A Reanalysis of Definiteness-Markers in Albanian Noun Inflection

Andreas Opitz[*]

Abstract

This paper addresses the distribution of morphological markers of the Albanian noun inflection. The patterns found are analyzed within the theoretical framework of Distributed Morphology. I will show that both underspecified morphological markers and marker-independent rules lead to an analysis that completely explains the syncretisms of noun inflection in Albanian. Moreover, I will argue for a further subanalysis of Albanian inflectional markers. This subanalysis straightforwardly leads to a new interpretation of traditionally so called 'definite markers' and establish a new coherent set of markers of definiteness instead.

1. Introduction

In this paper I want to provide a morphological analysis of Albanian noun inflection within the theoretical framework of Distributed Morphology (Halle & Marantz (1993)). In my analysis I will argue that both underspecified morphological markers and marker-independent rules – as basic components of the theory of Distributed Morphology – lead to an analysis that explains inner- and trans-paradigmatic syncretisms of noun inflection in Albanian. Moreover, I will show that the concept of fission (Noyer (1992)) provides support for a further subanalysis of Albanian inflectional markers. This subanalysis straightforwardly leads to a new interpretation of traditionally so called 'definite markers' (i.e. *a, u, i* in the nominative) and establish a new coherent set of markers of definiteness (i.e. *t, s, n*) instead.

2.   Distributed Morphology

Distributed Morphology (Halle & Marantz (1993)) conceives of inflection as a late, post-syntactic process. The output of syntax provides abstract, functional morphemes consisting of completely specified morpho-syntactic features but still lacking any phonological realization. Inflectional markers on the other hand are assumed to be vocabulary items that pair morpho-syntactic and phonological features. In a late, post-syntactic operation, these markers are inserted into functional morphemes. This insertion is mainly driven by the Subset Principle (Halle (1997); Noyer (1992); Wiese (1999); Müller (2006b)) that states:

(1) Subset Principle
     A vocabulary item *V* is inserted into a functional morpheme *M* iff:
     a.   The morpho-syntactic features of *V* are a subset of the morph-syntactic features of *M*.
     b.   *V* is the most specific vocabulary item that satisfies (1-a).

This definition permits underspecified markers to be inserted in a certain context. Vocabulary items whose features constitute a subset of each other are in competition for insertion. The vocabulary item whose features matches most specifically the features of the abstract morpheme will be inserted. Thus, a vocabulary item may be inserted which is not specified for *all* features of the context. This possibility of underspecified vocabulary items gives rise to syncretisms. For the analysis in this paper it is important to note that specificity is not understood as a simple feature counting criterion. Rather the quality of the features has to be taken in account too (see Müller (2006a); also Noyer (1992)).

(2) Specifity of vocabulary items
     A vocabulary item V_i is more specific than a vocabulary item V_j if there is a class of features **F** such that

     (i) V_i bears more features belonging to **F** than V_j does, and

     (ii) there is no higher-ranked class of features **F'** such that V_i and V_j have a different number of features in **F'**.

This means that different kinds of features (e.g. number, class, case) are hierarchically ordered. In the case of competing vocabulary items the one containing (most of) the higher ranked features on this hierarchy will be

inserted. As for the analysis of Albanian noun inflection I will assume a hierarchy class > case, number.[1]

A rather central assumption of Distributed Morphology is essential to make this theory capable of explaining systematic inflectional patterns such as syncretisms: Morpho-syntactic features of abstract morphemes can be altered post-syntactically. One of the most important operations is post-syntactic impoverishment by means of which features are deleted before vocabulary insertion takes place. As I will show, this operation is also involved in the Albanian noun inflection. Furthermore, a more specific version of this operation will be required: a post-syntactic rule that changes the value of a binary feature (here from [+singular] to [-singular]) instead of deleting it (for a more detailed discussion see below) (Noyer (1998)). Finally, in order to account for a sub-analysis of (some of) the multi-segmental inflectional markers in Albanian (that intuitively seem to bear information about case, number, class and definiteness on separate segments) an operation will be needed that splits off an abstract morpheme and makes some features accessible to the insertion of vocabulary items, but leaves other features of the morpheme unaffected and thus available for further vocabulary insertion. This operation is called fission and goes back to Noyer (1992). It can be formalized as follows (see Müller (2006b)):

(3) Fission:
     A morpheme M with the features α is fissioned by insertion of a vocabulary item V bearing the features β into two separate bundles of features: β and $(\alpha - \beta)$, so that:
     a. $(\alpha - \beta)$ is available for further insertion according to the subset principle.
     b. β is not available for further insertion.

Based on this theoretical background I want to establish an analysis of the Albanian nominal inflection.

---

[1] As I will show, it is the class feature that outranks the lower ranked features of case and number (and definiteness), while the relative order of these lower features is irrelevant for the analysis.

3.   Nominal Inflection in Albanian

Albanian is spoken by approximately 6 million people and constitutes its own independent branch of the Indo-European language family. There are no known closely related languages.

Albanian nouns are inflected by case (nominative, accusative, genitive, dative, ablative, and vocative) and number (singular and plural). In traditional grammars five declension classes are assumed (see Buchholz & Fiedler (1987)): Four declension classes in the singular but only one class in the plural, as shown in table 1.[2]

Sg	I		II		III		IV		Pl	
	indef	def	indef	def	indef	def	indef	def	indef	def
Nom	ø	i	ø	u	ø	a	ø	t/të	ø	t/të
Gen	i	it	u	ut	e	s/së	i	it	ve	ve
Dat	i	it	u	ut	e	s/së	i	it	ve	ve
Acc	ø	in/në	ø	un/në	ø	n/në	ø	t/të	ø	t/të
Abl	i	it	u	ut	e	s/së	i	it	sh	ve

Table 1. Albanian Inflection Classes Singular I-IV and Plural (traditional labels)

A noun's membership to one class is mainly determined by its gender but there is no one-to-one mapping. Most nouns of declensions I and II are masculine colloquial, but the two classes contain feminine nouns as well:

(4)   a. kodër   b. flutur
         hill.fem   *butterfly*.fem

On the other hand there are masculine nouns belonging to declension III typically containing feminine nouns:

(5)   a. babë -a   b. toskë -a
         father.masc   *Tosk*.masc

The decision to which class a certain noun belongs to is determined by a couple of factors such as grammatical gender, natural gender (sexus), phonological factors (termination, stress) etc. Therefore I decided not to label the declensions by 'masculine', 'feminine' and 'neuter' but to decompose the labels of the singular declensions by two independent binary features $[\pm \alpha, \pm \beta]$ (cf. Müller (2004)). Declensions I and II differ only in their vowel: *i*

---

[2]Note that all alternations with /ë/ are fully predictable and phonologically driven. They will be disregarded in the further analysis.

vs. *u*. This alternation is phonologically driven and fully predictable. This allows us to combine these two classes to one:

(6) Decomposition of declension classes

    I+II: $[\ +\alpha, -\beta]$
    III: $[\ -\alpha, +\beta]$
    IV: $[\ -\alpha, -\beta]$

Because among the six existing cases the vocative only occurs with a very limited number of words, it will not be considered in the following analysis. The remaining cases are also decomposed by two binary features ($\pm$ oblique, $\pm$ object) and, in addition, in the case of ablative by one privative feature [local], according to the actual usage of that case that now also covers function of the historical locative, (cf. Buchholz & Fiedler (1987): 212, 242-243):

(7) Decomposition of case features [3]

    Nominative   [− oblique, − object]
    Genitive   [+ oblique, − object]
    Dative   [+ oblique, + object]
    Accusative   [− oblique, + object]
    Ablative   [+ oblique, + object] [local]

A constitutive characteristic of Albanian noun inflection is its sensitivity to definiteness. As can be seen in table 1, inflectional markers differ with regard to the definiteness of the noun.

   Albanian nominal inflectional markers show numerous cases of syncretism. First of all, and already mentioned above, there are no distinct inflectional classes in the plural.[4] In addition, there are a number of syncretisms in the oblique cases: all indefinite nouns within each class and

---

[3]As the listing of the markers in table 1 show, the morphological forms of genitive, dative and ablative do not differ within each declension class. Thus one could assume one single morpho-syntactic case that includes all these three historically distinguished cases. This generalization is realized by their shared feature [± oblique]. In the tables given, the distinction between these cases by additional features ([±object], [local]) is maintained. Though the feature [±object] is not needed to distinguish the oblique cases from each other in the analysis provided here, it is needed to differentiate [−oblique] cases (i.e. nominative from accusative).

[4]The formation of the plural in Albanian is quite complex, among other things it includes phonological alternations of the stem. In the analysis on hand I will only consider the inflectional markers of the plural declension shown in table 1.

all definite nouns within each class show the same markers (so called inner-paradigmatic syncretism). Moreover, all indefinite nominative and accusative forms of all classes (including the plural declension) show the same form (i.e., ø, a inner- and trans-paradigmatic syncretism). The definite forms of nominative and accusative show complete syncretism only in class IV and in the plural declension. An interesting observation concerns markers of definite nouns. Although there are no total syncretisms between the declensions, a partial syncretism of single segments (phonemes) seems obvious. All markers of definite nouns in oblique cases of classes I, II and IV have at least the phoneme /t/ in common. Similarly, definite accusative markers of classes I, II and III all share the phoneme /n/. Furthermore, the markers of definite nouns in nominative and accusative forms of class IV and plural contain /t/. An adequate rearrangement of the paradigms clarifies the patterns of these syncretisms, as shown in table 2, here already labeled with decomposed features.

	[+sgl]		[+sgl]		[+sgl]		[-sgl]	
	$-\alpha + \beta$ (III)		$+\alpha$ -β (II+I)		$-\alpha$ -β (IV)			
	-def	+def	-def	+def	-def	+def	-def	+def
$-$obl, $-$obj	ø	a	ø	u	ø	t	ø	t
$-$obl, $+$obj	ø	n	ø	un	ø	t	ø	t
$+$obl, $-$obj	e	s	u	ut	i	it	ve	ve
$+$obl, $+$obj	e	s	u	ut	i	it	ve	ve
$+$obl, $+$obj [loc]	e	s	u	ut	i	it	sh	ve

Table 2. Inflection classes with decomposed features

One of the main assumptions of this paper is that all of the multi-segmental markers (except *ve*) can be decomposed into two different vocabulary items ($u + t$, $i + t$, $u + n$) each paired with a specific bundle of morpho-syntactical features. And as the specification of these features will show, there are no vocabulary items with identical phonological representation, i.e. each phonological representation (phoneme) is only used once as a vocabulary item. In other words, there are no homonymous vocabulary items. All occurrences of a phoneme (e.g. /u/,) represent the same bundle of morpho-syntactic features. And the specification of theses features (plus necessary impoverishment rules, see below) fully explains the distribution of the markers. If we examine the distribution of single markers, we can see that there are certain makers that occur only once in the whole paradigm (/a/, /sh/), these markers are highly specified and therefore block the insertion of less specified markers. Thus the most frequent marker /u/ is more

less-specified, showing up in definite and indefinite forms in different cases (sometimes combined with / n/ or /t/).

But not all of the syncretisms can be explained by underspecified vocabulary items alone. Before lexical insertion can take place, so called impoverishment rules apply. The first rule I suppose creates a unique context for the later insertion of /ø/ by deleting class features in [−obl, −def] contexts.

(8) $[\pm\alpha \pm \beta] \rightarrow$ ø / [−obl, −def] ___

Probably the most challenging part of this analysis is to explain the syncretisms of the markers of (definite) nominative and accusative of declensions IV ($[-\alpha - \beta]$) and the plural declension. The only way to avoid an insertion of (less specified) /i/ in the nominative and accusative singular of class IV would be to specify it for [+oblique] contexts (in contrast to /u/). In addition, the marker /t/ that appears there instead, would also show up in nominatives of classes $[-\alpha + \beta]$ and $[+\alpha - \beta]$. In order to block the occurrence of /t/ there, one would have to assume at least two differently specified markers /t/ in class $[-\alpha - \beta]$ (i.e. one for [−oblique], specified for class and one for [+oblique], the latter would be the same that shows up in class $[+\alpha - \beta]$). Moreover a third marker /t/ would be needed to cover the appearance in contexts of plural [−oblique]. I suppose a solution of this problem by changing the context of [−oblique] of class $[-\alpha - \beta]$ into a plural. But to establish a simple feature-changing rule, which would be adequate to a rule of referral, is undesirable here, because of the explanatory power of this kind of rules. The problem of literally every syncretism could be solved by this kind of descriptive rule. As Noyer (1998) shows, there are reasons to suppose feature changing impoverishment rules. But these rules are restricted in such a way that they may only change the value of a feature from more marked to less marked. Following this approach I propose a second rule:

(9) $[+\text{sgl}] \rightarrow [-\text{sgl}]$ / $[+\alpha + \beta, -\text{obl}]$ ___

Using the feature [± plural] instead of [± singular] would be a notational variant and fully compatible with the analysis established here. The reason I used the [± singular] feature is simply to emphasize the assumption that the plus-feature is marked and becomes unmarked by the rule (changing its value to minus). Thus, as a side-effect, in order to be consistent with Noyer's claim, my analysis has to suppose that the singular of Albanian nouns is more marked than the plural. This assumption is in conflict with the general and typologically based observation of (more) marked plural forms. There are four possibilities to solve this dilemma and simultaneously retain

an analysis of the Albanian data that is in accordance with Distributed Morphology:

(10) Solutions for the IV-plural-syncretism:
  a. Supposing different vocabulary items for /t/ in the nominative and accusative of declension IV and the nominative and accusative of the plural declension (with different but overlapping morpho-syntactical contexts). This would lead to homonymous vocabulary items and therefore to a less systematic analysis.
  b. Supposing that the number feature of declension IV [−plural] spreads to the plural and creates a singular context there (instead the other way around). This move clearly supports the assumption of a more marked plural form but leaves the problem of (10-a) unaffected: different markers for /t/ (at least in class [−α − β]) would still be needed.
  c. Weakening the restrictions for impoverishment rules to allow feature changing from a less to a more marked value which is undesirable theory-internally (see above).
  d. Postulating the plural as less marked (and thus declaring Albanian to be a typological exception).

I decided (10-d) to be less harmful to the general assumptions of Distributed Morphology but an alternative view on this topic is also possible.

The last rule I suppose to be at work concerning the Albanian noun inflection is a rule that deletes any class feature in plural context, leading to one single declension class:

(11) $[\pm\alpha \pm \beta] \rightarrow \o \ / \ [-\text{sgl}] \ \underline{\ \ }$

As shown above, all declension classes converge into one single plural declension. Another way to explain this fact, possible with this analysis, is to specify the morpho-syntactic context of each vocabulary item that contains a class specification additionally by the feature [+sgl]. Since the class feature in my analysis is considered to be ranked higher than case, number etc., an added feature ([+sgl]) would not change the order of vocabulary items determined by the Subset Principle. An analysis making use of an impoverishment rule instead is (at this point) in fact an abbreviated notational variant and leaves the rest of the analysis unaffected.

The impoverishment rules applying after syntax and before vocabulary insertion can so far be summarized as follows:

(12) Impoverishment rules of Albanian noun inflection
  a. $[\pm\alpha \pm \beta] \rightarrow \o \ / \ [-\text{obl}, -\text{def}] \ \underline{\ \ }$

    b. $[+\text{sgl}] \rightarrow [-\text{sgl}] \ / \ [-\alpha - \beta, -\text{obl}]$ __

    c. $[\pm\alpha \pm \beta] \rightarrow \o \ / \ [-\text{sgl}]$ __

Note that the rule in (12-b) has to apply before the rule in (12-c) in order to establish the context for rule (12-c) in the nominative/accusative of declension IV. So the contexts there become first $[-\text{sgl}]$ and – after the rule in (12-c) applies – their class features are deleted.

    Concerning the marker inventory, one last issue needs to be discussed. Following the assumptions just made above, /i/, /u/, and /t/ form distinct markers. This is also true for all *oblique* cases of declensions I+II and IV where the morpho-syntactic bundle of features is fissioned and therefore two different and underspecified markers can be inserted. The features of these vocabulary elements have to be the following:

(13)   Decomposition of /it/ and /ut/

      /i/ $\leftrightarrow \{[-\alpha]\}$
      /u/ $\leftrightarrow \{[+\alpha]\}$
      /t/ $\leftrightarrow \{[+\text{obl}], [+\text{sgl}], [+\text{def}]\}$

In a fully specified morpho-phonological context containing (amongst others) $\{[\pm\alpha], [+\text{obl}], [+\text{sgl}], [+\text{def}]\}$, these makers – vowel and /t/ – may *both* be inserted. First, the insertion of vowel discharges the value $[\pm\alpha]$ which is no longer accessible for further insertions but the other features in contrast still are. After that, the marker /t/, the features of which still form a subset of the (remaining) features of the morpho-syntactical context, can be inserted. As a consequence, the complex marker sequence /it/ or /ut/ results from this multiple insertion by fission.

    But this specification of the marker /t/ is not sufficient for an insertion in plural nominative and accusative contexts. Thus, another vocabulary item /t/ is necessary, representing the features $\{[-\text{obl}], [-\text{sgl}], [+\text{def}]\}$. Here we also seem to receive two homonymous vocabulary items, a fact already in (10-a) stated to be less desirable within the analysis:

(14)   Marker-homonymy of /t/
      a. $/t_1/ \leftrightarrow \{[+\text{obl}], [+\text{sgl}], [+\text{def}]\}$
      b. $/t_2/ \leftrightarrow \{[-\text{obl}], [-\text{sgl}], [+\text{def}]\}$

This problem is easily being solved using an alternative, more variable notation of the context. As obvious in (14), the context for insertion of both markers is quite similar, except the *values* of the [obl] and [slg] features. These values are exactly the opposite in both instances but are the same within one vocabulary item (both plus vs. both minus). One way to converge the contexts of t_1 and t_2 is to use the so called alpha-notation. In this

notation the values of (binary) features are abstractly formalized by another variable – usually α. As the letter α is already used within this analysis for the decomposition of declension classes, I will use $\aleph$ here instead. This leads to one single vocabulary item for /t/:

(15) /t/ $\leftrightarrow$ {[$\aleph$obl], [$\aleph$sgl], [+def]}

If $\aleph$ here represents the value plus, this notation covers the context of former t_1. If it stands on the other hand for minus, it will lead to the context for insertion of t_2 in (14).

The complete list of vocabulary items that I assume for the inflection of Albanian nouns is listed in (16).

(16) 1. /s/ $\leftrightarrow$ {[$+\beta$], [+obl], [+def]}
      2. /a/ $\leftrightarrow$ {[$+\beta$], [−obl, −obj]}
      3. /e/ $\leftrightarrow$ {[$+\beta$], [+obl]}
      4. /i/ $\leftrightarrow$ {[$-\alpha, -\beta$]}
      5. /u/ $\leftrightarrow$ {[$+\alpha$]}
      6. /n/ $\leftrightarrow$ {[−obl,+obj], [+sgl], [+def]}
      7. /sh/ $\leftrightarrow$ {[+obl, loc], [−sgl], [−def]}
      8. /t/ $\leftrightarrow$ {[$\aleph$ obl], [$\aleph$ sgl], [+def]}
      9. /ve/ $\leftrightarrow$ {[+obl], [−sgl]}
      10. /ø/ $\leftrightarrow$ {[elsewhere]}

This listing of items together with the impoverishment rules in (12) correctly generate the forms of the inflectional system of Albanian nouns shown in tables 1 and table 2, respectively. Note that the order of the items is completely explained by the specificity criterion, assuming that class features are ranked higher than any other features here.

This supports an analysis of some markers (i.e. the vowels) as theme vowels that have to be attached closer (in other words: earlier) to the stem than other inflectional markers. However, treating them as theme vowels leads to complications in class III $[-\alpha + \beta]$: no such vowel occurs in [+def] contexts (except /a/ in the nominative which is even distinct to the /e/ in [−def] contexts). Maybe the theme vowel of this class is more abstract (or phonologically underspecified), lacking concrete surface realization. The assumption of an underspecified vowel in class III would cover another (partial) syncretism: The phonological feature [-cont] of the vowel seems to alter /t/ (that appears in the other classes) to /s/ in oblique contexts. The questions whether there are theme vowels in Albanian and what is the nature and realization of a hypothetical theme vowel in class III is beyond the scope of this paper. Additional data and detailed analyses would be needed.

The last point I want to discuss concerns iconicity. In general, the Iconic-

ity Principle states that there are certain relations between form and function. In his 2004 analysis of Russian noun inflection Müller uses the following definition:

(17) Iconicity Principle
       Similarity of form implies similarity of function
       (in a domain Σ, and unless there is evidence to the contrary).

In the system of Albanian noun inflection, as it is analyzed in this paper, this principle is apparently at work. Besides the notion that the elsewhere marker is highly iconic (no morpho-syntactical features, no phonlogical realization), on closer inspection the listing of the items in (16) reveals a noteworthy correlation of the feature [± definiteness] with its phonological representations. Each context of a vocabulary item representing [+def] is paired with an inflection marker that exhibits the phonological feature [coronal]: /s/, /n/, /t/ (in contrast to e.g. /ve/). As mentioned above, the feature of definiteness plays an important role in Albanian noun inflection and an analysis within the framework of Distributed Morphology reveals that definiteness is always expressed by the phonological feature [coronal]. The remaining differences between the markers of definiteness (manner, nasality) correspond to different morpho-syntactic features (class, number, case). This is in sharp contrast to traditional analyses in which the vowels in the nominative and accusative (i.e. *a, u/i*) are assumed to be 'markers of definiteness'. This traditional view seems obvious if one looks at the data:

(18) a. Dalina kishte një **libër**
           Dalina had   a   **book.INDEF**
           'Dalina had a book.'

      b. **Libri**   ishte shumë i bukur.
           book.DEF was   very   of beautiful
           'The book was very beautiful.'

By contrast, in the analysis provided in this paper, these vowels bear no specification of definiteness although they show up in definite contexts. All markers that represent the morpho-syntactic feature [+def] are consonants, more precisely coronal consonants. This systematic relation is not obvious by a look at the markers in a classical paradigm that are not decomposed and not sub-analyzed. The problem is, that there are markers in 'definite' cells of such paradigms (and that therefore have to be assumed to bear a [+def] feature) that are not coronal: e.g. /a/, /u/, and /ve/. But under an analysis in the framework of Distributed Morphology these markers do *not* carry the feature [+def] and their appearance in definite contexts is explained by the principles mentioned above: Underspecification of markers