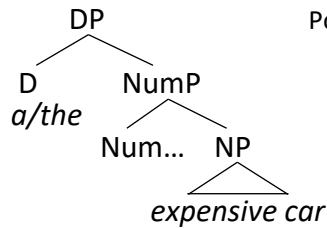


Extraction and the size of Xhosa nominal expressions

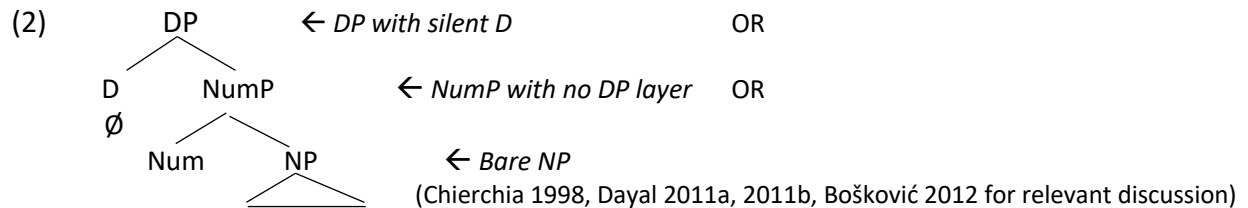
Vicki Carstens
University of Connecticut

1. Introduction

(1) DP Point of departure (Abney 1987 on DP; Ritter 1991, Carstens 1991 on NumP)



Absent articles, what's the size and structure of a nominal expression? Some hypotheses:



My claim: Xhosa nominals are DPs headed (a) by an affixal, dummy D or (b) otherwise, a null D.

2. Two kinds of evidence that nominals without articles are smaller

Semantic: gaps in available readings for nominals without articles argue that D is absent.
In particular, indefinite readings often rely on special strategies.

(3) ek laRkaa aur ek laRkii kamre meN aaye. laRkii baith-gayii. [Hindi; Dayal 2017:87]
one boy and one girl room in came. girl sat-down
'A boy and a girl came into the room. The girl sat down.'

Syntactic: nominals without articles are porous for extraction – attributed to smaller size.

*Left branch extraction: OK in Serbo-Croatian; *in English*

(4) Skupa/Ta_i je video [t_i kola]. [Bošković 2012:180]
Expensive/that is seen car
'That car/an expensive car is seen'

(5) a. *Expensive is seen car.
b. *That is seen car.

*Postnominal adjunct extraction: OK in Serbo-Croatian; *in English*

(6) Iz kojeg grada_i je Petar sreo [djevojke t_i]? [Bošković 2012:180]
from which city is Peter met girls
'From which city did Peter meet girls?'

(7) a. Peter met [_{DP} (the) girls from this city]
b. *From which city_i did Peter meet [_{DP} (the) girls t_i]?

3 Testing Xhosa

3.1 Ambiguity of nominals on semantic tests

Xhosa shows systematic ambiguity on tests of (in-)definiteness.¹

- (8) I-nkwenkwe nentombi ba-ngene e-gumb-ini. [Xhosa]
 AUG-9boy and.AUG-9girl 2SM-entered LOC-room-LOC.
I-n-tombi y-a-xelela i-nkwenkwe ukuba i-hlal-e phantsi.
 AUG-9-girl 9SM-ASP-tell AUG-9-boy that 9SM-sit-SBJ down

‘A boy and a girl entered the room. The girl told the boy to sit down.’

- (9) Lo m-fundi a-ka-z’u-ku-qhuba kakuhle kuba khange
 1this 1-student 1SM-NEG-AUG-15-perform well because NEG
 a-wu-gqib-e u-m-buzo.
 1SM-3OM-finish-PST AUG-3-question

Literally: ‘This student won’t do well because s/he didn’t finish question.’ OK: ∃>Neg

3.2 As in English, extraction of a post-nominal adjunct fails

- (10) a. Ndi-v-e [i-n-goma y-a-phesheya]. concord-bearing locative is a nominal modifier
 1SSM-hear-PST AUG-9-song 9-of-overseas
 ‘I heard [a song from overseas].’
 b. *Y-a-phesheya_i ndi-v-e [i-n-goma ____i].
 9-of-overseas 1SSM-hear-PST AUG-9-song
 Intended: From overseas I heard [a song ___].
 c. Ndi-v-e i-n-goma phesheya. concordless locative is a vP-modifier
 1SSM-hear-PST AUG-9-song overseas
 ‘I heard [a song] [overseas].’
 d. Phesheya_i ndi-v-e i-n-goma ____i.
 overseas 1SSM-hear-PST AUG-9-song
 ‘Overseas I heard a song.’

The nouns in examples above bear both an inner noun class prefix identified by #1-10 and an outer one, a.k.a. the *augment* (glossed AUG). *Augmentless* nominals function as strict negative dependents and post-verbal *wh*-words:

- (11) a. U-bon-e ∅-bani? b. A-ndi-bon-anga ∅-m-ntu!
 2SSM-see-PST 1who NEG-1SSM-see-NEG.PST 1-person
 ‘Who did you see?’ ‘I didn’t see anybody!’

Though augmentless nominals are morphologically barer/smaller in their lack of the outer prefix, this does not seem to have syntactic size correlates for the nominals they head in that these are no less opaque to movement.

¹ In glosses, AUG = augment; SM = subject marker; OM = object marker; SBJ = subjunctive; DISJ = disjoint, a VP-final verb form; REL = relative marker. Numerals indicate noun class (number+gender). Subject agreement in person/number is 1sSM/2sSM or 1PISM/2PISM (first or second singular or plural) as opposed to e.g. 1SM = SM in noun class 1. [+/-A] occasionally abbreviates augmented/less.

*Fronting of post-nominal adjunct from augmentless negative dependent

- (12) a. A-ndi-v-anga [∅-n-goma y-a-phesheya].
 NEG-1SSM-hear-NEG 9-song 9-of-overseas
 'I didn't hear [any song from overseas].'
- b. *Y-a-phesheya_i a-ndi-v-anga [∅-n-goma ___i].
 9-of-overseas NEG-1SSM-hear-NEG 9-song
 Intended: From overseas I didn't hear [any song __].

*Fronting of post-nominal adjunct from a wh-phrase with augmentless head

- (13) a. U-v-e [∅-yi-phi i-n-goma y-a-phesheya]?
 2SSM-hear-PST 9-which AUG-9-song 9-of-overseas
 'Which song from overseas did you hear?'
- b. *Y-aphesheya_i u-v-e [(i)-yi-phi i-n-goma ___i].
 9-of-overseas 2SSM-hear-PST AUG-9-which AUG-9-song
 Intended: From overseas_i which song did you hear ___i?

3.3 Extraction of AP fails (= another type of post-nominal adjunct in Xhosa)

- (14) a. U-Sabebelo w-a-bona [i-n-tombi e-n-tle] *Augmented*
 AUG-1Sabelo 1SM-PST-see AUG-9girl REL-9beautiful
 'Sabelo saw a beautiful girl.'
- b. *E-ntle_i U-Sabebelo w-a-bona [i-n-tombi ___i]
 REL-9beautiful AUG-1Sabelo 1SM-PST-see AUG-9girl
 Intended: Sabelo saw a beautiful girl.
- (15) a. U-Sabebelo a-ka-bon-anga [∅-n-tombi (e)-n-tle]. *Augmentless*
 AUG-1Sabelo NEG-1SM-see-NEG 9-girl REL-9-beautiful
 'Sabelo didn't see a(ny) beautiful girl.'
- b. *(E)-ntle_i U-Sabebelo a-ka-bon-anga [∅-n-tombi ___i]
 AUG.REL-9beautiful AUG-1Sabelo NEG-1SM-see-NEG 9-girl
 Intended: Sabelo didn't see a(ny) beautiful girl.

3.4 Left branch extraction fails

(16) Demonstratives may be pre- or post-nominal

- a. Ndi-fund-e [i-n-cwadi le].
 1SSM-read-PST AUG-9-book 9this
 'I read this book.'
- b. Ndi-fund-e [le i-n-cwadi].
 1SSM-read-PST 9this AUG-9-book
 'I read this book.'

(17) Demonstratives may not front

*Le ndi-fund-e i-n-cwadi.
9this 1SSM-read-PST AUG-9-book
Intended: I read this book

Absent a demonstrative, the noun is at the left edge of the nominal. It may not front, either.

(18) a. Nd-a-bona [a-ma-doda a-ma-de]
1SSM-PST-see AUG-6-men REL-6-tall
'I saw tall men.'

b. *A-ma-doda_i nd-a-bona [___i a-ma-de]
AUG-6-men 1SSM-PST-see REL-6-tall
Intended: I saw tall men.

(19) *∅-i-bhokisi a-ndi-vul-anga [___i (e-zin-kulu)]
10-boxes NEG-1SSM-buy-NEG.PST REL-10-big
Intended: I didn't open (big) boxes.

4. Adding an object marker: demonstratives extractible²

Object marker (OM) agrees in noun class with an augmented internal argument:

(20) a. Ndi-**yi**-fund-ile i-n-cwadi.
1SSM-9OM-read-DISJ AUG-9-book
'I read the book.'

b. Nd-a-**ba**-bona a-ba-ntwana.
1SSM-PST-2OM-see AUG-2-children
'I saw (the) children.'

(Augmentless nominals may not be agreed with, so I ignore them henceforth.)

(21) Licit fronting of a demonstrative with object marking on the verb

a. Le ndi-**yi**-fund-ile [t_{DEM} i-n-cwadi].
9this 1SSM-9OM-read-PST AUG-9-book
Literally: this I read book.

b. Lo ndi-**m**-bona [t_{DEM} u-m-twana].
1this 1SSM-1OM-read-PST AUG-1-child
Literally: this I saw child.

² In related Chichewa, OM has been proposed to liberalize the possibilities for fronting nouns and modifiers from an object nominal (Mchombo 2004, 2006, Branen & Davis 2019, 2022).

(22) A fronted demonstrative may be construed with a nominal in an embedded clause.

Le_i ndi-cinga [ukuba [ndi-**yi**-fund-ile [____i i-n-cwadi]]].
 9this 1SSM-think that [1SSM-9OM-read-PST AUG-9-book
 Literally: this I think that I read [__ book].

(23) The dependency is island-sensitive, supporting a movement analysis.

*Lo_i ndi-hamb-e [ndi-nga-**m**-bon-anga [____i u-m-ntwana]].
 1this 1SSM-leave-PST 1SSM-without-1OM-see-NEG AUG-1-child
 Literally: this I left without seeing [__ child].

But a noun may not front from the left edge of a nominal, despite object-marking.

(24) a. I-in-yosi z-a-luma [a-ba-zingeli a-ba-de]
 AUG-10-bees 10SM-PST-sting AUG-2-hunters AUG-2-tall
 'Bees stung the tall hunters.'

b. *A-ba-zingeli_i i-in-yosi z-a-(**ba**)-luma [____i a-ba-de]
 AUG-2-hunters AUG-10-bees 10SM-PST-sting AUG-2-tall
 Literally: Hunters bees stung [__ tall]

Adjectives and locative adjuncts cannot front from nominals despite object-marking

(25) a. *A-ba-de_i i-in-yosi z-a-(**ba**)-luma [a-ba-zingeli ____i]
 AUG-2-tall AUG-10-bees 10SM-PST-bite AUG-2-hunters
 Literally: Tall bees stung [__ hunters]

b. *Y-a-phesheya ndi-(**yi**)-v-e [i-n-goma ____i].
 9-of-overseas 1SSM-9OM-hear-PST AUG-9-song
 Intended: From overseas I heard [a/the song __].

5 Towards an analysis: why only demonstratives, and why is OM crucial?

I build on Bošković's approach to the English/Serbo-Croatian contrast.

(26) Skupa/Ta_i je video [t_i kola]. [Serbo-Croatian; Bošković 2012:180]
 Expensive/that is seen car
 'That car/an expensive car is seen'

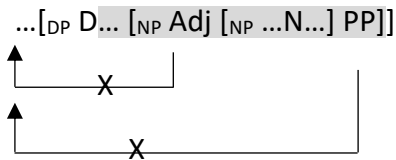
(27) a. *Expensive is seen car.
 b. *That is seen car.

(28) Phase Impenetrability Condition: In a phase α with head H, the domain of H is not accessible to operations outside α ; only H and its edge are accessible to such operations (Chomsky 2000:108).

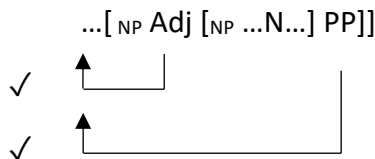
(29) [_{α} YP [H [_{β} ...WP...ZP...]]] Assuming H is a phase head, nothing can escape β

(30) [_{DP} YP [D [_{XP} ...WP...ZP...]]] Assuming D is a phase head = H in (28)/(29), nothing can escape XP

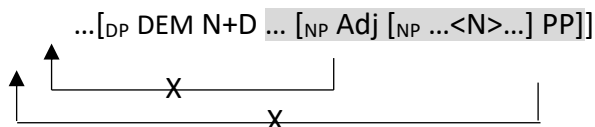
(31) Phasal DP blocks extraction from its interior (English) (Bošković 2012)



(32) Bare NP is porous (Serbo-Croatian)³



(33) Comparing Xhosa nominal dependents for accessibility



As for the role of OM, I adopt a proposal of Van Urk & Richards (2015) based on Dinka extraction phenomena: agreement of a probe P with a goal G renders the edge (and only the edge) of G transparent to other relations with P.

(34) a. $X_{u\Phi_i}$ [_{phase} YP [H [_{α} ...WP...ZP...]]]

Agree ($X_{u\Phi_i}$, Phase)

b. $X_{u\Phi_i}$ [_{phase} YP [H [_{α} ...WP...ZP...]]]

Phasal category values $u\Phi_i$ of X; YP at phase edge becomes accessible to X

It follows that the Xhosa head which bears object marking with an internal argument can probe and raise a demonstrative out of it, to its edge; from there the demonstrative proceeds.

(35) a. $v_{u\Phi_i}$ [_{DP} DEM [D [_{α} ...WP...ZP...]]]

DP values $u\Phi_i$ of v ; DEM at DP's edge becomes accessible to v

b. [DEM $v_{u\Phi_i}$ [_{DP} <DEM> [D ...]]]

DEM moves to the edge of $v\Phi$ and onwards

Summing up, owing to Phase Impenetrability and the unlocking role of agreement, an internal argument that is object-marked is the only possible source of demonstrative extraction, and only demonstratives may extract.

The account of the facts relies crucially on the assumption that Xhosa nominals are phasal DPs.

³ Bošković (2012) proposes that NP is a phase in Serbo-Croatian because it is the highest category in the extended projection of the noun, but modifiers adjoined to NP can escape it because not all segments of NP dominate them.

6. What heads Xhosa DP? Closing thoughts and questions

If Xhosa nominals are DPs, what is the head D? Following Visser (2008), De Dreu (2008), Adams (2010), Gambarage (2019): the augment is a D. Carstens (to appear a): augmentless nominals are headed by null D.⁴

- (36) a. (u-)m-fazi
AUG-1-woman
'(a/the) woman'
- b. $\begin{array}{c} \text{DP} \\ \swarrow \quad \searrow \\ \text{D} \quad \dots \text{NP} \\ \text{AUGMENT or} \\ \emptyset \end{array}$

Semantic content? A recent proposal: Gambarage (2019), Gambarage & Matthewson (2022) show that Nata augments similarly do not encode (in-)definiteness or specificity and argue that they are a kind of D which instead encodes speaker's belief in the existence of the nominal's referent.

- (37) **O**-mu-gáruka a-nga-βonek-ire Maria n-a-nga-tjɔmir-u. [Nata]
AUG-1-elder 1SM-COND-show.up-PST Maria FOC-1SM-COND-be.happy.FV
'If an elder showed up, Maria would be happy'

Accepted in context: A mother has a sick child and only elderly people know the traditional cure for the disease. There is a specific elder who knows the medicine for the disease. The mother says she will be happy if that elder shows up.

Rejected in context: A mother has a sick child and only elderly people know the traditional cure for the disease. She would be happy if any elder comes, but that's impossible, because there are no elders in this community.

- (38) \emptyset -mu-gáruka a-nga-βonek-ire Maria n-a-nga-tjɔmir-u.
1-elder 1SM-COND-show.up-PST Maria FOC-1SM-COND-be.happy.FV
'If any elder showed up Maria would be happy'

Accepted in context: A mother has a sick child and only elderly people know the traditional cure for the disease. She would be happy if any elder comes, but that's impossible, because there are no elders in this community.

Rejected in context: A mother has a sick child and only elderly people know the traditional cure for the disease. There is a specific elder who knows the medicine for the disease. The mother says she will be happy if that elder shows up

- (39) A-ka-βuy-a [ango \emptyset -mo-subhe/#o-mo-subhe a-ya-sek-a.
1SM-PST-say-FV that 1-man AUG-1-man 1SM-PST-laugh-FV
'S/he said that a man laughed' (I doubt it)

Gambarage (2019) and Gambarage & Matthewson (2022) claim Xhosa augments also encode speaker's belief in existence of the referent. However, there are differences between Nata and Xhosa that they do not address.

⁴ Swahili lacks augments yet its nominals are syntactically opaque and ambiguous for (in-)definiteness; see Carstens (to appear b). Carstens (1991) proposes a null D analysis.

(40) Ukuba ebenakubona *(u-)m-ntu u-m-dala, u-Mary a-nga-vuya.
 COMP 1SM.COND.see AUG-1-person AUG-1-elder AUG-1-Mary 1SM.COND-be.happy
 'If she could see an(y) elder, Maria would be happy'
 (Augment obligatory whether or not there are elders in the community)

(41) W-a-tsho (ukuba) ku-hlek-e *(u-)m-ntu.
 1SM-PST.say that 17SM-laugh-PST AUG-1-man
 'S/he said that a man laughed' (augment obligatory regardless of speaker's doubt)

Carstens (to appear a): the Xhosa null D may surface only where licensed via binding by negation or interrogative C_Q of *wh*-questions.⁵

(11) a. U-bon-e ∅-bani? b. A-ndi-bon-anga ∅-m-ntu!
 2SSM-see-PST 1who NEG-1SSM-see-NEG.PST 1-person
 'Who did you see?' 'I didn't see anybody!'

(42) Licensing null D of augmentless nominals in Xhosa

a. Neg... [DP ∅ ...] b. C_Q... [DP ∅ ...]

Otherwise the distribution reflects positional restrictions, including *[-A] preverbally

(43) a. *∅-bani u-fik-ile? b. *∅-m-ntu u-fik-ile /a-ka-fik-ile.
 1who 1SM-arrive-DISJ.PST 1-person 1SM-arrive-DISJ.PST/NEG-1SM-arrive-DISJ.PST
 Intended: Who arrived? Intended: Nobody arrived.

c. Ku-fik-e ∅-bani? d. A-ku-fik-anga ∅-m-ntu
 17SM-arrive-CONJ.PST 1who NEG-17SM-arrive-NEG 1-person
 'Who arrived?' 'Nobody arrived.'

(44) a. U-fun-a ukuba [*(u-)bani a-phek-e]?
 2SSM-want-FV COMP 1who 1SM-come-SBJ
 'Who do you want to cook?'

b. A-ndi-fun-i ukuba [*(u-)m-ntu a-phek-e].
 NEG-1SSM-want-NEG COMP AUG-1-person 1SM-come-SBJ
 'I don't want anyone to cook.'

The Xhosa augment

- No discernible meaning
- Obligatory wherever the null D is prohibited
- Conclusion: it is a dummy determiner, needed to fulfill a requirement that arguments be DPs in Xhosa (Stowell 1989, Longobardi 1994 among others, for discussion of such requirements).

The open question: how are definite and indefinite readings derived, in Xhosa? Both are available, hence if by typeshifting, no ranking effects ($\iota > \exists$) as in Dayal's account of Hindi.

For future research!

⁵ No speakers in this study accept [-A] in conditionals regardless of context. Of 6 speakers interviewed for Carstens & Mletshe (2016), 2 accepted them in conditionals such as 'If you see anybody, tell me' and in polar questions like 'Did you read any books this weekend?'. No relationship was apparent to belief in the existence of a referent.

References

- Abney, Steven. 1987. The English noun phrase in its sentential aspect. MIT PhD dissertation.
- Bošković, Željko. 2012. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories*, ed by G. Grewendorf and T. E. Zimmermann, 179-242.
- Branan, Kenyon & Davis, Colin. 2019. Agreement and unlocking at the edge. *Proceedings of the 2019 LSA*.
- Branan, Kenyon & Davis, Colin. 2022. Edges and extraction: Evidence from Chichewa. In *Descriptive and theoretical approaches to African linguistics*, p.149-164.
- Carstens, Vicki. 1991. The morphology and syntax of Determiner Phrases in Kiswahili. UCLA PhD dissertation.
- Carstens, Vicki. To appear a. On the size and category of Bantu nominal expressions. In Achiri-Taboh, B. (ed.), *The Bantu noun phrase: issues and perspectives*. London: Routledge.
- Carstens, Vicki. To appear b. Nguni bare nouns: licensing without Case. Forthcoming in *Linguistic Inquiry*. https://doi.org/10.1162/ling_a_00487
- Chierchia, Genaro. 1998. Reference to kinds across language. *Natural Language Semantics* 6.4, pp. 339-405.
- Chomsky, Noam. 2000. Minimalist Inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (eds.) *Step by Step*, 89-155. Cambridge, MA: The MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz, ed., *Ken Hale: A life in language*, 1-52, Cambridge, MA: The MIT Press.
- Dayal, Veneeta. 2011a. Hindi pseudo-incorporation. *Natural Language and Linguistic Theory* 29.1: 123-167.
- Dayal, Veneeta. 2011b. Bare noun phrases. *Semantics: an international handbook of natural language meaning*. Berlin: Walter de Gruyter, pp.1087-1108.
- De Dreu, Merijn. 2008. The internal structure of the Zulu DP. University of Leiden MA thesis.
- Gambarage, J.J. 2019. Belief-of-existence determiners: evidence from the syntax and semantics of Nata augments. UBC PhD dissertation.
- Gambarage, J.J. and L. Mathewson. 2022. The Bantu-Salish connection in determiner semantics. *Glossa: a journal of general linguistics* 7.1: <https://doi.org/10.16995/glossa.7685>
- Mchombo, Sam. 2004. *The syntax of Chichewa*. Cambridge University Press.
- Mchombo, Sam. 2006. Linear order constraints on split NPS in Chichewa. *ZAS Papers in Linguistics*, 43, pp.143-160.
- Ritter, Elizabeth. 1991. Two functional categories in noun phrases: evidence from Modern Hebrew. *Perspectives on phrase structure: heads and licensing*. Leiden: Brill Publications.
- Stowell, Timothy. 1989. Subjects, specifiers and X-bar theory. In Baltin, Mark & Kroch, Anthony (eds.), *Alternative Conceptions of Phrase Structure*, 232–262. Chicago and London: University of Chicago Press.
- Van Urk, Koppe & Norvin Richards. 2015. Two components of long-distance extraction: successive cyclic movement in Dinka. *Linguistic Inquiry* 46.1: 113-155.
- Visser, Marianna. Definiteness and specificity in the isiXhosa determiner phrase. *South African Journal of African Languages* 28.1: 11-29.