

Bengali Referring Expressions¹

Dustin A. Chacón – University of Minnesota: Twin Cities, ILES
Cornell Undergraduate Linguistics Colloquium, 2008

1. The Givenness Hierarchy and Introduction

- We can refer to entities using multiple referring expressions, and these referring expressions may refer to multiple entities in turn.
- How can a listener correctly decide on a speaker's intended referent, and how can a speaker signal his referent?
- Givenness Hierarchy, or "GH," (Gundel, Hedberg, Zacharski, 1993) is a cognitively-based framework for describing and explaining the form and interpretation of referring expressions.
- The Givenness Hierarchy consists of six cognitive status categories, which restrict the listener in selecting a mental representation to associate with the referent. (1)

(1) *Givenness Hierarchy*

in focus >	activated >	familiar >	uniquely >	referential >	type >
it	IT, that, this, this N	that N	identifiable the N	this N INDEF	identifiable a N

- There is no simple one to one correlation between use of a given type of referring expression and a given status – how does the GH account for this?
 - Unidirectional entailment relationship built into definitions. Any higher status entails lower statuses. The categories are NOT mutually exclusive.
 - Gricean maxim of quantity (Grice, 1975) and scalar implicatures. Use of a referring form that explicitly signals a given status may implicate the referent does not have a higher status.
 - This allows the GH to be restrictive enough to be explanatory, without being too restrictive to represent actual distribution of referring expressions.
 - A correlation of a referring expression with a category on the GH means that that expression signals the referent has *at least* that status in the mind of the addressee, and may implicate that it is not a higher status. Any given status on the GH in turn can be signaled using the associated status, or with any lower forms, given the entailment relationship.

2. Bengali referring expressions

2.1 Morphosyntax

¹ Special thanks to Jeanette Gundel for her patient, consistent help and guidance in doing this project. Also, thanks to Aniqua Hasan, Samira Choudhury, and Nandita Rahman for their valuable input, and Ryan Johnson for his suggestions. This project was funded by the CLA Honors Research Scholarship Program at the University of Minnesota: Twin Cities. All remaining errors are my own.

- Bengali has the following elements from which the referring expressions are constructed:

Word for one: *ek* (allomorph: *æk*), often used as indefinite article.

Demonstratives: *e, o, se* (allomorph: *ta*)

- May function as pronouns, or as determiners after affixing *-i* to demonstrative. (Dasgupta, 1992)

- I refer to these in this paper using the deictic terms “proximal,” “medial,” and “distal”. Dasgupta (1992) describes *se* as a “sequential” demonstrative, because of its use in certain syntactic constructions (e.g., correlative clauses)

∅ pronoun, no distinct personal pronouns

- Classifier use can be relevant to signal Givenness.

- Relation of classifier to NP is sufficient for distinct readings in terms of “specificity”, or presuppositionality (Bhattacharya, 1999: 85-86). His examples are shown in (2) and (3).

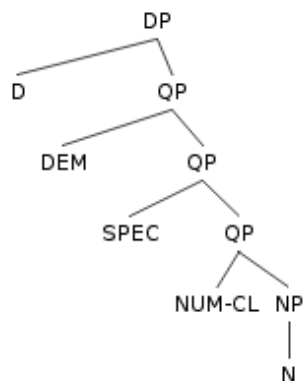
(2) *Ami du-to boi dekhechi*
 I two-CL book have.seen
 ‘I have seen two books’ [-specific]

(3) *Ami boi_i du-to t_i dekhechi*
 I book two-CL have.seen
 ‘I have seen the two books’ [+specific]

- CL licenses NP to become specified for [specific]. Movement of NP to [Spec, QP] allows a [+specific] feature to check with head Q (Numeral and classifier combined head).

- The final structure is shown in (4). (Bhattacharya, 1999)

(4) *Bengali DP model: (Bhattacharya, 1999)*



- CL may be used in an ‘antidefinite’ specific readings, according to Dasgupta (1983) These may refer to any kind of vegetable, not just lemons or chilis.:

- (5) *poʃʃir kache lebuʃa lɔnkaʃa cee newa*
 neighbor from lemon-CL chili-CL wanting take
 ‘borrowing lemons or chilis from a neighbor’

- CL may combine with *ek / æk* “one”, a bare nominal, a nominal preceded by a demonstrative, or a demonstrative pronoun.

2.2 Correlations with GH

- A children’s book and published personal diary were analyzed, as well as portions of an online forum, and a non-scholarly article on Asperger’s syndrome. Elicitations from native speakers were used to supplement the data. The relevant DPs were coded within the guidelines given in Gundel (revised, 2004). Referring expressions were correlated with the lowest status that had enough exemplars to be considered significant in relation to how many tokens were found in total for each expression. Exceptionally rare referring expressions were assigned a cognitive status only after several elicitations from native speakers in conjunction with discussion with said speakers.

- Examples were elicited from native speakers by replacing an already coded referring expression found in the corpus study with another referring expression, and then asking a native speaker to assess the grammaticality of the new sentence. An example of this is shown in (6). All context deemed necessary was also made accessible to the native speaker, often by reading the previous 5 to 10 sentences, depending on the referring expression in question. For example, in (6), the previous sentence had the referent in the object position of a dependent clause, and was read in conjunction with (5).

- (6) *ei kɔʃa ʃune kak* #er *khōj korte uʃe gælo.*
 this word hearing crow his finding to.do flying went
 (Activated) – originally bare nominal
 ‘Having heard these words, the crow flew away to find him.’

- In addition, either single sentence examples or paragraph examples with context were constructed and tested. Many of these were based off of examples from previous studies in the GH, or based on casual conversation and later coded. (7) is an example of a casual conversation-based elicitation.

- (7) *hæto* ∅ *sekhane* sei ganʃa *likhlo.*
 perhaps there that song-CL wrote
 (In-focus) – (Familiar) – not mentioned for 5 sentences
 Subject and
 topic of last 5
 sentences
 ‘Perhaps he wrote that song there.’

- The following correlations in (8) were found between the cognitive statuses on the GH and Bengali referring forms. The associated cognitive status is the lowest status available that the form may be used for, hypothesizing that that cognitive status is part of the conventional meaning of that referring form. The actual distributions of referring expression in relation to cognitive status are given in the chart in (9).

(8) *Correlations between GH and Bengali referring expressions*

in focus: \emptyset , $e(\text{ta})$ “this(-CL)”, $se(\text{ta})$ “that(-CL)” distal

activated: $o(\text{ta})$ “that(-CL)” medial, $ei\ N\text{ta}$ “this N-CL”

familiar: $oi\ N(\text{ta})$ “that N(-CL)”, $sei\ N(\text{ta})$ “that N(-CL)”

uniquely identifiable: $N\text{ta}$ “N-CL”

referential: $ei\ N$ “this N”

type identifiable: N , $ek(\text{ta})\ N$ “one(-CL) N”

(9) *Cognitive status and distribution of referring expressions in Bengali*

	In-focus	Activated	Familiar	Un. Id.	Ref.	Type Id.
\emptyset	16					
e	3					
$e\text{ta}$	4					
se	34					
$se\text{ta}$	2					
o	18	2				
$o\text{ta}$		2				
$ei\ N$	3	3	1	2	1	
$ei\ N\text{ta}$		1				
$oi\ N$		1	1			
$oi\ N\text{ta}$		1	1			
$sei\ N$	1	5	1			
$sei\ N\text{ta}$		1	1			
$N\text{ta}$		2	1	2		
N	13	36	26	55		14
$ek\ N$					2	1
$\text{æ}k\text{ta}\ N$					6	8

2.3 Comparisons with other Languages

- In Bengali, as do Chinese, Japanese, Russian, and Spanish, the null pronoun requires In-Focus status. Although not specifically built into the GH, there seems to be a correlation between referring expressions with minimal conceptual content encoding higher statuses in all languages.

- The proximal demonstrative determiner in Bengali, followed by a classifier, signals at least Activated status, whereas the other demonstrative determiners encode at least Familiar. Similar phenomena are seen in Chinese, English, Japanese, and Spanish. (Gundel et al, 1993)

- \emptyset , $e(ta)$, $se(ta)$ – the proximal and distal demonstrative pronouns – encode the In-focus status, as do personal pronouns in other languages. The medial demonstrative pronoun $o(ta)$ encodes the Activated cognitive status, as do other demonstrative pronouns in other languages (Gundel et al, 1993). (9) shows that $se(ta)$ and $e(ta)$ cannot be used with an activated cognitive status, whereas $o(ta)$ can.²

- (10) *abfese se porifranto hoe kacchop̃tir afae tholer kache phire.elo.*
 finally he exhausted became tortoise in.hope bag near returned
- kintu* \emptyset

<i>tholer</i> / bag's	<i>o(ta)r</i> ³ / that's	<i>#e(ta)r</i> / this's	<i>#sefar / tar</i> that's
--------------------------	--	----------------------------	-------------------------------

gae judhumatro ækta boro
 but he body.in only a big
 (Activated)
- gorto dekhlo*
 hole saw
 'Finally, he came back to the bag hoping for the tortoise, exhausted, but he only saw a hole in the body of the bag / this / that'

- Like English, Bengali has an 'indefinite this', a referring expression that minimally encodes the Referential cognitive status. In using a form that explicitly encodes this status, the speaker assumes the addressee will either access or construct a mental representation of the intended referent by the end of the sentence – thus, the intended referent of an "indefinite this" may be qualified later in the sentence to allow the listener to appropriately create or select the mental representation, as in (11). In Bengali, the indefinite this must not include a post-posed classifier, as seen in (12) and (13).

- (11) I was walking down the street the other day, and I ran into **this guy** who said he was your boss. Does he live somewhere by me?

- (12) *ami*

<i>ei</i> this	<i>lok(#ta)-ke</i> person(CL)-ACC
-------------------	--------------------------------------

dekhechilam ar se bolechilo je se
 I saw and he said that he
 (Referential)
- tomar boss chilo*
 your boss was
 'I saw this person, and he said he was your boss.'

² For a discussion of the differences between the three demonstrative pronouns in terms of Binding Theory, see Dasgupta (1992).

³ My native speaker reported that *ota* was slightly better here in that the classifier seemed to exclude a possible animate reading. In fact, all of my examples of demonstrative pronouns without classifiers were animate. However, there does not seem to be a correlation between inanimacy and classifier use in other cases.

- (13) *kicchukhən* *əpekha.kərar* *por* *tara*

<i>ei</i>	<i>afəŋka(#-ta)</i>
this	fear(-CL)

korlo
a.while waiting after they did
(Referential)
je *se* *nifcoei* *kon* *fikarir* *jale* *dhəra.pəteche*⁴
that he certainly some hunter's net.LOC was.caught
‘After waiting a while, they feared (“did this fear”) that he certainly fell into some
hunter’s net’

- Contrasting to the classifier-less indefinite ‘this’, the referent of sequence *ei Nta* “this N-CL” must be at least Activated. It is infelicitous in at most Familiar environments, as shown in (13). This further shows that there is a distinction between the sequence *ei N* “this N” and *ei Nta* “this N-CL”, since *ei Nta* seems to be unable to be used for referents below Activated, whereas *ei N* appears as Familiar, Uniquely Identifiable, and Referential. (The questionability of the *ei N* sequence in (14) is presumably due to the scalar implicature effect.)

- (14) (Mouse not mentioned for 10 sentences)
- | | | | | | | |
|------------|-----------------|-----------|---------------|----------------------|-----------------------|-------------------|
| <i>tai</i> | <i>taṭataṭi</i> | <i>se</i> | <i>īdur</i> / | <i>oi īdur(ṭa)</i> / | <i>sei īdur(ṭa)</i> / | <i>?ei īdur</i> / |
| so | quickly | he | mouse | that mouse | that mouse | this mouse |
- (Familiar)
- | | | | | |
|-------------------|----------|------------------|--------------|------------------|
| <i>#ei īdurṭa</i> | <i>o</i> | <i>kəcchoper</i> | <i>kache</i> | <i>phire.elo</i> |
| this mouse-CL | and | tortoise | near | returned |
- ‘So, he quickly returned to that mouse and tortoise.’

- Bhattacharya (1999) referred to the semantic key involved in the movement that yields the sequence N-CL to be “specificity”. I presume that ‘specific’ means at or above Referential – the cognitive status that refers to a unique entity, and that assumes that the listener has a mental representation of it before the end of the sentence activated. Note that ‘specific’ and Referential do not necessarily imply ‘definite.’ However, all the instances of Referential cognitive status that were found involved either no classifier, or a pre-posed classifier (*ei N*, *ek N*, and *əkṭa N*). This shows that a NP may be specific without a classifier, and if there is a classifier, without NP movement to [Spec, QP]. All the instances of post-posed classifiers were above Uniquely Identifiable, the distribution associated with the pretheoretic term ‘definite’ (Gundel, Hedberg, Zacharski, 2001). In fact, *Nta* cannot be used in at most Referential environments, as shown in (15).

⁴ This example was from elicitation. I had an informant loosely translate the sentence given in (8). She gave both this form as well as *ek lok-ke* “a person-ACC.” She stated that the use of the indefinite pronoun was noticeably better. She also found (10) to be unusual, although it was taken from a corpus. It is possible that she may be losing this form in her speech, or that there is another semantic feature.

- (15) *jəkhon amar baccar bæes tin chilo, təkhn amra* *bondhu(#ta)r*
friend(CL)'s
when my baby's age three was, then we
(Referential)
baṛite chilam, ar amar bondhu amake boleche je...
house.LOC were, and my friend me have.said that...
'When my baby was three, we were at a friend of mine's house, and my friend told me..'

- Dasgupta (1983)'s "antidefinite" specific may be a specific case of 'generic definites' (Gundel et al, 2001). These are Uniquely Identifiable forms in which the listener can be expected to build or access a unique mental representation of the referent based on the information coded in the nominal. In other words, the addressee is able to access a unique mental representation of the different kinds of tourists in (16), distinguishing them from other classes.

- (16) Montreal has something to offer **the religious tourist** and **the night life tourist**, **the art connoisseur**, **the sports fan**, **the intrepid walker** and **the avid consumer**.

-I speculate that the post-posed classifier signals either definiteness, or [+specific] in addition to further semantic content. If such additional content were found, it may also explain the apparent free variation in the use of classifiers with demonstratives, apart from the "definite" and "indefinite this" cases.

- As mentioned before, it seems that the classifier has no effect in terms of cognitive status for referring expressions that signal cognitive statuses above Uniquely Identifiable.

3. Conclusion

- The correlations between cognitive status and Bengali referring forms were determined.

- Bengali proximal demonstrative pronouns and distal demonstrative pronouns pattern like personal pronouns in other languages, and the medial demonstrative pronoun patterns like demonstrative pronouns in other languages.

- Bhattacharya (1999)'s [specific] feature does not seem to be adequate to motivate a post-posed classifier in Bengali. It may be a step in the right direction, provided additional semantic data associated with the classifier is assessed. Ideally, this additional information should account for the supposed variability in higher statuses. Otherwise, the sequence NP-CL may indicate definiteness.

References

- Ariel, Mira. 1988. Referring and accessibility. *Journal of Linguistics* 24, 67-87.
- Bhattacharya, Tanmoy. 1999. Specificity in the Bangla DP. *Yearbook of South Asian Languages and Linguistics* 2, Rajendra Singh (ed.), 71-99. Sage Publication: London.
- Dasgupta, Probal. 1992. Pronominality and Deixis in Bangla. *Linguistic Analysis* 22.1-2, 61-77.
- . 1983. On the Bangla classifier *ta*, its penumbra, and definiteness. *Indian Linguistics* 44.1-4, 10-26.
- Grice, H. Paul. 1975. Logic and Conversation. *Syntax and Semantics* 3, Peter Cole and Jerry L. Morgan (eds.), 41-58. New York: Academic Press.
- Garrod, S.C., and A. J. Sanford. 1982. The mental representation of discourse in a focussed memory system: Implications for the interpretation of anaphoric noun phrases. *Journal of Semantics* 1, 21-41.
- Gundel, Jeanette K., Michael Hegarty and Kaja Borthen. 2003. Cognitive Status, Information Structure, and Pronominal Reference to Clausally Introduced Entities. *Journal of Logic, Language and Information* 12, 281-299. Netherlands: Kluwer Academic Publishers.
- Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 2004. Coding protocol for statuses on the Givenness Hierarchy. http://www.sfu.ca/~hedberg/Coding_for_Cognitive_Status.pdf
- . 2001. Definite descriptions and cognitive status in English: why accomodation is unnecessary. *English Language and Linguistics* 5.2, 273-295. Cambridge University Press.
- . 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69.2, 274-307
- Imam, Janahara. 1986. *Ekattorer Dinguli* [The Days of '71]. Dhaka: Shandhani Prakashani.
- Junghare, Indira. 1983. Markers of Definiteness in Indo-Aryan. *Indian Linguistics* 44.1-4, 43-53.
- Radice, William. 1994. *Teach Yourself Bengali*. Ohio: The Mc-Graw Hill Companies, Inc.
- Sahoo, Kalyanamalini. 1999. The Syntax of Definiteness in Oriya. *Yearbook of South Asian Languages and Linguistics* 2, Rajendra Singh (ed.), 101-117. London: Sage Publications.
- Sharma, Elizabeth. 1985. *Car bondhu* [The Four Friends]. Perivale: Tiger Books Unlimimted.
- Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. *The Oxford Handbook of Comparative Syntax*, Guglielmo Cinque and richard Kayne (eds.), 806-838. Oxford University Press: Oxford.