

# Familiar Definite Marking in Magahi

Aidan Sharma

Rutgers University

FASAL 14 at Stony Brook University

April 4-6, 2024



## Unique vs Anaphoric Definites

- Two main uses of definites – uniqueness and anaphoricity<sup>1</sup>

(1) John bought **a book** and a magazine. [Anaphoric]  
**The book** was expensive.

(2) *Context:* in a room with exactly one table [Immediate Situation]  
**The table** is covered with books.

(3) *Context:* in America. [Larger Situation]  
**The president** gave a speech last night.

(4) a. John bought **a book**. [Bridging]  
**The author** is French.

b. John's hands were freezing as he was **driving** down the street.  
**The steering wheel** was bitterly cold and he had forgotten his gloves.

<sup>1</sup>examples slightly modified from Schwarz 2013, 2019

# Unique vs Anaphoric Definites

- Schwarz (2009) notes a distinction between these two types of definites in German
- The unique/weak determiner can contract following a preposition but the anaphoric/strong determiner cannot:

(5) *Hans ging zum Haus*  
 Hans went **to\_the**<sub>weak</sub> house

‘Hans went to the house’.

(6) *Hans ging zu dem Haus*  
 Hans went **to the**<sub>strong</sub> house

‘Hans went to the house’.

## German Unique Definites

- *Context*: A cabinet meeting with one chancellor and several ministers<sup>2</sup>

(7) *In der Kabinettsitzung heute wird ein neuer Vorschlag **vam***  
 in the cabinet.meeting today is a new proposal **by\_the**<sub>weak</sub>  
*{Kanzlet / #Minister} erwartet*  
**chancellor/minister** expected

‘In today’s cabinet meeting, a new proposal by the chancellor/minister is expected’.

(8) *#In der Kabinettsitzung heute wird ein neuer Vorschlag **van dem***  
 in the cabinet.meeting today is a new proposal **by the**<sub>strong</sub>  
*Kanzlet erwartet*  
**chancellor** expected

Intended: ‘In today’s cabinet meeting, a new proposal by the chancellor is expected’.

<sup>2</sup>Schwarz 2019

# German Anaphoric Definites

- (9) *In der New York Bibliothek gibt es ein Buch über Topinambur.*  
 in the New York library exists EXPL a book about topinambur  
*Neulich war ich dort und habe #im / in dem Buch*  
 recently was I there and have in\_the<sub>weak</sub> / in the<sub>strong</sub> book  
*nach einer Antwort auf die Frage gesucht, ob man Topinambur*  
 for an answer to the question searched whether one topinambur  
*grillen kann.*  
 grill can

‘In the New York public library, there is a book about topinambur. Recently, I was there and searched in the book for an answer to the question of whether one can grill topinambur’.

# Syntax/Semantics of German Definites

- Schwarz claims the difference between the anaphoric and unique definites is the presence of an index in the former:

(10) a.  $[DP[\mathbf{the}_{weak} \ s] \ NP]$

b.  $[[\mathbf{the}_{weak}]^g = \lambda s_r \lambda P_{\langle e, st \rangle} . \iota x [P(x)(s_r)]$

(11) a.  $[DP \ i[[\mathbf{the}_{strong} \ s] \ NP]]$

b.  $[[\mathbf{the}_{strong}]^g = \lambda s_r \lambda P_{\langle e, st \rangle} \lambda y . \iota x [P(x)(s_r) \ \& \ x = y]$

# Unique and Anaphoric Definites in Other Languages

- Since Schwarz's (2009) initial proposal, differences between unique and anaphoric definites have been reported in Korean (Cho 2017), Thai (Jenks 2015), Lithuanian (Šereikaitė 2019), among others
- One of these languages, Akan, has an (overt) strong determiner **nɔ** but no (overt) weak counterpart (Arkoh & Matthewson 2013)<sup>3</sup>

<sup>3</sup>but see Bombi (2018), Owusu (2022) for other analyses of this pattern

# Akan Familiar Definite Article

- **nɔ** can be used as a **definite determiner**, a 3sg animate object pronoun, or dependent clause marker (A&M 2013:4)
- Its use as a definite determiner is shown here:

(12) a. *Mɔ̀tɔ̀-ɔ̀*                      *èkùtú* *bí*  
 1SG.SUBJ-buy-PAST orange REF

‘I bought a (certain) orange’.

b. *Èku*    *`tú*    *nɔ́* *yè*    *dèw* *dʒè*  
 orange **FAM** be nice so

‘The orange is/was so nice’.

# Akan Familiar Definite Article

- *nɔ* does not strictly require linguistic antecedence

*Context:* a parent talking to his/her spouse about their children

(13) *Mbofra nó wó dan nó mu*  
 children **FA** be room FAM in

‘The children are in the room’.

(A&M 2013:7, from Saah 1994:152)

## Akan Familiar Definite Article

- However, *nɔ* is incompatible with larger situation unique definites and these are expressed with bare nouns:

(14) *Kwámì nyá-à kràtàá fí-ì Ègyá krónkrón póp hó*  
 Kwame get-PAST letter from-PAST father holy pope there

‘Kwame got a letter from the holy father Pope’.

(15) *Ámstónj nyí nyímpá áà ó-dzí-ì kán tú-ù kó-ò*  
 Armstrong is person REL 3SG.SUBJ-eat-PAST first fly-PAST moon  
*òsirán dɔ̀*  
 top

‘Armstrong was the first person to fly to the moon’.

## Akan Familiar Definite Article

- Arkoh & Matthewson adopt Schwarz's semantics for the strong definite article in German, repeated here:

$$(16) \quad \llbracket n\bar{v} \rrbracket^g = \lambda s_r \lambda P_{\langle e, st \rangle} \lambda y. \iota x [P(x)(s_r) \ \& \ x = y]$$

- Akan, however, shows an interesting pattern in that it marks strong/anaphoric definites but not weak/unique ones
- I will show that Magahi (Indo-Aryan) also marks anaphoric, but not unique, definites and propose a syntax and semantics of familiar definite marking
- Magahi has around 12 million speakers, mostly in Bihar, India<sup>4</sup>
- Uncited data comes from 5 native speaker consultants from Bihar<sup>5</sup>, 3 currently residing in the US and 2 still in Bihar
  - Also, 10 native speaker respondents to an online survey on Google Forms asking for English to Magahi translations

<sup>4</sup>Verma 2006 from Grimes 1996

<sup>5</sup>elicitations were done in English with 4/5 consultants and in Hindi with the remaining consultant

# Magahi Familiar Definites

- Magahi has a morpheme *waa* that attaches to nouns to encode familiarity
- *waa* has allophones *iyaa* and *aa* generally after vowels and *maa* generally after nasals, but they are not totally in complementary distribution (Alok 2022):

- (17) a. *ghar-waa/\*iyaa/\*aa/\*maa*  
house-FM
- b. *kaado-iyaa/\*waa/\*aa/\*maa*  
mud-FM
- c. *sapana-maa/waa/\*iyaa/\*aa*  
dream-FM
- d. *am-maa/waa/iyaa/\*aa*  
mud-FM

- Nouns marked with *waa* can be interpreted as either singular or plural

## Magahi Indefinites, Kinds, and Generics

- First, it is important to show that Magahi indefinites and kinds are incompatible with *waa* and are expressed as bare nouns or numeral-classifier noun expressions

(18) *Hamra (e-go) kutta-(#waa) chahi*  
 I one-CL dog-FM want

‘I want a dog’.

[Indefinite]

(19) *Dainasor-(#waa) bilupt ho gelai*  
 dinosaur-FM extinct be went

‘Dinosaurs are extinct’.

(Alok 2012) [Kind]

(20) *chiRai-(#waa) uda hai*  
 bird-FM fly be

‘Birds fly’.

(Alok 2012) [Generic]

# Magahi Familiar Definites

- Like the German strong article, *waa* is required on definites with a linguistic antecedent. Bare nouns are disallowed.

(21) *Kal ham e-go kutta dekhaliai. (Uu) kutt-#(waa) baRi sundar*  
 yesterday I one-CL dog saw DEM dog-FM very beautiful  
*halai*  
 was

‘Yesterday I saw a dog. The dog was very beautiful.’

## Magahi Familiar Definites

- Again similar to *nɔ* in Akan, *waa* is licensed by a weaker definition of familiarity – by being perceptually visible to both speaker and addressee or being generally familiar to all participants in the discourse (cf. Roberts' 2003 notion of weak familiarity)

*Context:* Ram and Rakesh are outside and a cute dog comes up to them. Ram says to Rakesh:

(22) *Ham (ii) kutt-#(waa) ke dulare le chaho hie*  
 1SG DEM dog-waa ACC pet give want be

'I want to pet the dog.'

- Note that even in object position *waa* is obligatory

# Magahi Familiar Definites

*Context:* Ram and John are from the same town, which has a single hospital and everyone in the town is familiar with the hospital. John tells Ram he's not feeling well, and Ram replies:

(23) *Jaa aspatal-iyaa in dekhwaala*  
 go hospital-FM in examine

‘Go and get examined in the hospital’

(Alok 2022)

## Magahi Larger Situation Unique Definites

- Magahi larger situation unique definites are expressed with bare nouns. *waa* is disallowed.

(24) *Suuraj-(#waa) poorab meN uugo hai*  
 Sun-FM east in rise be

‘The sun rises in the east.’

(25) *America ke raaspati-(#yaa) paagal hai*  
 America GEN president-FM crazy is

‘America’s president is crazy.’

(26) *duniya ke sab se tej aamdmi-(#yaa) se milai ke mun hai*  
 world GEN all OBL smart man-FM OBL meet PRT want be

‘I want to meet the world’s smartest man.’

## *waa* and Honorificity

- *waa* is known to interact with honorificity in Magahi
- For instance, *waa* cannot be used on the proper name/title of a social superior.
  - If your younger brother's name is Ratan, you can call him Ratanmaa. But if your older brother's name is Ratan, you cannot.
- *waa* can be used to show disrespect in such an instance:

*Context:* Ram is socially superior to the speaker and addressee. However, the speaker had a fight with Ram and is angry with him. The speaker says:

(27) *Ram-maa aawait hai*  
 Ram-FM coming be.NHON

‘Ram is coming’. + The speaker disrespects Ram. (Alok 2022)

- There is a question then, whether the bare nouns used for larger situation unique definites are a result of a weak/strong definite distinction or an honorificity distinction<sup>6</sup>

---

<sup>6</sup>Thanks to an anonymous reviewer for pointing this out

## *waa* and Honorificity

- However, even in examples where the speaker is disrespecting the unique referent, *waa* is disallowed:

(28) *Pahla aadami(-#yaa) je chand par ud ke gelai buRbak hai*  
 first man-FM REL moon on fly PRT went idiot is

‘The first man who flew to the moon is an idiot’.

(29) *Obama ke baad ke halai raaspati? Uu raaspati(-#yaa) buRbak halai*  
 Obama GEN after who was president DEM president-FM idiot was

‘Who was the president after Obama? That president was an idiot’.

- However, in a scenario where the speaker is talking about multiple solar systems “*hamar suuraj-waa*” (our sun-FM) is possible.

## *waa* is not a Demonstrative

- In terms of determining the status of *waa*, besides the fact that it co-occurs with demonstratives, there is another good reason to think it is not a demonstrative. Compare the Magahi and English examples below with respect to Löbner's (1985) diagnostic:

(30) # *laik-waa sutit hai au laik-waa sutit hai*  
 boy-FM sleep be and boy-FM sleep be

Intended: 'That boy is sleeping and that boy is not sleeping'. (Alok 2012)

- (31) a. # *The boy is sleeping and the boy is not sleeping.*  
 b. *That boy is sleeping and that boy is not sleeping.*

## Semantics of *waa*

- My proposal assumes that *waa* has a presupposition of non-honorificity and focuses on its role as a familiar definite marker.
- Regarding the relationship of *waa* to both definiteness and nonhonorificity, a first pass could be the Taboo of Directness (Wang 2023) – essentially, the idea that unmarked semantic features (e.g. plural, 3rd person, indefiniteness) are used to convey honorificity.
- The semantics I'm using for *waa* is then the same as Schwarz's semantics for the German strong determiner with the additional presupposition of non-honorificity:

$$(32) \quad \llbracket \text{waa} \rrbracket^g = \lambda s_r \lambda P_{\langle e, st \rangle} \lambda y : \exists! x (P(x)(s_r) \ \& \ x = y). \iota x [P(x)(s_r) \ \& \ x = y]$$

## Syntax of *waa*

- My proposal is that *waa* is a definite determiner. Besides its association with definiteness, it shares another properties of determiners: it is not used in vocatives<sup>7</sup> (Longobardi 1994)

*Context:* You see a boy you know walking around during the day when he should be at school, so you call out to him to find out what's going on:

- (33) *arre laRaka-(#waa), enne aao to!*  
 hey boy-FM here come TOP  
 'Hey boy, come here!'

- Nouns marked with *waa* are also not used predicatively:

- (34) *toN laRaka-(#waa) ho*  
 you boy-FM are  
 'You are a boy'.

<sup>7</sup>This claim is less robust for proper names.

## Against *waa* as a Determiner

- Alok (2022) argues against calling *waa* a determiner by appealing to Chierchia's (1998) Blocking Principle:

### (35) **Blocking Principle**

For any type shifting operation  $\pi$  and any X:

\* $\iota(X)$  if there is a determiner D such that for any X in its domain  $D(X) = \iota(X)$ .

- Alok argues that the existence of bare noun definites such as the following example indicate that *waa* as a determiner would violate the Blocking Principle (BP).

### (36) *Ham kitaab paRhali*

I book read

'I read a/the book(s)

## Against *waa* as a Determiner

- However, the fact is *waa* has the semantics of a familiar/strong determiner. Therefore, the existence of bare noun definites does not violate the BP, as bare noun definites are used in contexts where the familiar determiner *waa* is not licensed.
- These are larger situation definites and definites that are not familiar to both speaker and addressee.
- Alok also gives the following examples and calls them indefinites with *waa*:

- (37) a. *chaar go kitab-waa*  
 four CL book-FM  
 ‘four of the books’
- b. *kuchh kitab-waa*  
 some book-waa  
 ‘some of the books’

- These constructions require an analysis, but I do not think this partitive expressions indicate *waa* is not a determiner.

## Syntax of *waa*


- There are other expressions involving *waa* and the numeral-classifier that will be helpful in determining the syntax involved (Kumar 2019)<sup>8</sup>:

(38) a. *kitab-waa tin-go de diha*  
 book-FM three-CL give give

‘Give me the three books’

b. \**tin-go kitab-waa de diha*  
 three-CL book-FM give give

Intended: ‘Give me the three books’.

<sup>8</sup>The aggregative suffix is preferred to the classifier here (with no raising). 

## Bangla and Odia Definites

- This looks similar to the raising seen in Bangla (Dayal 2012) and Odia (Das 2022) definites:

(39) Bangla (Dayal 2012)

a. *boi Ta*  
book CL

‘the book’

b. \**Ta boi*  
CL book

c. *ek Ta boi*  
one CL book

‘A/One book’

d. \**boi ek Ta*  
book one CL

# Syntax of *waa*

- I'm adopting the analysis of Jenks & Konate (2022)<sup>9</sup>, and propose that a [+anaphoric] feature on *waa* attracts the NP to its Spec.

---

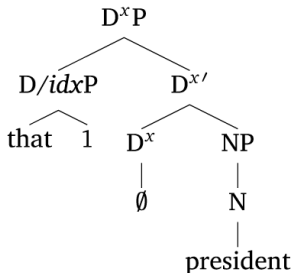
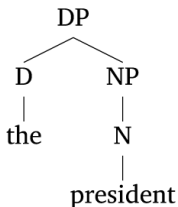
*Plain/unique definite*

---

*Indexed definite*

---

Syntax



Semantics  $[[DP]]^g = \iota x.[\text{Pres}(x,s)]$

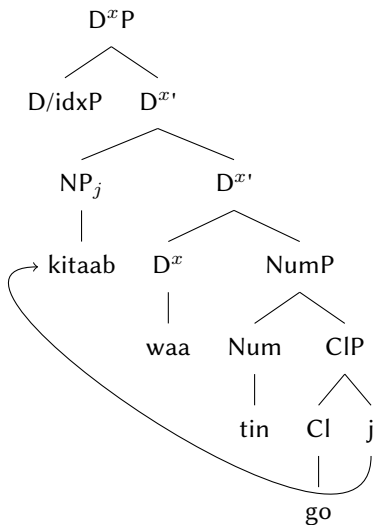
$[[D^xP]]^g = \iota x.[\text{Pres}(x,s) \wedge x = g(1)]$

---

<sup>9</sup>For an alternative see Hanink (2021)

Syntax of *waa*

(40)



## Conclusion and Future Prospects

- This derives the correct word order, but more work on indefinite and kind expressions and the uses of numeral, classifiers, and bare nouns needs to be done to determine the entire compositional semantics of the Magahi noun phrase.
- The aggregative marker on numerals should be explored further.
- Plural marking is also possible in Magahi, but I have avoided discussion here because the reported data is messy and inconsistent.
- Attributive adjectives have not been mentioned, as they have different forms in definite expressions with *waa* as compared to attributive adjectives that modify bare nouns. A suffix *-kaa/-kii* shows up on the adjectives in DPs with *waa*, and it has been claimed to be a type of determiner spreading<sup>10</sup>
- A complete analysis of the Magahi data should also account for the (non-)honorificity and proper name data.

---

<sup>10</sup>Kumar 2022

## Conclusion and Future Prospects

- Ahn (2017) extends the typology to include exophoric definites based on evidence from Korean, Romanian, and American Sign Language:

### Definiteness spectrum

Semantics	[uniqueness]	[familiarity]	[exophoric]
English	<i>the</i> [unique, anaphoric]	<i>that</i> [anaphoric, deictic]	
German	<i>the<sub>w</sub></i> [unique]	<i>thes</i> [anaphoric]	demonstratives [deictic]
Korean	bare	<i>ku</i>	<i>ce</i>
Romanian	bare, def	short dem	long dem
ASL	bare	IX <sub>locus</sub>	IX <sub>ref</sub>

- Where exactly Magahi fits into the typology of unique/anaphoric definite marking can also be further explored – currently it seems to use bare nouns for uniqueness, *waa* for anaphoricity, and demonstratives for deixis, but the co-occurrence of *waa* and the demonstrative has not been extensively studied

# Acknowledgements

## Thank You!

I would like to thank my consultants for their time and patience, as well as Maria Kouneli, the other members of her Fall 2023 syntax seminar, Deepak Alok, Sreekar Ragotham, Shannon Bryant, Indira Das, Dorothy Ahn, and Chandan Kumar for feedback and advice on earlier versions of this work.

# References

- Ahn, D. (2017). Definite and demonstrative descriptions: a micro-typology. *Proceedings of GLOW in Asia XI*, 1, 33-48.
- Arkoh, R., Matthewson, L. (2013). A familiar definite article in Akan. *Lingua*, 123, 1-30.
- Alok, D. (2012). A language without article: The case of Magahi. Unpublished M.phil thesis. Jawaharlal Nehru University.
- Alok, D. (2022). The Discourse Particle-waa in Magahi Nominal Domain. *Language Diversity, Linguistic Theory, and Cultural Identity in South Asia*. Yale University. April 29-30.
- Bombi, C. (2018). Definiteness in Akan: Familiarity and uniqueness revisited. In *Semantics and Linguistic Theory* (pp. 141-160).
- Cho, J. (2017). The acquisition of different types of definite noun phrases in L2-English. *International Journal of Bilingualism*, 21(3), 367-382.
- Das, I. (2022). Unique and Anaphoric Definites in Odia. Ms. Rutgers University.
- Dayal, V. (2012). Bangla classifiers: Mediating between kinds and objects. *Rivista di Linguistica*, 24(2), 195-226.

# References

- Grimes, Barbara F. (1996) *Ethnologue*, Houston: Summer Institute of Linguistics.
- Hanink, E.A. (2021). DP structure and internally headed relatives in Washo. *Nat Lang Linguist Theory* 39, 505–554. <https://doi.org/10.1007/s11049-020-09482-y>
- Jenks, P. & Konate, R., (2022) “Indexed definiteness” , *Glossa: a journal of general linguistics* 7(1). doi: <https://doi.org/10.16995/glossa.5717>
- Jenks, P. (2015). Two kinds of definites in numeral classifier languages. *Proceedings of SALT* 25. 103–124.
- Kumar, C. (2022). Multiple Determiners in Magahi: A Case Beyond Agreement. *Indian Journal of Languages and Linguistics*.
- Kumar, C. (2019). Typology of Nominal modifiers in Major languages of Bihar-Magahi, Bhojpuri, Maithili, Angika and Bajjika. [Unpublished Doctoral Dissertation]. Jawaharlal Nehru University, India.
- Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and philosophy*, 26, 287-350.
- Saah, K. (1994). *Studies in Akan syntax, acquisition, and sentence processing*. PhD Dissertation, University of Ottawa.

# References

- Schwarz, F. (2009). Two types of definites in natural language. University of Massachusetts Amherst.
- Schwarz, F. (2013). Two kinds of definites cross-linguistically. *Language and linguistics compass*, 7(10), 534-559.
- Schwarz, F. (2019). Weak vs. strong definite articles: Meaning and form across languages. In Ana Aguilar-Guevara, Julia Pozas Loyo & Violeta Vázquez-Rojas Maldonado (eds.), *Definiteness across languages*, 1–37. Berlin: Language Science Press.
- Šereikaitė, M. (2019). Strong vs. weak definites: Evidence from Lithuanian adjectives. In Ana Aguilar-Guevara, Julia Pozas Loyo & Violeta Vázquez-Rojas Maldonado (eds.), *Definiteness across languages*, 83–111. Berlin: Language Science Press.
- Owusu, A. P. (2022). *Cross-Categorial Definiteness/Familiarity* (Doctoral dissertation, Rutgers The State University of New Jersey, School of Graduate Studies).
- Verma, S. (2007). Magahi. In *The Indo-Aryan Languages* (pp. 547-565). Routledge.
- Wang, R. (2023). Honorifics without [hon]. *Natural Language & Linguistic Theory*, 41(3), 1287-1347.