A Distributed Morphology Approach to Argument Encoding in Kambera

Doreen Georgi

Abstract

In the Malayo-Polynesian language Kambera definite arguments are encoded on the verb via clitics expressing person, number, and case. These markers show many instances of syncretism, especially if they are subanalysed. The aim of this article is to derive the distribution of the homonymous inflection markers in the framework of Distributed Morphology and to motivate that syncretic forms do not appear accidentially.

The main insight of this procedure is that we find a correlation between the items' specificity and their sonority – a principle called iconicity by Wiese (1999). This strengthens the hypothesis that it qualifies as a meta-grammatical principle for the architecture of morphological systems.

1. Introduction

Kambera is a Malayo-Polynesian language spoken in the east of Sumba, an Indonesian island. Definite arguments are encoded on the verb by clitics whose form varies between person, number and case. These inflection markers exhibit many instances of syncretism below the clitic level that means that the clitics are subanalysed into smaller units. The aim of this article on argument encoding in Kambera is to account for the distribution of the homonymous forms, appearing especially among the four cases, by decomposing grammatical categories into more abstract binary features. As a result, natural classes can be formed, when members of a category share the same abstract features and feature values to which inflection markers refer. Consequently, markers can be inserted in more than one context. The feature-based framework that offers the relevant mechanisms like underspecification and late insertion of phonological material is Distributed Morphology and will be adopted throughout the paper.

The assumption in the background is that the distribution of syncretic forms might not be arbitrary but systematic, as formulated in the meta-

Subanalysis of Argument Encoding in Distributed Morphology, 1-21
Gereon Müller & Jochen Trommer (eds.)
Linguistische Arbeits Berichte 84, Universität Leipzig 2006

grammatical Syncretism Principle in Müller (2004):

Syncretism Principle: Identity of form implies identity of function.

In the following analysis I will strengthen the validity of the Syncretism Principle by showing that syncretism in Kambera does not occur accidentally, but can be derived systematically and furthermore that additional rules needed to account for cases that might seem to be exceptional are motivated well, too. Finally, Kambera's argument encoding markers exhibit a correlation between sonority and specificity providing evidence for iconicity as another meta-grammatical principle.

2. Background information

2.1. Distributed Morphology

The theory of Distributed Morphology has been developed by Morris Halle and Alec Marantz at the beginnings of the nineties and aims to account for inflectional as well as derivational morphology. It is assumed that terminal nodes are organised in hierarchical structures built up in syntax. The terminal nodes simply consist of grammatical features but lack phonological features. These are added after syntax by insertion of vocabulary items relating phonological information and morpho-syntactic features, which constitute their insertion context. Before vocabulary insertion, rules may apply in order to change the feature bundles used in syntax. Features can be deleted or added, whole terminal nodes can fuse to one or divide into several ones. As a consequence the set of vocabulary items that theoretically could be inserted changes. Which of the competing vocabulary items that are at disposal is inserted, is regulated by three concepts.

Subset Principle:[1] A vocabulary item V can only be inserted into a terminal node iff (i) and (ii) hold: (i) The morpho-syntactic features of the Vocabulary Item V are a subset of the features of this terminal node and (ii) there is no other vocabulary item Y that fulfils (i) and is more specific than V.

[1]The definitions are based on formulations by Halle & Marantz (1993), Halle & Marantz (1994) amd Harley & Noyer (1999).

Specificity: A vocabulary item V is more specific than a vocabulary item Y if it shares more features with the terminal node it can be inserted in according to the Subset Principle than Y does.

Underspecification: A vocabulary item does not need to match every feature of the terminal node. Vocabulary items are often underspecified so that more of them are in competition for insertion into a single terminal node.

Underspecification provides the possibility to derive syncretisms because underspecified vocabulary items do not define a unique insertion context, but encode features that several terminal nodes possess. A relevant operation for the following analysis that takes place before insertion is called Impoverishment and has been introduced by Bonet (1991). It eliminates features of the terminal node so that the expected vocabulary items do not fit anymore and a less specific item wins the competition.

2.2. Kambera

Kambera has a high dialectal variation among the 150,000 speakers and no standard variant is defined. The dialect on which the analysis will be based is spoken in the village Kataka, described in the grammar of Klamer (1998). For the analysis of the verbal agreement morphology it will be necessary to take a closer look at Kambera's syllable structure, phonological processes and some facts of alignment first.[2]

Kambera has a simple syllable structure. Syllables consist of at least one vowel, an onset consonant is optional. All syllables are open and roots have to be built up by exactly two of them. If there is an onset then it cannot be complex; the language allows only one initial consonant: (C)V. Nevertheless Kambera has 5 prenasalised segments consisting of a nasal that assimilates its place of articulation to the following plosive, affricate or semivowel. It is important to note that these prenasalised segments are considered to be one phonological unit so that the restrictions on the syllable structure are kept. Furthermore, the nasal in a prenasalised structure itself triggers assimiliation of the following consonant concerning voice.

[2]For verification of the facts, the cited examples and more detailed information about Kambera see Klamer (1998), especially chapters 1, 2, 3, and 5.

Kambera is a head-marking language. The subject and object(s) are marked on the verb by clitics that are cumulative exponents expressing number and person of the arguments as well as case. Only definite arguments trigger agreement on the verb and can be optionally realised, as pro-drop is possible in Kambera.[3] In contrast to this, indefinite arguments are never marked on the verbal head and cannot be omitted. The arguments themselves do not show any marking at all. The language distinguishes the four cases nominative, accusative, dative, and genitive. Exponents expressing nominative are realised in preverbal position as proclitics and exponents expressing one of the other cases are enclitcs appearing in postverbal position. In addition, the verb bears markers for aspect and mood that show up between the verb stem and the agreement marker for non-nominative exponents. But these won't be taken into account here. Kambera has two numbers, singular and plural. In addition to first, second, and third person it distinguishes between first person inclusive and exclusive in the plural.

Agentive arguments like the A-argument of transitive verbs and the S-argument of intransitives usually get nominative case. The prototypical object, the P-argument of transitives, receives accusative case.[4] The genitive is characteristic for possessors in nominal phrases and dative case marks recipients, goals and sources. This is shown in the examples (1) and (2) below. The terminology already implies that the language predominantly follows the accusative alignment pattern, especially in syntax,for example concerning controlled embedded clauses, raising constructions, and relative clauses. Nevertheless some minor parts of its morphology show ergative alignment, briefly discussed in footnote 7.

(1) (I Ama) `na-` kei `-nja` ri
ART father 3S.NOM- buy -3P.DAT vegetable
'Father bought them vegetables.'

[3]All of the definite arguments are marked on the verb as long as no constraint on possible combinations of two object clitics concerning their person is violated. Clusters of two third person object clitics, two times a first or second person object clitic, as well as a third person followed by a first or second person clitic are disallowed. All other sequences are grammatical.

[4]These letters have been introduced by Dixon (1972) as a notation for verbal arguments. S stands for the sole argument of intransitives, regardless of whether it is actively or inactively involved in the event described, since the grammar treats them equally (as it is the case in Kambera). A represents the argument realizing the agent in transitive contexts. There are two ways in which the non-agentive argument of transitives have been named: Dixon called it O, whereas Comrie(1989) used P for 'patient' and kept A and S for the other arguments.

(2) (Na tau wutu) **na-** palu **-ka** (nyungga)
ART person be fat 3S.NOM- hit -1S.ACC I
'The big fat man hit me.'

In (1) the definite arguments *Ama* and *nja* trigger agreement on the verb whereas the indefinite *ri* does not. The goal shows up in the dative, the agent in the nominative. In (2) both arguments are marked by clitics on the verb with P receiving accusative case.

3. Analysis

3.1. Data

Table 1 gives an overview of all the clitics that mark arguments on the verb.[5] Segments in brackets can be optionally realised depending on each speaker, mostly the shorter forms are used in speech. In the analysis, I will take into account the complete forms.

T1:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-ŋgu	-ka	-ŋga
2nd singular	(m)u-	-mu	-kau	-ŋgau
3rd singular	na-	-na	-ja	-nja
1st incl. plural	ta-	-nda	-ta	-nda
1st excl. plural	ma-	-ma	-kama	-ŋgama
2nd plural	(m)i-	-mi	-ka(m)i	-ŋga(m)i
3rd plural	da-	-da	-ha	-ndʒa

By looking at these agreement markers it is important to note that the forms get more and more complex from left to right. The dative markers consist of two syllables, the nominative and genitive forms of only one syllable and the accusative markers are mixed, only the first person exclusive and the second person plural are polysyllabic. The markers for nominative and genitive in the 1st person exclusive and 2nd person plural resemble the corresponding markers in the accusative and dative, so the first conclusion is that at least

[5]Some forms differ from those listed in Klamer (1998) because I use phonetic transcriptions of the clitics instead of the orthographic variant in the grammar.

the dative and accusative markers are not only built up of one vocabulary item, but that vocabulary insertion has taken place several times.

Another striking fact is that apart from the third person plural the dative forms seem to be derived from their accusative counterparts by adding a nasal triggering assimilation of voice of the following consonant that itself triggers assimilation of the nasal concerning its place of articulation. So progressive and regressive assimilation take place at the same time. All the dative markers begin with a nasal, fused to a prenasalised consonant if the following segment is a plosive or a semivowel. A similar situation occurs in the genitive, but here only first person singular and first person inclusive plural show the nasal and assimilation processes.

To generalize the facts, I hypothesise that the nasal is the same abstract phoneme in the dative and also *all* genitive markers and that it changes its actual form due to phonological processes. Secondly, the nasal is in fact a vocabulary item saturating a feature bundle that these two cases share. Consequently dative and genitive form a natural class with regard to this feature bundle. Immediately the question arises, why only some of the genitive forms realise this nasal. Therefore it is necessary to justify this hypothesis. Assuming it would be right and all the dative and genitive markers would begin with a nasal what, indeed, they do already in the dative. What would the situation look like in the genitive forms? The markers for first person exclusive plural, second person singular and plural and third person singular would start with a consonant cluster consisting of two nasals. But the language has a constraint that rules out onset clusters. So the reason why we cannot see the nasal in these forms is due to a phonotactic constraint that repairs the ill formed sequences either by eliminating the first nasal or directly by inserting it only in an environment where it is not followed by a nasal. In all the other forms the problem does not occur and the nasal can be added. Where it is followed by a plosive, affricate or semivowel both segments are fused to a prenasalised consonant counting only as one segment in phonetics so that the restriction against complex onsets is not violated. Now there is still one marker left, where a prenasalised segment would be expected but obviously does not occur: the third person plural genitive is just built up of a plosive and a vowel. I will argue in section 3.3.3 that this is due to further rules applying before vocabulary insertion.

By separating the nasal from the clitics and turning back the assimilations described, we get even more syncretic forms. The result is shown in table 2.

T2:[6]

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-N+ku	-ka	-N+ka
2nd singular	(m)u-	-mu	-kau	-N+kau
3rd singular	na-	-na	-ja	-N+ja
1st incl. plural	ta-	-N+ta	-ta	-N+ta
1st excl. plural	ma-	-ma	-kama	-N+kama
2nd plural	(m)i-	-mi	-ka(m)i	-N+ka(m)i
3rd plural	da-	-da	-ha	-N+dʒa

The whole argumentation can only be maintained if the adding of the nasal really triggers assimilation of voice of the following consonant, or more precisely, of the following plosive, because the glide /j/ in /nja/ already has the same voice as the nasal; and the consonant must trigger assimilation of the nasal concerning its place of articulation. If all these processes did not take place, there would yet be more markers, for example /ku/ and /gu/ in the first person singular nominative and genitive as well as /n/ and /ŋ/. Because of their articulatory resemblances it would be nice to put them down to one abstract marker /ku/ and /N/ and in that way reducing the number of vocabulary items. Even if assimilations are one of the most regularly applying phonological processes in the languages of the world, is there any reason justifying its application here? Kambera shows a process that derives a certain class of intransitives from transitives by prenasalising the word-initial plosive. This always triggers progressive assimilation of the plosive that gets voiced and the nasal regressively assimilates to the plosive's place of articulation. Examples are pairs like

pata 'break X' → mbata 'be broken'
tutu 'roll X' → ndutu 'roll over'
kodang 'move X' → ŋodang 'be loose (e.g. tooth)'

As the process is not productive anymore, it cannot be applied to all transitives of the relevant class today.
I would like to suggest that /kama/ and /kami/ in the first person exclusive and the second person plural of the accusative, respectively, should be subanalysed further because obviously /ma/ and /mi/ are markers for the

[6]In what follows, I will write /N/ as the abstraction of all nasals independent from their place of articulation, because it is not clear which of the actually realised nasals represents the base the other forms are derived from via assimilation.

corresponding person and number features independent from case so that /ka/ has to add something that characterizes accusative and dative as it only shows up in these two cases. Again I hypothesize that consequently accusative and dative have to form a natural class concerning a feature bundle realised by the vocabulary item /ka/ and that, in order to receive a unified picture of the nasal /N/ and /ka/, originally all the clitics realising these two cases should show /ka/ immediately in front of the last syllable. Of course now there cannot be a phonological reason explaining why /ka/ is not realised in the third person and the first person inclusive. The concept of Fission will provide reasons for the lack of /ka/, outlined in section 3.3.3.

After cutting off /ka/, too, we get the following vocabulary items in table 3.

T3:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-N+ku	-ka(+ø)	-N+ka(+ø)
2nd singular	(m)u-	-mu	-ka+u	-N+ka+u
3rd singular	na-	-na	-ja	-N+ja
1st incl. plural	ta-	-N+ta	-ta	-N+ta
1st excl. plural	ma-	-ma	-ka+ma	-N+ka+ma
2nd plural	(m)i-	-mi	-ka+(m)i	-N+ka+(m)i
3rd plural	da-	-da	-ha	-N+dʒa

Ignoring the items /N/ and /ka/ which have argued to be separable for the time being, the remaining markers are listed in table 4. Among these markers there are actually no syncretisms between different persons, but many instances of syncretism show up between cases, only /ha/ and /dʒa/ appear just once. Furthermore the syncretic forms show a relatively unified distribution.

T4:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-ku		
2nd singular	(m)u-	-mu	-u	-u
3rd singular	na-	-na	-ja	-ja
1st incl. plural	ta-	-ta	-ta	-ta
1st excl. plural	ma-	-ma	-ma	-ma
2nd plural	(m)i-	-mi	-(m)i	-(m)i
3rd plural	da-	-da	-ha	-dʒa

The markers for first and second person plural are the same for all cases.

The corresponding vocabulary items thus clearly do not refer to features that cases are built up of. In all the other persons two cases group together and receive one form. These are always nominative and genitive on the one hand, and accusative and dative on the other.

3.2. Feature decomposition

To sum up what has been stated so far, nominative and genitive should form a natural class because of the grouping of the markers in T4, accusative and dative must have a feature in common due to the same reason and the fact that the vocabulary item /ka/ distributes only in these two cases. Finally, genitive and dative should form a natural class marked by /N/. There is no evidence for genitive and accusative or nominative and dative to share feature bundles. All this is reflected in the following case decomposition:

Case
Nominative [–obl, –obj]
Genitive [+obl, –obj]
Accusative [–obl, +obj]
Dative [+obl, +obj]

The abbreviations [obl] and [obj] stand for 'oblique' and 'object'.[7] [+oblique] designates cases that are typically neither used for agent nor patient arguments, which receive nominative and accusative respectively in Kambera, so that genitive and dative constitute the oblique cases. [+object] describes cases that can never mark agent arguments. These are accusative and dative. In contrast to this, nominative and genitive can both mark agent arguments. The nominative is the prototypical case for sole argument of intransitive and the agent argument of transitive verbs in Kambera. The genitive has the same function in special contexts. Look at the examples below:

(3) Na kabela [$_{CP}$ na pa$_1$- piti- na na tau nuna t$_1$]
ART machete ART REL- take- 3S.GEN ART person DEI.3S
na- ruhak
3S.NOM- be broken
'The machete that this man took is broken.'

[7] The decomposition of cases into smaller features go back to Jakobson (1962a) and Bierwisch (1967).

(4) Mbada laku `-na` -ka
already go -3S.GEN -PRF
'He has already gone.'

(5) Mary's picture of John

Sentences in which the genitive clitic is used to mark A or S as in (4) do not express actions but rather describe "a circumstantial state of affairs" (Klamer(1998), p. 64/65). This means that the information in (4) does not emphasize the action that someone has gone, but it is a fact needed to describe the circumstances of some other action that is the topic of conversation. Furthermore, sentences in which an P is relativised mark the embedded subject with the genitive instead of nominative. This is shown in example (3), where *na kabela* is the patient of the action described by *piti*. The prefix *pa-* indicates that the patientive argument is relativised, an agentive one would demand another prefix. Now the subject *na tau* is marked on the verb by *na*, the marker for third person genitive. Accusative or dative clitics, however, are not used to mark A or S.[8] That the genitive marks agents is not extraordinary. A similar situation is found in English within NPs. In example (5) Mary can be the agent, the one who took the picture

[8]There do exist constructions in which the dative and accusative forms mark A or S-arguments. These are the parts of Kambera's morphology that show ergative alignment. I will give a short overview of the most important constructions discussed in more detail in Klamer (1998, chapter 5). Only declarative sentences are surveyed.
In the 'double-S' construction, the single argument of an intransitive verb is expressed twice by the expected nominative clitic and the corresponding accusative clitic. Speakers *can* apply it to express moods and certainty about an event. Its application is restricted mainly to religious and poetic texts. The use of this construction decreases and most examples are considered to be archaic today.
In the 'absolutive' construction, an accusative clitic encoding the S-argument represents an impersonal pronoun or shows up when the verb is left-dislocated. In this context, the action expressed by the verb is emphasized with regard to the agent argument. But these constructions are less frequent and mostly not obligatory.
The'accusative'-S construction marks S with the accusative clitic and indicates that the argument is less involved in the action expressed by the verb; it describes rather a situation than an action. Furthermore it is a marked construction; the S-argument can also be encoded by the nominative clitic in these contexts, depending on semantics.
These special cases are a consequence of the language's change from the ergative-absolutive to the nominative-accusative pattern. We find markings from an earlier stage in Kambera's history that do not obey the generalizations mentioned above, but as they show up less frequently, require very special contexts and depend on semantics, and some of them are archaic, these constructions do not falsify the generalization that nominative and genitve form a natural class excluding and accusative.

with someone's camera and is marked in a nominal phrase with the genitive. All these facts justify the proposed decomposition of cases, in which nominative and genitive form a natural class. This procedure will be repeated with person and number decomposition below. In order to decompose four persons, the two features [±1] and [±2] are necessary. The two numbers only need one feature. No possible feature combination is left unused.

Person	**Case**	**Number**
1st person [+1, –2]	Nominative [–obl, –obj]	singular [+sing]
2nd person [–1, +2]	Genitive [+obl, –obj]	plural [–sing]
3rd person [–1, –2]	Accusative [–obl, +obj]	
inclusive person [+1, +2]	Dative [+obl, +obj]	

On this basis, (6) provides the features that the segments /N/ and /ka/ have to realise.

(6) /N/ ↔ [+obl] /__ [–nasal]
/ka/ ↔ [+obj]

The feature accusative and dative share is [+obj], whereas genitive and dative share [+obl] with the restriction that /N/ can only be inserted in a context where it is followed by a segment with the feature [–nasal].

3.3. Vocabulary insertion

3.3.1. Vocabulary items

Here are all the vocabulary items relating phonological information and morpho-syntactic features; most of them are underspecified. I will discuss these vocabulary items and what conclusions can be drawn from them in section 3.4.

(7) a. /ha/ ↔ [–1, –2, –sing, –obl, +obj]
b. /dʒa/ ↔ [–1, –2, –sing, +obj]
c. /da/ ↔ [–1, –2, –sing]
d. /ta/ ↔ [+1, +2, –sing]
e. /ja/ ↔ [–1, –2, +obj]
f. /ku/ ↔ [+1, +sing, –obj]
g. /mu/ ↔ [+2, +sing,–obj]
h. /na/ ↔ [+sing, –obj]

i. /ma/ ↔ [–2, –sing]
j. /mi/ ↔ [+2, –sing]
k. /u/ ↔ [+2]
l. /ka/ ↔ [+obj]
m. /N/ ↔ [+obl] /__ [–nasal]

3.3.2. Assumptions

The important concepts necessary for vocabulary insertion – Subset Principle, Specifity, and Underspecification have been defined in section 2.1.

A fourth assumption concerns the concept of fission. According to the definitions above, only one vocabulary item can be inserted in a terminal node, but fission ensures that several ones can be put in. It is important to keep in mind that this analysis will be based on a notion of fission that differs from the standard way in which this concept was defined at the beginnings of Distributed Morphology by Halle & Marantz (1993). I will apply fission in the sense of Noyer and Frampton.

Fission (Noyer(1992), Frampton (2002)):
If insertion of a vocabulary item V with the morpho-syntactic feature β takes place into a fissioned morpheme M with the morpho-syntactic feature α, then α is split up into β and $\alpha - \beta$, such that (i) and (ii) hold:
(i) $\alpha - \beta$ is available for further vocabulary insertion.
(ii) β is not available for further vocabulary insertion.

This means that the insertion of a vocabulary item deletes the corresponding features of the terminal node and consequently no other vocabulary item with at least one of the deleted features can be inserted afterwards, because it could never fulfil the Subset Principle. But the features that have not been deleted in the terminal node after the first insertion are free for insertion of other vocabulary items as long as there are still undeleted features.

3.3.3. Insertion and further rules

First, the syntactic positions in which the items will be inserted need to be defined. The whole forms consist of the stem and a clitic in preposition to the stem. This clitic is subanalysed in up to three units: /N/, /ka/, and /X/, where X stands for any of the other items. None of these subparts has

to be obligatorily present in all of the clitcs, but at least one of them has to be realised. The stem constitutes the verbal head in the syntactic structure and insertion of the other items takes place into functional heads W, Y, and Z, which the verb raises to. This results in the structure [$_W$/N/[$_Y$/ka/[$_Z$/X/ stem]]].

Note that the following insertion of vocabulary items does not need any feature hierarchy or ranking of rules. The right distribution of markers follows automatically. I will begin with the most specific item; finally /N/ and /ka/ will be inserted.

The only fully specified marker /ha/ is inserted in the third person plural accusative. Afterwards /dʒa/ would fit in the same person accusative and dative. Since /ha/ is more specific than /dʒa/, it blocks its insertion in the accusative. /da/ marks all third person plural contexts, but the dative and accusative are already occupied, so only nominative and genitive are left over for /da/. /ta/ stands for first person inclusive in all cases and is not blocked by any other vocabulary item. Next /ja/ is inserted in all remaining third person accusative or dative contexts – this concerns only the third person singular since the plural is already filled. /ku/ marks the first person and /mu/ the second person singular nominative and genitive without restrictions. Afterwards /na/ fills all terminal nodes in nominative and genitive singular contexts that are left. It cannot overwrite /ku/ and /mu/ because both are more specific. Then /ma/ is inserted in all first person exclusive plural contexts and /mi/ in all second person plural contexts. The last vocabulary item /u/ can only show up in the second person singular accusative and dative. Although /u/ is the least specific vocabulary item, it really only fits in two terminal nodes. In all the others the feature [+2] has already been deleted by a more specific item or the relevant feature has the opposite value. Up to this point there is no item for the first person singular accusative and dative. The result is shown in table 5; all the other markers that would have fit as well but were less specific are added in brackets.

T5:

	Nominative	Genitive	Accusative	Dative
1st singular	ku (na)	ku (na)		
2nd singular	mu (na, u)	mu (na, u)	u	u
3rd singular	na	na	ja	ja
1st incl. plural	ta (mi, u)g	ta (mi, u)	ta (mi, u)	ta (mi, u)
1st excl. plural	ma	ma	ma	ma
2nd plural	mi (u)	mi (u)	mi (u)	mi (u)
3rd plural	da (ma)	da (ma)	ha (dʒa, da, ja, ma)	dʒa (da, ja, ma)

Now the vocabulary item /ka/ can be inserted in terminal nodes that still have the feature [+obj]. All dative and accusative contexts originally had this feature but where the items /ha/, /dʒa/ and /ja/ (all in the third person) have been inserted, /ka/ cannot fulfil the Subset Principle because these items referred to the feature [+obj] and deleted it by their insertion. That is why /ka/ can only appear in first and second person contexts. However, it does not appear in the first person inclusive plural, although the feature [+obj] is still at disposal. As there does not seem to be another plausible reason in Kambera's grammar for this, the impoverishment rule (8) is needed. It substitutes the feature [+obj] by in the context [+1, +2] before vocabulary insertion,that means that [+obj] is deleted without vocabulary insertion.[9] As a consequence, the feature bundle of /ka/ is not a subset of the features of the relevant terminal node and therefore fails to be inserted.

Impoverishment rules:

(8) [+obj] ↔ / [+1, +2]

(9) [+obl] ↔ / [–1, –2, –sing, –obj]

Finally, only the nasal is left over. It phonetically realises the feature [+obl] that is present in all genitive and dative terminal nodes. As none of the other vocabulary items already referred to this feature, /N/ can freely be inserted if the context is right, that means if the constraint against complex onsets remains respected. In the third person plural genitive a prenasalised consonant would be expected but does not occur. Again there seems to be no phonological reasons that could block the appearance of the nasal. Therefore impoverishment must have been at work. Rule (9) deletes the feature [+obl] in the third person plural genitive and the nasal cannot be inserted.

What we get after the complete vocabulary insertion is again table 3, repeated under 6. Where impoverishment has taken place, the clitics are

[9]For an alternative approach in the framework of Minimalist Distributed Morphology see Trommer (2003), who argues that there is only one operation, vocabulary insertion, which the other operations in Distributed Morphology can be reduced to. Impoverishment that leads to suppression of phonological material corresponds to insertion of a high specific empty vocabulary item that deletes features. In the relevant example it could be ↔ [+obj] / [+1, +2].

set in boldface.

T6:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-N+ku	-ka(+ø)	-N+ka(+ø)
2nd singular	(m)u-	-mu	-ka+u	-N+ka+u
3rd singular	na-	-na	-ja	-N+ja
1st incl. plural	ta-	-N+ta	***-ta***	***-N+ta***
1st excl. plural	ma-	-ma	-ka+ma	-N+ka+ma
2nd plural	(m)i-	-mi	-ka+(m)i	-N+ka+(m)i
3rd plural	da-	***-da***	-ha	-N+dʒa

Afterwards the phonological processes described in section 2.2. apply, changing voiceless obstruents to voiced obstruents, that means only plosives in the case of clitics in Kambera's argument encoding system.

Phonological rule:

(10) [+cons, –cont, –voiced] → [+voiced] / [+nasal] __

This results in table 1 from the beginning of the analysis, repeated here in table 7.

T7:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-ŋgu	-ka	-ŋga
2nd singular	(m)u-	-mu	-kau	-ŋgau
3rd singular	na-	-na	-ja	-nja
1st incl. plural	ta-	-nda	-ta	-nda
1st excl. plural	ma-	-ma	-kama	-ŋgama
2nd plural	(m)i-	-mi	-ka(m)i	-ŋga(m)i
3rd plural	da-	-da	-ha	-ndʒa

At this point, I come back to the fact of increasing complexity of the clitcs from nominative to dative. As the dative is the only case that is decomposed into the features [+obl] and [+obj] and /ka/ and /N/ refer to the positive values of these features, both vocabulary items can be inserted, resulting in the complex dative clitics. In contrast to this, the corresponding features of the nominative both have negative values and none of the two vocabulary

items shows up in the nominative clitics, representing the least complex forms. Genitive and accusative case both have one of these features with a positive value, so that only /ka/ or /N/ can appear, if not blocked by other rules.

3.4. Remarks on vocabulary items

For ease of exposition the vocabulary items are repeated here:

a. /ha/ ↔ [–1, –2, –sing, -obl, +obj]
b. /dʒa/↔ [–1, –2, –sing, +obj]
c. /da/ ↔ [–1, –2, –sing]
d. /ta/ ↔ [+1, +2, –sing]
e. /ja/ ↔ [–1, –2, +obj]
f. /ku/ ↔ [+1, +sing, –obj]
g. /mu/ ↔ [+2, +sing,–obj]
h. /na/ ↔ [+sing, –obj]
i. /ma/ ↔ [—2, –sing]
j. /mi/ ↔ [+2, –sing]
k. /u/ ↔ [+2]
l. /ka/ ↔ [+obj]
m. /N/ ↔[+obl] /__ [–nasal][10]

[10]Bobaljik (2002) argues that a morpheme's features can condition allomorphy of a morpheme that is either farther away from or closer to the root than these features and that the direction of the this sensitivity is systematically derivable from the kind of features that trigger the allomorphy. If the morpheme is closer to the root, it is said to be *inwards sensitive* and if it is further away from the root it is called *outwards sensitive*. The crucial point is that outwards senitivity can only be triggered by morpho-syntactic features, whereas inwards sensitivity is exclusively triggered by syntactically irrelevant features such as phonological or class features. According to the syntactic structure presented in this section, the insertion context of /N/ supports this analysis because the allomorphy of /N/ – its realization or its absence, as well as its place of articulation if it is realized – depends on the features of a morpheme that is closer to the root, e.g. /ka/ or any /X/. This inwards sensitivity is clearly conditioned by their phonological features [–nasal] and [αplace], where the latter is an abreviation for any feature of place the following item has and the nasal assimilates to. However, Kambera also challenges the strict formulation of this correlation. As described in section 3.1, the initial consonant of the item /ka/ assimilates its voicing to the more peripheral /N/ with the feature [αvoice]. This sensitivity is outwards but depends on phonological features, or put different, a phonological feature determines the form of a more inwards morpheme and thus clearly

Altogether thirteen vocabulary items are needed, with which all syncretisms can be derived. This is the minimal number, no marker occurs twice. One vocabulary item is completely specified, all the others are underspecified. Furthermore, no hierarchy among the features is necessary, the insertion takes place without conflicts between competing vocabulary items. It might seem redundant to specify the item /ta/ for number, because the feature bundle [+1, +2] unambiguously marks the inclusive person that only appears in the plural. But without [–sing] /ta/ would compete with /mi/ for insertion in the inclusive person. As both have two features, none of them could be chosen by the Specifity Principle - they are equally specific.

Besides, if /ta/ realises three features, it fits well in a pattern in the list of vocabulary items, apart from /ka/ and /N/. On the whole, vocabulary items beginning with an obstruent or glide (a. – f.) are more specific than the items starting with a nasal (g.– j.), except for /mu/. The least specific one is a bare vowel. This leads to the following correlation between specifity of vocabulary items and sonority: The more specific an item, the less sonorous it is, according to the hierarchy of sonority.

Hierarchy of sonority: obstruents ≫ nasals ≫ vowels

This phenomenon is known as iconicity in the sense of Wiese (1999): There is a correlation between form and function of the vocabulary items. Their order in terms of specifity corresponds to their order concerning sonority. Obstruents are more sonorous than nasals, these are more sonorous than vowels and this is reflected in their specifity, that means the number of features every item has. First of all it supports the assumption that iconicity might be a meta- grammatical principle which grammars obey and are organised after. And second, as iconicity only shows up among the markers except for /ka/ and /N/, it provides another motivation for assuming that these two markers have a special status, realizing only the two features [+obj] and [+obl], so that they can be treated separately from the rest of the other vocabulary items, as they do not fit in the patterns of syncretism in table 4.

Furthermore, there is no elsewhere marker without any features that would fit in everywhere as long as there does not exist a more specific marker: ↔ [ ]. Perhaps it would be attractive to assume such a phonologically empty vocabulary item for the first person singular accusative and

---

contradicts Bobaljik's assumption that "'morphophonological diacritics ... can only serve as the context for allomorphy of a more peripheral affix."' (Bobaljik (2002), p. 14).

dative in table 6, where apart from /ka/ or /N/ none of the other items can be inserted. But of course nothing in the analysis would change if it were there. I rejected this possibility for the following reasons. A conflict appears in connection with the adopted definition of Fission. If the elsewhere marker fits in everywhere, where features of the terminal node are not yet deleted by insertion of another vocabulary item, it should be inserted in every terminal node except for the third person plural accusative as /ha/ is not underspecified. This seems to lead to a flood of phonological empty categories, whose existence cannot be substantiated. At the same time the question arises, what exactly should be inserted for, that means, which features should be deleted after its insertion. Only one feature, resulting in up to three elsewhere markers for one terminal node, for example the second person singular accusative, where only two features are deleted during vocabulary insertion, or should it be inserted for the whole remaining feature bundle of a terminal node? So in general, one who wants to have this sort of elsewhere marker could use it, but will have to introduce several restrictions on its use, e.g. that it could only be inserted once per terminal node. As I cannot find any motivation for this marker, I do not assume that such a vocabulary item exists. Its existence would cause a lot of additional problems, but as the analysis works well without it, I chose the simpler variant. On the other hand, I do not see any serious problem by claiming that not all features of terminal nodes need to be deleted during vocabulary insertion. They are simply left over. It would be interesting to investigate if other processes refer to exactly these features that are predicted not to be deleted.

The appearance of the vocabulary item /ka/ could be described in another way by making use of alpha notation, where variables range over concrete feature values. The exact value of every single feature does not matter, it is important if two or more features share the same value or not, expressed by variables.[11] For the distribution of /ka/ in this example, this yields /ka/ $\leftrightarrow$ [α1, $-\alpha$2, +obj], referring to contexts in which the person features [1] and [2] have different values, like [+1, –2] or [–1, +2]. As the third and first person inclusive are decomposed in [α1, α2], /ka/ cannot be inserted in the corresponding terminal nodes. An advantage of this analysis is that no impoverishment rule for the inclusive person is needed to get the right distribution. But in my opinion this is just a way out to refer to a class that actually is not a real natural class. Natural classes originate di-

[11] The alpha notation has been introduced by Chomsky (1965).

rectly from feature decomposition. By using alpha notation, all persons and cases referred to in this analysis could form a natural class in that way, for example nominative and dative – there is no reason for this. Why should it just be necessary for person features and only relevant for one vocabulary item /ka/? Its application seems quite arbitrary.

The last problem that needs to be solved concerns the linear order of the vocabulary items /N/ + /ka/ + /X/. The order does not follow from any principle yet. A simple answer would be to fix the order of features in the terminal node to [±obl, ±obj, ...] which the inserted vocabulary items keep, or assume a feature hierarchy obl ¿ obj ¿ ... , but there seems to be no deeper explanation for that. A more attractive approach could be increasing complexity from left to right, what resembles Behagel's Law of growing constituents (Behagel (1932)) for syntax, claiming that from a phonological point of view, items increase in complexity from left to right: /N/ is less complex than the other items. We can see the same with incresing specifity of vocabulary items in the syntactic structure – the more specific an item, the closer it is situated to the stem: $[_W$/N/$[_Y$/ka/$[_Z$/X/ stem]]]. All markers except /N/, /ka/ and /u/ realise more features than these, for example /mi/ realises two features, /ka/ only one. In conclusion /ka/ precedes /mi/ and every other /X/. There is a conflict between these items and /u/ as all of them only represent one feature, and their contexts overlap in the second person singular accusative and dative. Probably the language just keeps the order that results in the clear cases as above with /mi/ for reasons of uniformity, so that /ka/ appears in front of /u/, too. The nasal must occur before /X/ for the same reasons of complexity.

As vocabulary insertion starts with the most specific items and ends by inserting the least specific ones and always puts them into the head to the left of the structure that has already been built up, insertion takes place in a strictly cyclic way: /X/ is inserted at first and preceeds the stem, /N/ and /ka/ are inserted afterwards and preceed /X/. But then the question is what happens, when /N/ and /ka/ show up both in some dative forms, because, putting aside the context restriction on /N/, they are equally complex. Why does /N/ precede /ka/? First of all, it is phonologically light in contrast to /ka/, and light material often tends to precede heavy material. Second, if it followed /ka/ we would create forms in which again two nasals were next to each other, like *[kanmi] for the second person plural dative. As Kambera does not have codas, the nasal cannot be part of the first syllable. On the other hand, it cannot form the onset together with /m/ as complex onsets are forbidden as well. To conclude, the order /ka/+/N/+/X/ would result in an impossible syllable structure and the opposite order wins: /N/+/ka/+/X/.

4. Conclusion

To conclude, I will summarize the main insights of this article. Kambera provides a lot instances of syncretism among the clitics encoding definite arguments on the verb. All of them can be derived with a small set of mechanisms provided by Distributed Morphology, especially Underspecification of vocabulary items, Fission, allowing vocabulary insertion to take place several times per terminal node, and Impoverishment that deletes features and helps to account for absent forms that would be expected to appear. This suggests that syncretisms do not exist accidentially but are systematic: the contexts that the homonymous forms show up in share features that vocabulary items refer to. First, phonological processes had to be turned back to discover all sycretic forms, and the majority of clitics has been subanalysed into smaller units.

The most impotant insight is that these units show a correlation between sonority and specifity. This supports the assumption that iconicity could be another meta-grammatical principle, especially, as this correlation has until now only been stated for morphological systems of Indo-European languages, e.g. by Wiese (1999) for pronominal inflection in German, by Müller (2004) for Russian noun inflection and by Müller (2005) for Icelandic noun declensions. Kambera, a Malayo-Polynesian language, shows that this phenomenon is not restricted to European languages and offers another argument that morphological systems are not built up by chance, but obey deeper grammatical principles.

Bibliography

Behaghel, Otto (1932): Deutsche Syntax IV. Carl Winters, Heidelberg.

Bierwisch, Manfred (1967): Syntactic Features in Morphology: General Problems of So-Called Pronominal Inflection in German. In: *To Honour Roman Jakobson.* Mouton, The Hague/Paris, pp. 239-270.

Bobaljik, Jonathan David (2002): The Ins and Outs of Contextual Allomorphy. In: K. Grohmann & C. Struijke, eds., *University of Maryland Working Papers in Linguistics.* Vol. 10, University of Maryland, College Park, pp. 35-71.

Bonet, Eulália (1991): Morphology after Syntax. PhD thesis, MIT, Cambridge, Mass.

Chomsky, Noam (1965): Aspects of the Theory of Syntax. Cambridge, Mass.: MIT Press.

Comrie, Bernard (1989): Language Universals and Linguistic Typology. 2 edn, Blackwell, Oxford.

Dixon, R.M.W. (1972): The Dyirbal Language of North Queensland. Cambridge University Press, Cambridge.

Frampton, John (2002): Syncretism, Impoverishment, and the Structure of Person Features. In: M. Andronis, E. Debenport, A. Pycha & K. Yoshimura, eds., *Papers from the Chicago Linguistics Society Meeting.* Vol. 38, Chicago, pp. 207-222.

Halle, Morris & Alec Marantz (1993): Distributed Morphology and the Pieces of Inflection. In: K. Hale & S. J. Keyser, eds., *The View from Building 20.* MIT Oress, Cambridge, Mass., pp. 111-176.

Halle, Morris & Alec Marantz (1994): Some Key Features of Distributed Morphology. In: A. Carnie, H.Harley & T. Bures, eds., *Papers on Phonology and Morphology.* Vol. 21 of MIT Working Papers in Linguistics, MITWPL, Cambridge, Mass., pp. 275-288.

Harley, Heidi & Rolf Noyer (1999): Distributed Morphology, *GLOT International* 4/4, pp. 3-9.

Jakobson, Roman (1962a): Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus. In: *Selected Writings.* Vol. 2, Mouton, The Hague and Paris, pp. 23-71.

Klamer, Marian (1998): A Grammar of Kambera. Mouton de Gruyter, Berlin, New York.

Müller, Gereon (2004): A Distributed Morphology Approach to Syncretism in Russian Noun Inflection. In: O. Arnaudova, W. Browne, M. L. Rivero & D. Stojanovic, eds., Proceedings of FASL 12. University of Ottawa.

Müller, Gereon (2005): Syncretism and Iconicity in Icelandic Noun Declensions: A Distributed Morphology Approach. In: G. Booji & J. van Marle, eds., *Yearbook of Morphology 2004.* Springer, Dordrecht, pp. 229-271.

Noyer, Rolf (1992): Features, Positions, and Affixes in Autonomous Morphological Structure. PhD thesis, MIT, Cambridge, Mass.

Trommer, Jochen (2003): Feature (non)-insertion in a minimalist approach to spellout. Proceedings of CLS39: Main Session. Chicago.

Wiese, Bernd (1999): Unterspezifizierte Paradigmen. Form und Funktion in der pronominalen Deklination, *Linguistik Online*4. (www.linguistik-online.de/3_99)