Predication and NP Structure in an Omnipredicative Language: The Case of Khoekhoe

Michael Hahn

2014

1 The Phenomenon

Launey (1994, 2002) has proposed the concept of *omnipredicativity*, describing languages where members of all major open word classes can function equally and without derivation as predicates, and in which the predicative use is primary and referential use is derived syntactically by relativization. Omnipredicativity is different from the lack of word classes: in Classical Nahuatl, the language studied by Launey, nouns and verbs are clearly distinguished by their morphological properties, but on the syntactic level, nouns and verbs can both be used predicatively and referentially in the same ways.

While there has been interest in the analysis of nonverbal predication in HPSG (e.g., Bender, 2001, Henri & Abeillé, 2007), to our knowledge, predication in languages exhibiting omnipredicativity has not been addressed in HPSG. In this paper, we examine noun phrases and predication in Khoekhoe, a Khoisan language spoken in Namibia and South Africa. We show that it exhibits features typical of 'omnipredicative' languages and present a formal HPSG analysis, in which members of all open word classes enter the syntax as predicates and in which all argument NPs are derived in a uniform manner as projections of pronominal elements, modified by relative clauses. Despite the radical differences between Khoekhoe and European languages in the relevant areas, our analysis will crucially build on standard components of HPSG analyses. No special rule licensing predicative use of nouns is required, and referential use will be derived based on Sag's (1997) analysis of English relative clauses.

In Khoekhoe, there are three open word classes: verbs, nouns, and adjectives. They are clearly distinguished, first, by the derivation morphemes applicable to them: only verbs and adjectives allow valencey-changing suffixes (passive, reflexive, reciprocal, applicative, pronominal object markers), while only nouns can form diminutives. Second, only nouns can have inherent gender. Third, adjectives and nouns have a fixed order within NPs: adjectives can modify nouns, but not the other way around. Nonetheless, the three classes show striking similarities in their syntactic behavior.

Khoekhoe is an SOV language. The V slot may be occupied by a word from any of the three open word classes: a verb (1a), an adjective (1b), or a noun (1c-d). Both commons nouns (1c) and proper nouns (1d) can be used. Even deictic elements can act as predicates (1e-f). While the choice of the TAM marker depends on the predicate, the syntactic behaviour of the different predicates is entirely parallel.

	a.	saa=ts	ge	ra	khii	b.	om=s	ge	a	kai
		you=2MS	DECL	TAM	come		house=3FS	DECL	TAM	big
		'You are con	ning.'				'The house ¹	is big.'		
	c.	saa=ts	ge	a	gao-ao	d.	saa=ts	ge	a	Petru
(1)		you=2MS	DECL	TAM	king		you=2MS	DECL	TAM	Petrus
		'You are a ki	ing.'			'You are Petrus.'				
	e.	om=s	ge	а	nee	f.	tii=ta	ge	а	saa
		house=3FS	DECL	TAM	this		I=1s	DECL	TAM	you
	'The house is this one.'					'I am you.'				

¹Khoekhoe NPs are not marked for definiteness and the choice of definiteness in the translations is arbitrary.

Unlike languages like Russian and Arabic, it is not possible to simply analyze these clauses as copulative structures. The crucial point is that the predicative element in (1c-f) is not an NP as it would occur in an argument position. In Khoekhoe, argument NPs generally end with a person-gender-number (PGN) morpheme (=b in 2a-b, =ts, =ta and =s in 1), which is not found when a noun is used as a predicate. Compare the NP gao-ao=b in (2) with the predicative noun gao-ao in (1c):

	a.	[gao-ao=b]	ge	ra	khii	b.	tii=ta	ge	[gao-ao=b]	!oa	ra	mîî
(2)		king=3MS	DECL	TAM	come		1s=1s	DECL	king=3MS	to	TAM	speak
		'The king is	coming.	,			ʻI am sp	eaking to	o the king.'			

On the other hand, there is a close parallelism between argument NPs headed by a noun and free relative clauses functioning as arguments. Fre relative clauses consist of a clause containing a gap (or a resumptive pronun) and a final PGN morpheme indicating the index features of the referent:

	a.	[khii	ra]=b	b.	[mûû=ta	ra]=b
(3)		come	tam=3ms		see=1s	tam=3ms
		'the on	e (m.) who is coming'		'the one (m.) I am seeing'

Formally, a major difference between the relative clauses in (3) and the argument NPs in (2) seems to be that the NPs do not contain a TAM marker. But this just reflects a usage preference: an NP may carry a TAM marker (*gao-ao* a=b king TAM=3MS 'the king', compare this with (1b)) and a verbal predicate may occur without a TAM marker (*saa=ts ge |khii* 'you come', cf. (1a)), but these options are dispreferred, possibly because nouns generally denote permanent properties for which TAM marking within an argument NP would add no information.

Argument NPs headed by a noun may also be marked for negation in a way completely parallel to predicates:

	a.	∥îî=b	ge	a	[Petru	tama]	b. Petru	tama=b
(4)		3s=3мs	DECL	TAM	Petrus	NEG	Petrus	NEG=3MS
		'He is no	ot Petru	s.'			'the or	he who is not Petrus'

Nonverbal predicates are also subject to the same word order alternations as verbal predicates. In particular, both may be fronted to the position immediately in front of the subject, which then can only be expressed by an enclitic PGN marker:

	a.	khii=ts	ge	ra	b.	gao-ao=ts	ge	a	c.	gao-ao=b	ge	a
(5)		come=2MS	DECL	TAM		king=2MS	DECL	TAM		king=3MS	DECL	TAM
		'You are co	ming.'			'You are a	king.'			'He is a kin	ıg.'	

Since the TAM marker *a* and the declarative clause type marker *ge* are optional in clauses like (5b-c), this has the consequence that expressions that look like noun phrases, such as gao-ao=b in (5c), may constitute clauses. This is reminicent of the situation in Nahuatl, where a noun phrase may constitute an utterance, which Launey considers typical of 'omnipredicative' languages.

The parallelism extends to more complex NPs with modifiers, since nouns may form complex predicates with modifiers such as adjectives (6a) and relative clauses (7a):

(6)	a.	Petru=b Petrus=3MS 'Petrus is a f	ge DECL riendly	(a) TAM v perse	[khoe: friend on.'	xa khoe] ly persor	b	. khoexa friendly 'the frie	k p nd	hoe=t erson lly pei	son'					
	a.	Petru=b	ge	a	[‡an	tama=ta	hââ]	khoe	b.	[[‡an	tama=ta	hââ]	khoe=b]	ge	go	khii
(7)		Petrus=3MS	DECL	TAM	know	NEG=1S	TAM	person		know	NEG=1S	TAM	person	DECL	TAM	come
		'Petrus is a r	berson	that I	don't l	know.'				'A pe	rson that	I don'	t know ca	ime.'		

The same behaviour is exhibited by all members of closed word classes that may, when combined with a PGN marker, form an argument NP. For instance, argument NPs that have traditionally been regarded as personal pronouns such as *saa=ts* 'you (m.s.)' actually consist of a deictic element that may also function as a predicate (as in 1f), and a PGN marker. The only lexical elements that can function referentially but not as predicates are the PGN markers

themselves, which are also used as enclitic subject pronouns. Any more complex NP can be analyzed as consisting of a relative clause or a predicate and a PGN marker.² As this is exactly the structure of free relative clauses in Khoekhoe, we claim that all NPs (except for the bare PGN markers) are free relative clauses – as argued for Classical Nahuatl by Launey (Launey, 2002, 117). As Khoekhoe nouns are essentially predicates and phrases only become referential by the addition of PGN markers, we assume that the PGN marker always is the head, which is compatible with the general head-final word-order of Khoekhoe. We will henceforth refer to Khoekhoe argument 'NPs' as DPs.

2 Analysis

We assume that not only verbs, but also nouns, adjectives, and some other words including deictics enter the syntax as predicates, with a non-empty SUBJ list, and that any phrase with an empty COMPS list and a nonempty SUBJ list may combine with a subject DP in a *head-subj-phrase* to form a clause. For instance, we assume the following entries for the noun *khoe* 'person' and the adjectivee *khoexa* 'friendly':

(8)	Ē.	$\begin{bmatrix} SUBJ & \langle DP_i \rangle \end{bmatrix}$		$\begin{bmatrix} \text{SUBJ} & \left\langle DP_i \right\rangle \end{bmatrix}$
	CAT	COMPS $\langle \rangle$	CAT	COMPS $\langle \rangle$
		HEAD noun		HEAD <i>adjective</i>
		INCONT <i>person</i> '		[INCONT <i>friendly</i> '
	CONT	PARTS $\left\langle person', \alpha(x_i) \right\rangle$	CONT	PARTS $\left< friendly', \alpha(x_i) \right>$
	_ /	$\land \{ \alpha \succ person' \}$	- /	$\land \{ \alpha \succ friendly' \}$

where DP is an abbreviation for a saturated structure with HEAD *pgn-marker*. The left structure is very similar to those resulting from a lexical rule licensing predicative nouns in English assumed, for instance, by Ginzburg & Sag (2001), but nouns do not have corresponding non-predicative lexical entries in Khoekhoe.

The semantic contribution is formalized using the underspecified framework of Lexical Resource Semantics (LRS, Richter & Sailer, 2003). INCONT is the core semantic contribution, while PARTS contains all subterms of the overall semantics that are contributed by the word. Informally, the description states that *khoe* contributes the term *person*' and that some term containing *person*' is applied to x_i , the index of the subject. If the predicate consists only of the noun *khoe*, its overall semantic representation will be $person'(x_i)$, while it will be more complex if modifiers are added. It is not possible to simply state that the representation should be $person(x_i)$, as non-intersective modifiers may be added. We may note that exactly the same semantic analysis can be assumed in an LRS analysis for English, where the noun would carry the index *i* in its INDEX feature, not on its SUBJ list.

Complex nominal predicates While the predicative word classes behave in very similar ways when used as predicates, they show differences in their abilities to modify other elements. We assume that the choices are formalized by lexical rules. For instance, adjectives may have their MOD value set to the SYNSEM value of a noun stem so that adjectival modifiers can combine with nouns to form complex predicates like *khoexa khoe* in (6a). We assume the structure in (9) for modifying adjectives like *khoexa* 'friendly' in *khoexa khoe* 'friendly person'. If the predicate consists only of the adjective and the noun *khoe* 'person', α will have to be identified with *friendly*', β with *person*', and γ with *friendly*'(*person*'), so that the overall semantic representation will read *friendly*'(*person*', x_i), which is the standard Montagovian analysis for ' X_i is a friendly person.' (Montague, 1973). The semantic analysis is again the same as for English.

²There are two exceptions to this claim. First, two or more coreferent NPs with identical PGN marking can be serialized, yielding a single NP. Second, NPs may have a possessive marker consisting of ââ and the PGN marker adequate for the possessor, which follow the PGN marker of the NP. Both cases can be accounted for easily by adding two phrasal types, the first one recursively licensing NPs consisting of two coreferent NPs, and the second one licensing NPs consisting of an NP and a possessive marker.

$$(9) \begin{bmatrix} SUBJ & \langle \rangle & & \\ COMPS & \langle \rangle & & \\ & \begin{bmatrix} adjective & & \\ & & \begin{bmatrix} adjective & & \\ & & & \\ MOD & \begin{bmatrix} HEAD & noun \\ & SUBJ & & \langle DP_i \rangle \\ & CONT | INCONT & person' \end{bmatrix} \end{bmatrix} \land \begin{pmatrix} \alpha \succ friendly' \\ \beta \succ person' \\ \gamma \succ \alpha(\beta) \end{pmatrix}$$

$$CONT \begin{bmatrix} INCONT & friendly' \\ & PARTS & \langle friendly', \alpha(\beta)_{\langle et \rangle}, \gamma(x_i)_t \rangle \end{bmatrix}$$

To ensure that nouns form a complex predicate with modifiers and complements and to prevent other elements like the TAM marker from intervening between the elements of the complex predicate, we use the boolean-valued LEX feature, essentially the way it is used to enforce formation of verbal complexes (Müller, 2002). In this sense, complex nominal predicates have the same status as complex verbal predicates. The precise constraints will be discussed in the full paper.

Relative clauses and DPs There are two basic types of relative clauses: *modifying relative clauses*, which modify a noun and form with it a complex nominal predicate, and *independent relative clauses*, which modify a PGN marker and form with it an argument DP. As in Sag (1997)'s analysis of English, modifying relative clauses are clauses whose MOD value is the SYNSEM value of the projection of a noun, allowing them to combine with a noun to form a complex nominal predicate (10a). Similarly, predicates or clauses in free relative clauses select a PGN marker via MOD (10b):

(10)	mod-rel-cl]	indep-rel-cl		
	SLASH	$\langle DP_i, \rangle$	SLASH	$\langle DP_i, \ldots \rangle$	
		[HEAD noun]		HEAD	pgn-marker
	MOD		MOD	INDEX	i
	L	$\begin{bmatrix} SOBJ & DF_i \end{bmatrix}$		LEX	+

Thus, DPs are constructed from clauses or predicates by adjunction to a PGN marker via the ordinary *head-adjunct-phrase* type. Spurious ambiguities resulting from recursive application are prevented by the LEX feature. As there are PGN markers for all persons, the analysis correctly predicts the availability of non-first-person DPs, which Launey considers typical of the 'omnipredicative' type. An example is *sa* ||nao=ta 'I, your uncle' in (11):

(11) [sa ||nao=ta]=s ta $\pm \hat{u}\hat{u}$? your uncle=1S=2FS TAM eat

'Are you (trying to) eat me, your uncle?'

As adjunction is in principle optional, PGN markers may also form complete DPs by themselves, but independently required constraints on the LEX feature enforce that this is only possible in the subject position in sentences like those in (5), as shown in Hahn (2013).

PGN Markers On the semantic level, the PGN marker contributes a box in the sense of Discourse Representation Theory (DRT, Kamp & Reyle, 1993), which binds the variable representing the DP's referent. Its contents are filled in by the contributions of the other constituents. We assume the following entry for the first-person PGN marker =ts as in (1a):



Khoekhoe does not appear to have generalized quantifiers, as Jelinek (1995) has argued for Straits Salish. Numerals and elements like *huu* 'every' are actually predicates.

Example Analysis Figure 1 shows the analysis of (6a). For simplicity, we show a flat structure on the clause level, which may be replaced with one of the more complex structures that have been proposed for Khoekhoe clause structure (Washburn, 2001, den Besten, 2002, Hahn, 2013).

The head of the S node is the complex predicate *khoexa khoe* '(be a) friendly person', whose head is the noun *khoe* '(be a) person', modified by the adjective *khoexa* '(be) friendly'. The complex predicate subcategorizes for a subject, which is realized by the DP *Petru=b*. It consists of the predicative proper name *Petru* and the PGN morpheme *=b*, with which it forms a *head-adjunct-phrase*. As it is headed by the PGN morpheme, the syntactic parts of its SYNSEM information are inherited from it, with the exception of the LEX feature.

On the semantic level, the representation of the complex predicate is $friendly'(person, x_i)$, as discussed above. The DP contributes the DRT box and the entry $x_i = petrus$, which is the contribution of the predicate *Petru*. The resulting semantic representation for the sentence is then given by:



Figure 1: Analysis of (6a).

References

Bender, Emily M. 2001. Syntactic variation and linguistic competence: The case of AAVE copula absence: Stanford University dissertation. http://faculty.washington.edu/ebender/dissertation/.

- den Besten, Hans. 2002. Khoekhoe syntax and its implications for L2 acquisition of Dutch and Afrikaans. *Journal of Germanic Linguistics* 14.1. 3 56.
- Ginzburg, Jonathan & Ivan A. Sag. 2001. *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. Stanford, California: CSLI Publications. http://cslipublications.stanford.edu/site/1575862786.html.
- Hahn, Michael. 2013. Word Order Variation in Khoekhoe. In Stefan Müller (ed.), *Proceedings of HPSG 2013*, Stanford: CSLI Publications.
- Henri, Fabiola & Anne Abeillé. 2007. The syntax of Copular Construction in Mauritian. In Stefan Müller (ed.), Proceedings of HPSG 2007, 130–149. Stanford: CSLI Publications.
- Jelinek, Eloise. 1995. Quantification in Straits Salish. In E. Bach, E. Jelinek & B. H. Bartee (eds.), *Quantification in natural language*, vol. 54 Studies in Linguistics and Philosophy, 487–540. Springer.
- Kamp, Hans & Uwe Reyle. 1993. From Discourse to Logic. Dordrecht: Kluwer.
- Launey, Michel. 1994. Une grammaire omnipredicative: Essai sur la morphosyntaxe du nahuatl classique. Paris: CNRS Editions.
- Launey, Michel. 2002. Compound nouns vs. incorporation in Classical Nahuatl. Studia typologica 4. 113–134.
- Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. In K. J. J. Hintikka, J. M. E. Moravcsik & P. Suppes (eds.), *Approaches to natural language*, 221–242. Dordrecht: Reidel.
- Müller, Stefan. 2002. Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German (Studies in Constraint-Based Lexicalism 13). Stanford: CSLI Publications. http://hpsg.fu-berlin.de/ stefan/Pub/complex.html.
- Richter, Frank & Manfred Sailer. 2003. Basic Concepts of Lexical Resource Semantics. In Arnold Beckmann & Norbert Preining (eds.), *Esslli 2003 course material i*, vol. 5 Collegium Logicum, 87–143. Wien: Kurt Gödel Society.
- Sag, Ivan A. 1997. English relative clause constructions. Journal of Linguistics 431-484.
- Washburn, P. 2001. A Minimalist approach to Khoekhoe declaratives. Cornell Working Papers in Linguistics 18. 28-56.