
- and Suhua HU 2010. An overt determiner in a classifier language, *Proceedings of GLOW-in-Asia VIII*, Beijing, China.
- KAPLAN, David 1989. Demonstratives. In J. Almog, J. Perry, & H. Wettstein (eds.). *Themes from Kaplan*, Oxford University Press.
- KRATZER, Angelika 1998. Scope or pseudoscope? are there wide-scope indefinites. In Susan Rothstein (ed.). *Events and Grammar*, Kluwer Academic Publishers.
- KRIFKA, Manfred 1995. Common nouns: a contrastive analysis of Chinese and English. In ed. Gregory Carlson & Francis J. Pelletier *The generic book*. Chicago: University of Chicago Press.
- LANDMAN, Fred 2004. *Indefinites and the Type of Sets*. Malden, MA: Blackwell Publishing.
- LI, Xu-Ping 2011. *On the Semantics of Classifiers in Chinese*. Bar-Ilan University. PhD dissertation.
- LONGOBARDI, Giuseppe 1994. Reference and proper names. *Linguistic Inquiry* 25. 609–665.
- MASICA, Colin 1976. *Defining a Linguistic Area: South Asia*. Chicago: University of Chicago Press.
- MONTAGUE, Richard 1974. *Formal Philosophy: Selected Papers of Richard Montague*, R.H. Thomason (ed.). New Haven: Yale University Press.

- PARTEE, Barbara 1986. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk et al. (eds.). *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Dordrecht: Foris.
- REINHART, Tanya 1997. Quantifier scope: how labor is divided between QR and choice functions. *Linguistics and Philosophy* 20. 335-397.
- ROBINSON, Heather 2005. *Unexpected (In)definiteness: Plural Generic Expressions in Romance*. Rutgers University. PhD dissertation.
- SWARZSCHILD, Roger 2002. Singleton indefinites. *Journal of Semantics* 19.
- SUTRADHAR, Joyshree 2006. Classifiers in South Asian languages: A typological survey. Unpublished Jawaharlal Nehru University ms.
- TRINH, Tue 2011. Nominal reference in two classifier languages. *Proceedings of Sinn & Bedeutung* 15.
- WINTER, Yoad 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20. 399–467.
- WOLTER, Linsey 2006. *That's That: The Semantics and Pragmatics of Demonstrative Noun Phrases*. UCSC. PhD dissertation.
- YANG, Rong 2001. *Common Nouns, Classifiers, and Quantification in Chinese*. Rutgers University. PhD dissertation.
- ZWEIG, Eytan 2009. Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32.353–407.

Department of Linguistics,
Rutgers University,
18 Seminary Place,
New Brunswick, NJ 08901,
U.S.A.
dayal@rci.rutgers.edu