Multi-categorial items as underspecified lexical entries

The case of Kambera *wàngu*⋆

Marian Klamer
University of Leiden

1. Introduction

Grammaticalization processes are often described as a particular item going through distinct diachronic stages. At the same time, it is generally agreed upon that grammaticalisation is a gradual process, [1] i.e., such descriptions are abstractions from what happens in reality: grammaticalisation takes place along a continuum, not as a sequence of discrete stages. We also know that the items involved in such grammaticalisation continua are often semantically vague and structurally ambiguous, and often undergo changes in 'word class (affiliation)' or '(lexical) category'.

Most authors assume that a change of category at least involves a reanalysis of the underlying structure of the grammatical context in which an item is used.[2] Since the reanalysis causes a change in the lexical properties of an item, the item appears in the sentence with a different word class: it is now part of a different constituent type, and it has a different interpretation than it originally had.

Note that such a characterisation of the different stages of category change is not a description of the actual *process* of change: it remains unclear exactly which transitions took place in the lexicon, and why. The reason is that a 'category (label)' is not a cognitive unit, and a 'category (label) change' is not a lexical process. The former is a theoretical construct to characterise the combined distributional properties of words, and the latter is a coarse description of the outcome of a sequence of many (often small) changes in the distributional properties of an item.

In order to formally characterise processes of category change more precisely, it is necessary to start by analysing the characteristics of the original lexical item involved and the items derived from it: exactly *which* semantic and/or grammatical

features are involved when an item changes the way it combines with other linguistic elements? And how are these features manipulated during the change: are semantic and/or grammatical features added, lost, or both? Which contexts are favorable to changes in the combinatorial possibilities of an item, and which contexts simply allow for it? Only by studying the lexical item and its contexts in very much detail will we be able to characterise the changes that took place in the original lexical item to arrive at its different word class affiliation.

Since category change does not take place overnight nor in clearly distinct stages, but is a slow process along a continuum, we expect to find cases of category change 'on the go'. Cases to look for are those where in one synchronic stage a single item has various functions, interpretations and combinatorial properties, and occurs in distinct but co-existing structures, *as well as* in a sizeable number of ambiguous contexts.

This paper studies such a instance of category change 'on the go' in Kambera, an Austronesian language spoken on the island of Sumba in Eastern Indonesia (Klamer 1998). In this language, there is a word *wàngu*, with various synchronic functions and combinatorial properties. It appears in a continuum which includes the following six distinct syntactic contexts:

(1) Points in the (synchronic) continuum of contexts where *wàngu* occurs[3]
Corpus: 12 hours of spoken and transcribed texts, N = 138[4]
   a. As an independent, main instrumental verb: 5,1% (7/138)
   b. As the second member of verbal compound: 50,7% (70/138)
   c. Ambiguous between verb in a compound or a preposition: 11,6% (16/138)
   d. As the head of a PP: 5,1% (7/138)
   e. As an ambiguous clause linker (P, C, V?) 12,3% (17/138)
   f. As a matrix ('raising') verb in biclausal constructions: 15,2% (21/138)

Observe that the ratio of the distinct occurrences of *wàngu* shows a large variation. For example, as an independent instrumental verb it only occurs in about 5% of the cases, while its use as the second member in a verbal compound accounts for more than 50% of its occurrences.

The sentences (2)–(4) illustrate three major points in the continuum: from a main instrumental verb translatable as "use", in (2), via a prepositional function translatable as "with", as in (3), *wàngu* becomes an (untranslatable) matrix verb heading an embedded clause, as in (4):

(2) *Nda ku-**wàngu**-a huru ba ku-ngangu*
NEG 1SN-wàngu-MOD spoon CNJ 1SN-eat
"I don't use a spoon when I eat" (elic.)

(3) *Ku-palu-ha da ahu-mu nyumu* ***wàngu*** *ài*
1SN-hit-3PA ARTdog-2SG you wàngu wood
"I hit your dogs with a stick" (elic.)

(4) *Talanga la anda-ka nyungga hi na-****wàngu*** *pa-urang*
while LOC road-PRF I CNJ 3SN-wàngu SR-rain
"While I was on the road it began to rain"

Significantly, however, in stage (c) and (e) in (1) the context of *wàngu* is ambiguous and its categorical status unclear. This ambiguity is not to be ignored as 'noise' in the data: the ambiguous contexts make up about 25% of the contexts where *wàngu* occurs, and are about as frequent as the unambiguous stages (i.e., (d) and (f) together). In other words, *wàngu* belongs to two word classes synchronically: Verb – with the subclasses of main (1a, b) and matrix (1f) verb – and Preposition (1d), while we must also account for the fact that in about 24% of its occurrences, it cannot be assigned to any particular class at all, as in (1c, e).[5]

The main issues addressed in this paper is therefore how a multifunctional (or multicategorial), ambiguous element should be represented in the lexicon. Analogous to the way phonologists study the feature make-up of particular phonological segments by analysing processes of (synchronic and diachronic) change in which those segments are involved, it is also possible to study the lexical-conceptual features of an item by looking at its synchronic and diachronic variation. Since there are no written records from earlier stages of Kambera, this paper only analyses the synchronic patterns of variation of *wàngu*. However, since synchronic variation is the source of diachronic change, there should be no principled distinction between the formal representation of an item that is synchronically multicategorial, and one that changes its category over time.

This paper is organised as follows. In Section 1 I present some background information on Kambera morpho-syntax. In Section 2, the various stages in the grammaticalization chain of *wàngu* are discussed in more detail, by linking the distinct functions of *wàngu* to distinct syntactic configurations. In Section 3 I present a scenario for how *wàngu* may have become the item it is now.

In the remainder of this introduction I address issues relating to the lexical representation of multifunctional items such as *wàngu* that appear to occur in a so-called 'grammaticalisation chain'. Standard theories of lexical representation generally assume that, in principle, a lexical element should belong to one word class only. In case we find the same form in different contexts and with different functions, we assume either that we are dealing with homophonous words (whose similarity in form is accidental), or with polysemous words (where the similarity in form goes back to one lexical unit). In the former case, there is no problem to assign the words to different categories, since there is no semantic connection between them. In the case of polysemy, the situation is rather more complex. An

example of a polysemous word in present-day English is the adjective[6] *like*. This adjective (stripped from its adjectival inflections, Maling 1983: 277) allows a preposition (P), complementiser (C), or adverb (Adv) interpretation, as in the following illustrations (from Klamer 2000: 96):

(5)

Cry like a baby	like = A or P?
He ran like crazy	like = A, P or C?
It looks like he will win	like = C
I wouldn't mind, it's just	like I prefer not to like = C
It goes like "bang"	like = C, P, or Adv?

The distinction between A *like*, C *like,* P *like* and Adv *like,* if it exists, is clearly gradual and entirely determined by the different syntactic contexts of *like*. It is therefore not immediately clear if, and how, *like* could be categorized in a classic lexical model: if we assign the different *like*s to distinct categories, we lose the generalisation that they are synchronically related; and if we lump all its functions and treat it as one item, we fail to account for all its distributional properties. Examples like this can be multiplied for English, and indeed occur in every language. For Kambera *wàngu* the classic model provides roughly two options. First, we may assume that there are three (homophonous) items *wàngu*: one an instrumental verb, one a preposition, and one an untranslatable matrix ('raising') verb. This option has two drawbacks: it denies the obvious semantic connections between the various occurrences of this item (they are more often than not polysemous), and it does not account for the fact that in about 25% of its occurrences it is fundamentally ambiguous and cannot be assigned to any particular category. The second option is to assume that the instrumental verb, the preposition, and the matrix verb *wàngu* are polysemous items that go back to one 'basic' form, in this case presumably the main instrumental verb. The basic lexical representation of *wàngu* would then be as a main instrumental verb. The problem with this option is that it does not take into account how *wàngu* is actually used: the ratios given in (1a) above show that *wàngu* is used as an independent instrumental verb in only 5% of the cases, while at the same time, it is *not* used as a verb in about 43% of its occurrences (i.e., contexts (1c–f)). If in almost half of its actual occurrences it is not used as a verb, it seems odd to characterise its distribution as basically verbal in the lexicon. In sum, a lexicon which employs discrete categories does not seem to allow for synchronically multifunctional items such as Kambera *wàngu*.

Heine (1992, 1993: 79, 112–116) has explicitly pointed out this problem in relation to multifunctional items in 'grammaticalization chains', and argues that we need a new type of categorization to account for items that occur in such chains. He proposes to analyse a 'grammaticalization chain' as a distinct type of category, based on the taxonomic principles of 'family resemblance' logic. A family resem-

blance category is a set of items that share a form, while every member shares at least one attribute with one or more of the other members. A family resemblance category is somewhat similar to a 'prototype category' (cf. Rosch 1978, and later work), since not every member is equally representative of its category, and the category has fuzzy boundaries. However, a family resemblance category is crucially distinct from a prototype category, because the latter has one prototype member – an item that combines *all* attributes that define the membership of the category –, whereas a family resemblance category does not have such a prototype. Also, in a prototype category, all members of the category share at least one attribute, whereas there is no such shared attribute in a family resemblance category (Heine 1993: 114).

In short, Heine suggests that items in a grammaticalization chain are a special type of category, based on the taxonomic principles of family resemblance, and that this type of category should be part of the lexicon *alongside* the classic discrete lexical categories. His proposal therefore introduces a new set of taxonomic principles into the lexicon, in addition to the old set. If we want to keep our theoretical model of grammar as simple as possible, this is a less attractive move. In this paper I argue that it is also unnecessary, since the formal categorization of items in 'grammaticalization chains' can be incorporated into a lexical model with discrete categories, if we allow for a properly articulated theory of lexical representation that can handle the variable, gradient properties of individual lexical items. Also addressing the issue of category change, Haspelmath (1998) argues that we can express gradient word class membership and word class changes by using graded notations. (For instance: $V_{1.0}$ for ordinary verbs, $V_{.7/}$ $P_{.3}$ for preposition-like verbs, $V_{.2/}$ $P_{.8}$ for verb-like prepositions, etc., see Haspelmath 1998: 330.) The proposal presented here is similar to Haspelmath's in that it assumes that gradient word class membership can and must be formally expressed, but it expresses the gradience in a different way. Though the categories used in my proposal are discrete lexical categories, I do allow categories to be more or less specified. The variable, gradient properties of individual lexical items are then expressed by manipulating (i.e., adding, deleting, or changing) the lexical-conceptual features of that item. The features referred to here are very general notions that in some way or other recur in all theories about argument and lexical conceptual structure: the number of arguments of an item, their hierarchical organisation (internal/external), their syntactic category, and their semantic role (cf. Jackendoff 1990; Rappaport Hovav & Levin 1998; Bresnan 2001). The following hypotheses concerning lexical representation are specifically explored in this paper:

1. Lexical items are formally expressed as predicate-attribute combinations. Attributes are variable in type and number. For example, the lexical argument

structure of a verb may vary in the number of its arguments (1, 2, 3) and their type (Agent, Patient, Beneficiary, etc).

2. The semantic bleaching of a verb may imply:
   - A change in attribute number (e.g., loss or addition of one or more argument(s), see Vincent 1999 and references cited there, and Klamer 2000), and/or
   - A change in attribute type (e.g., loss of a specific semantic role).

   The result of semantic bleaching is a lexical item with a less complex/specified argument structure: an "underspecified" lexical item.
3. Underspecified lexical items exist side-by-side with fully specified lexical items.
4. An underspecified lexical item can get a specific interpretation through itssyntactic context, and various contexts may invoke distinct interpretations of one single underspecified item.

With respect to the analysis of the specific Kambera item *wàngu*, it is proposed that originally *wàngu* must have been an instrumental verb, which, as a result of the well-known 'semantic bleaching' has been stripped of part of its argument structure and semantic roles: it lost its external argument slot, the Agent role of that argument, and the Instrument role of the internal argument. As a result, the original verb is now an underspecified lexical item represented as *X <x>*. The different interpretations of this item when it functions as either an instrumental verb, a preposition or a matrix verb are caused by the different constructions in which it occurs. In other words, we witness that a full verb has grammaticalised into a semantically bleached, category neutral lexical item *X <x>*, which allows different contextually induced interpretations. In this way, we have characterised about 95% of the actual occurrences of *wàngu*. The remaining 5% are occurrences of *wàngu* in its original function as independent instrumental verb, but we will see that even as independent instrumental verb its lexical and morpho-syntactic properties are reduced.

2. Preliminaries: Kambera argument marking

Before going into the further details of multifunctional *wàngu,* I outline some aspects of Kambera morpho-syntax that are relevant to interpret the data. Kambera is a head-marking language; verbal arguments are commonly marked on the verb by pronominal clitics. The Agent argument of a simple declarative sentence and the single argument of an intransitive predicate are canonically marked with a nominative proclitic, while a definite Patient is canonically marked with an accusative, as illustrated in (6). The coreferent NPs are optional adjuncts.

(6) *Na tau wútu na-palu-ka nyungga*
ART person be.fat 3SN-hit-1SA I
"The big man hit me"

Definite Beneficiaries/Recipients are canonically marked with a dative clitic, as the contrast between (7a–b) shows: in (7a), the object clitic marks the Patient and is accusative, in (7b) it marks the Recipient and is dative. Grammatical definiteness is signalled by the presence of one of the definite articles: *na* (singular), *da* (plural) or *i* (proper name). Only grammatically definite objects are marked on the verb; indefinite Patients or Recipients are never crossreferenced, compare (7a–b) with (7c).

(7) a. *Da-ngàndi-ya na uhu*
3PN-take-3SA ART rice
"They take the rice"

(8) b. *Da-ngàndi-nya na uhu i Ama*
3PN-take-3SD ART rice ART father
"They bring father the rice"

(9) c. *Da-ngàndi uhu*
3SN-take rice
"They take (some) rice"

The dative clitic paradigm is a hybrid category, since it marks both Recipients and Patients: any transitive verb that ends in *-ng* or *-ngu* must mark its object with a dative clitic; even if it is semantically a Patient. An example is the object of *píngu* "know" in (10). Observe that *-ngu* and the dative clitic are in complementary distribution.

(10) *Nda ku- pí -nya na laku-mu*
NEG 1SN-know -3SD ART go-2SG
"I didn't know that you'd gone"

Because the item under study in this paper, *wàngu*, also ends in *-ngu*, its object clitic must always be dative, even if it expresses a Patient. Indeed, we will see below that the semantic role of the dative object of *wàngu* may vary from Instrument, Recipient/Beneficiary, to Patient, while it may also be semantically ambiguous. A dative object whose role is ambiguous is illustrated in (11a), where *-nya* refers to 'the hymn', which may be the Patient of "sing" or the Instrument of *wà-* "use":

(11) *Ku-rongu-kau ba u-ludu* ***wà-nya*** *na ludu hali*
1SN-hear-2SA CNJ 2SN-sing wàngu-3SD ART song be.holy
"I heard you sing the hymn" (lit. "I heard you and you sang (using) the hymn")

This sentence also illustrate the use of conjunction *ba* "and, as, while, because" in Kambera. Kambera only has coordinating conjunctions such as *ba* in (11), other examples are *hi* "and, so", *ka* "so that", and *jàka* "if, when". The language does allow clauses to be embedded, but employs particular non-finite morphemes to mark those (for example, there are special markers for relative clauses and for controlled clauses); while nominalised clauses are marked with a genitive subject. In short, Kambera has no subordinating complementisers, a fact we will return to in the analysis of the clause-combining function of *wàngu* in Section 2.5 below.[7]

3. The distinct functions of *wàngu*

As was mentioned above, the occurrences of *wàngu* in a 12-hour corpus of transcribed spoken texts indicate that *wàngu* occurs in a continuum of grammatical contexts, in which we can distinguish at least six distinct stages. They are given in (1) and repeated here briefly: (a) as independent instrumental verb (5,1%), (b) as second member of verbal compound (50,7%), (c) ambiguous between verb in a compound, or preposition (11,6%), (d) as the head of a PP (5,1%), (e) as an ambiguous clause linker (12,3%), (f) as a matrix verb in biclausal constructions (15,2%). The following subsections present analyses of these six contexts.

3.1 *Wàngu* as an independent instrumental verb

Sentences (12) and (13) illustrate that *wàngu* can be used as an independent, instrumental verb. Note, however, that neither of them expresses the two arguments of *wàngu* at the same time: in (12), the subject of *wàngu* is marked, but the object is not; in (13), the subject is unmarked, while the object is.

(12) *Nda ku-wàngu-a huru ba ku-ngangu*
NEG 1SN-wàngu-MOD spoon CNJ 1SN-eat
"I don't use a spoon when I eat"

(13) *Njadi-mbu jàka wà-nya yena?*
be.possible-also if wàngu-3SD this.one
"Is (it) also possible with this one?"/"Is (it) also possible if (I/we) use this one?"

There are indications that synchronically, *wàngu* is losing its function as independent instrumental verb. First, it functions as a main instrumental verb in only 5.1% of its occurrences – Kambera speakers also use the loan word *paki* (from Indonesian *pakai* "use"). There are communicative reasons for this: *wàngu* has (by now) developed a rather generic semantics (including "use", "with", "because of",

"while", etc., see the subsections below), while *paki* is more specific. In addition, *wàngu* is homophonous with yet another generic verb: the quotative verb *wà(ng)* "report/say", which is used extremely frequently in Kambera discourse.[8] It is also significant to note that the corpus does not contain any examples of *wàngu* used as an independent instrumental verb with *both* its arguments expressed overtly (i.e., it occurs with a subject clitic or with an object clitic, but never with both). I take this to indicate that the morpho-syntactic properties of the verb in its function as a main, instrumental verb are breaking down as a result of the fact that elements of its argument structure have also been lost. (This will be further discussed in the subsections below.) But if *wàngu* is hardly functioning as an independent instrumental verb anymore, why then claim that this is its 'basic' or 'original' function? The reason is that more than 50% of its occurrences are (still) typically verbal, and instrumental, when *wàngu* appears as the second member of a compound verb. This is discussed in the next subsection.

3.2 *Wàngu* as the second verb in a verbal compound

In the majority of cases, instrumental *wàngu* is the second verb in a verbal compound. Compound verbs are a productive morphological category in Kambera. They are derived by combining (any) two verbs.[9] Illustrations are *tila wàrung* "kick & dispose of" > "kick (s.o./s.t.) away" and *palai nyara* "run & chase" > "run after (s.o./s.t.)"

In a similar vein, the instrumental verb *wàngu* combines with any other verb to derive a compound verb, whereby *wàngu* is always the second verb in the compound. Some examples are:

(14)	*taku*	"scoop X"	*taku wàngu*	"scoop X using Y"
	riki	"laugh"	*riki wàngu*	"laugh about/because of Y"
	hayidi	"play games"	*hayidi wàngu*	"play games on Y"
	tanda	"know X"	*tanda wàngu*	"know X because of Y"
	pabanjar	"talk"	*pabanjar wàngu*	"talk about Y"; "talk using language Y"
	nggidik	"shake"	*nggidik wàngu*	"shake because of Y", "be worried because of Y"
	meti	"die"	*meti wàngu*	"die because of Y"

In all these compounds, the object Y is an instrument: a thing to scoop with, or the instrument by which one can have a laugh, play a game, know something, talk or say something, shake, become worried or die. The Kambera compound verb is lexically and syntactically a unit (Klamer 1998: Chapter 7). This is an important distinction between *wàngu* in compound verbs and *wàngu* in other positions in

the sentence. Syntactically, the two verbs that constitute a compound are one verbal unit; more particularly, in this configuration, *wàngu* never projects a separate syntactic constituent. The compound verb as a whole is the head of the Predicate Phrase (PredP). A Kambera PredP may also contain adverbs (which are separate words), and is the domain of attachment for the clitic cluster.[10] In (15) the PredP of the second clause contains the compound verb *panuang wàngu* "continue doing/with something" and the reduplicated adverb *ju-juang* "only". Note that the clitics *-ma-ki-a-da-ka* attach to the PredP as a whole, including the adverb *ju-juang*, and not to an individual verb.

(15) *…kei-ma-danya-ka uda hawiang,* [[***panuang wangu***]$_{V}$
buy-EMP-3PCONT-PRF EMP.3P some continue wàngu
ju-juang $_{Adv}$]$_{PredP}$ *-ma-ki-a-da-ka*
RDP-only -EMP-MOD-MOD-3PG-PRF
"… they are buying some, and they simply continue (with it)…"[11]

In (15) the shared subject of *panuang wàngu* "continue with s.t." is expressed with a genitive enclitic *-da*, and there is no overt object. In (16), however, the compound verb *meti wàngu* "die because of something" has both an overt subject (*ku-*) and an overt object (*-nya*).[12] Since the first member of the compound is an intransitive verb (*meti* "die"), the object of the compound verb must be the complement that originally belonged to *wàngu.*

(16) *Ai ndia, puli-bia-ngga bùdi, jàka ku-**meti***
EXC no release-MOD-1SD in.fact if 1SN-die
***wà**-ma-nya-i una, nda nggàra ehi-a*
wàngu-EMP-3SD-ITER EMP-3S NEG what content-MOD
"O no, just let me go, if I die as a result of it – I don"t care"

Notice, however, that neither the compound verb, nor *wàngu* alone, forms a syntactic constituent with this object: the clitic *-nya* is separated from the PredP by an emphatic clitic *-ma*. In fact, more modal clitics could be added here: in the clitic cluster pronominal clitics are preceded by modal and emphatic clitics, and the compound verb can never form a syntactic constituent with its complement clitic to the exclusion of the other clitics.

Lexically, the two verbs are also a unit, since they have a single, merged, argument structure. If we assume that an instrumental verb has two arguments, X and Y, X would be commonly considered the external argument, and semantically be an Agent, while the Y would be the internal argument, and canonically an Instrument. When instrumental compound verbs are derived, we expect that the argument structure of the verb *wàngu* merges with the argument structure of

the verb it combines with, and this is sketched in (17a) for intransitive base verbs, and (17b) for transitive base verbs:

(17) a. Intransitive base verb + *wàngu* => Transitive compound verb
Example: *riki* <x> "X laugh" + *wàngu* <x <z>> "X use Z"
=> *riki wàngu* <x <z>> "X laugh about Z"
b. Transitive base verb + *wàngu* => Ditransitive compound verb
Example: *taku* <x <y>> " X scoop Y" + *wàngu* <x <z>> "X use Z"
=> *taku wàngu* <x <y, z>> "X scoop Y using Z"

The argument merger results in a structure where the Agent of both verbs is shared, and the compound verb always has at least *one* internal argument. (If the base verb is transitive, it has two). The sentence in (18) consists of two clauses, both of which contain an instrumental compound with a shared Agent and a shared Patient:

(18) *…ka ku- **langu li wà** -ki-nya na aya-nggu,*
CNJ 1SN- word word wàngu -MOD-3SD ART older.sibling-1SG
"…so I can say (s.t.) about my brother,
*ka ku- **paní wà** -nya na aya-nggu*
CNJ 1SN- tell wàngu -3SD ART older.sibling-1SG
so I can tell (s.t.) about my brother."

In case the compound verb has two shared internal arguments, only one of those is usually morpho-syntactically expressed (the other one being implied). There is no structural preference for one argument in particular: either one of the shared arguments may be expressed, pragmatic and/or discourse considerations usually decide which one is. For example, in (19), the object clitic refers to the Patient argument originally belonging to the verb "take" (wife), while the Instrument (dowry) is implied. In (20), the object clitic must refer to the argument of *wàngu* since there is no interpretation available where the statue could be the Patient of "tell".[13]

(19) *…hi mài pa-**piti** wà-**nya**$_j$ [na kuru uma-nggu]$_j$ hu dita la*
CNJ come SR-take wàngu-3SD ART wife-1SG DIR up LOC
Jawa
Java
"…and (I) came to get a wife with (it, i.e. dowry) up there in Java"

(20) *…hi kiri wà-ma nyuma na paní **wà-nja**$_j$ [da katoda*
CNJ begin wàngu-1PG we ART tell wà-3PD ART pole
kawindu]$_j$
yard
"…and we began to tell (s.t. to) the *katoda kawindu*"

In many cases, the referent of the object clitic is ambiguous, and may be interpreted as the argument of either of the two verbs. For example, the object of the compound verb in (21) allows two interpretations: as Patient of "discuss" or Instrument of *wàngu*.

(21) *Mu- hili lua-ki ná haromu, apu, wà-na-nya*
2SN- again go-MOD DEI tomorrow granny report-3SG-3SD
" "You go again tomorrow, gran," he told her,
muda-a una nuna nú, u- ***báta wà*** *pàku*
be.easy-MOD EMP-3s that.one DEI 2SN- discuss wàngu in.fact
-nya
MOD-3SD
"it's easy, you should just discuss it (with words)" "

To conclude, the argument structure of instrumental *wàngu* in a compound verb merges with that of the other verb, and together they also constitute one syntactic unit. In the latter respect, the compound verb configuration is distinct from configurations where *wàngu* heads an independent constituent in the clause. Such configurations are discussed in the next section.

3.3 *Wàngu* as either a verb or a preposition

In this section we will see that *wàngu* can also occur outside the Predicate Phrase and form an independent, separate syntactic constituent, while it is *semantically* still part of the compound verb.

The contrast is illustrated in (22a–b). In (22a) *wàngu* occurs in the 'original' configuration: as part of a verbal compound, heading the PredP. In (22b), the same (semantic) compound verb is used, but *wàngu* is now moved to a position in the clausal periphery (following the adjunct NP "your dogs"). This position is normally occupied by temporal or locative adjuncts, the latter of which are often PPs, as illustrated in (23). In other words, the constituent with *wàngu* in (22b) has the positional properties of a preposition.

(22) a. *Ku-palu* ***wà****-nja ài da ahu-mu nyumu*
1SN-hit wàngu-3PD wood ART dog-2SPOSS you
"I hit your dogs with a stick"

b. *Ku- palu-ha da ahu-mu nyumu* ***wà****-nja ài*
1SSUBJ- hit-3PA ART dog-2SG you wàngu-3PD wood
"I hit your dogs using a stick (on them)"

(23) *…hi ku-pu-puha-bia-ya una* [*na karobu tunu*]$_{NP}$ [***la***
CNJ 1SN-RDP-drop-MOD-3SA EMP3S ART pumpkin roast LOC
mbombang]$_{PP}$ …
ground[14]
"…and I just dropped the roasted pumpkin on the ground"

Observe, however, that while the object is cliticised once in (22a), it occurs twice in (22b): once on "hit", and once on *wàngu.* The two major distinctions between verbs and prepositions in Kambera are that (i) object clitics attach to PredP's *only*, and (ii) that predicates may consist of a single verb, but never of a single preposition. In other words, since the extraposed *wàngu* in (22b) has an object clitic, and object clitics never attach to prepositions, *wàngu* cannot be a preposition in that context. Morpho-syntactically, then, it is a verb, but note that this verb occurs in the peripheral position that is typically reserved for PPs.

While *wàngu* in (22b) can still be called a verb for morpho-syntactic reasons, and the interpretation of the sentence suggests that this verb is still a semantic unit (a compound) with the other verb in the clause, there are also contexts where *wàngu* is truly ambiguous between a verb and a preposition. In such contexts, *wàngu* appears without an object clitic and is directly adjacent to both the verb (with which it would form a compound) and the object noun (with which it would form a PP). In such contexts, *wàngu* can be analysed as in (24a), or as a preposition, as in (24b). Illustrations of such configurations are given in (25) and (26).

(24) a. [V *wàngu*] $_{V}$ [N]
b. [V] [*wàngu* N]$_{PP}$

(25) *Ngangu kokuru jua-a, ngangu* ***wàngu*** *tolung*
eat coconut only-MOD eat wàngu meat
"Eat coconut only, with meat"

(26) *Na kaweda-na na tau nuna lundu na-bei* ***wàngu***
ART be.old-3SG ART person that.one until 3SN-crawl wàngu
kamata kaba
top husk
"Because of the age of that person s/he crawls using coconut husks"[15]

Note that the context where this ambiguity arises is quite restricted. *Wàngu* must be uninflected, it must have an indefinite object, the object noun must be bare, and *wàngu* and its object must be linearly adjacent. In other words, *wàngu* in (27) is not ambiguous: the object *na iyang* "fish" is adjacent to *wàngu* but not a bare noun (since it has a definite article), and *kawàdak* "money" is indefinite and a bare noun, but is not linearly adjacent to *wàngu*:

(27) *Ee, kawàdak ta-kei wàngu na iyang wà-nggu*
EXCL sacrifice 1PN-buy wàngu ART fish report-1SG
"Oh my, money to buy the fish with, I said"

Despite these restrictions, we still find that in 11.6% of the cases examined in the database, *wàngu* occurs in such ambiguos contexts, which thereby outnumber the percentage of contexts (5.1%) where an unambiguous Preposition-interpretation is available (see the next subsection). Ambiguity is clearly one of major characteristics of *wàngu*.

3.4 *Wàngu* as the head of a prepositional phrase

In 5.1% of the contexts examined, *wàngu* is interpreted as an unambiguous preposition since it appears in the clausal periphery along with a bare complement noun, just like a canonical preposition would. (Semantically, however, it is still a unit with the verb in the Predicate Phrase, as in the compound verb constructions discussed in 2.2.) Sentences (28a–b) illustrate the preposional function of *wàngu*. They contrast with (22a–b), where *wàngu* is interpreted as an (inflected) verb.

(28) a. *Ku- palu-ha da ahu-mu nyumu* ***wàngu*** *ài*
1SN-hit-3PA ART dog-2SG you wàngu wood
"I hit your dogs with a stick"
b. ***Wàngu*** *ài ba ku-palu -ha da ahu -mu nyumu*
wàngu wood CNJ 1SN-hit-3PA ART dog-2SG you
"With a stick I hit your dogs"

Additional examples of the prepositional use of *wàngu* are (29), where we know that *wàngu* occurs outside the PredP because it follows the object clitic that is attached to the edge of the Predicate Phrase. As a preposition, it governs a complex NP:

(29) *…ka u-dunda-nya* [***wàngu*** [*ngahu-mu dangu pulu-mu*]
CNJ 2SN-call.together-3SD wàngu spirit-2SG and word-2SG
"…call him with your spirit and your word"

In (30), *wàngu* projects a prepositional phrase which follows the peripheral, temporal adjunct *ni kawài* "just now":

(30) [[*…pabanjar-di-manya-i*] *duma*] [*ni kawai*] [***wàngu*** *hilu*
speak-EMP-1PCONT-ITER EMP.1P DEI just.now wàngu language
Humba]
Sumba
"… we have also been speaking Sumbanese just now…"

In sum, if *wàngu* occurs outside the PredP, with a nominal complement but without an object clitic, it has the structural properties of a canonical Kambera preposition.

3.5 *Wàngu* as ambiguous item with a clausal complement

Although complements of prepositions are canonically referential nominals, prepositions also have a proposition – a clause – as their complement. Kambera has one general locative preposition *la*, and this preposition also governs subordinate clauses. In biclausal constructions, *la* introduces the second, subordinate, clause as the explicit purpose or goal of the matrix clause. Examples are (31) and (32). Observe that the subordinate clauses are marked with the subordinating (SR) morpheme *pa-*, and that their subject is controlled by the matrix subject. (Arguments for this analysis are presented in Klamer 1998: Chapter 8.) *La* is preceded by an intonational break.

(31) *Hi ku-njadi* ***la*** ***pa-*** *piti-ya*
CNJ 1SN-be.able LOC SR- take-3SOBJ
"So I'll be able to take her/him/it"

(32) *Da-puru-ka uda nú, ngàndi-danya bi kabela bi*
3PN-descend-PRF EMP.3P DEI bring-3PCONT DER machete DER
nímbu nú
spear DEI
"They came down, bringing machetes and spears
la ***pa-****pa-meti wà-nya da ular dàngu da wuya*
LOC SR-CAU-die wàngu-3SD ART snake and ART crocodile
(in order to) to kill the snakes and the crocodiles with (it)"

In a similar vein, the form *wàngu* may have a clausal complement, and link a subordinate clause to a matrix clause, as illustrated in (33), (34) and (35). Like *la, wàngu* is preceded by an intonational break.

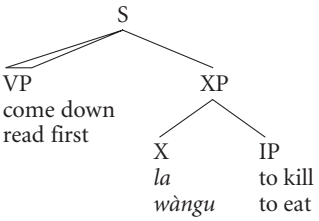
(33) *Mandapung-danya ndalihu* ***wàngu*** *pa-pa-ndalihung*
sit-3PCONT segment wàngu SR-CAU-segment-ng
"The segments (joints) of the corn stalk appear and get their shape"
(lit. "the segments settle while becoming segments")

(34) *…patiang ana mandài-ndài* ***wàngu*** *pa-buta ana rumba*
wait DIM RDP-be.long.time wàngu SR-pluck DIM grass
"…(we) wait for some time, weeding some grass in the meantime"

(35) *Mbaca mangilu* ***wàngu*** *pa-ngangu*
read first wàngu SR-eat
"(We) read before eating"

Wàngu occurs in biclausal configurations such as these in 12,3% of the cases studied. Semantically, it has scope over the second clause, and in (34) and (35) it indicates that the events expressed in the subordinate clause happen simultaneously with, or in immediate sequence to, the event of the matrix clause. In their clause-linking function, then, the canonical preposition *la* and the item *wàngu* pattern alike:

(36)



"Come down to kill" / "Read before eating"

Yet, there are significant syntactic differences between subordinate clauses introduced by *la* and those introduced by *wàngu*. First, clauses with *la* cannot be preceded by a conjunction, while clauses with *wàngu* can. This applies to Kambera conjunctions in general and is illustrated with the conjunction *hi* in (37)–(39):

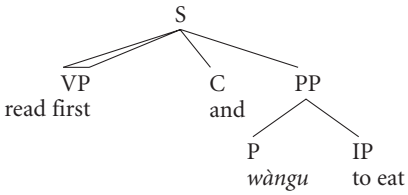
(37) *Hi ku-njadi* ****hi*** ***la*** *pa- piti-ya*
CNJ 1SN-be.able CNJ LOC SR- take-3SOBJ
"So I'll be able to take her/him/it"

(38) *Mbaca mangilu* ***hi*** ***wàngu*** *pa-ngangu*
read first CNJ wàngu SR-eat
"We read and then we eat"

(39) *Talanga la anda-ka nyungga* ***hi*** ***wàngu*** *pa-urang*
while LOC road-PRF I CNJ wàngu SR-rain
"While I was on the road it began to rain"

The fact that a *coordinating* conjunction such as *hi* precedes *wàngu* and a *subordinated* clause is remarkable because clause coordinators by definition mark paratactic clause relations. Moreover, under a prepositional reading of *wàngu*, the coordinator would subcategorise for a PP, which would result in the syntactically illformed configuration in (40) (compare (36)):[16]

(40)



"(We) read first (and then we) eat"

In Section 2.6 we will see that in such contexts *wàngu* is often interpreted as a verb, since it gets a subject proclitic. In other words, the configuration where a conjunction is followed by an uninflected, categorically ambiguous *wàngu* governing a subordinated clause is not very 'stable', because the syntactic context excludes any prepositional interpretation of *wàngu*, while forcing its verbal reading.

The second syntactic contrast between clause-linking prepostion *la* and ambiguous *wàngu* is that subordinate clauses with *la* must appear with a matrix clause, while clauses with *wàngu* may occur independently. Sentence (41) illustrates this. Structurally, the *wàngu* clause in this sentence is a subordinate clause (since it is marked with the subordinating morpheme *pa-*), but it appears without a matrix clause: it is linearly preceded by an interjection/exclamation ("good heavens"), and, before that, by a clause with a nominal predicate that is not its matrix clause. In other words, grammatically *wàngu* would link two clauses, but one of these is absent. Such 'reduced' clausal sequences are impossible with the preposition *la.*

(41) *Yena ama, jàka wài huhu mini katoba-ya,*
this.one father if water milk male be.crazy-3SA
"Here, dad, if this is 'mini katoba"s milk,
ka nggiki-na wà-mu, ba ***wàngu*** *pa-nda padàdu-mbu-nya*
CNJ how-3SG report-2SG CNJ wàngu SR-NEG endure-also-3SD
good heavens!, it is hard to stand
na wau-na na wài huhu nuna.[17]
ART odour-3SG ART water milk that.one
the stench of that milk…"

Thirdly, the semantic relation between clauses linked by *wàngu* is less specified than between clauses that are linked by *la*: while *la* always expresses a purpose or goal, the semantic function of *wàngu* is much less clear, as illustrated in (41). In many cases, it expresses a temporal notion (in particular, "happen/do simultaneously, or subsequently" as in (33), (34), (35), (42), see below).

In sum, compared to clauses introduced by the preposition *la*, subordinate clauses with *wàngu* enjoy a relative syntactic freedom: they may be preceded by an optional coordinating conjunction, as well as an optional matrix clause. Moreover,

the semantic relation between a *wàngu* clause and the matrix clause (if there is one) is vague and open to various interpretations, depending on the context.

In general, an *increase* of syntactic and semantic freedom of an item is formally expressed as a *decrease* in its lexical specifications and/or restrictions. For the lexical representation of *wàngu* this suggests that it must have fewer (syntactic) subcategorization restrictions and fewer (semantic) attributes (arguments, semantic roles) than, for example, a canonical verb, or a preposition such as *la*. In particular, it seems that *wàngu* has no external argument of its own (anymore), and that the semantic content of its internal argument (Instrument) is lost.

3.6 *Wàngu* as a matrix verb

When *wàngu* occurs in biclausal contexts, and especially when it is preceded by a conjunction, it is interpreted as a (matrix) verb. Its verbal function is overtly marked by the presence of a subject proclitic.[18] This matrix verb subcategorises for a subordinate clause (a clause marked with *pa-*), as illustrated in (42) and (43):

(42) *Talanga la anda-ka nyungga hi* ***na-wàngu*** *pa-urang*
while LOC road-PRF I CNJ 3SN-wàngu SR-rain
"While I was on the road it began to rain"

(43) *…jàka ningu-ka banda pa-ngàndi-nggu nú haromu,*
if be-PRF cattle REL-bring-1SG DEI tomorrow
"…if I bring dowry tomorrow
ka ***u-wàngu*** [*pa-puru toma-ka*]
CNJ 2SN-wàngu SR-descend meet-1SA
you will come down to meet me (because of it/in order to get it)"

Wàngu functions as such a matrix verb with an overt subject marker in 15,2% of its total occurrences. What is the meaning of this matrix verb? Its semantics are extremely vague: it may express a temporal relation "happen/do simultaneously, subsequently", as in (42) and (44), it may express a general instrumental or causal relation, as in (43), and it may mean "use", as in (45).

(44) *Ndedi ana luhu nú*,
not.yet DIM exit DEI
"He hadn't left (the woods) yet,
nda ***na-wàngu*** *pa-pàdang mapini ngga-nggàra, ndia-ma*
NEG 3SN-wàngu SR-sense hear RDP-what NO-EMP
(and) he didn't hear a thing, nothing at all"

(45) *Jàka* ***ta-wàngu*** *paní hau, 'nyuna na kapala hakola',*
if 1PN-wàngu word one.CLF he ART head school
wà-da-nya
report-3PG-3SD
"In (lit. If we use) other words, "he's the head of the school", they tell him"

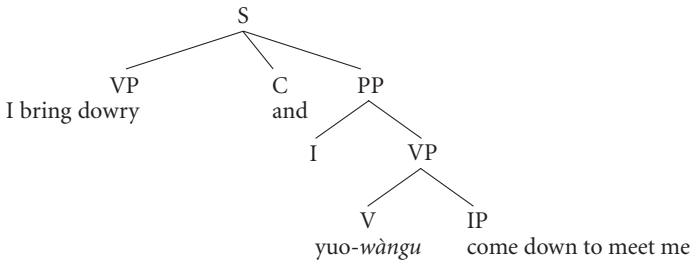
In addition, it links clauses expressing two simultaneous states of affairs: in (46) the state of affairs of not yet having pulled out the knife is simultaneous to the hitting, in (47) the state of affairs of having finished the cutting is simulteneous to the trying to lift it up.

(46) *Ee, ndedi-ma na-butuh-ya na kirih-na*;
EXCL not.yet-EMP 3SN-pull.out-3SA ART k.o. knife-3SG
"He hadn't yet pulled out his knife…
na-wàngu *pa-pangàmbah tú-tú -ma-a-na-nja nú*
3SN-wàngu SR-hit.with.fists RDP-put -EMP-MOD-3SG-3SD DEI
he just hit them like this (gesture of hitting with fists)"

(47) *Hàla hi* ***na-wàngu*** *pa-katiri-ya na bi ai,*
be.finished CNJ 3SN-wàngu SR-cut-3SA ART DER wood
"After he had cut the wood
hi na- kama pa-pajàjak-nya
CNJ 3SN- try SR-lift.up-3SD
he tried to lift it up…"

The matrix-verb-interpretation of *wàngu* is clearly related to the context discussed in the previous subsection, where it connects two clauses while being preceded by a conjunction. We saw that this resulted in illformed configurations like the one in (40). However, this illformed construction can easily be turned into a well-formed one by analysing *wàngu* in (40) as a verb rather than a preposition. The resulting configuration is represented in (48): *wàngu* is now a verb, and projects its own clause which is coordinated with the previous one (*I bring dowry*) by the conjunction (*and*). At the same time, *wàngu* is also the matrix verb of an embedded, non-finite clause (*come down to meet me*), whose subject is controlled by the subject of *wàngu*.

(48)



"... I bring dowry and you will come down to meet me"

Observe that, though the subject of *wàngu* controls the embedded subject, its *referent* is the Agent of the embedded verb. (For example, the subject of *wàngu* in (43) and (48) refers to the Agent of 'come down to meet'). In other words, it appears that *wàngu* has a syntactic subject, but no Agent argument. This is what we expect to be the case if *wàngu* is indeed a reanalysed ambiguous lexical item (category neutral between P and V) in the contexts discussed here (cf. Sections 2.4, 2.5), since prepositions never have an Agent argument. At the same time, the embedded verb has an Agent. But since subordinate clauses marked with *pa-* cannot have overt subjects, this Agent can only be expressed as the subject of the matrix verb. In sum, lacking its own Agent, *wàngu* expresses the Agent of the embedded verb as its subject.

In conclusion, pressure to avoid an ill-formed configuration forces the ambiguous item *wàngu* to function as a matrix verb in those biclausal contexts where it is preceded by a coordinator. Since *wàngu* no longer has an external argument of its own, its syntactic subject expresses the external argument of the *embedded* verb.[19]

4. The *wàngu* "grammaticalization chain" is in fact only one lexical item

On the basis of the evidence presented in Section 2, I analyse *wàngu* as an underspecified lexical item: *X <x>*. Its function as instrumental verb in compound constructions, its prepositional function, its use as a semantically vague matrix verb, and its numerous ambiguous occurrences are all expressions of one, generic lexical item.

The internal argument of *wàngu, <x>*, is always expressed as a syntactic complement, but in different configurations. It may be expressed as: (i) (one of) the object(s) of a compound verb, (ii) the nominal complement of a preposition, or (iii) a clausal complement of an ambiguous item or of a matrix verb.

I propose the following scenario for the development of *wàngu* from a main instrumental V into the underspecified lexical item X:

1. Independent instrumental verb, with two semantically specified arguments:
   V '*wàngu*' < Agent < Instrument >>
2. For various functional reasons, including homophony with the quotative verb *wà(ng)* and the availability of Indonesian loan *paki*, *wàngu* falls into disuse as independent instrumental verb, and loses some of its morpho-syntactic properties.
3. As an instrumental verb, it is still frequently used in compound verbs. Since Kambera compound verbs have a merged argument structure, *wàngu* and the other verb share their Agents. As a result, the Agent of *wàngu* is never expressed as a separate morpho-syntactic entity. Since the context of compound verbs do not allow it ever to be visible at the surface, it is lost from the lexical argument structure of *wàngu* altogether:
   V '*wàngu*' < Instrument >
4. The internal argument of *wàngu* is also part of the merged argument structure of compound verbs, and as a result (some of) its semantic features are also lost, so that the interpretation of the complement of *wàngu* becomes more generic, 'object'-like:
   V '*wàngu*' < Instrument/Goal/Beneficiary/Patient/… >
5. Without semantic role for its only argument, *wàngu* is no longer a typical verb (since prepositions and complementisers also have complements). *Wàngu* has *de facto* become an underspecified lexical item with a single, unspecified, internal argument:
   X '*wàngu*' < x >
6. Given the appropriate syntactic context, it can now be interpreted as either a main verb, a preposition, a matrix/raising verb, or remain ambiguous between any of these interpretations.

5. Conclusions

By analysing the grammaticalization chain of *wàngu* as different surface manifestations of one single, impoverished lexical item *X <x>*, we account for almost all of its synchronic occurrences and functions. Though I assume that *wàngu* used to be a fully fledged instrumental verb this is not crucial to the analysis. The central idea of the analysis presented here is that synchronically, *wàngu* has an extremely simple lexical argument structure (no external argument, and one semantically generic internal argument), while its various functions are entirely determined by the different syntactic configurations in which it occurs. As contextually derived functions, they do not need to be lexically represented.

In this view, the lexicon contains underspecified items as well as items with a fully specified predicate and argument structure. The lexicon is a dynamic entity, where items may lose and gain argument structure and semantic content as a result of both functional constraints and syntactic pressure.

I hope to have shown that in a slightly more articulated theory of lexical representation, combined with an analysis of the interaction between different linguistic modules (here, lexicon and morpho-syntax), there is no need to treat items in grammaticalization chains as a 'special species' in the lexicon. Multicategorial items, or items in grammaticalization chains can be categorized using the same discrete categories and taxonomic principles we use for other lexical items. The only thing in which they differ from canonical lexical items is that they have fewer lexical features, so that their function and interpretation is largely dependent on the syntactic context in which they appear.

Notes

* I would like to thank Elly van Gelderen for discussion about the syntactic analysis of *wàngu*, and the anonymous reviewer for insightful and helpful comments on the prefinal draft of the paper. **Abbreviations**: 1, 2, 3 = person, A = Accusative, ART = Article, CAU = Causative prefix, CLF = Classifier, CNJ = Conjunction, D = Dative, DEI = Deictic, DER = Derogatory marker, DIM = Diminutive, DIR = Directional particle, EMP = Emphasis, EXCL= Exclamation, G = Genitive, ITER = Iterative aspect clitic, LOC = Locative preposition, MOD = Modal clitic, N = Nominative, NEG = Negation, p = plural, PRF = Perfective aspect clitic, RDP = Reduplication, s = singular, SR = Clitic marking subordinated clauses.

1. See for example Lightfoot (1991: 162–163) and Haspelmath (1998: 329).

2. See the discussion of reanalysis in Harris and Campbell (1995: Ch. 4). Some classic examples of processes where V changes into P can be found in Lord (1973), Heine and Reh (1982), Hopper and Traugott (1993), and Harris and Campbell (1995). A recent study of category change (P changing into C in English) within the generative (Minimalist) framework is Van Gelderen (1998).

3. There is another, homophonous verb *wà(ng)* 'report (to); do', which is used very frequently in Kambera discourse to report direct/indirect speech, visions, and sounds. The (near) homophony of this verb and instrumental *wàngu* suggests a shared ancestor. In addition, because both verbs have very generic semantics, it is tempting to try to analyse them as one and the same verb from a synchronic point of view too. This is what I set out to do when I started the research for this paper, and what an anonymous reviewer suggested after reading the previous draft of the paper. However, homophony and semantic vagueness are not sufficient evidence to assume that synchronically quotative *wà(ng)* and instrumental *wàngu* are the same verb, especially since there are a number morpho-syntactic distinctions between them: (i) quotative *wà* is an intransitive root verb (the form *wà-ng* is its applicative

derivation), while instrumental *wàngu* is a transitive root verb; (ii) *wà* only takes a genitive subject, *wàngu* a nominative or a genitive subject; (iii) applicative *wà-ng* takes subject and object pronominal clitics simultaneously, instrumental *wàngu* does not allow simultaneous subject and object clitics; (iv) quotes always precede the quotative verb, instrumental objects always follows the instrumental verb; (v) 'derived' functions of the quotative verb include being used in naming strategies ("to call someone X") and as a discourse particle ("you know"), while the 'derived' functions of instrumental *wàngu* are that it can be used as a preposition or as a matrix verb in biclausal constructions. To me, it does not seem appropriate to derive all these functions from one generic lexical item; for one, because it would be unclear what the argument structure of that item would be: quotative *wà* is fundamentally intransitive (and becomes transitive by applicative derivation), while *wàngu* is transitive in all its occurrences. For a full exposition of the characteristics of quotative *wà* I refer to Klamer (2000), where the syntactic and lexical properties of *wà(ng)* are compared with quotative verbs in two genetically related languages (Buru and Tukang Besi), and to Klamer (2002), which discusses the various (derived) functions and grammaticalization properties of *wa(ng)*. These articles extend the analysis of quotative *wà* in Klamer (1998: Section 8.2.4). The present paper supplements and extends the discussion of instrumental *wàngu* in Klamer (1998: Section 7.2.1).

4. Most examples in this paper were taken from the same corpus of conversations and narratives that was used to calculate the frequencies. None of the examples is elicited through an intermediate language, but some examples were brought up as isolated sentences during discussions with native speakers about the use of *wàngu*. All examples have been checked by native speakers.

5. Among other things, this implies that synchronically speakers may have two, three or more interpretations of the same form available, varying per context (contra Haspelmath 1998: 341). It does not imply, however, that speakers are necessarily aware of the (historical) relation between the competing interpretations, since a historical scenario is a linguist's construct and does not necessarily reflect a speaker's perspective (cf. Joseph 2002, (26)).

6. Despite the homophony, verbal *like* (OE *lician*, cf. Allen 1995) and adjectival *like* (OE *gelic*, Maling 1983) are not cognates.

7. For more information on Kambera complementation, see Klamer (1998: Chapter 8), and Klamer (2000, 2002).

8. See Note 4 for a motivation why I assume that the semantically generic verbs *wà* and *wàngu* are not the same verb (synchronically).

9. Note that Kambera compound verbs are not serial verbs. For arguments, see Klamer (1998: Chapter 8).

10. The Kambera clitic cluster that attaches to the PredP contains up to 9 clitics: modal, pronominal and aspectual. The PredP and the clitic cluster cannot be separated from each other, i.e. adjuncts always follow the clitic cluster or precede the Pred P.

11. Context: Person talking about how easy it is for the people who live in town and have a constant supply of fish.

12. The object clitic *-nya* refers to the proposition in the preceding clause ('you should just let me go').

13. Since a *katoda kawindu* is a religious statue in the yard, used to address prayers to, in the context of (20) the argument of *wàngu* is interpreted as a Goal rather than an Instrument.

14. *Mbombang* is a specific term for the ground under a house on stilts, the place where domestic animals are kept.

15. *Kamata* is the top part of a fruit, the location where a stalk connects it to the tree. *Kaba* is the husk of a coconut. Disabled Sumbanese people crawl using the top part of a coconut husk to protect their hands.

16. An alternative would be to analyse *wàngu* as it appears in (40) as a conjunction introducing a (subordinated) CP. However, in Section 2.6 we will see that Kambera speakers tend to re-analyse *wàngu* in contexts like (40) as a (main, matrix) verb, which goes against the subordinating conjunction analysis. Moreover, it is unclear whether the latter analysis would be available to Kambera speakers in the first place, since all of the Kambera conjunctions are coordinating (Section 1). Lacking a class of subordinating conjunctions, the CP analysis of a *wàngu* clause would presuppose the introduction of a new functional category in Kambera.

17. Context: Main character of folk narrative is assigned with finding a bottle of the (stinking) milk of a mythological human-like person (*mini katoba*).

18. Note that if prepositions and conjunctions took pronominal marking (which they do not in Kambera), they would be expected to have an *object*, not a *subject*.

19. This use of *wàngu* may be compared with English raising verbs such as *seem* in *He seems to lie* (< *It seems that he lies*).

References

Allen, Cynthia L. (1995). *Case Marking and Reanalysis*. Oxford: Clarendon Press.

Bresnan, Joan (2001). *Lexical-functional Syntax.* Oxford: Blackwell Publishers.

Gelderen, Elly van (1998). "The Future of *for to*." *American Journal of Germanic Linguistics & Literatures, 10*(1), 45–71.

Harris, A. C. & L. Campbell (1995). *Historical Syntax in Cross-linguistic Perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin (1998). "Does Grammaticalization Need Reanalysis?" *Studies in Language, 22*(2), 315–351.

Heine, Bernd (1992). "Grammaticalization Chains." *Studies in Language, 16*(2), 335–368.

Heine, Bernd (1993). *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: Oxford University Press.

Heine, Bernd & Mechtild Reh (1982). *Patterns of Grammaticalization in African languages*. Arbeiten des Kölner Unviersalienprojekts, 47. University of Cologne: Institute of Linguistics.

Hopper, Paul & Elizabeth Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Jackendoff, Ray (1990). *Semantic Structures*. Cambridge, MA: The MIT Press.

Joseph, Brian D. (2002). "Rescuing Traditional Historical Linguistics from Grammaticalization "Theory"". Invited paper presented at the conference *New Reflections on Grammaticalization 2*, University of Amsterdam, April 2002.

Klamer, Marian (1998). *A grammar of Kambera*. Berlin: Mouton de Gruyter.

Klamer, Marian (2000). "How Report Verbs Become Quote Markers and Complementisers." *Lingua, 110*, 69–98.

Klamer, Marian (2002). "Report Constructions in Kambera (Austronesian)." In Tom Güldemann & Manfred von Rucador (Eds.), *Reported discourse* (pp. 323–340). Amsterdam: John Benjamins.

Lightfoot, David (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.

Lord, Carol (1973). Serial Verbs in Transition. *Studies in African Linguistics, 4*, 269–296.

Maling, Joan (1983). "Transitive Adjectives: a Case of Categorial Reanalysis." In Frank Heny & Barry Richards (Eds.), *Linguistic Categories: Auxiliaries and Related Puzzles* (pp. 253–289). Dordrecht: D. Reidel Publishing Company.

Rappaport Hovav, Malka & Beth Levin (1998). "Morphology and Lexical Semantics." In Arnold Zwicky & Andrew Spencer (Eds.), *Handbook of Morphology* (pp. 248–271). Oxford: Blackwell Publishers.

Rosch, Eleanor (1978). "Principles of Categorization." In E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Vincent, N. (1999). "The Evolution of C-Structure: Prepositions and PPs from Indo-European to Romance." *Linguistics, 37*(6), 1111–1153.