

---

# Proper names, predicates, and the parts-of-speech system of Santali

FELIX RAU<sup>1</sup>

## 6.1 Introduction

The present paper gives preliminary observations on the syntax and semantics of proper names in the North Munda language Santali. Munda languages, such as Mundari, Santali, and Kharia, have been reported to exhibit a large degree of flexibility in their parts-of-speech systems. The focus of this paper is on proper names in predicate position as an extreme case of flexibility, because proper names are generally thought to be referring expressions that cannot easily be given a predicative reading. Given these characteristics, the combination of predication and proper names provides a good test case for the flexibility of lexemes and generality of this characteristic.

I will argue for an account of the parts-of-speech system of Santali emphasizing the flexibility of categories. Contrary to claims that this flexibility is restricted and displays semantic irregularities, I will try to show that in Santali flexibility is indeed regular and a general characteristic of the language, and that it is carried to extremes that defy derivational explanations. By focusing on a small phenomenon—instead of analysing the whole parts-of-speech system of Santali—I hope to contribute to the discussion on the parts-of-speech systems in Munda languages that has been revived in recent years (cf. Evans and Osada 2005; Peterson 2005; Hengeveld and Rijkhoff 2005), broaden the empirical base, and thus enhance our understanding of Santali and hopefully the Munda languages in general.

The claims I make are about Santali only, but I will often contrast my findings with the statements made by Evans and Osada (2005) about Mundari. I believe that what

<sup>1</sup> I would like to thank Eva van Lier, Jan Rijkhoff, Arlo Griffiths, and the two anonymous reviewers for their very helpful comments.

is said here could have some bearing on the analysis of Mundari and probably other Munda languages. The general direction of my analysis has been argued for by Peterson (2005, *this volume*). It is also relevant to note that not all Munda languages seem to display such a flexibility in their parts-of-speech systems, and most South Munda languages such as Sora, Gorum, and Gutob—with which I am much more familiar—are quite different in this domain and display much less flexibility than the North Munda languages and Kharia.

The statements made in this paper are based mainly on data from the texts of Bodding (1926, 1927, 1929b). In the grammatical analysis of Santali I will generally follow Neukom and make it explicit if I deviate from his analysis.<sup>2</sup> I also use a slightly modified version of his spelling conventions for the examples taken from Bodding.<sup>3</sup>

## 6.2 Previous accounts

The parts-of-speech systems of Munda languages (especially the North Munda languages Santali and Mundari) have been subject to several analyses. There are two main lines of argument: the first focuses on the variety of syntactic contexts a given lexeme can occur in and emphasizes the general validity of this property. This approach is taken by majority of scholars, beginning with Hoffmann for Mundari and Bodding for Santali through to Hengeveld (1992b) and Hengeveld and Rijkhoff (2005) for Mundari, Neukom (2001) for Santali, and Peterson (2005) for Kharia. The second approach, mainly represented by Evans and Osada (2005), focuses on inequalities in the distribution of lexemes, and argues for a more ordinary parts-of-speech system with nouns, verbs, and adjectives. This approach is in fact also taken by most of the lexicographic work on North Munda languages (e.g. Bodding 1929–1936).

### 6.2.1 *The flexibility analysis*

The traditional analysis regards the North Munda languages Mundari and Santali as clear cases of flexible languages with only one major word class, in which every lexeme can occur in every syntactic position (see Hengeveld et al. (2004) for a comprehensive typology of parts-of-speech systems).<sup>4</sup> The traditional approach is perhaps best outlined by Hoffmann (1903):

'Instead, then, of Parts of Speech with well-defined functions and a precise but rich denotative and connotative power, we meet in Mundari with words of great functional elasticity, and therefore of a vague signifying power—words which, whilst *denoting* living beings, actions,

<sup>2</sup> Several of the sentences I have taken from Bodding (1926, 1927, 1929b) are also used in Neukom (2001). In these cases, I have only cited the original source.

<sup>3</sup> I use the IPA characters *ɔ* and *ɨ* instead of the indological style *o* and *i* to write the retroflex phonemes.

<sup>4</sup> Only Bhat (1997) regards Munda languages as omnipredicative and thus as rigid languages with only one word class.

qualities, and relations, do generally not by themselves *connote* the manner in which the mind conceives the things signified. That *connotation* is generally left to the context of the proposition or the circumstances under which it is uttered; [...]' (Hoffmann 1903: xx–xxi; original emphasis)

An opposing analysis is that taken by Evans and Osada (2005), which I concentrate on in Section 6.2.2. This analysis is unique in its data coverage of a specific Munda language (Mundari) and rejects the flexibility analysis.

### 6.2.2 *The analysis of Evans and Osada (2005)*

Evans and Osada (2005) propose an analysis of the Mundari parts-of-speech system which claims 'that Mundari clearly distinguishes nouns from verbs, though (like English, Chinese, and many other languages) it has widespread zero conversion' (Evans and Osada 2005: 384). They posit three criteria as requisites for a flexibility analysis: equivalent combinatorics of all lexemes; compositionality of semantics in all functions; and bidirectionality of flexibility (Evans and Osada 2005: 366). In their view, what seems to be flexibility of categories in Mundari does not meet their criteria and is in fact conversion or zero derivation. Unfortunately they do not specify what exactly they understand by conversion, but as a lexical process, it should apply to lexemes only, and its semantic effect should be lexeme-specific and not regular. This last claim is stated explicitly in Evans and Osada (2005: 374). Another property of a lexical derivational process such as conversion would be that there is no requirement for it to apply to all lexemes. In fact, Evans and Osada (2005: 384) claim that the process does not extend over the whole lexicon and is thus not general, but they estimate half of the lexicon to be subject to the process of conversion.

Peterson (2005) and Hengeveld and Rijkhoff (2005) have argued convincingly against this analysis, although they differ in what they accept as requisites for flexibility. Most importantly, Hengeveld and Rijkhoff (2005) reject the requirement of semantic compositionality. However, the claims made by Evans and Osada (2005) that flexibility is not a general property of the lexemes and that many alleged examples of flexibility show non-compositional lexeme-specific idiosyncrasies still stands.

## 6.3 Arguments and predicates in Santali

In this section, I will outline two aspects of the grammar of Santali relevant to the present purpose, before going on to the actual focus of this paper, i.e. predication and proper names. After introducing the main predicate position of a Santali sentence and some of its syntactic and morphological characteristics, I will sketch the behaviour of arguments and adjuncts in Santali syntax. For a more comprehensive account of the grammar of Santali see Neukom (2001), Ghosh (2008), and Bodding (1929a).

6.3.1 *Main predicate position*

A sentence in Santali can contain several predicates, but there can only be one main predicate. This main predicate has a unique morphological device that identifies it: the so-called indicative suffix *-a*; see Neukom (2001: 145).<sup>5</sup> Other verbal morphology, such as TAM suffixes and person markers, is not confined to the main predicate position, but may occur on all predicates.

The main predicate position in Santali syntax can be best delimited by two morphosyntactic clues: the aforementioned indicative suffix *-a*, which occurs at the right edge of the predicate position, and the subject clitic, which according to Neukom (2001: 113) is ‘normally attached to the word that immediately precedes the verb’ but can also follow the indicative suffix *-a* in some situations. The wording ‘immediately precedes the verb’ is somewhat unfortunate, since there are several examples in which the subject clitic does not precede the lexeme one would like to call a verb, but rather the predicate-constituting syntagma that is predicated, as in (1). The clitic precedes the whole ArgP that forms the main predicate in this sentence.

- (1) *ape-pe* [mǝŕẽ hɔɾ-a]  
you(PL)-2PL.S five people-IND  
‘You are five (people).’ (Bodding 1929b: 350)

Sentences such as (1) show how the main predicate is morphosyntactically delimited. The left boundary is marked by the subject clitic, which attaches to the material preceding the predicate, while the right boundary is situated directly after the indicative marker. Hence, the structure of the main predicate in a sentence can be represented as in (2).

- (2) NON\_PREDICATED\_PART-SUBJ [MAIN\_PREDICATE-IND]

The subject clitic may also stand in a different position, directly after the indicative marker *-a*. This is the case in at least two circumstances. First, being an enclitic, the subject clitic obviously requires some material to its left to which it can attach. If no material precedes the predicate, the subject clitic is attached to the right of the predicate. A single-lexeme sentence as in (3) is a typical example of this behaviour.

- (3) *dal-et<sup>2</sup>-kan-a-e*  
strike-IPFV.ACT-IPFV-IND-3SG.S  
‘He is striking.’ (Neukom 2001: 64)

<sup>5</sup> This is only true for some sentence types. Most non-asserted sentences such as imperatives and presupposed clauses lack this marker. Sentences of this kind will not be discussed in this paper. However, subordinated clauses, which do not take this marker either, will be briefly discussed together with arguments.

A second, more complicated, case—which at closer inspection may nevertheless turn out to be quite similar to the first—is the rare situation in which more than one lexeme constitutes the sentence, while the subject clitic, contrary to expectation, follows the indicative marker, as in (4). In some cases this might indicate that the sentence should be regarded asthetic, which again would result in a situation where there is no material preceding the main predicate. Yet in other cases this seems unlikely and the motivation for the placement of the subject clitic in sentence-final position remains unclear.

- (4) *ale-dɔ*                    *lelha*   *bhucun*   *koŋka*   *Bhūiə*   *kan-a-le*  
 we(PL.EXCL)-TOP   stupid   ignorant   foolish   Bhuya   COP-IND-1PL.EXCL.S  
 ‘We are foolish, stupid, witless Bhuyas.’ (Bodding 1929b: 350; cf. Neukom 2001: 114)

Virtually every lexeme can form a main predicate in Santali. Besides event-denoting lexemes such as *dal* ‘to strike’ in (3) above, entity- and property-denoting lexemes such as *raj* ‘king’ and *maran* ‘big’ in (5) and (6), respectively, can also be used as predicates. The main predicate does not even have to be an actual lexeme, as is the case with the onomatopoeic *ãã* in (7).

- (5) *adɔ-e*                    *raj-en-a*  
 then-3SG.S   king-PST.MV-IND  
 ‘So he became king.’ (Bodding 1929b: 8)

- (6) *bəjun-dɔ-e*                    *maran-a*  
 Bajun-TOP-3SG.S   big-IND  
 ‘Bajun is big (senior).’ (Bodding 1927: 2 [translation FR])

- (7) *bar*   *pɛ*                    *dhao-e*                    *ãã-y-en-a*  
 two   three   time-3SG.S   groan-y-PST.MV-IND  
 ‘It (i.e. the buffalo) groaned two or three times.’ (Neukom 2001: 15)

Furthermore, the main predicate can consist of non-lexical units such as phrases. The postpositional phrase *kombɾo tuluj* ‘with thieves’ in (8) is an example of this kind of main predicate. The point that phrasal and other complex units can function as predicates in Munda languages has been made before (Peterson 2005: 396f.), and it is in itself a very good argument against analysing the flexibility as a lexical derivational process.

- (8) *alo-m*                    *kombɾo*   *tuluj-ok'-a*  
 PROH-2SG.S   thief   with-MV-IND  
 ‘Don’t keep company with thieves.’ (Neukom 2001: 15)

The common property of all these predicates is that they always get a predicative interpretation and are never referential. Also, main predicates are always asserted and never presupposed. These semantic characteristics of the main-predicate position are crucial for the present argument.

6.3.2 *Argument position*

The argument positions of a sentence in Santali are occupied by argument phrases (Neukom 2000: 18), such as the definite argument phrase in (9). Argument phrases (ArgP) consist at least of a head element—such as *koṛa gidrə* ‘boy child’—which may be preceded by modifiers, such as the reduplicated *huḍiṇ* ‘small’ in (10), and adnominal demonstratives, such as *uni* in both of the sentences below.

- (9) *uni koṛa gidrə-dɔ*  
 that(AN) boy child-TOP  
 ‘that boy (child)’ (Bodding 1929b: 86)

- (10) *uni huḍiṇ huḍiṇ gidrə-dɔ*  
 that(AN) small small child-TOP  
 ‘that very small child’ (Bodding 1929b: 84)

Predicates and whole clauses can also be part of an argument phrase. In (11), a complete clause, consisting of a compound verb with tense affix and subject clitic, occupies the modifier position in the ArgP. The subject clitic is attached to the demonstrative that is part of the ArgP and precedes the clause.

- (11) *uni-y[-e bujhəu-ŋɔk'-ket'] hɔɾ*  
 that(AN)-y-3SG.S understand-little-PST.ACT person  
 ‘the man who understands little.’ (Bodding 1926: 18)

ArgPs can be marked for case. This marking is done by suffixes such as *-then* in (12). These suffixes are also attached to subordinated clauses as in (13), where the same suffix *-then* functions as a subordinator. These clauses are syntactically and morphologically complete, except that they cannot contain the indicative marker *-a*. Nevertheless they display case marking and behave syntactically like any other case-marked ArgP.

- (12) *alge hɔɾ-then-dɔ alo-pe læi-a*  
 outside person-DAT-TOP PROH-2PL.S tell-IND  
 ‘Don’t tell it to outsiders.’ (Neukom 2001: 24)

- (13) *gapa-dɔ am-ge si-ok'-then ɖangra-dɔ*  
 tomorrow-TOP you(SG)-FOC plough-MV-DAT bullock-TOP  
*laga-əgu-kin-me*  
 drive-bring-3DU.O-2SG.S  
 ‘Tomorrow you shall drive the bullocks to where I am ploughing.’ (Bodding 1926: 100)

The structure and morphology of the ArgP is independent of the category of the elements that form its constituents. There seem to be no general restrictions on

what category an element must belong to so as to function as a head or modifier in an ArgP.<sup>6</sup>

Thus, in summary, the relevant facts about arguments and predicates in Santali are the following. ArgPs in argument position are referential or quantifying expressions with a specific structure and morphology. Predicates can be part of and even head of an ArgP, but lack the indicative marker *-a* in this position. Semantically, predicates or clauses that take the form of an ArgP are not asserted, but presupposed. On the other hand, nominals as well as complete ArgPs and PPs can be placed in the main predicate position, just like event- or property-denoting lexemes. Any material that forms the main predicate is non-referential, regardless of its lexical semantics.

#### 6.4 Nominal sentences

In light of the main concern of this paper, namely the semantics and syntactic properties of the predicate position in Santali, I will concentrate in the following on a special case of predication, the so-called nominal sentence. Nominal sentences are a group of sentence types that are defined by a common syntactic property: the predicating element is formally not a verb, but an ArgP (or an AP or PP). In some languages, such sentences involve the use of a special copular verb, while in other languages the predicate ArgP and the subject ArgP are juxtaposed without an element of this kind. Regardless of their language-specific form, nominal sentences can be divided into different types according to their meaning.

There are different accounts of the types of nominal sentences, starting with Higgins (1979), and no consensus has been reached on their number or what their characteristics are. I will confine my discussion of nominal sentences to a very simple typology of three types (following Mikkelsen (2005)): predicational, specificational, and equative. Mikkelsen (2005: 58) gives clear exemplary sentences for these three types. Example (14) is unambiguously a predicational sentence, (15) is unambiguously equative, and (16) is clearly, although not unambiguously, a specificational sentence.

(14) *The winner is Republican.* (Mikkelsen 2005: 58, example 4.18)

(15) *He is McGovern.* (Mikkelsen 2005: 58, example 4.20)

(16) *The winner is Nixon.* (Mikkelsen 2005: 58, example 4.19)

<sup>6</sup> The only exception seems to be bi- or trivalent lexemes, which do not occur in head position without having their argument positions satisfied by either detransitivation or the use of arguments. Given the connection between the existence of transitive lexemes and the presence of a noun-verb distinction made by Rijkhoff (2003), this fact might be evidence against the idea that the Santali parts-of-speech system is entirely flexible. A detailed study of transitivity in Santali is still a desideratum; hence, nothing more can be said about this topic here.

Predicational sentences such as (14) ascribe a property, denoted by the complement, to the subject (Declerck 1988; Mikkelsen 2005), and are thus closer in their semantic structure to verbal sentences than the other two types. Equative sentences actually predicate the fact that the referents of the two nominals involved are the same entity. Sentences such as (15) are exceptional in that they have a referential nominal—here the proper name *McGovern*—as their complement. Finally, specificational sentences are special in two respects: they have a referential nominal in the complement—the proper name *Nixon* in example (16)—and a property-denoting (or predicational)<sup>7</sup> expression in the subject position. However, sentences like (16) are ambiguous because definite ArgPs such as *the winner* can also be interpreted as referential, depending on the context; this would render the sentence an equative one. Specificational sentences are highly restricted by discourse and are frequently ambiguous, even in a concrete context, between specificational and equative readings.

The three different types of sentences can thus be characterized by the distribution of referential and non-referential nominals in subject and complement position. Table 6.1 lists the different types and the referential status of their subjects and complements (after Mikkelsen 2005: 50).

TABLE 6.1 Subjects and complements

<i>Sentence type</i>	<i>Subject</i>	<i>Complement</i>
predicational	referential	non-referential
specificational	non-referential	referential
equative	referential	referential

I will disregard specificational sentences in Santali for the time being, because my data are insufficient to present a substantiated discussion of this problematic sentence type. The semantic difference between the predicates in predicational and equative sentences is sufficient for the purpose of this paper and corresponds with significant differences in syntactic structure.

In Santali, nominal sentences are composed of a subject, a complement, and often an element *kan*, which has been analysed as a copula. The difference between predicational and equative sentences is minimal at first sight. Predicational sentences have the subject clitic attached to the subject, for example *-e* in (17), while the clitic

<sup>7</sup> I follow Mikkelsen (2005: 51) in using the terms *property-denoting* and *predicative* interchangeably. Formally, this requires viewing properties as functions from individuals to functions from worlds to truth values.



occurs at the end of the whole sentence in equational sentences. Example (18) shows the latter variant.<sup>8</sup>

- (17) *mit'-dɔ-e*            *bhut*    *kan-a*  
 one-TOP-3SG.S    ghost    COP-IND  
 'One was a ghost.' (Bodding 1926: 8)

- (18) *nui*            *ma*    *iŋ-ren*    *hoɽ*    *kan-e*  
 this(AN)    MOD    1SG-GEN.AN    person    COP-3SG.S  
 'This is my wife.' (Bodding 1926: 8)

The placement of the subject clitic is relevant for the delimitation of the main predicate. Thus on a closer inspection of the structure of the predicational sentence (17), we see that the nominal in the complement is situated inside the main-predicate position. Yet at first sight, it does not appear to function as the predicate on its own, but is accompanied by the copula. This situation seems to support a claim made by Evans and Osada that nominals cannot constitute a predicate on their own. To support their claim, they present nominals in predicative position (Evans and Osada 2005: 371). As their examples (24b) and (25b)—here given as (19) and (20), respectively—seem to denote more complex events than one would like to see arising from the lexical semantics of the nominal, Evans and Osada take these sentences as evidence that a more complex derivational process must be assumed to account for the semantics of these sentences.

- (19) *dasi-aka-n-a=ko*  
 servant<sup>9</sup>-INIT\_PROG-INTR<sup>10</sup>-IND=3PL.S  
 '(They) are working as servants.' (Evans and Osada 2005: 369, example (24b))
- (20) *soma=eq*            *baRae-aka-n-a*  
 Soma=3SG.S            baRae-INIT\_PROG-INTR-IND  
 'Soma has become a baRae.' (Evans and Osada 2005: 370, example (25b))

Evans and Osada rightly focus on the compositionality of the semantics of these sentences and state: 'First, one would still need to find an aspect allowing *mastaR*, *baRae*, *baa*, etc. to be used in the exactly composed meaning "be a teacher", "be a blacksmith", "be a servant", etc.' (Evans and Osada 2005: 371). Although they are right to expound the problems of these sentences and mention the potential influence of

<sup>8</sup> The example lacks the indicative suffix *-a*, because the modal particle *ma* is present (see Neukom (2001: 162), where sentence (18) also appears). This particle is used when the speaker assumes the statement to be known, but it has no influence on the placement of the subject clitic.

<sup>9</sup> The lexeme *dasi* is glossed 'serve' by Evans and Osada (2005: 369). Since they themselves use the gloss 'servant' in example (24a) of their article, and since it does mean 'servant' in referential use and 'work as a servant' and not 'to serve' in predicational use, I have changed their gloss for clarity.

<sup>10</sup> The gloss INTR for *-n* is missing in Evans and Osada (2005: 369).

the TAM categories, the examples given by them contain a suffix *-aka* glossed 'initiated progressive', whose contribution to the semantics is not discussed. With this problem in mind, let me come back to our Santali example (17) and the copula *kan*.

The element *kan* in Santali can be interpreted as a copula or an imperfective suffix (cf. the remarks in Neukom, 2001: 17). There are no morphological properties that distinguish the two in constructions such as example (17). The imperfective suffix *-kan* is part of the TAM morphology and is used to express ongoing and habitual actions as well as durative aspect. However, if *-kan* can be regarded as part of the morphological paradigm, it might be interesting to study the predicative use of alleged nominals in other forms of this paradigm.<sup>11</sup> Neukom (2001: 110) presents the lexeme *ʃuər* 'orphan' in three TAM forms. The first form is the zero-marked stative<sup>12</sup> in example (21), which exhibits a genuine predicational meaning. The second occurrence of *ʃuər* 'orphan' is part of a complex predicate (Neukom 2001: 142) and carries a completive past active suffix; it has basically causative semantics. The last example (23) displays middle voice marking, which yields inchoative semantics.

- (21) *alaŋ-dɔ-laŋ* *ʃuər-ge-a*  
 we(DU.INCL)-TOP-1DU.INCL.S orphan-FOC-IND  
 'We are orphans' (Neukom 2001: 110, example (18a), originally from Bodding)
- (22) *huɖiŋ gidrə-i* *ʃuər-oʃo-kad-e-a*  
 small child-3SG.S orphan-leave-COMPL.PST.ACT-3SG.O-IND  
 'She left a child motherless.' (Neukom 2001: 110, example (18b))
- (23) *khange uni gidrə-dɔ-e* *ʃuər-en-a*  
 then that(AN) child-TOP-3SG.S orphan-PST.MV-IND  
 'Then the child became an orphan.' (Neukom 2001: 110, example (18c), originally from Bodding)

These examples show that lexemes such as *ʃuər* 'orphan' show a regular behaviour in predicate position, similar to that of event-denoting lexemes and most closely resembling that of state- or property-denoting lexemes, which all display a 'to be *x*' meaning in stative usage, causative semantics such as 'to make *x*' in active voice, and an inchoative 'to become *x*' reading in middle voice (cf. Neukom 2001: 109). The semantic differences between sentences with the imperfective suffix (or copular) *-kan*, such as example (17), and sentences such as (21), which lack a TAM suffix, seem to be minimal and probably restricted to aspectual subtleties. The instances of such

<sup>11</sup> There is, however, a past-tense copula *tahɛkan*, which behaves more like an independent lexeme than *kan* (Neukom 2001: 171ff.).

<sup>12</sup> Zero-marked forms are normally analysed as active non-past in Santali (Neukom 2001: 62).

structures in the texts of Bodding (1926, 1927, 1929b) are not very telling with respect to the semantic differences between the nominal sentences in active non-past and imperfective form. These structures need to be tested with speakers using sophisticated tests for aspect and *Aktionsart*.

The data presented in this section demonstrate how general the predicative semantics of the main predicate position are. The referential complement of an equative sentence is realized as an argument, while the non-referential complement of a predicational sentence is realized as the main predicate of the sentence. Furthermore, the predicative usage of these supposed nominals seems to be remarkably regular in its semantics and closely parallels the semantics of event-denoting lexemes. This evidence clearly favours an analysis that assumes the flexibility of the lexemes involved and explains the different semantics via the syntactic positions in which they occur. Nevertheless, one could still argue for the existence of two different words such as a nominal orphan<sub>N</sub> and verbal orphan<sub>V</sub>, where the latter is derived from the nominal *tʉər* 'orphan'. To show that this lexical line of argumentation runs into problems, I will focus in the next section on lexemes that have no affinity to a property interpretation, namely proper names.

### 6.5 Predicate position and proper names

Proper names are a special kind of nominal expression. They are used to name and to refer to individuals and are thought to be directly referring expressions and inherently definite. Like indexicals, proper names depend highly on the context to determine their referent. Usually, they constitute a lexical subclass of a category *noun* or a small independent category among the nominals (cf. J.M. Anderson 2007; van Langendonck 2007), but are generally neglected in the discussion of lexical categories and parts-of-speech systems. This is probably due to their marginal status among the *nominals* and their inherent tendency to occur in referential usage only.

Their dominantly referential function and their context-dependence makes proper names a good test case for flexibility of the parts-of-speech system of Santali. It is difficult to imagine a language with a fully productive mechanism for deriving a verbal lexeme with predictable semantics from a proper name, given the latter's characteristics as sketched above.

The precise semantics of proper names are disputed, but there are two main lines of reasoning, which are referred to respectively as rigid-designator theory and definite-description theory. The rigid-designator theory, as advocated by Kripke (1972) and many others, assumes that proper names are directly referring rigid designators (or indexicals). In this analysis, proper names cannot occur as predicates, since they are directly referring. Proper names that do seem to occur as predicates have to be explained as either not actually being predicated or not being used in the predicate, but merely mentioned. This line of reasoning is not very enlightening for

my present purpose, so I will follow the other line of argumentation, stemming from Frege (1893) and Kneale (1962), which sees them as definite descriptions of some kind. By equating them semantically to definite descriptions, proper names are not referential by themselves in this theory, but quantificational (at least from a Russellian perspective). This makes proper names compatible with predication, although they can be—and most frequently are—used to refer.

The definite-description theory has been defended by Geurts (1997) and Matushansky (2009) under the name *quotation theory*, because it assumes that the proper name is quoted in its semantics. Under this theory, a proper name *N* means *the individual N* following Geurts (1997), while Matushansky (2009), by reference to Recanati (1993), adds a naming convention which results in something like: *x is referent of [N] by virtue of the naming convention R*. For Bach (2002), in his nominal description theory, a proper name *N* means *the bearer of 'N'*, but in contrast to the quotational theory, Bach (2002: 76) views the meaning of proper names not as quotational but as reflexive. This avoids some of the complications that come with the notion of quotation. Since there is no consensus on the exact semantics of a proper name within the definite-description theory, I will for now assume a meaning along the lines of *N = the individual named N*. Thus I would like to keep the naming relation explicit in the semantics of proper names, but do not commit myself to a precise formal analysis or the quotational or reflexive character of its semantics.

Even under a definite-description analysis, non-referential use of proper names is a marginal case, but at least a possibility, and one would not expect proper names to have the same distribution as state- or property-denoting lexemes.

In nominal sentences, proper names are expected to occur mainly in the complement position of equative sentences, as *McGovern* does in example (15), and in the subject position of predicative sentences, as in example (6), repeated here as (24).

- (24) *bajun-dɔ-e*                      *marəŋ-a*  
 Bajun-TOP-3SG.S              big-IND  
 'Bajun is big (senior).'                                      (Bodding 1927: 2 [translation FR])

In the light of the discussion of the semantics of the predicate position in Section 6.4, the function and semantics of proper names are unlikely to be compatible with the predicate position. The claim made by Evans and Osada (2005: 371) that 'proper names, such as Ranci 'Ranchi' [a toponym—FR], are unavailable for predicate use' does not therefore seem to constitute a serious objection to a flexibility analysis of Munda languages. After all, proper names are normally used referentially, and the predicate position in Santali forces a property-denoting and thus non-referential reading.

The only types of sentence in which referential expressions are expected to occur in complement position are equative sentences. Since the complements are referential, they should not occur in predicate position in Santali. And in fact, equative

sentences like (25) have the structure SUBJ COMP [COP]<sub>PREDD</sub> with the proper name in the complement argument position and only a copula in predicate position.<sup>13</sup>

- (25) [Context: Two jackals that feature in the story turn out to be a god in disguise]  
*unkin-dɔ*                      *Chando-ge*<sup>14</sup>[-*kin*                      *tahẽkan-a*]  
 that(AN).DU-TOP      Chando-FOC-3DU.S      COP.PST-IND  
 ‘They were Chando himself.’ (Bodding 1926: 160)

There are also examples where the copula is missing, as in (26), but the crucial point is that the proper name still has the form of a complement and does not carry the indicative suffix, nor is it preceded by the subject clitic. The predicate position of these sentences has to be analysed as empty.

- (26) *iŋ-dɔ*      *Sitəŋi*      *jugi*  
 I-TOP      Sitari      yogi  
 ‘I am yogi Sitari.’<sup>15</sup> (Bodding 1929b: 24)

There are, however, some contexts in which proper names are used predicatively (Neukom 2001: 14) and thus non-referentially. The act of naming is probably the most prominent example of an arguably predicative use. Interestingly, this is the only usage in which proper names occur in the predicate position in Bodding (1926, 1927, 1929b). In a naming construction, the proper name occurs as an applicative<sup>16</sup> in the predicate position. In (27) the ArgP headed by *ŋutum* ‘name’ is the subject of the sentence, and the ArgP headed by *hɔpɔn* ‘son’ functions as the object. The proper name *Turtə* functions as a transitive predicate that is combined with the two arguments. The sentence structure can be mimicked by the ‘literal’ translation ‘The boy’s name is Turta-ing the woman’s son’. This is in fact given in Neukom as an alternative translation.<sup>17</sup>

<sup>13</sup> Specificational sentences also have a referential complement, but as noted in Section 6.4 it is unknown how this sentence type is formed in Santali. As stated there, I exclude them from the discussion here.

<sup>14</sup> The focus marker *ge* is not part of the equative structure, but is due to the contrastive context in which this sentence stands.

<sup>15</sup> Bodding translates this sentence as ‘My name is Sitari jugi.’ This is probably a more idiomatic translation in the context of this story, but the sentence structure and a comparison to other sentences strongly favour a translation as given above.

<sup>16</sup> For a description of the applicative in Santali, see Neukom (2001: 120).

<sup>17</sup> There is an alternative naming construction in Santali, which involves a proper name in complement position (like *Bitnə* in the following sentence) and the lexeme *ŋutum* meaning ‘name’ in predicate position:

*ona-te-ge*                      *uni-dɔ*                      *Bitnə-ko*                      *ŋutum-kad-e-a*  
 that(INAN)-INST-FOC      that(AN)-FOC      Span-3PL.S      name-COMPL-3SG.O-IND  
 ‘Therefore they called (named [FR]) him Span.’ (Bodding 1927: 150)

- (27) *uni buɟhi-ren hɔpɔn-tɛt' koɽa-w-ak' ɲutum-dɔ*  
 that(AN) old.woman-GEN.AN son-3.PPOSS boy-W-NML.INAN name-TOP  
*Turtə-w-a-e-a*  
 Turta-W-APPL-3SG.O-IND  
 'The old woman's son's name was Turta.' (Bodding 1926: 110)

The argument *ɲutum* 'name' is not necessary for the naming reading. Sentence (28) is a case in point (context: in a village there lived two people, mother and son). The structure of this sentence could be mimicked—along the lines of Neukom's 'literal' translation—as 'Her son was Anua-ed'.

- (28) *hɔpɔn-tɛt'-dɔ ɲnuə-a-e-a*  
 son-3PPOSS-TOP Anua-APPL-3SG.O-IND  
 'The name of the son was Anua.' (Bodding 1926: 98)

The semantics of the proper names in these naming sentences can reasonably be analysed as compositional, if we assume semantics along the lines of a definite-description theory, and if the naming relation is regarded as a part of its semantics. In sentence (28), the description individual named Anua is made into a transitive property by means of the applicative suffix *-a* and ascribed to *hɔpɔn* 'son' by some unnamed agent, which could be paraphrased as 'S is made the individual named Anua (by X)'. Since being made the individual named N is actually being the object in the act of naming, the naming semantics seems to fall out naturally if we assume a meaning *the individual named N* for proper names. Since there is independent evidence for this assumption (cf. Matushansky 2009), these sentences and their semantics are a strong case in favour of the regular semantic properties of the predicate position and the flexibility of lexemes in Santali, as the sentence meaning can be derived compositionally from the lexical semantics of the proper name, the applicative, and the predicative semantics of the main-predicative position.

There are also complex syntactic units that contain proper names and occur in predicate position. Sentence (29) contains two coordinated proper names in its main-predicate position and could be paraphrased as 'They were Kara-and-Guja-ed.' Although the semantics of this sentence are complicated by the necessarily distributive reading of the predicate, it is essentially identical in structure with the examples given above. It is clear that an example such as (29) cannot possibly be analysed on a lexical level.

- (29) *unkin-dɔ Kaɽa ar Guja-w-a-kin-a*  
 that(AN).DU-TOP Kara and Guja-W-APPL-3DU.O-IND  
 'Their names were Kara and Guja.' (Bodding 1927: 162)

In light of the remarkable ability of proper names to be used as transitive predicates in naming constructions, it would be interesting to see how these lexemes behave syntactically and semantically as intransitive predicates. There is no instance



Despite the limitations of the data, I hope to have shown convincingly that proper names can occur as predicates in the main predicate position in Santali, and that their behaviour is parallel to that of other lexemes with descriptive semantics and even state- and event-denoting lexemes.

## 6.6 Conclusion

With this paper I intended to contribute to the analysis of the parts-of-speech system of Santali, and in focusing on the predicate position, I hope to have shown that lexemes in this language are in fact flexible to a remarkable degree. While the use of seemingly nominal lexemes as predicates in as general a way as in Santali is per se noteworthy, the use of proper names as predicates makes a conversion analysis problematic. As a matter of fact, the semantics of all these sentences is uniform, with the differences in meaning arising from the interaction between the semantics of the syntactic positions and the lexical semantics. Hence I am confident in claiming that they can be explained without having to resort to any lexical mechanism.

Obviously, more research is needed regarding the lexical and grammatical properties of proper names in Santali. There are also many open questions about the parts-of-speech system of Santali, as the analysis presented in this paper does not serve as an argument for bidirectionality of flexibility. However, the regular usage of proper names as predicates is a case in point for the generality of flexibility. If closely related languages with a very similar grammar, such as Mundari, really cannot use proper names in predicate position, it would be an interesting task for further research to determine where exactly the differences in the lexical and grammatical structure lie.