

Mundari as a flexible language

Kees Hengeveld (University of Amsterdam)
Jan Rijkhoff (University of Aarhus)

1. Introduction

In this reply to Evans & Osada (this issue, henceforth E&O), we intend to show that Mundari is one of the languages without distinct classes of verbs and nouns as far as its basic, non-derived vocabulary is concerned. Our contribution is organized as follows. Section 2 briefly presents a typology of parts-of-speech (PoS) systems, followed by a critical evaluation of the three criteria E&O use to establish the lack of word class distinctions in a language (section 3). In section 4 we present evidence to support our claim that Mundari has ‘flexible’ basic lexemes (i.e. there is no fundamental distinction between nouns and verbs), thus disputing E&O’s claim that Mundari has clearly definable classes of verbs and nouns. The last section is concerned with a set of grammatical features which correlate with the presence of flexible lexemes in a language. It is concluded that Mundari displays all of the predicted features of a language that does not clearly distinguish between separate classes of nouns and verbs.

2. Parts-of-speech systems

We agree with E&O that there are no languages “without word classes” (cf. the title of their article). As a matter of fact, we believe that all natural languages have at least one lexical word class. This is shown in Figure 1 below, which shows that there are two basic types of PoS system: FLEXIBLE and RIGID systems. In languages with a flexible PoS system some or all of the functions of a lexeme in an actual linguistic expression (i.e. verbal, nominal, adjectival, adverbial function) are performed by the same group of lexemes (Types 1-3). In languages with a rigid PoS system (Types 4-7) these functions are distributed over distinct, non-overlapping groups of lexemes. For a detailed discussion of this typology of PoS systems we refer to Hengeveld et al. (2004).

		head of predicate phrase	head of referential phrase	modifier of head of referential phrase	modifier of head of predicate phrase
Flexible PoS systems	Type 1	contentive			
	Type 2	verb	non-verb		
	Type 3	verb	noun	modifier	
Rigid PoS Systems	Type 4	verb	noun	adjective	adverb
	Type 5	verb	noun	adjective	
	Type 6	verb	noun		
	Type 7	verb			

Figure 1. Typology of parts-of-speech systems (adverb = manner adverb)

The rigid word classes in Figure 1 (verb, noun, adjective, manner adverb) are defined in terms of their functions in constructing predications in the following way:

- a verb (V) is a lexeme that can be used as the head of a predicate phrase ONLY;
- a noun (N) is a lexeme that can be used as the head of a referential phrase;
- an adjective (A) is a lexeme that can be used as a modifier of the head of a referential phrase;
- a manner adverb (MAdv) is a lexeme that can be used as a modifier of the head of a predicate phrase¹.

The reason we restrict ourselves to manner adverbs is that other kinds of adverbs, such as temporal and spatial ones, do not modify the head of the predicate phrase, but rather the sentence as a whole. Note also that our definitions permit all lexemes to be used as the main predicate (see also E&O p.13). There is, however, an important difference between verbs and other lexemes in that the predicative use is the only unmarked function of verbs, while in the case of other classes of lexemes the question whether or not they have a predicative use does not affect their classification.

On the basis of the definitions of specialized word classes above, flexible lexemes may be defined in terms of the combination of functions they may fulfill in constructing a linguistic expression. Full flexibility means that members of a single lexeme class, called CONTENTIVES, may be used, without applying any derivational process, as head and as modifier of both predicate and referential phrases; NON-VERBS may be used in all non-predicative functions in contrast with a class of true verbs; and MODIFIERS are lexemes that may be used to modify the head of a predicate or referential phrase.

The typology of parts-of-speech systems presented in Figure 1 shows that there are two kinds of language without a noun/verb distinction: one flexible (Type 1), characterized by the fact that its lexemes may be used in all possible functions; the other rigid (Type 7), characterized by the fact that all lexemes are verbs, so that all non-predicative functions can only be realized indirectly through independent predications. Languages with PoS systems of Type 1 and Type 7 are rare, but Samoan (flexible) and Tuscarora (rigid)² are languages that come rather close to the types at either end of the spectrum (see Hengeveld et al. 2004 for discussion). Since we believe Mundari is a flexible language of Type 1 as far as its basic lexeme inventory is concerned, we will concentrate on this type of language in what follows.

Note, incidentally, that our Type 1 languages are the same as E&O's precategorial languages, but that our Type 7 languages do not coincide with their omnipredicative languages, i.e. languages in which all lexemes can be used predicatively. This definition ignores any non-predicative (nominal etc.) functions a lexeme may fulfill, which is a crucial element in our definition of lexical word classes. Interestingly, E&O do not seem to regard their division between precategorial and omnipredicative languages to be very crucial, since they subsume both types under the label "monocategorial" (E&O p.11). By contrast, we will demonstrate in section 5 that it is important to distinguish between the two monocategorial subtypes in our classification of PoS systems (Types 1 and 7), as each type is associated with a different set of grammatical features. E&O furthermore

¹ We use the notion 'predicate phrase' not in the sense of VP but to refer to just the predicate and its modifiers. Thus, in *He read the book* the predicate phrase consists of one element, *read*.

² There is no consensus among the experts on the PoS system in Tuscarora and other Iroquoian languages. For example, Mithun (2000) puts more emphasis on certain grammatical differences between lexemes, whereas Sasse (1993a) is more impressed by the similarities. Whoever is closest to the truth, there seems to be no doubt that verbs, or rather predications, take over large part of the job of nouns in other languages. We have classified Tuscarora as displaying an intermediate system of 6/7.

distinguish “Broschartian” languages, but Tongan, the language Broschart (1997) discusses, is no different from E&O’s precategorical languages, the only difference being that Broschart presents a detailed semantic classification of Tongan lexemes to explain their grammatical behaviour. We will return to this issue in section 3 below.

We would like to emphasize that there are probably no pure languages, and that each distinct type in this classification of PoS systems should rather be considered as a reference point on a continuum (Hengeveld 1992: 58). Any classification of parts-of-speech systems has to take the diachronic dimension into account. Changes in flexible parts of speech first seem to lead to semantic specialization, then to morphological specialization, and ultimately to a more rigid type of PoS system. Therefore languages at best approximate a certain type. There is nothing controversial about this. For example, in syntactic typology a language is classified as an SOV language, even if this is only the dominant or most frequent pattern. Similarly, in order to capture the nature of a parts-of-speech system, we have to allow for some fuzziness at the edges. This is accounted for in a more detailed presentation of our typology of PoS systems in Figure 2 below, which also contains intermediate types.

A flexible language is considered to have an intermediate PoS system when its lexeme classes are compatible with two contiguous systems within the hierarchy. This situation obtains when derived lexemes have fewer functional possibilities than basic lexemes within a language. As we will show below, Mundari represents such a system: it is a fully flexible type 1 language if its basic lexemes are taken into consideration, but it also has derivational processes that produce lexemes that can be used in all slots apart from the predicate slot, a type 2 feature. Mundari is therefore classified as a type 1/2 language (section 4.2.2). A rigid language is classified as having an intermediate PoS system when the last class of lexemes that is relevant for that language consists of a small, closed class of items. The full classification of PoS systems, including intermediate ones, is given in Figure 2 (cf. Smit 2001).

		head of predicate phrase	head of referential phrase	modifier of head of referential phrase	modifier of head of predicate phrase
Flexible PoS systems	Type 1	contentive			
	Type 1/2	contentive	non-verb		
	Type 2	verb	non-verb		
	Type 2/3	verb	non-verb	modifier	
	Type 3	verb	noun	modifier	
	Type 3/4	verb	noun	modifier	adverb
Rigid PoS Systems	Type 4	verb	noun	adjective	adverb
	Type 4/5	verb	noun	adjective	(adverb)
	Type 5	verb	noun	adjective	
	Type 5/6	verb	noun	(adjective)	
	Type 6	verb	noun		
	Type 6/7	verb	(noun)		
Type 7	verb				

Figure 2. Typology of parts-of-speech systems, including intermediate systems (adverb = manner adverb)

The highest degree of flexibility is displayed by languages that use the same set of lexemes as heads of predicate phrases and as heads of referential phrases. In order to show what that means for the language system, and to illustrate our definitions at the same time, let us contrast this with the situation obtaining in a language with specialized nouns and verbs.

In several languages that distinguish between nouns and verbs, members of both word classes can be used as the main predicate of the clause, but only nouns (and not verbs) can immediately be used as the head of an NP without modification. This is for instance the case in Dutch, a Type 3/4 language. Bare nominal lexemes such as *soldaat* ‘soldier’ can be used as the main predicate:

Dutch (Indo-European)

- (1) *Die jongen is soldaat_N.*
 that boy be:3SG.PRES soldier_N
 ‘That boy is a soldier.’ (lit. “That boy is soldier.”)

The reason to treat *soldaat* in (1) as a bare nominal predicate is that it cannot be modified in any way; note that the same form is also used with a plural subject:

Dutch (Indo-European)

- (2) *Die jongen-s zijn soldaat_N.*
 those boy-PL be:PL.PRES soldier_N
 ‘Those boys are soldiers.’ (lit. “Those boys are soldier.”)

Of course, the nominal lexeme may also be the head of a predicative NP, but in this case we have a phrasal predicate rather than a bare nominal predicate:

Dutch (Indo-European)

- (3) *Die jongen-s zijn [twee goede soldaat-en]_{NP}.*
 those boy-PL be:PL.PRES [two good soldier-PL]_{NP}.
 ‘Those boys are two good soldiers.’

Verbal lexemes such as *zing*– ‘sing’ are only used as main predicates (4), not as the head of a noun phrase (5). For referential purposes the nominal lexeme *lied* ‘song’ must be used (6):

Dutch (Indo-European)

- (4) *Iemand_N zing_{V-t}.*
 someone_N sing_V-3SG.PRES
 ‘Someone is singing.’
- (5) **De/het zing_V was prachtig_A.*
 the:C/NT sing_V be:SG.PAST beautiful
 ‘The sing was beautiful.’
- (6) *Het lied_N was prachtig_A.*
 the:NT song_N be:SG.PAST beautiful
 ‘The song was beautiful.’

By contrast, in a language with a class of contentives (Type 1), such as Samoan, a lexeme can be both the head of a predicate phrase and the head of a referential phrase (note that “noun phrase” would be a misnomer here, as Samoan does not have a distinct class of nouns):

Samoan (Austronesian; Mosel & Hovdhaugen 1992: 80, 73, 74)

- (7) *'Ua mālosi le lā*
 PERF strong ART sun
 'The sun is strong.' (lit. "The sun strongs.")
- (8) *'Ua lā le aso.*
 PERF sun ART day
 'The sun is shining today.' (lit. "The day suns.")
- (9) *E alu le pasi i Apia.*
 GENR go ART bus DIR Apia
 'The bus goes to Apia.'
- (10) *le alu o le pasi i Apia.*
 ART go POSS ART bus DIR Apia
 'the going of the bus to Apia' (lit. "the go of the bus to Apia")

It is only for ontological (i.e. non-linguistic) reasons that certain Samoan contentives tend to be used more in one or the other function (we will return to this in section 3 below).

"Although certain full words seem to be used more as verb or more as a NP nucleus for semantic reasons, there are no lexical or grammatical constraints on why a particular word cannot be used in the one or the other function." (Mosel & Hovdhaugen 1992: 73)

In Samoan, morphosyntactic clues indicate what particular function a lexeme fulfils. If it serves as the head of the predicate phrase, it will occur in clause-initial position and combine with tense-aspect-mood particles; if it serves as the head of a referential phrase (NP) it will occur later in the clause and appear with an article or a preposition.

Mundari is of the same type as Samoan as far as its basic vocabulary is concerned, as we will argue in section 4. Before doing so, we have to address the issue of which criteria should be used to establish flexibility of a PoS system. This is the aim of the next section.

3. Establishing flexibility

In their article E&O (p.1, p.27) propose three general criteria that must be met before any language can be said to lack a noun-verb distinction: (1) distributional equivalence, including bidirectional flexibility; (2) strict semantic compositionality for argument and predicate uses; and (3) exhaustiveness through the lexicon. In this section we will argue that only the third criterion is useful the way it is formulated by E&O, but that the other two require some modifications.

3.1. Distributional equivalence and bidirectional flexibility

E&O (p.13) state that "members of what are claimed to be merged classes [note that the use of the term "merged class" suggests that flexible lexemes have originated as nouns or verbs, KH&JR] should have identical distributions in terms of both morphological and syntactic categories. Rigorously applied, all members [...] should be equally acceptable in both primary syntactic functions as arguments and predicates".

The problem with this criterion is that its rigorous application overlooks an important consideration that influences the use of a flexible lexeme in one or the other function. Lexemes are put to use to lexicalize conceptualizations of the world required in certain communicative situations in a certain socio-cultural context. Since there is no reason to assume that all our concepts are equally symmetrical with respect to predicating and

referring functions in a particular language, we may expect certain flexible lexemes to occur more as predicates than as arguments, whereas other lexemes are used more often as arguments than as predicates. This is explicitly stated to be the case in Samoan, which has a class of contentives (Type 1 in Figure 1). Even though there are no lexical or grammatical constraints on the use of a contentive as an argument or as a predicate (see above), there can be differences in terms of distribution and frequency:

“Not all roots occur with the same frequency as verbs and nouns. Some roots predominantly function as verbs, whereas others are more likely to be found in the function of nouns. Until now we have not, for instance, found *alu* “go” in a nominal function [but they do provide an example themselves, see (10) above, KH&JR] or *mea* “thing” in a verbal function [...]. But we hesitate to say that *alu* is inherently a verb and *mea* inherently a noun for two reasons. Firstly, we cannot find any functional explanation why *alu* should not be used as a noun and *mea* as a verb, whereas, for instance, *gaoi* “thief, to steal” and *tagata* “person, to be a person” are bi-functional. And, secondly, previous experience taught us to be careful with classifications. The more texts we analyzed, and included in our corpus, the more items were unexpectedly found in nominal or verbal function.” (Mosel & Hovdhaugen 1992: 77)

Statements to the same effect can be found regarding languages with a flexible class of ‘non-verbs’ (Type 2 in Figure 1), such as Ngiyambaa. Here the difference relates to morphological properties rather than frequency of use in a particular function: some members of the class of what are called nominals in Australian linguistics (“non-verbs” in Figure 1) may appear in reduplicated form whereas others never do. As in Samoan, the difference is attributed to non-linguistic factors:

“Semantically, nominals are divided into two groups; those which are not subject to productive reduplication and those which are. ... Nominals which do not reduplicate are normally translated by English nouns, and those which do undergo reduplication are normally translated by adjectives. The possibility of productive reduplication could be advanced as a formal criterion for similarly dividing Ngiyambaa nominals into two subclasses, nouns and adjectives. But in Ngiyambaa there are no known further differences, morphological or syntactic, as between non-reduplicating and reduplicating nominals. Syntactically, for instance, any nominal which can be a constituent part of a NP can also be the sole representative of a NP ... *gi:djan* may translate either ‘green’ or ‘(a/the) green one’. To introduce the terms ‘noun’ and ‘adjective’ as synonyms for ‘non-reduplicating’ and ‘reduplicating’ would serve no descriptive purpose elsewhere in the grammar.” (Donaldson 1980: 70f.)

3.2. Compositionality

The criterion of compositionality says that “any semantic differences between the uses of a putative “fluid” lexeme in two syntactic positions (say argument and predicate) must be attributable to the function of that position” (E&O p.13). The discussion of this second requirement is in our view somewhat muddled by the fact that at this point the authors already take it for granted that Mundari has distinct classes of verbs and nouns, and that any semantic differences are due to lexical derivation by zero-conversion. Here are a few illustrative quotations (E&O p.19-20):

“The fact that some nouns, when they become verbs, have only minor semantic additions, [...]”

“In the Mundari case, it is common for primary nouns to be used as predicates, [...]”

“In fact, though we can use verbs freely as arguments, [...]”

The main problem in relation to the compositionality criterion concerns the coercion effect of the syntactic slot on the lexeme, i.e. the placement of a contentive in an argument slot or in a predicate slot is enough to give the flexible lexeme a distinct verbal or nominal sense, respectively. E&O (p.16) state that in the case of truly flexible lexemes “all semantically comparable words in the same slot should undergo the same semantic augmentation”. Due to the possible interference of aspectual and other inflectional affixes on the Mundari predicate, E&O (p.16) then formulate an amended version of the criterion of compositionality: “there should be the same semantic change in all lexemes placed in that position, except for semantic interactions attributable to inflections borne by it, e.g. aspect”. But even then, they claim, the problem does not disappear, since the ‘semantic increments’ of lexemes - when used in nominal and verbal functions - are too varied to be attributable to the syntactic slot. Ideally “there should be isomorphic changes in all lexemes placed in a given functional position”.

There are at least two problems with this statement, each of which we will discuss at some length. The first concerns E&O’s assumption that a flexible lexeme acquires some meaning components which it did not have before (“semantic increment”, “semantic augmentation”) when it is used in an actual linguistic expression. The second problem deals with what E&O (p.18) call “the arbitrariness” of this increment, i.e. the observation that the semantics of (what we regard as) flexible lexemes is “far from constant” when we compare their verbal and nominal senses.

3.2.1. Vague vs. specific meaning

It is important to note that E&O (p.16) assume that each Mundari lexeme has a *distinct* verbal or nominal sense (which is then changed through the invisible process of zero-conversion). There is, however, another way to deal with the semantics of Mundari words, which we regard as flexible lexemes. In Hengeveld et al. (2004: 539–541) it is argued that flexible lexemes are semantically *vague*. A semantically vague lexeme has a general meaning which covers all the more specific possibilities (Cruse 1986: 51). Several tests have been proposed to distinguish between vagueness and ambiguity or polysemy. One linguistic test involves anaphoric reference with ‘so’. Compare the following sentences:

(11) *Duffy discovered a mole, and so did Clark.*

(12) *Judith is your cousin, and so is Bill.*

The noun ‘mole’ is *ambiguous* between the senses ‘small burrowing mammal’ and ‘long dormant spy’. This ambiguity is reflected in the fact that anaphoric *so* in (11) can only be used with the same sense as the one selected for the *mole* in the first clause. So a situation in which Duffy saw a small burrowing mammal and Clark a long dormant spy could not be described using (11). By contrast, if a noun has a *vague* meaning, such as *cousin* in (12), its sense allows for different specifications in the case of anaphoric reference with *so* (we can find an abstract definition that covers both interpretations: male cousin and female cousin).³

In Hengeveld et al. (2004) we have applied the notion of vagueness to parts-of-speech systems and proposed that flexible lexemes are vague, not polysemous. Obviously, when

³ See for further discussion of ambiguity tests e.g. Cruse 1986: 49–83; Tuggy 1993; Geeraerts 1993; Dunbar 2001.

applied to word class distinctions, we are concerned with a more fundamental and abstract kind of vagueness: categorial vagueness, which holds across the various functions a lexeme may fulfil in a linguistic expression.

Inspired by Wilkins's account of noun semantics in Arrernte (Wilkins 2000), we presented an outline of what may happen when a flexible lexeme with a vague sense is used in an actual linguistic expression. Let us suppose that the vague meaning of a flexible lexeme consists of meaning components A B C D E etc. (where A B C etc. can be, for example, features, definitions, descriptions, knowledge structures). By placing the flexible item in a particular syntactic slot the speaker highlights those meaning components of the flexible item that are relevant for a certain lexical (verbal, nominal, etc.) function, downplaying other meaning components. A very simplified representation of this process may look as follows:

MEANING COMPONENTS OF SAMOAN *LĀ*: A B C D E (see examples 7 and 8):

	A	B	C	D	E	...	Highlighted properties of <i>lā</i> :
Slot: head of clause	+		+		+		A C E => verbal meaning (<i>be_sunny</i>)
Slot: head of 'NP'			+		+	+	B D E => nominal meaning (<i>sun</i>)
Slot: modifier of 'noun'		+	+	+			B C D => adjectival meaning (<i>sunny</i>)

In other words, it is the use of a vague lexeme in a certain context (syntactic slot) that brings out certain parts of its meaning, giving the category neutral lexeme a particular categorial (verbal, nominal, etc.) flavour. Both the verbal and nominal sense of a flexible lexeme are contained in its (vague) semantics and the context only highlights the meaning components that are already there, giving the flexible item its verbal or nominal flavour. Such an analysis does not lead to any semantic 'increment' added by the context (syntactic slot).

In a way our vagueness account also solves the problem of semantic variation between the verbal and nominal sense of flexible items in use: since each flexible lexeme has its own unique set of meaning components that are highlighted or downplayed when used in a particular slot, we actually expect there to be semantic variation (arbitrariness) between nominal or verbal senses of a flexible lexeme in use. The 'problem' of semantic arbitrariness will also be addressed in the next section, but from a different perspective.

3.2.2. Semantic variation

E&O find it problematic that 'coercion' from the syntactic slot does not always have the same semantic impact on a lexeme, as in their opinion "there should be the same semantic change in all lexemes placed in a given functional position". They provide some examples to show that this is clearly not the case. For example (p.18):

"Let us give some examples from transitive uses. A common meaning for basically nominal roots (sic – KH&JR) used as transitive verbs is 'cause a(n) N to exist: examples are *bir* 'forest; plant a forest'; *lad* 'pancake; make pancakes'; *maNDi* 'food, make food'. But frequently conversions (sic – KH&JR) of this type take on an additional metaphorical meaning that can no longer be precisely paraphrased as causative of existence. In the case of (1) [(13) below, KH&JR] and (2) [(14) below, KH&JR], for example, [...] the semantic increment 'gather, (so as to resemble a ...)' does not mean simply 'cause to be a mountain', or at most 'cause to become a mountain', but means more specifically 'to heap up' Even though the metaphor it appeals to is rather obviously based on a caused existence meaning, by likening a large

group to a mountain, it is nonetheless one specific addition, instead of other imaginable additions (e.g. ‘cause to be tall’, ‘cause to be outstanding’), and must therefore be treated as lexicalized.”

Mundari (Austroasiatic; E&O 18)

- (13) *buru=ko* *bai-ke-d-a*⁴
 mountain=3PL make-COMPL-TR-PRED⁵
 ‘They made the mountain.’
- (14) *saan=ko* *buru-ke-d-a*
 firewood=3PL mountain-COMPL-TR-PRED
 ‘They heaped up the firewood.’

In our opinion E&O’s discussion of the problem of semantic arbitrariness fails to appreciate the difference between conceptualization and lexicalization and, as a result, they seem to treat metaphor as something different from normal language. The conversion from conceptualization into lexicalization is a complex creative process, which involves, among others, metonymy (based on associative relations) and metaphor (based on analogy by transfer of features). All languages are littered with metaphors and metonyms and nowadays these forms of so-called “figurative speech” are generally regarded as an integral part of human language and categorization (cf. Lakoff & Johnson 1980, Lakoff 1987, Heine et al. 1991, Heine 1997).

Since there are many phases between a new metaphor like *surfing the internet* and an old (lexicalized) metaphor such as *shuttle* (nobody makes the link to sewing machines when *shuttle* is used to refer to the commuter air service between two major cities; Saeed 2003: 15), it is impossible to draw a hard and fast line between literal and non-literal language. In other words, metaphor and metonymy (and other manifestations of so-called figurative use of language such as synecdoche) are part and parcel of normal, contextualized speech in any natural language, irrespective of its parts-of-speech system. How exactly concepts and lexemes are connected in the process of language production and perception is an entirely different matter that goes beyond the grammatical system as such.

3.3. Exhaustiveness

The principle of exhaustiveness concerns the size of the word class: it does not make sense to claim that a language has a Type 1 PoS system if it only has a minor class of flexible lexemes. We agree with E&O (p.22) that exhaustiveness is a relevant criterion, but again only to the extent that there is semantic compatibility between a certain meaning of a lexeme and the occurrence of that lexeme in a certain syntactic slot. We furthermore agree with them that one has to be aware of the possible existence of minor word classes and subclasses within the major word class(es), such as animal or plant names. Taking these caveats into account, it seems that there are indeed languages that meet the requirement of exhaustiveness, such as Samoan, where according Mosel & Hovdhaugen (1992: 77) basically all lexemes (“roots”) are flexible (see also the quotation from their Samoan grammar in section 3.1).

An important problem to be addressed is the distinction between systematic flexibility and “rampant zero-conversion”. E&O are not the first to argue for zero-conversion to

⁴ In this and following examples we follow the orthography and glosses of the original unless otherwise indicated.

⁵ Note that we gloss the morpheme *-a* as PRED ‘predicative’ rather than as IND ‘indicative’, for reasons to be specified in section 5.

explain properties of multifunctional lexemes (see e.g. Vonen 1994 on Tokelau), but their article does not address some basic questions about this process. For example, is there a principled way to distinguish across languages between zero-conversion and polysemy? Or, how can we decide unequivocally on the directionality of the conversion? There are no references to studies on zero-conversion that address these issues, although some work has been done in this area (e.g. Marchand 1964, Don 2004). In sum, those who advocate the zero-conversion analysis will have to provide more evidence to substantiate their hypothesis. In section 5 we will provide supporting evidence for our own alternative proposal by testing a number of very specific hypotheses concerning languages with a flexible word class.

4. The case of Mundari

In this section we will give some arguments in favour of an analysis of the Mundari parts-of-speech system as a flexible one. We would like to emphasize that we do not have first-hand knowledge of Mundari, so that, as is common in typological work, we have to base ourselves on the published sources. Our arguments will therefore not go beyond those sources and are meant to show what kind of argumentation one would need to establish flexibility. Before turning to this complex issue, we would like to make a few remarks about the history of the subject.

4.1. Some bibliographical notes and corrections

At various places in their article, E&O claim that Hoffmann changed his view on the flexibility of Mundari lexemes in between the writing of his Mundari grammar and of his *Encyclopaedia Mundarica* (EM), e.g.

“Significantly, when Hoffmann compiled his massive *Encyclopaedia Mundarica* several decades later, he reverted to an analysis in terms of standard word classes, though the dictionary entries exhibit highly productive zero-conversion.”

“But by the time he came to work on the sixteenth-volume, 4889-page *Encyclopaedia Mundarica* (henceforth EM), with decades of further work on the language behind him, Hoffmann had retreated from his earlier and more radical position: he lists words with word class labels, and in places states explicitly that certain words must be used in particular functions.”

“... Hoffmann himself retreated from the original position (stated in his 1903 grammar) once he was forced, in his gigantic masterpiece, the *Encyclopaedia Mundarica*, to make an exhaustive analysis of thousands of lexemes.”

This is a rather serious misrepresentation of the facts. Hoffmann clearly states his own position in the introduction to the EM, when he writes:

“I have nothing to add to what I stated in the Introduction to my Mundari Grammar on the subject. There I have shown that Mundari words have still such a great vagueness or functional elasticity that there can be no question of distinct parts-of-speech in that language. And therefore I stick to the terminology proposed there, which speaks of the function then and there performed by a word in a given sentence.” (EM 8-9)

In the preface to EM, Hoffmann also refers to this particular feature, when he explains how words borrowed from other languages are made to fit the Mundari parts-of-speech system:

“But the Mundas do not merely Mundarize the outer form or the body of the borrowed words, they also Mundarize their soul, i.e., their function or signifying power. ... The Munda ... makes the borrowed words perform the function of nouns, adjectives, transitive or intransitive predicates ...” (EM ix)

This means that when Hoffmann gives the meaning of e.g. a “noun” in EM, what he actually is describing is the meaning of a lexical item when used as the head of a referential phrase, without committing himself to classifying the item in question as a noun. This is an unfortunate but not uncommon practice, one which he applied and explained in his grammar, and one that was later adopted by e.g. Cook (1965). But this also means that the word class labels that Hoffmann used in EM cannot be taken in the way that E&O interpret them. Hoffmann simply lists the meanings of lexemes that he observed, and indicates in which function the given meaning is available. Only rarely does he explicitly indicate that the use of a lexeme in a certain function is excluded, and we must assume that he did so in a systematic way. This is for instance the case with the lexeme *an* ‘dawn’ which occurs with the comment “never used sbstly” (EM: 120).

E&O also refer to earlier work by both of us, in which we mention Mundari in passing, indeed quoting Hoffmann’s (1903) grammar. They do not refer to a more recent article (Hengeveld et al. 2004), in which Mundari is for the first time actually part of the sample under investigation and therefore fully analyzed. In this article we also used Cook’s (1965) grammar, not mentioned by E&O. Cook (1965) is a description of Mundari that on the one hand systematizes Hoffmann’s grammar, but on the other incorporates and analyzes data gathered by the author during an eight year stay among the Mundas. In several respects this grammar nicely complements the information in Hoffmann (1903), Sinha (1975) and Osada (1992).

4.2. The classification of Mundari

In Hengeveld et al. (2004), we classify Mundari as a type 1/2 language, which is characterized as follows:

“In order for a flexible language to qualify as having an intermediate parts-of-speech system, its lexeme classes should be compatible with two contiguous systems within the hierarchy at the same time. This situation may obtain, for instance, when derived lexemes have fewer functional possibilities than basic lexemes within a language. Thus, Mundari is a fully flexible type 1 language if its basic lexemes are taken into consideration, but it also has a derivational process that produces lexemes that can be used in all slots apart from the predicate slot, a type 2 feature. In order to account for these facts Mundari is classified as a type 1/2 language.” (Hengeveld et al. 2004: 538)

We will first go into the nature of basic lexemes, and then deal with derived lexemes separately.

4.2.1. Basic lexemes

Not only Hoffmann (1903), but also Cook (1965) and Osada (1992) emphasize the flexibility of Mundari basic lexemes:

“The problem in Mundari morphology is that not only in syntax, but even in the morphological formation of words, a stem occurs now as one part of speech, now another.” (Cook 1965: 108)

“In Mundari a prototypical lexical verb can be used syntactically as a noun without any morphological change whereas a prototypical noun can be verbalized with verbal ending. For example, *buru* ‘mountain’, *jom* ‘to eat’ are apparently good candidates for a prototypical noun and verb respectively, but *buru* can be used as a verb ‘to heap up’ and *jom* as a noun ‘food’.” (Osada 1992: 43)

Both grammars then go on, much in the vein of the EM, to organize the grammar in terms of the traditional word classes. Thus Cook (1965: 108) continues the previous quote by saying that a single lexeme “... might occur now with the noun suffixes, now with the verb suffixes, now as an independent particle”, and goes on to assign words to lexical classes on the basis of frequency. Osada (1992: 43) also classifies lexemes in terms of traditional word classes on the basis of morphological and syntactic criteria. Both authors actually revert to the same strategy as Hoffmann in his EM, when they call something a “noun” when they want to describe the morphological and syntactic behaviour of a lexeme when used as the head of a referential phrase, and a “verb” when talking about a lexeme used as the head of a predicate phrase, etc. In doing so, they confuse the category of a lexeme with the morphosyntactic properties of a certain functional-syntactic slot.

The arguments that E&O bring up against a precategorial analysis are interrelated with their implicit assumption that lexemes denoting physical entities (individuals or objects) are nouns, and that lexemes denoting temporal entities (actions, processes, situations) are verbs. Using their terminology, a major problem they see is that their “verbs” cannot systematically be used referentially denoting a physical entity. This problem evaporates, however, when one accepts that lexemes denoting temporal entities may do so irrespectively of whether they are used predicatively or referentially, as seems to be the general rule in Mundari. E&O give the following examples:

Mundari (Austroasiatic; E&O 21)

- (15) *her=ko caba-ja-n-a*
 sow=3PL finish-INCEP-INTR-PRED
 ‘They have finished sowing.’
- (16) *dub=ko laga-ja-n-a*
 sit=3PL be.tired.of-INCEP-INTR-PRED
 ‘They are tired of sitting.’

E&O (p.21) note with respect to these examples: “When used in complement clauses, by contrast, [they] do not need to take aspectual or transitivity markers, and can appear as the sole element of a NP [...]” Subsequently they discard these examples as instances of infinitive-like behaviour of ‘verbs’. The reason E&O claim that this counts as an argument against the precategorial analysis of Mundari lexemes is that here the lexeme is not used in a prototypical function (i.e. it is not used to refer to a physical object). In our approach the denotation of a lexeme is completely irrelevant for its classification as a member of a specific wordclass. It is only the function that the lexeme fulfils in building up a predication that counts as an argument for its classification. The lexemes *her* ‘sow’ in (15) and *dub* ‘sit’ in (16) are used as the head and sole element of a referential phrase, and they can be so used without the application of any derivational process. They thus comply just as well with our definition of nouns as they do with our definition of verbs, which means that they are flexible elements.

From this we may conclude that some lexemes in the Mundari lexicon have meaning components that only relate to properties of second order (temporal) entities (see section

3.2.1 on the semantic representation of vague lexemes). Other lexemes in Mundari contain meaning components that only relate to properties of first order (spatial) entities. However, since these semantic features do not appear to impose any restrictions on the referential or predicative use of the Mundari lexeme, there is no reason to recognize distinct classes of verbs and nouns.

E&O also consider it a problem that a lexeme that serves as the main predicate sometimes requires the presence of a copula. This, however, is due to the fact that not only bare lexemes but also referential phrases may act as predicates, as illustrated for Dutch above. Given the fact that a lexeme, without affixes or determiners, may constitute a complete NP by itself in Mundari (cf. E&O p.20), *hoRo* in (17) can also be analyzed as a phrasal constituent rather than just a bare predicate, as in (18):⁶

Mundari (E&O p.17)

(17) *ne dasi hoRo tan-iq*
 this servant Munda COP-3SG.S
 ‘This servant is a Munda’

(18) *ne dasi hoRo-a=eq*
 this servant Munda-PRED=3SG
 ‘This servant speaks Munda’

The phrasal analysis of (17) is supported by the fact that in this construction the predicative expression may be modified, as in the following example, in which the predicate carries the plural subject marker:

Mundari (Austroasiatic; Langendoen 1967)

(19) *en hoɽo-ko munɕa-ko ta-n-ko*
 those man-PL headman-3PL.S COP-INTR-3PL.S
 ‘Those men are headmen.’

We conclude, therefore, that in Mundari a copula is only required when the predicate is phrasal, so that the presence of this copula cannot be used as an argument in classifying the lexeme occurring within that phrase as a noun. The phrasal analysis of (17) and (19) goes well with the fact that these constructions have a classifying reading, as opposed to the property-assigning reading of (18), which is typical of bare predicates in general (Hengeveld 1992). This also explains the fact that there are subclasses of lexemes denoting physical entities with a limited availability for predicate use, such as the proper names, kinship terms, and animal and plant names discussed by E&O. In many cases these lexemes are less likely to be semantically compatible with the property assigning reading of bare predicates.

A last major argument that E&O bring up against a precategory analysis of Mundari is the fact that their “verbs” have to be “converted into headless clauses before being placed in an argument slot”. Examples they provide are the following:

⁶ In E&O’s article, the lexeme *hoRo* is glossed as ‘Munda’ in one example and ‘speak.Munda’ in the next. Since we fail to see the rationale behind this decision, we have provided *hoRo* with the same gloss in both examples, viz. ‘Munda’.

Mundari (Austroasiatic; E&O 20)

- (20) *om-ke-n=iq*
 give-COMPL-INTR=3SG.SUB
 ‘the one who gave’
- (21) *susun-ta-n=iq*
 dance-PROGR.OR-INTR=3SG.SUB
 ‘the one who is dancing’

What E&O actually seem to say, in our terminology, is that a flexible lexeme whose meaning components only relate to properties of temporal entities must occur in a special construction type when used to denote a physical entity (an agent). This is not surprising, since a bare lexeme used referentially denotes a temporal entity, the activity, itself, as shown above in examples (15) and (16). The headless relative construction is an appropriate solution to produce an agentive meaning. The lexeme in question is used predicatively in order to provide a description of a physical entity (the agent) in terms of the activity this physical entity is engaged in.

Once again, there is only a problem if one accepts E&O's conception of prototypical nouns as denoting physical entities. On the other hand, the same facts can be easily accounted for if Mundari lexemes are regarded as flexible items with a vague meaning.

On a more detailed level, E&O's arguments run into trouble on page 24, where they claim that *haga* ‘brother’ gets the extension to ‘be in the same clan’ in predicative use. EM, however, gives ‘clan member’ as one of the meanings of *haga* in substantive function, so it is not surprising that this meaning also surfaces in predicative use.

4.2.2. Derived lexemes

All our previous remarks on the classification of Mundari concerned the basic inventory of lexemes. In several languages with a flexible parts-of-speech system, derived lexemes show a degree of flexibility that is one step lower on the hierarchy than that of basic lexemes (Smit 2001). This is also true of Mundari. The major derivational process involved is illustrated in the following examples:

Mundari (Austroasiatic; Cook 1965: 144; Osada 1992: 62)

- (22) a. *dal* ‘strike’ *da-n-al* ‘a blow’
 b. *dub* ‘sit’ *du-n-ub* ‘a meeting’
 c. *ol* ‘to write’ *o-n-ol* ‘the writing’

The infix *-n-* in (22) “transforms the verb root into an abstract inanimate noun stem, which is no longer capable of verb inflection” (Cook 1965: 144).⁷ Hence we may conclude that there is at least one class of (derived) lexemes that cannot be used predicatively and should therefore be classified as non-verb in Mundari, which shows that the verb/non-verb opposition is at least relevant to the analysis of Mundari in the domain of derived lexemes.

5. Mundari: a flexible language

In previous work we found that there are strong correlations between the parts-of-speech system a language employs, and other grammatical features of that language. These correlations hold especially for languages with flexible lexemes. We will briefly discuss these features below, showing that Mundari behaves exactly the way one would expect a

⁷ Notice that Cook assigns the term ‘verb’ rather loosely to a lexeme that has the predicative use as its most frequent one and that ‘inanimate’ does not refer to grammatical but natural gender (i.e. to a property of the referent rather than the lexeme; see section 5.3.2 below).

flexible language to behave.⁸ We take this to be further evidence for our analysis of Mundari as a Type 1/2 language in section 4. Note that the predictions we make could never hold for Type 7 languages, which have a single class of verbs rather than contentives.

5.1. Identifiability

Since flexible lexemes are not inherently specified for a particular syntactic category (such as “verb” or “noun”) their occurrence in a linguistic expression is potentially confusing. For example, how is the addressee to know whether a flexible lexemes serves as the head of the predicate phrase (verbal function) or the head of the referential phrase (nominal function)? We hypothesized that the grammar of a flexible language contains certain features that help the hearer correctly identify the intended function of a flexible lexeme in a linguistic expression:

“The existence of a specialized lexical class in a language, i.e. a lexical class whose members are tied to one syntactic slot, makes it less necessary for this language to mark this slot and the phrase within which this slot occurs syntactically or morphologically; conversely, the existence of a flexible lexical class in a language, i.e. a lexical class whose members may occur in various syntactic slots, makes it more necessary for this language to mark these slots and the phrases within which these slots occur syntactically or morphologically.” (Hengeveld et al. 2004: 546)

In brief, when lexical specialization is absent (as is the case with flexible lexemes) additional disambiguating strategies are invoked, i.e. there is a trade-off between lexical type and morphosyntax.

More specific hypotheses tested against data from a representative sample of the world language show that all flexible languages have special word order constraints or special morphological markers to indicate the function of a flexible lexeme (for details we refer to Hengeveld et al. 2004).

At the sentence level all flexible languages in our sample turn out to be either (main) predicate final or (main) predicate initial (or in traditional terms that would be inappropriate for truly flexible languages: SOV or VSO). Apparently, flexible languages exploit the highly recognizable sentence-initial and sentence-final position for identificational purposes. Obviously any other position for the main predicate (such as predicate medial or ‘SVO’) would be problematic for a language without a clear noun–verb distinction if no other clues are available (see below). Significantly, there are no such constraints in languages with rigid PoS systems, which from a cross-linguistic perspective can have the main predicate in any position in the sentence. In other words, speakers of a flexible language such as Mundari know that the last contentive is always the main predicate, whereas speakers of Samoan (also type 1) expect the first lexeme to serve as the main predicate.

If a flexible language allows for word order variation, there is always a morphological element to mark the role of a lexeme. For example, in Samoan, placement of any referential phrase in sentence initial position, i.e. before the predicate phrase, is accompanied by the addition of a morphological marker, the presentative particle *’o*. Mundari goes one step further by not only applying rigid predicate-final word order, but also systematically marking every predicate with a morpheme *-a* ‘PRED(icative)’ that has no other function but to mark a

⁸ For a detailed discussion we refer to Hengeveld (1992), Rijkhoff (2002, 2003), Hengeveld, Rijkhoff & Siewierska (2004), Hengeveld (forthcoming), Hengeveld & Valstar (forthcoming).

lexical unit as the predicate of the sentence.

Mundari (Austroasiatic; Osada 1992: 95)

- (23) *diku=ñ* *itu-a-d-ko-a*
 Hindi=1SG teach-ASP-TR-3PL-PRED
 ‘I have taught them Hindi.’

Similar things about the disambiguating role of morphosyntax in a flexible language can be said about the phrase level. In flexible languages without a separate class of nouns (types 1-2/3), the potential functional ambiguity arising from the nature of their parts-of-speech system concerns the interpretation of a lexical element as a head or a modifier within a referential phrase, since the interpretation of a non-verb as the head or a modifier of a referential phrase may interfere with its potential interpretation as the head or the modifier of the same or a contiguous referential phrase. For this reason we expect the order of head and modifier in referential phrases to be fixed, unless the language employs a special morphological marker which uniquely identifies the head-modifier relation within the referential phrase (as in e.g. Tagalog; cf. Hengeveld et al. 2004: 553–4.). Mundari indeed displays this rigid ordering of heads and modifiers within phrases, having the modifier preceding the head under all circumstances.

No such constraints were found in the languages with a rigid PoS system, in that word order variation among noun and adjective (or among verb and manner adverb for that matter) is not unusual and is not connected with the appearance of special morphological marker.

We have already discussed matters concerning the copula. Here, too, Mundari behaves as expected. Since there is no fundamental distinction between verbs and other lexical word classes, there is no need to mark the use of a bare, non-verbal predicate with a copula: all bare predicates are simply contentives. As we saw in section 4.2.1, only phrasal predicates are given special treatment in Mundari.

5.2. Applicability

In order for a lexical class to be applicable in various syntactic slots, it should have FORMAL INTEGRITY, i.e. be formally independent of morphological material specific to a certain syntactic slot, and it should display MORPHOLOGICAL UNITY, i.e. there should be no of intrinsic subclasses triggering specific morphological processes.

The formal integrity of a lexeme, i.e. its formal independence of morphological material specific to a certain syntactic slot, increases its applicability in various syntactic slots. Flexible lexemes are therefore not expected to show morphologically conditioned stem alternation. In other words, flexible languages are expected to be agglutinative or isolating, but never fusional (Hengeveld forthcoming). The grammars of Mundari show that this language is indeed strictly agglutinating.⁹ The restrictions on stem alternation also mean that suppletion is absent in flexible languages, and the Mundari data confirm this too.

The morphological unity of a lexical class also increases its applicability across syntactic slots. We therefore do not expect morphologically determined intrinsic declension and conjugation classes in a truly flexible language (Hengeveld & Valstar unpubl.). This prediction turns out to be true as well for Mundari (Hoffmann 1903: 5) and all other languages with the relevant flexible word classes we are aware of. For example, a division into grammatical genders is absent in Samoan (type 1), Quechua, Turkish and Hurrian

⁹ Just one exception to this is reported in Cook (1966: 163), who notes that the stem *bano*’, ‘not to be, not to exist, not to be present, to be dead’ “has an allomorph *bang-* used with all animate pronouns.” Arguably, however, this stem has auxiliary status.

(type 2; see Rijkhoff 2002: 60). Mundari does show differences in the morphosyntactic treatment of animate and inanimate entities, but this is a case of natural gender (a.k.a. “semantic gender”, “biological gender”, or “sex gender”), which concerns properties of the referent in the external world rather than lexical features of a lexeme. One way of expressing natural gender in Mundari is by adding the equivalents of “male” (*sandi*) and “female” (*enga*) to the noun denoting e.g. an animal:

Mundari (Austroasiatic; Hoffmann 1903: 7)

(24)	a. <i>sim</i>	b. <i>sandi sim</i>	c. <i>enga sim</i>
	‘fowl’	‘cock’	‘hen’

This strategy is, again, typical of flexible languages.

5.3. Differentiation

The degree of differentiation a parts-of-speech system displays in terms of the syntactic slots lexical items may occupy is reflected in the degree of semantic differentiation within or between the lexical classes in terms of basic semantic features coded within them. Members of a flexible word classes are not expected to be specified for category specific features such as transitivity and number. These predictions are borne out both in the case of Mundari and other languages without distinct classes of verbs or nouns.

5.3.1. Transitivity

Let us first deal with the complex issue of transitivity. We use the common sense notion of transitivity as implicitly or explicitly used in grammatical descriptions, where lexemes are regarded as being transitive when they designate a (dynamic) relationship between two obligatory participants: an agent/subject and a patient/object. All languages with a major, distinct class of verbs have a set of basic transitive lexemes in the lexicon (Rijkhoff 2003).¹⁰ By contrast, a basic set of transitive lexemes is absent in flexible languages, such as Samoan:

“With the exception of a very small class of locative verbs [...], Samoan verbs do not require more than one argument, i.e. S or O. If we define obligatory transitive verbs as bi-valent verbs which express transitive actions and which require two arguments referring to the agent and the patient, then Samoan does not have obligatory transitive verbs.” (Mosel 1991: 188)

“If we compare Samoan verbs with transitive and intransitive verbs in other languages where these two categories are distinguished in terms of the number of obligatory arguments, then there are no cardinal transitive verbs in Samoan, i.e. bi-valent verbs expressing transitive actions. Except for a very small class [...], all Samoan verbs (including ergative verbs) maximally require one argument, namely S or O, both of which are expressed by absolute noun phrases in basic verbal clauses.” (Mosel & Hovdhaugen 1992: 724)

¹⁰ Rijkhoff (2003) argues that a language can only have distinct classes of verbs, nouns, and adjectives if the basic meaning of lexical items somehow encodes the prototypical properties of temporal and spatial entities (events and things). The prototypical event is an activity that involves an agent and a patient; the prototypical thing is a concrete object. Thus, a language can only have major, distinct classes of verbs, nouns and adjectives if the lexicon contains (a) items that designate a dynamic relationship between an agent and a patient, and (b) items that designate a property that is specified as having a boundary in the spatial dimension.

Note furthermore that due to the mono-valency of Samoan contentives, “valency changing derivations do not result in a valency-increase or decrease, but only in valency-rearrangement changing the grammatical relations” (Mosel and Hovdhaugen 1992: 729).

The connection between the lack of a rigid noun-verb distinction and the absence of transitive lexemes has also been observed by Jelinek & Demers in their work on the Salishan languages, spoken in the American Northwest (cf. also Kuipers 1968):¹¹

“The feature of Straits Salish syntax that permits the lack of constraints on the distribution of lexical roots is the fact that the feature of transitivity is not a lexical property of a subset of roots.” (Jelinek & Demers 1994: 700)

Sasse (1993b: 654), referring to work by Broschart (1987, 1991) also suggests that there is a connection between non-transitivity and lexical flexibility, when he writes that lexemes in Salishan languages “denote ‘oriented’ [...] states of affairs, i.e. they characterize an individual in terms of participant role it plays in a state of affairs, e.g. as an actor or undergoer. It is by virtue of this property that they are able to occur both in argument and in predicate position.” In other words, it is the lack of transitivity that makes it possible for lexemes to be flexible, to be used in verbal and in nominal function in languages like Salish and Samoan.

Although there is much research to be done with respect to the argument structure of Mundari basic lexemes (see note 12), it seems that there is no good reason to assume that (a proper subset of the) bare Mundari lexemes are divided into ‘transitive’ or ‘intransitive’ subclasses. The valency of a predicate in an actual utterance is said to be explicitly signalled by an extra morphological marker added to all tense-aspect markers: /-d/ for transitive and /-n/ for intransitive (Sinha 1975: 77).¹²

Mundari (Austroasiatic; Osada 1992: 89):

(25) *dub-aka-n-a-e?*

sit-ASP-INTR-PRED-3SG.S

‘He is still sitting.’

(26) *hon dub-aka-d-i-a-e?*

child sit-ASP-TR-3SG.OBJ-PRED-3SG.S

‘He has caused a child to sit down.’

¹¹ In the Salishan languages, which are also deemed to have no categorial distinction between nouns and verbs (Czaykowska-Higgins & Kinkade 1998: 35), it is, however, possible to derive transitive lexemes (verbs): “When there is no overt TRAN element, the sentence is [–TRAN]” (Jelinek & Demers 1994: 700).

¹² Hoffmann (1903: 164-165) hypothesizes that the markers for transitivity and intransitivity do not carry much semantic weight: “That Mundari should have two differently ending Suffixes for Transitive and Intransitive Predicates is itself very remarkable. Transitivity and Intransitivity are not so much objective qualities of actions as subjective modes of conceiving actions. Even such actions as pre-suppose a terminus distinct from the agent may be conceived *exclusively in their relation to agent, i.e., as Intransitive*. Now these subjective modes of conceiving actions are not in other languages expressed by special distinctive formative elements *ad hoc*. It would therefore be very extra-ordinary if the Mundas had recourse to a special root for the purpose of directly denoting transitivity, if, in a word, the consonant *d* represented an original root which could both in *form* and in *meaning* differ from *n*. The consonant *d* must, I think, be considered as a mere alternative for the consonant *n*, and the substitution of *d* for *n* has been caused solely by phonetic exigencies.”

Notice that the distinction between transitive and intransitive is neutralized in three of the four present tense-aspects and in both future markers (cf. Sinha 1975: 82). Clearly this is an area where much more research is needed (especially regarding valency changing operations, passive formation etc.; cf. Foley 1998).

- (27) *kumbuŋu-kij=ko sab-aka-d-ki-a*
 thief-DL=3PL catch-ASP-TR-3DL.OBJ-PRED
 ‘They have caught two thieves.’
- (28) *kumbuŋu-kij sab-aka-n-a-ki*
 thief-DL catch-ASP-INTR-PRED-3DL.S
 ‘Two thieves have been caught.’

5.3.2. Number

In many (but by no means all) languages the unmarked form of a noun that is used to refer to a concrete object is inherently specified for number in that it denotes a *singular* object. Thus, the Dutch noun *fiets* ‘bicycle’ refers to a single bike in

Dutch (Indo-European)

- (31) *Waar is je fiets?*
 where is your.SG bicycle
 ‘Where is your bike?’

If reference is made to more than one bicycle the plural suffix must be added:

Dutch (Indo-European)

- (32) *Waar zijn jullie fiets-en?*
 where are your.PL bicycle-PL
 ‘Where are your bikes?’

Not all languages, however, employ ‘count nouns’ of this type. As a matter of fact, cross-linguistic evidence suggests that the majority of languages use transnumeral nouns for concrete objects, i.e. nouns not inherently specified for some number value; Rijkhoff 2002: 146-156). Since codification of grammatical number in the lexeme would make a lexeme ‘unflexible’, i.e. unsuitable for direct insertion in other syntactic slots, one may hypothesize that flexible lexemes in PoS systems of types 1 and 2 (Figure 1) are always transnumeral. This turns out to be true for all languages with the relevant PoS systems that we are aware of (Rijkhoff 2002: 42).

At first sight one may be lead to believe that Mundari has a plural marker for animate entities (*-ko*; literally ‘they’; notice that the same form also appears in the predicate complex). This element is, however, probably better analyzed as a clitic third person plural pronoun attached to the phrase rather than the lexeme.¹³

Mundari (Austroasiatic; Hoffmann 1903: 8, 100)

- (33) *hoŋo-ko*
 man-PL
 ‘men’ (lit. ‘man-they’)
- (34) *bau-ing-te-ko*
 senior_brother-1SG-BEN-PL / they
 ‘my senior brothers’

¹³ This is not uncommon. Another language using this strategy is, for example, Bambara (Brauner 1974: 26–28; Kastenholz 1989: 21).

This seems to indicate that Mundari lexemes are transnumeral, which means that, in this respect, too, it behaves just like other languages without a distinct class of nouns.¹⁴

6. Conclusion

We disagree with Evans and Osada (p.26) when they claim that “there are no attested cases of languages lacking a noun-verb distinction”. In this contribution we have tried to demonstrate (in section 4) that there are languages with parts-of-speech systems at the extreme end of the flexibility scale (Figure 1: type 1 or 1/2) and that Mundari is one of them. In addition we have shown that the grammar of Mundari displays all the additional morpho-syntactic and semantic features that correlate with a very flexible parts-of-speech system (section 5).

Correspondence address Hengeveld:

Dept. of Theoretical Linguistics, University of Amsterdam, Spuistraat 210,
NL-1012 VT Amsterdam, the Netherlands; e-mail: p.c.hengeveld@uva.nl

Correspondence address Rijkhoff:

Dept. of Linguistics, University of Aarhus, Building 1410, Ndr. Ringgade,
DK-8000 Aarhus C, Denmark; e-mail: linjr@hum.au.dk

Abbreviations: 1/2/3 1st/2nd/3rd person; ART article; ASP aspect; BEN beneficiary; C common gender; COMPL completive aspect; COP copula; DIR directional; DL dual; GNR general aspect-tense-mood marker; INCEPT inceptive; IND indicative; INTR intransitive; N noun; NP noun phrase; NT neuter gender; OBJ object; PAST past; PERF perfective; PL plural; POSS possessive; PRED predicative (see fn. 3); PRES present; PROGR.OR progressive oriented; S subject; SG singular; SUB subordinator; TR transitive; V verb.

¹⁴ It might be interesting to add that according to Hoffmann (1903: 21) *-ko* “is one of the plural suffixes in Chinese” and that Osada (1992: 131), referring to Emeneau (1980: 114), mentions the use of numeral classifiers, which are also found in Chinese and which often correlate with transnumeral nouns.

References

- Brauner, Siegmund (1974). *Lehrbuch des Bambara*. Leipzig: Enzyklopädie.
- Broschart, Jürgen (1987). *Noun, verb, and participation*. akup (= Arbeiten des Kölner Universalien-Projekts) 38. Köln: Institut für Sprachwissenschaft.
- Broschart, Jürgen (1991). Noun, verb, and participation (a typology of the Noun/Verb-distinction). In Hansjakob Seiler & Waldfried Premper (eds.), *Partizipation: Das sprachliche Erfassen von Sachverhalten* (Language Universals Series 6), 65-137. Tübingen: Gunter Narr.
- Broschart, Jürgen (1997). Why Tongan does it differently: categorial distinctions in a language without nouns and verbs. *Linguistic Typology* 1: 123–165.
- Czaykowska-Higgins, Ewa & M. Dale Kinkade (1998). Salish languages and linguistics. In Ewa Czaykowska-Higgins & M. Dale Kinkade (eds.), *Salish languages and linguistics: theoretical and descriptive perspectives*, 1-68. Berlin: Mouton de Gruyter.
- Cook, W.A. (1965). *A descriptive analysis of Mundari: a study of the structure of the Mundari language according to the methods of linguistic science*. Doctoral Dissertation, Georgetown University.
- Cruse, D.A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Don, Jan (2004). Categories in the lexicon. *Linguistics* 42-5: 931-956.
- Donaldson, Tamsin (1980). *Ngiyambaa: the language of the Wangaaybuwan of New South Wales*. Cambridge: Cambridge University Press.
- Dunbar, George (2001). Towards a cognitive analysis of polysemy, ambiguity, and vagueness. *Cognitive Linguistics* 12-1: 1–14.
- Emeneau, M.B. (1980). *Language and linguistic area. Essays by Murray B. Emeneau*. (selected and introduced by Anwar S. Dil). Stanford: Stanford University Press.
- Evans, Nicholas & Toshiko Osada (this issue). Mundari: the myth of a language without word classes. *Linguistic Typology* 9-3.
- Foley, William A. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper read at the 3rd LFG conference, Brisbane (Australia), 30 June – 3 July.
- Geeraerts, Dirk (1993). Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics* 4: 223–272.
- Heine, Bernd (1997). *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer (1991). *Grammaticalization: a conceptual framework*. Chicago: The University of Chicago Press.
- Hengeveld, Kees (1992). *Non-verbal predication: theory, typology, diachrony*. Berlin – New York: Mouton de Gruyter.
- Hengeveld, Kees (forthcoming). Parts-of-speech systems and morphological typology.
- Hengeveld, Kees, Jan Rijkhoff & Anna Siewierska (2004). Parts-of-speech systems and word order. *Journal of Linguistics* 40-3: 527-570.
- Hengeveld, Kees & Marieke Valstar (no date). Parts of speech systems and lexical subclasses. Ms. Department of Linguistics, University of Amsterdam.
- Hoffmann, John (1903). *Mundari grammar*. Calcutta: The Secretariat Press.
- Hoffmann, John & Arthur van Emelen (1928-78). *Encyclopedia Mundarica*. 16 volumes. Patna: Government Superintendent Printing [Reprint: New Delhi: Gyan Publishing House, 1990]
- Jelinek, Eloise & Richard A. Demers (1994). Predicates and pronominal arguments in Straits Salish. *Language* 70-4: 697–736.
- Kastenholz, Raimund (1989). *Grundkurs Bambara (Manding) mit Texten*. Köln: Köppe.
- Kuipers, Aert (1968). The categories verb-noun and transitive-intransitive in English and Squamish. *Lingua* 21: 610–626.

- Lakoff, George & Mark Johnson (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, George (1987). *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Langendoen, D. Terrence (1967). The copula in Mundari. In John W.M. Verhaar (ed.), *The verb 'be' and its synonyms* (Foundations of Language Supplementary Series 1), 75–100. Dordrecht: Reidel.
- Marchand, H. (1964). A set of criteria for the establishing of derivational relationship between words unmarked by derivational morphemes. *Indogermanische Forschungen* 69-1: 10-19.
- Mithun, Marianne (2000). Noun and verb in Iroquoian languages: multicategorisation from multiple criteria. In Vogel, Petra M. & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 397–420. Berlin - New York: Mouton de Gruyter.
- Mosel, Ulrike (1991). Transitivity and reflexivity in Samoan. *Australian Journal of Linguistics* 11: 175–194.
- Mosel, Ulrike & Even Hovdhaugen (1992). *Samoan reference grammar*. Oslo: Universitetsforlaget AS.
- Osada, Toshiki (1992). *A reference grammar of Mundari*. Tokyo: Institute for the Languages and Cultures of Asia and Africa.
- Rijkhoff, Jan (2002). *The noun phrase* (Oxford Studies in Typology and Theoretical Linguistics). Oxford: Oxford University Press.
- Rijkhoff, Jan (2003). When can a language have nouns and verbs? *Acta Linguistica Hafniensia* 35: 7-38.
- Sasse, Hans-Jürgen (1993a). Das Nomen - eine universale Kategorie?. *Sprachtypologie und Universalienforschung (STUF)* 46-3, 187-221.
- Sasse, Hans-Jürgen (1993b). Syntactic categories and subcategories. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: an international handbook of contemporary research*. 2 vols., 646–686. Berlin: Walter de Gruyter.
- Sinha, N.K. (1975). *Mundari grammar* (Central Institute of Indian Languages Grammar Series 2). Mysore: Central Institute of Indian Languages.
- Smit, Niels (2001). *De rol van derivatie bij lexicale specialisatie*. MA thesis, Department of Linguistics, University of Amsterdam.
- Tuggy, David (1993). Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4: 273–290.
- Vonen, Arnfinn M. (1994). Multifunctionality and morphology in Tokelau and English. *Nordic Journal of Linguistics* 17-2: 155-178.
- Wilkins, David P. (2000). Ants, ancestors and medicine: a semantic and pragmatic account of classifier constructions in Arrernte (Central Australia). In Gunter Senft (ed.), *Systems of nominal classification* (Language, Culture and Cognition 4), 147–216. Cambridge: Cambridge University Press.