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 whose form varies between person, number and case. These markers show many instances of syncretism below the clitic level. The aim of this article is to derive the distribution of the homonymous inflection markers and thereby to motivate that syncretic forms do not appear accidentally, as stated in the *Syncretism Principle* in Müller (2004). The main insight is that we find a correlation between the items' specificity, that means the number of features they realize, and their sonority – a principle called *Iconicity* by Wiese (1999)– that has until now exclusively been stated for some Indo-European languages. This fact strengthens the hypothesis that iconicity qualifies as a meta-grammatical principle for the architecture of morphological systems.

1 Introduction

In this paper I will survey the morphology of argument encoding in the Malayo-Polynesian language Kambera. In this language arguments are encoded on the verb by clitics that show many instances of syncretism, especially if they are subanalysed into smaller units below the clitic level, for which I will argue. The goal of this procedure is to derive the distribution of homonymous inflection markers and thereby to strengthen the hypothesis that syncretic forms do not distribute arbitrarily but systematically, as formulated in the meta-grammatical Syncretism Principle in Müller (2004):

Syncretism Principle: Identity of form implies identity of function.

Furthermore, I will argue for the need of subanalysis in order to detect a correlation between specificity of markers and their sonority, which was called *Iconicity* by Wiese (1999). His definition of this term deviates from the traditional notion. I will proceed as follows: first, I provide basic information about

Kambera and the theoretical framework of Distributed Morphology and then I go through the data and the analysis.

2   Background information

2.1   Kambera

Kambera is a Malayo-Polynesian language spoken in the east of the Indonesian island Sumba. There is 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), from which the language facts and examples are cited.

Only definite arguments, subject and object(s), are encoded on the verb by clitics expressing number and person of the arguments as well as case.[1] The arguments are not marked for any category. Head-marking, no case marking on arguments and the fact that the verb and the pronominal markers can form a complete sentence without the full NPs lead to the conclusion that grammatical relations in Kambera are expressed by cross-reference and not by agreement (cf. the distinction in Nichols (1986).

The language distinguishes the cases nominative, accusative, dative, and genitive. As one can conclude from these labels, Kambera exhibits nominative accusative alignment with respect to case marking. Nominative marks the sole argument of an intransitive verb and the agent argument of a transitive one. The patient argument of transitives receives accusative. Dative realizes goals, recipients, and sources, whereas the genitive marks possession.

Exponents expressing nominative case show up in preverbal position as proclitics and clitics expressing one of the other cases are enclitics.[2] These facts are illustrated by the following examples (Klamer (1998), p. 41).

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

---

[1] 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. The clusters are grammatical iff the first clitic is 1st or 2nd person and the second clitic is 3rd person. If a cluster does not correspond to this requirement, only the indirect object is cliticised (Klamer(1998), p.64).

[2] In addition, the verb bears markers for aspect and mood that show up between the verb stem and the non-nominative clitics. But these won't be taken into account here.

(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 verb *kei* has three arguments, but only two of them are encoded by clitics, because *ri* is indefinite. The recipient is marked by a dative enclitic and the agent by a nominative proclitic. If the arguments are encoded on the verb, they are usually dropped and are only used for emphasis and disambiguation. In (2) we see that the patient of a transitive verb gets the accusative case.

    T_1 provides an overview of all pronominal clitics.[3] Segments in brackets can be optionally realised, mostly the shorter forms are used in speech. I will take into account the complete forms in the analysis.

T_1:

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

As we can see in T_1, the language distinguishes two numbers, singular and plural, and first person inclusive and exclusive in the plural. Furthermore, there are already some instances of syncretism, e.g. in the third person singular nominative and genitive or in the corresponding plural cases, but I will argue there to be a lot more syncretic forms.

2.2   Distributed Morphology

In order to derive the distribution of the homonymous forms, a theoretical framework is needed that makes use of underspecification. One is Distributed Morphology whose basic assumptions are outlined below.

    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 (Halle & Marantz (1993), Halle & Marantz (1994)). It is assumed that phonological material, vocabulary items abbreviated as VI, is inserted post-syntactically into terminal nodes. That means

---

[3] Some forms differ from those listed in Klamer (1998), p. 62, because I use phonetic transcriptions of the clitics according to the notation in table 2.1. in Klamer (1998), p. 10.)

syntax operates on feature bundles. A vocabulary item as in (3) relates phonological information, represented by /Y/, and morpho-syntactic features [A, B, C], which constitute their insertion context. These features arise through decomposition of grammatical categories, e.g. case, person, and number[4].

(3) $/Y/ \leftrightarrow$ [A, B, C]

Before vocabulary insertion, rules may apply in order to manipulate the output of syntax. A mechanism that is needed to derive Kambera's argument encoding system is deletion of features in terminal nodes, which I will define more precisely at the relevant step of the analysis. As a result, the set of vocabulary items that theoretically could be inserted changes.

Most of the VI are underspecified with respect to their insertion context, that means that they do not have to realize all the features of a terminal node, but only a subset. As a consequence, one VI may encode features that several terminal nodes possess and logically could fit in all of them: this results in syncretism.[5]

Another consequence of underspecification is that also several VIs could be inserted into a single terminal node. The competition among the items is resolved by two concepts:

Subset Principle[6]: 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.

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 into according to the Subset Principle than Y does.

---

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

[5] I would like to emphasize that the analysis is not framework-specific. We need a mechanism to derive (as we will see *partial*) syncretism and the possibility to change the output of syntax or a method to exchange an expected VI. In Paradigm Function Morphology (Stump (2001)), the former can be derived by rule blocks and the latter by rules of referral. In an optimality-theoretic approach like Wunderlich (2004), the deletion of features can be handled with the help of a faithfulness constraint, whose violation may not be fatal in a special context as a consequence of interaction with higher ranked constraints.

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

These definitions ensure that the inserted VI is the one that shares more features of the terminal node than any other VI does[7].

3 Analysis

3.1 Data

First of all, the relevant vocabulary items have to be isolated. Consider T_1.

T_1:

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

The dative forms seem to be derived from their accusative counterparts by adding a nasal. If this nasal is separated, we get forms like *ka* for first person singular accusative and *ga* for the same context dative. Because of their articulatory resemblance it would be nice to put them down to one vocabulary item and so to minimize the number of VIs. Therefore, it has to be proven that assimilations took place. Evidence comes from a process that derives intransitive verbs from transitives by adding a nasal as in (4) (Klamer (1998), p.263).

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

In this configuration the nasal always assimilates its place of articulation to the following plosive, affricate or semivowel, whereas the nasal triggers assimilation of the following consonant with respect to voice. Such clusters of nasal+consonant are called prenasalised segments. T_2[8] shows the segments of

---

[7] Note that these definitions do not ensure that one and only one VI wins the competition. There may still be several VIs for one terminal node, if they obey the *Subset Principle* and are equally specific. But this situation won't occur in the following analysis.

[8] In what follows, I will write *N* for the nasal to abstract away from its actual place of articulation, because it is not clear from the data which one constitutes the underlying form.

each clitic before the assimilations have taken place.

T$_2$:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-N+gu	-ka	-N+ga
2nd singular	(mu)-	-mu	-kau	-N+gau
3rd singular	na-	-na	-ja	-N+ja
1st in. plural	ta-	-N+da	-ta	-N+da
1st ex. plural	ma-	-ma	-kama	-N+gama
2nd plural	(m)i-	-mi	-ka(m)i	-N+ga(m)i
3rd plural	da-	-da	-ha	-N+dʒa

A similar situation occurs in the genitive, but here the nasal appears only in first singular and first person plural inclusive. As *N* in genitive and dative should be one and the same vocabulary item, but it is quite difficult to describe exactly all the contexts, where *N* shows up in terms of features, another explanation for its distribution has to be found. The idea is that the nasal realizes a feature bundle that these two cases share and should actually show up in all the corresponding clitics. To explain why it does not, Kambera's syllable structure has to be taken into account. Syllables consist of at least one vowel and are always open[9]. There can be an onset consonant: (C)V. It is important to note that due to a phonotactic constraint the onset cannot be complex. The prenasalised segments count as one phonological unit and therefore do not violate this restriction. If the nasal was inserted in all the genitive forms, syllables with two initial consonants would be created, e.g. in the second person singular, in contrast to what is required by the phonotactic constraint. Either these forms are ruled out at spell-out or they are not even created by including the restriction in the insertion context of the relevant vocabulary item. The second strategy will be applied in section 3.3.

Another sequence that can be separated is *ka* in some accusative and dative contexts, as in T$_3$. Again I hypothesize that this VI realizes a feature bundle concerning which accusative and dative form a natural class. I will account for its absence in third person and first person inclusive plural in section 3.3. Ignoring the items *N* and *ka*, the forms in T$_4$ are left.

---

[9] It might seem that Kambera has closed syllables, too, e.g. if morphemes are added during derivational processes, like for applicatives. But epenthesis of a paragogic [u] takes place after the consonants /l, r, h, t, k, /, so that the CVC syllable structure created by derivational processes changes into two open syllables CV.CV. This [u] sometimes gets lost in rapid speech and therefore pretended closed syllables arise (see the discussion in Klamer (1998), section 2.2.1.).

T$_3$:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-N+ku	-ka+Ø	-N+ka
2nd singular	(mu)-	-mu	-ka+u	-N+ka+u
3rd singular	na-	-na	-ja	-N+ja
1st in. plural	ta-	-N+ta	-ta	-N+da
1st ex. 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

T$_4$:

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

Now a lot more instances of syncretism can be seen. Only *dʒa* and *ha* in the third person plural dative and accusative appear just once. In the other plural contexts, the markers do not distinguish between cases at all: *ta, ma*, and *mi* realize only person and number features. In the singular a grouping of nominative and genitive versus accusative and dative can be observed.

Note that the accusative and dative forms in T$_4$ are the same except for third person plural. This can be explained by Kambera's phonology. Given that the dative is derived from the accusative, we cannot get the expected clitic **nha*, because this would yield a complex onset, which is excluded in Kambera. The nasal and the fricative cannot form a prenasalised segment – this is only possible with plosives, affricates or semivowels. Of course the result could be a syncretism of accusative and dative, both realized as *ha*. This possibility is excluded for other reasons concerning the available distributions of syncretisms, see footnote 13. As a consequence, the clitic must consist of different segments. The unexpected distinction also provides further evidence that the dative is derived from the accusative. Assuming it to be the other way around, the accusative is derived from the dative by deleting N. But then *dʒa* should show up in the accusative and there is no reason why this should not be an option. The unexpected distinction also provides further evidence that the dative is derived from the accusative.

3.2   Feature decomposition

After this subanalysis, the context for the insertion of the clitics' subparts has to be described in terms of features. 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 T_4, 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:

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

In order to decompose two numbers, a binary feature [±pl] is needed. The four persons arise from cross-classification of two binary features [±1] and [±2]. Third person is neither first nor second person and the inclusive person is characterized as a combination of first and second person, because the speaker as well as the addressee is involved.

The cases are decomposed in [±obl] and [±obj]. [+obl] cases are typically neither used to mark the core arguments of the verb – genitive and dative in Kambera. [+obj] cases do not mark agent arguments – accusative and dative[10]. We can conclude that ka realizes [+obj] and N realizes [+obl]. Above it was outlined that nominative marks agent arguments, and genitive does not, but genitive is decomposed in [–obj], which means that it is able to do so. This assumption is justified by the fact that under certain circumstances a genitive instead of the expected nominative clitic shows up to mark an agent argument (Klamer (1998), p. 322):

---

[10]   There do exist constructions in which the dative and accusative clitics or combinations with clitics of other cases mark the sole argument of intransitive verbs (called S by Dixon (1972) and Comrie (1989)). These are parts of Kambera's morphology that show ergative alignment: in the 'double-S' construction S is expressed by the expected nominative clitic and the corresponding accusative clitic. The 'absolutive' construction marks S with an accusative clitic, and in the 'continuative aspect' construction, S is encoded by a genitive and dative clitic. For discussion in more detail see Klamer (1998), chapter 5. On the whole, these peculiarities are explained by the language's change from the ergative-absolutive to the nominative-accusative pattern of case marking. We find remains of an earlier stage in Kambera's history. But as they show up less frequently, the expected nominative clitic can often be used as well, some of them are considered to be archaic today, and their use sometimes depends on pragmatics, these constructions do not seriously challenge my generalization that nominative and genitive form a natural class excluding dative and accusative.

(5) Na   nggula [$_{CP}$ na- pa$_1$-   pa.ngàndi   -nggu
    ART   sugar   ART RmO- CAUS.take.to   -1S.GEN
    -nggau t$_1$]
    -2S.DAT
    'The sugar that I sent (to) you.'

Sentences in which the patient argument of an embedded verb, here *ngàndi*, is relativised, mark the agent of this verb with genitive instead of nominative, indicated by the clitic *-nggu*.

3.3   Vocabulary insertion

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] in order to fulfil the phonotactic restriction.

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

The remaining VIs are provided below. Note that except for *ha* all of them are underspecified and I assume that they can refer to negative specifications[11].

---

[11]   An alternative analysis would be to make use of radical underspecification and to refer only to positive feature specifications (cf. Bierwisch (1976), Wiese (1999) vs. Blevins (1993)). The list of VIs in (7) would look as provided below. This analysis has several disadvantages. First, we need the extrinsic order of insertion rules as listed, which the analysis with negative feature values does not need. Furthermore, the existence of an elsewhere marker makes it impossible to get the gaps in first person singular accusative and dative, shown in T$_4$. Apart from that, *ku* is inserted before the elsewhere marker distributes. Besides, this analysis needs the same Impoverishment rules as the analysis presented in this article. But even then, it cannot account for the right distribution of *ka*, which would not show up in the second person singular accusative and dative. Finally, the system is not iconic in the sense outlined in section 4.

a. ma ↔ [+1, +pl] b. mi $ [+2, +pl]
c. ta ↔ [+1, +2]
d. dÿa ↔ [+pl, +obj, +obl]
e. ha ↔ [+pl, +obj]
f. da ↔ [+pl]
g. u ↔ [+2, +obj]
h. ku ↔ [+1]
i. ja ↔ [+obj]
j. mu ↔ [+2]
k. na ↔ [ ]

(7) a. /dʒa/ ↔ [−1, −2, +pl, +obl, +obj]
     b. /ha/ ↔ [−1, −2, +pl, +obj]
     c. /da/ ↔ [−1, −2, +pl]
     d. /ta/ ↔ **[+1, +2, +pl]**
     e. /ku/ ↔ [+1, −pl, −obj]
     f. /ja/ ↔ [−1, −2, +obj]
     g. /mu/ ↔ [+2, −pl,−obj]
     h. /na/ ↔ [−pl, −obj]
     i. /ma/ ↔ [−2, +pl]
     j. /mi/ ↔ [+2, +pl]
     k. /u/ ↔ [+2]

Consider the syntactic structure into which the VIs will be inserted:
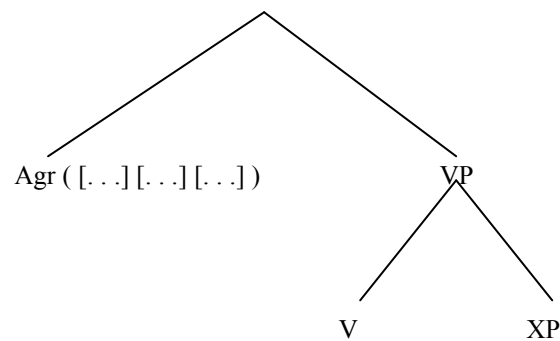
Kambera clitics attach to phrases, in the case of pronominal clitics to VP, which is built up of the verbal head and an XP that is either V's complement or an adjunct (see Klamer (1997), p. 915). A functional category is merged as the sister of VP into which the clitics are inserted. It is called Agr here, because it contains the agreement feature bundles including the inherent phi-features of the arguments and case features. Therefore, Agr can be thought of corresponding to v, which is the sister of VP (cf. Larson's (1989) VP-shell) and a functional category bearing agreement features (Chomsky (2000)). Via Agree (Chomsky (2000, 2001b), Adger (2004)), the phi-features of the object(s) are valued on v that bears the accusative feature, and the phi-features of the subject are valued on T bearing nominative case. These features of T are given to v by Agree, when T values tense on v, because the features will be realized there, attached to VP.

   Furthermore, it is necessary to ensure that more than one VI can be inserted in a terminal node. This is achieved by the concept of *Fission* in the sense of Noyer and Frampton (Noyer (1992), Frampton (2002)).

Fission: 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 α − β, such that (i) and (ii) hold:

(i) α – β 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 after the first application are available for insertion of other VIs as long as there are still undeleted features. The list of VIs has to be scanned top-down for each terminal node so that the most specific VI is inserted at first. I will give a short illustration of the insertion of the competing VIs *ha*, *ʤa*, and *ja* in T_5.

T_5:

	Nominative [-obl, -obj]	Genitive [+obl, -obj]	Accusative [-obl, +obj]	Dative [+obl, +obj]
1st singular [+1, -2, -pl]				
2nd singular [-1, +2, -pl]				
3rd singular [-1, -2, -pl]			-ja	-ja
1st in. plural [+1, +2, +pl]				
1st ex. plural [+1, -2, +pl]				
2nd plural [-1, +2, +pl]				
3rd plural [-1, -2, +pl]			-ha	-ʤa

ʤa is fully specified for third person plural dative. *ha* fits in third person plural dative and accusative, but as *ʤa* is more specific than *ha*, the latter is only inserted into the accusative context. *ja* matches the features of third person dative and accusative, irrespective of number. Again, the more specific VIs *ha* and *ʤa* are inserted into the relevant plural contexts, so that *ja* is realized only in the singular. The same procedure applies to the remaining VIs and results in T_6. The markers in brackets would have matched the terminal node's features as well, but were less specific than the actual item.

T_6:

	Nominative [-obl, -obj]	Genitive [+obl, -obj]	Accusative [-obl, +obj]	Dative [+obl, +obj]
1st singular [+1, -2, -pl]	ku (na)	ku (na)		
2nd singular [-1, +2, -pl]	mu (na, u)	mu (na, u)	u	u
3rd singular [-1, -2, -pl]	na	na	-ja	-ja
1st in. plural [+1, +2, +pl]	ta (mi, u)	ta (mi, u)	ta (mi, u)	ta (mi, u)
1st ex. plural [+1, -2, +pl]	ma	ma	ma	ma
2nd plural [-1, +2, +pl]	mi (u)	mi (u)	mi (u)	mi (u)
3rd plural [-1, -2, +pl]	da (ma)	da (ma)	-ha (da, ja, ma)	-ʤa (ha, da, ja, mas)

Afterwards, *ka* and *N* are inserted, because they are less specific than the other VIs. Starting with /ka/ ↔ [+obj], it is important to note that all dative and accusative contexts originally had this feature, but where the items *ha, ʤa* and *ja* have been inserted, [+obj] has been deleted (see the insertion above). That is the reason why *ka* never shows up in third person contexts. Nevertheless, it should be realized in first person inclusive plural, as *ta* did not refer to [+obj]. To explain this, another concept of Distributed Morphology is needed. The relevant feature in the terminal node has to be deleted after syntax but before vocabulary insertion. 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. The mechanism is called *Impoverishment* (Bonet (1991)) and deletes features in terminal nodes (without inserting a VI)[12].

Impoverishment rule:

(8)   [+obj] → Ø / [+1, +2]

The result is shown in T_7.

T_7:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-ku	-ka	-ka
2nd singular	(mu)-	-mu	-ka+u	-ka+u
3rd singular	na-	-na	-ja	-ja
1st in. plural	ta-	-ta	-ta	-da
1st ex. plural	ma-	-ma	-ka+ma	-ka+ma
2nd plural	(m)i-	-mi	-ka+(m)i	-ka+(m)i
3rd plural	da-	-da	-ha	-ʤa

The insertion of N $\leftrightarrow$ [+obl] / [–nasal] proceeds in the same way. All genitive and dative contexts still share this feature, as none of the other VIs refered to it. *N* is inserted if the phonotactic constraint is respected, expressed by the requirement that the following segment must have the feature [–nasal]. In spite of that, it does not show up in third person plural genitive. As I cannot find a phonological reason that blocks the appearance of the nasal, Impoverishment must have been applied once again. Rule (9) deletes the feature [+obl] in the third person plural genitive and the nasal cannot be inserted according to the Subset Principle[13]. This results in T_3 of section 3.1.

Impoverishment rule[14]:
[+obl] $\rightarrow$ Ø / [–1, –2, +pl, –obj]

T_3:

	Nominative	Genitive	Accusative	Dative
1st singular	ku-	-N+ku	-ka+Ø	-N+ka
2nd singular	(mu)-	-mu	-ka+u	-N+ka+u
3rd singular	na-	-na	-ja	-N+ja
1st in. plural	ta-	-N+ta	-ta	-N+da
1st ex. 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+ʤa

---

[13]   The alternative in Minimalist Distributed Morphology is again insertion of a highly specific zero VI: Ø $\leftrightarrow$ [+obl] / [–1, –2, +pl, –obj].

[14]   The reason for this deletion seems to be arbitrary, but it might be motivated by the observation that the language avoids syncretism on the clitic level between persons in the whole paradigm. Syncretic forms only show up between cases. Furthermore, these homonymous case forms are distributed in a way that they show up in different positions relative to the verbal stem – nominative clitics are in preverbal position and genitive clitics in postverbal position. There is only one syncretic form for genitive and dative that are both realized postverbally. In this way, Kambera strictly minimizes ambiguity of forms. It is nevertheless an economic system, because the clitics are systematically built up by few segments, which results in syncretism below the clitic level, as shown in the analysis. If the third person plural genitive was *nda*, as expected, the result would be an ambiguity with first person inclusive in the same case, which the language obviously wants to avoid.

Afterwards, the phonological processes described in section 3.1 apply, changing voiceless plosives to voiced plosives after a nasal and assimilating the nasal's place of articulation to the following consonant.

Phonological rules:

(10) [+cons, –cont] $\rightarrow$ [+voiced] / [+nasal] ___[15]

     [+nasal] $\rightarrow$ [αplace]/ ___ [+cons, αplace]

3.4 Remarks

In this section, I will discuss some questions that might have been arisen during the analysis. First, consider the linear order of the (up to three) vocabulary items that constitute a clitic. It is N+ka+X, where X can be replaced by one of the other VIs.

---
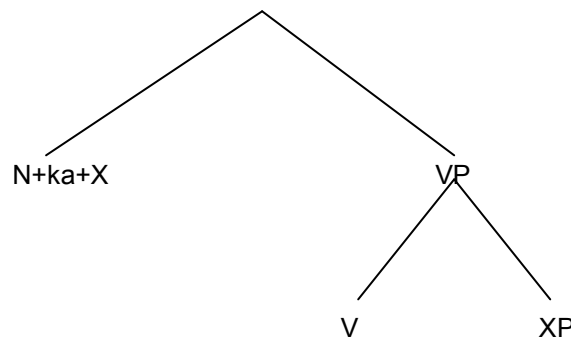
[15] Consider the list of vocabulary items, especially the insertion context of N. Bobaljik (2002) argues that a morpheme's features can condition allomorphy of a morpheme that is either farther away from or closer to the root (VP in the case of Kambera) than these features and that the direction of the this sensitivity is systematically derivable from the kind of features that trigger the allomorphy. If the features that trigger allomorphy are closer to the root, the morpheme is said to be inwards sensitive and if it the features are further away from the root the morpheme 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 order of insertion presented in this section, the insertion context of *N* supports this analysis because its place of articulation and even its realization depend on the phonological features of a morpheme that is closer to the root, namely *ka* or X. 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, as *N* is situated further away from VP, but it depends on phonological features, or in other words, a phonological feature determines the form of a more inwards morpheme and thus clearly contradicts Bobaljik's assumption that "'morphopho-nological diacritics ... can only serve as the context for allomorphy of a more peripheral affix."' (Bobaljik (2002), p. 14).

The more specific an item, the closer it is situated to the stem. X is more specific than *N* and *ka* and will therefore be inserted first. If a VI always attaches to the left edge of the existing structure that has already been built up, we can derive the order and furthermore, insertion proceeds strictly cyclic.

The question is what principle regulates the order of *N* and *ka*, because putting aside the context restriction of the nasal, they are equally specific. Here Kambera's syllable structure solves the problem. There are two options available: *N* proceeds *ka* or *ka* proceeds *N*. In the latter case, we would create forms in which two consonants were next to each other, e.g. like *[kaNmi]* for the second person plural dative. As Kambera does not have codas, *N* cannot be part of the first syllable. On the other hand, it cannot form the onset together with the following consonant 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[16].

Next, the appearance of the vocabulary item *ka* could be described in another way by making use of alpha notation[17]. To derive the distribution of *ka*, this yields /ka/ $\leftrightarrow$ [α1, -α2, +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 directly 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 for which no evidence can be found. The application of alpha notation seems quite arbitrary.

---

[16]   An alternative may be to insert the paragogic [u] after *N* (see footnote 9). In this way ka+Nu+X would conform to the restrictions on Kambera's possible syllable structure. But this idea is not unproblematic. Epenthesis of [u] only takes place after velar nasals and there is no reason to assume that all nasals after *ka* and before X are velar. Besides, the language treats epenthesis as a kind of repair mechanism that is only applied if no other possibility is available to correct ill-formed sequences. We see this, when we consider the prenasalised segments. Either the nasal and the following consonant assimilate with respect to voice and place of articulation and afterwards form one phonological unit, or, if this strategy is not applicable, the nasal is not inserted (see the context restriction on the VI *N*). Both strategies result in the syllable structure CV and avoid *CCV. Although [u] could be inserted between the nasal and the following consonant instead, the language does not choose this solution. Epenthesis seems to be a kind of last resort and is necessary to repair structures like *CVC, where none of the strategies outlined above is applicable, but here the exchange of *N* and *ka* is sufficient.

[17]   Alpha notation has been introduced by Chomsky (1965). In this notation, 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, which is expressed by variables.

But of course this analysis could be applied, if one preferred it – none of the insights and conclusions would change.

Consider the VI *ta*. It might seem redundant to specify it for number, because the feature bundle [+1, +2] unambiguously encodes the inclusive person that only exists in the plural. But without [+pl] *ta* would compete with *mi* for insertion in the inclusive person. As both have two features, none of them could be chosen by the *Specificity Principle*. To solve the problem, another Impoverishment rule or a feature hierarchy could be established. For reasons of complexity, I decided not to introduce any of them. The less rules are needed, the simpler is the analysis. Furthermore, if *ta* realizes three features, it fits well in a pattern that will be discussed in section 4.

Furthermore, there is no elsewhere marker, a VI that is phonologically empty and does not realize any features18. It would fit in everywhere as long as there does not exist a more specific marker: Ø ↔ [ ]. It might seem attractive to assume such a VI for the first person singular accusative and dative in T4, where apart from ka or N none of the other items can be inserted. Nothing in the analysis would change if it existed. Nevertheless, I rejected this possibility for the following reason. 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 VI, it should be inserted in every terminal node except for the third person plural accusative, as ha is not underspecified. Things get even worse: as the elsewhere marker does not delete features, the terminal node's features will never be used up and consequently the elsewhere marker would incessantly be inserted. The marker is not needed for the analysis of Kambera's pronominal clitics and its existence would cause a lot of additional problems. Consequently, it is not integrated. This implies that not all the features of a terminal node need to be deleted during vocabulary insertion. They are simply left over.

In addition, following a proposal of Klamer (1997, p. 916/917), the fact that only definite arguments trigger agreement on the verb can be included in the analysis. It is just necessary to add the feature [+def] to all VIs. In this way, nothing is inserted, if the arguments are not definite.

In this article I primarily considered the clitics' morphological and phonological shape, respectively, but I was neither concerned with the linearization of

---

18   Note that the alternative VIs in footnote 12 and 13, the highly specific zero markers, are not elsewhere markers in this sense, because they realize morpho-syntactic features that are deleted in a terminal node when inserted into it.

mood, aspect, and pronominal clitics nor did I deal with the linear order of the pronominal clitics with respect to each other. Klamer (1997) discusses these problems in detail and concludes that Kambera clitic placement cannot be easily derived by syntactic head movement (at least this causes a lot of additional problems), by lexical word formation processes or by phonological characteristics alone. Besides, the morphological form of some clitics depends on information that is not available until the syntactic derivation is completed, e.g. what other clitics are in the structure or pragmatic aspects. Due to this, she prefers a post-syntactic approach to Kambera clitics with late insertion of phonological material, which is linearized afterwards (Klamer (1997), p. 915). Distributed Morphology offers the relevant mechanisms and this is another reason why the framework has been chosen for the analysis.

Finally, a restriction on Kambera clitic clusters can be accounted for which has been formulated as follows in Klamer (1997, p. 903): "If there is a pronominal clitic following genitive, it must be dative." This idiosyncrasy of Kambera can be captured by a special type of Impoverishment, feature changing Impoverishment (Noyer (1998)), in which features are not deleted, but their values are changed. In the relevant rule, we need to change the value of the feature [–obl] to [+obl], whenever there is a feature bundle with the specification for genitive, [+obl, −obj], somewhere in the Agr node, too. As the restriction is imposed on the combination of object clitics, the value of [obj] is already positive and does not need to be transformed. The resulting feature bundle contains [+obl, +obj], encoding dative.

(11) $[- \text{obl}, +\text{obj}] \rightarrow [+\text{obl}] / [_{\text{Agr}} \ldots [+\text{obl}, -\text{obj}] \ldots]$

4 Implications

All instances of syncretism can be derived by the minimal number of under-specified) VIs, no marker occurs twice. Furthermore, no feature hierarchy is needed to get the right distribution.

More interestingly, there is a correlation between specificity of vocabulary items and their phonological weight and sonority, respectively: Ignoring *N* and *ka*, the heaviest syllable *dʒa*, consisting of an affricate and a vowel, realizes the most features, the lightest syllable, the bare vowel, realizes the least features. The remaining VIs all consist of two segments but realize a different number of features. To distinguish between these, sonority plays a crucial role: the more specific an item, the less sonorous it is, according to the hierarchy of sonority.

Hierarchy of sonority: vowels > nasals > obstruents

On the whole, vocabulary items beginning with an obstruent (b.-f.), except for the glide, are more specific than the items starting with a nasal (g-j.):

(12) a. /dʒa/ ↔ [–1, –2, +pl, +obl, +obj]
b. /ha/ ↔ [–1, –2, +pl, +obj]
c. /da/ ↔ [–1, –2, +pl]
d. /ta/ ↔ [**+**1, +2, +pl]
e. /ku/ ↔ [+1, –pl, –obj]
f. /ja/ ↔ [–1, –2, +obj]
g. /mu/ ↔ [+2, –pl,–obj]
h. /na/ ↔ [–pl, –obj]
i. /ma/ ↔ [–2, +pl]
j. /mi/ ↔ [+2, +pl]
k. /u/ ↔ [+2]

This phenomenon is called iconicity by Wiese (1999): there is a correlation between form and function of the vocabulary items. Their order in terms of specificity corresponds to their order with respect to sonority. The term iconicity is applied in a somewhat different way than it usually is in the typological literature. The definition refers to a feature-based concept of iconicity: function is measured by counting the number of features an underspecified VI realizes, which presupposes an abstract analysis relying on feature decomposition and underspecification. This correlation is only found among the VIs except for *N* and *ka*. This provides further motivation that these two markers have a special status as they only realize the case features [+obj] and [+obl].

5 Conclusions

To conclude, I will summarize the main insights of this article. Kambera's pronominal clitics provide a lot instances of syncretism below the clitic level. All of them can be derived by underspecification and feature decomposition that
allow vocabulary insertion to take place several times per terminal node. This suggests that syncretisms do not appear accidentally but are systematic: the contexts that the homonymous forms show up in share features that vocabulary items refer to. Furthermore, the analysis supports the hypothesis that iconicity might be a principle that morphological systems tend to be organised after, because similar correlations have been found by Wiese (1999) for pronominal inflection in German, by Müller (2004) for Russian noun inflection, and by Müller (2005) for Icelandic noun declensions. All these are Indo-European languages, but Kambera shows that the phenomenon is not restricted to this language family. These insights support the view that morphological systems are not built up by chance but obey deeper grammatical principles.
Besides, it is important that the correlation between specificity and sonority could only be discovered by assuming that morphological (argument encoding) markers are composed of smaller units (subanalysis) in order to detect instances of partial syncretism (or block syncretism in Paradigm Function Morphology,

see Stump (2001) vs. Baerman et al. (2005)), supporting the need for abstract analyses.

References

Adger, David (2004): Core syntax - a minimalist approach. Oxford Universitiy Press, Oxford.

Baerman, Matthew, Dunstan Brown & Greville Corbett (2005): The Syntax- Morphology Interface. A Study of Syncretism. Cambridge University Press, Cambridge.

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 Press, Cambridge, Mass.

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

Chomsky, Noam (2000): Minimalist Inquiries: The framework. In: Roger Martin, David Michaels, and Juan Uriagereka, eds., *Step by Step*. MIT Press, Cambridge, Mass., pp. 89.155.

Chomsky, Noam (2001b): Derivation by phase. In: Michael Kenstowicz, ed., Ken Hale: *a Life in Language*. MIT Press, Cambridge, Mass., pp. 1-52.

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 Press, 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 (1997): Spelling out clitics in Kambera. In: *Linguistics 35*. Mouton de Gruyter, Berlin, New York, pp. 895-927.

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

Larson, Richard (1988): On the Double Object Construction. In: *Linguistic Inquiry 19*. MIT Press, Cambridge, Mass., pp. 335-391.

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.

Noyer, Rolf (1998): Impoverishment theory and Morphosyntactic Markedness. In: S. Lapointe, D. Brentare & P. Farrell, eds., *Morphology and its Relation to Phonology and Syntax*. CSLI, Palo Alto, pp. 264U285.

Stump, Gregory (2001): Inflectional Morphology. Cambridge University Press, Cambridge.

Trommer, Jochen (2003): Morphology consuming Syntax' Resources. Procceedings of ESSLI Workshop on Resource Logics and Minimalist Grammars, Nijmegen.

Wiese, Bernd (1999): Unterspezifizierte Paradigmen. Form und Funktion in der pronominalen Deklination, Linguistik Online4. (www.linguistik-online.de/3 99).

Wunderlich, Dieter (2004): Is There Any Need for the Concept of Directional Syncretism? In: G. Müller, L. Gunkel & G. Zifonun, eds., *Explorations in Nominal Inflection*. Mouton de Gruyter, Berlin, pp. 373-395.

Leipzig, September 2007   Doreen Georgi

(Universität Leipzig, Institut für Linguistik, Beethovenstraße 15, D-04107 Leipzig)