

- Matzinger, Joachim
1998 Albanisch *unë* 'ich' im System der albanischen Personalpronomina. *Indogermanische Forschungen* 103: 185–201.
- Matzinger, Joachim
2006 *Der altalbanische Text [E] Mbsuame e Krështerë (Dottrina cristiana) des Lekë Matrënga von 1592. Eine Einführung in die albanische Sprachwissenschaft*. Dettelbach: Röhl.
- Matzinger, Joachim
2007 Altalbanisch <Bieh> /ujë/ „Wasser“ und die Kategorie der Massennomina bei Buzuku. In: Bardhyl Demiraj (ed.), *Nach 450 Jahren. Buzukus »Missale« und seine Rezeption in unserer Zeit. 2. Deutsch-Albanische kulturwissenschaftliche Tagung in München vom 14. bis 15. Oktober 2005*. Wiesbaden: Harrassowitz, 169–190.
- Newmark, Leonard, Philip Hubbard, and Peter Prifti
1982 *Standard Albanian. A reference grammar for students*. Stanford: Stanford University Press.
- Sasse, Hans-Jürgen
1991 *Arvanitika. Die albanischen Sprachreste in Griechenland*, Teil 1. Wiesbaden: Harrassowitz.
- Schumacher, Stefan
2004 *Die keltischen Primärverben. Ein vergleichendes etymologisches und morphologisches Lexikon. Unter Mitarbeit von Britta Schulze-Thulin und Caroline aan de Wiel*. Innsbruck: Institut für Sprachen und Literaturen der Universität.
- Schumacher, Stefan and Joachim Matzinger
2013 *Die Verben des Altalbanischen. Belegwörterbuch, Vorgeschichte und Etymologie. Unter Mitarbeit von Anna-Maria Adaktylos*. Wiesbaden: Harrassowitz.
- Topalli, Kolec
2011 *Gramatikë historike e gjuhës shqipe* [Historical grammar of Albanian]. Tirana: Albano-logjike.

Joachim Matzinger, Vienna (Austria)

Stefan Schumacher, Potzneusiedl (Austria)

97. The syntax of Albanian

- | | |
|---|----------------------|
| 1. Introduction | 4. Word order |
| 2. Nominal morphosyntax and
adpositional phrases | 5. Sentential syntax |
| 3. Verbal syntax | 6. References |

1. Introduction

Albanian is the stepchild of Indo-European linguistics, being perhaps the least investigated and least understood of the separate major branches of Indo-European. Moreover, within Indo-European historical investigations, syntax is perhaps the least explored component of grammar, and less is known about the syntax of Proto-Indo-European (PIE)

than about its phonology or morphology. Putting these two facts together means that obtaining a clear picture of Albanian historical syntax and the emergence of Albanian syntactic structures out of PIE is especially challenging.

This task is complicated by another factor, one that, at the same time, offers some important opportunities for insights into the extent and mechanisms of contact-induced language change. This factor is Albanian's involvement in the Balkan Sprachbund. That is, due to intense and sustained bi- and multi-lingual contact among speakers of various languages in the Balkans – most notably Aromanian, Bulgarian, Daco-Romanian, Greek, Macedonian, Romany, to a lesser extent Turkish, and of course Albanian (both major dialects, Geg and Tosk) – these languages have converged on common structures and common characteristics at all linguistic levels: phonology, morphology, lexicon, semantics, and syntax. Moreover, the syntactic parallels extend to nominal, verbal, and sentential syntax. Therefore, careful analysis is needed to differentiate those features of Albanian syntax inherited from PIE from those acquired by contact with neighboring languages; it is therefore always crucial to take the Balkan Sprachbund, and thus the possibility of contact-induced characteristics, into account whenever any discussion of Albanian is undertaken, especially when historical concerns are paramount.

In many ways, Albanian syntax is unremarkable from an Indo-European perspective, since among the key areas to consider, such as nominal case usage, subject-verb agreement, noun-adjective agreement, behavior of weak pronouns (“clitics”), presence of preverbs, occurrence of prepositions, the use of middle voice verb forms for reflexives and passives, impersonal verb forms, and the like, many represent, for the most part, familiar syntactic properties found in other branches of the family. Moreover, some aspects of Albanian syntax look rather like those found in “standard average European” languages, for instance several of the periphrastic tenses, and in that way they do not seem particularly “exotic” or unusual even if not dating to PIE.

Still, there are interesting and important characteristics to note about Albanian syntax, both synchronically and diachronically, with a mix of inherited elements from PIE usage and innovative constructions and uses involving both internally motivated and externally caused change. In what follows, various properties of Albanian nominal, verbal, and sentential syntax are surveyed, and what is interesting both from a general and from an Indo-European perspective is highlighted.

2. Nominal morphosyntax and adpositional phrases

Within the sphere of the syntax and internal structure of the Albanian noun phrase, especially noteworthy are the various “little words” or “particles” that occur with nominal forms and within the noun phrase. They are mostly found with various modifying elements, whether other nouns in possessive structures, markers of definiteness and specificity, or adjectives.

2.1. Modifiers within nominal phrases

To start with modifiers, they typically follow the noun, and with genitive case forms indicating a possessor of the head noun, there is an obligatory connective element linking

them to the noun they follow. This linking element is referred to variously as a “connective particle”, “adjectival article”, “*nyje* particle” (from the Albanian for “knot”), among other labels. The *nyje* element has either the form *i*, *e*, *të*, or *së*, depending on the case, gender, number, and definiteness of the modified noun (though *të* versus *së* is based on the final sound of the noun form, with *së* occurring after the noun ending *-s[ë]*). Some examples are given in (1):

- (1) a. *libr-i i Çimit*
book-the/NOM NYJE Çimi/GEN
‘Çimi’s book / the book of Çimi’s’
- b. *mora libr-in e Çimit*
took/1SG book-the/ACC NYJE Çimi/GEN
‘I took Çimi’s book / the book of Çimi’s’
- c. *mora një libër të Çimit*
took/1SG a book/ACC NYJE Çimi/GEN
‘I took a book of Çimi’s’

With adjectives, the connective element may or may not occur, with its presence or absence being a matter of morphological and lexical idiosyncrasy, depending on the derivation of the adjective: most basic adjectives are “articulated” (i.e., require the *nyje*) and certain suffixes always yield articulated adjectives while others (especially but not exclusively, those of foreign origin) always yield unarticulated ones. Some examples of each type are given in (2) and (3) respectively:

- (2) a. *libr-i i madh*
book-the/NOM NYJE big
‘the big book’
- b. *mora libr-in e madh*
took/1SG book-the/ACC NYJE big
‘I took the big book’
- c. *mora librat të mëdhenj*
took/1SG books/ACC NYJE big/MASC.PL
‘I took the big books’
- (3) a. *libr-i edukativ*
book-the/NOM educational
‘the educational book’
- b. *kam një mik djaloshar*
have/1SG a friend/ACC youthful
‘I have a youthful friend’.

The occurrence of the connective with basic adjectives and its general absence with forms of foreign origin suggest that this is an old trait within Albanian, but it is not one that predates Common Albanian.

Diachronically, the *nyje* forms continue old demonstrative elements (the *t/s* alternation from a formal standpoint, though not a distributional one, reflecting in some way the *t/s* alternation found for instance in Sanskrit *sa* ‘this/MASC’ vs. *tad* ‘this/NTR’, though other demonstrative elements might be involved), so that the original syntagm may have involved multiple markings for deixis and definiteness, reinterpreted as a linking element.

2.2. Definiteness within the nominal phrase

Definiteness is signaled by means of a postpositive marker, e.g. *katund* ‘village (NOM)’ / *katund-i* ‘the village (NOM)’, *vajzë* ‘girl (ACC)’ / *vajzë-n* ‘the girl (ACC)’. The definiteness marker, usually called an article, is actually postpositive (enclitic) within the noun phrase as a whole, attaching to the noun itself when the modifier has its usual position after the noun, e.g. *vajzë-n të shkretë* ‘the miserable girl (ACC)’, but attaching to the adjective when it precedes the noun, for emphasis or contrast, e.g. *të shkretë-n vajzë* ‘the miserable girl (ACC)’.

The postpositive article, like the connective, has its origins in PIE demonstrative elements (the *-n* of the accusative singular, for instance, probably reflects the outcome of the PIE accusative **-m* with a postposed demonstrative, that is **-m=tom > =n=tom > -nnV > -n[ë]*). It is a feature shared with other Balkan languages, in particular Macedonian, Bulgarian, Aromanian, and Daco-Romanian, though each language uses its own native material. Although it is likely to have diffused into these languages through contact, in this case, the postpositive placement may be a substratum feature of an autochthonous Balkan language predating Albanian (so Hamp 1982: 79, based on an analysis of the place name *Drobeta* as “a Latin misunderstanding or misparsing in Moesia Inferior of **druḡā-tā*, a definite noun phrase with postposed article”).

2.3. Nominal cases

As the examples above with a variety of nominal cases show, thematic and grammatical relations are indicated by case-forms of nouns. Besides the nominative, accusative, genitive, and dative exemplified above, there is also an ablative case, e.g. *zogj pulash* ‘birds from-hens (i.e. chicks)’. The ablative is distinct from the dative only in the indefinite plural forms, and it is used somewhat infrequently now, being increasingly replaced in many of its functions by various prepositional phrases or by the dative case.

In Old Albanian (e.g. in the 1555 Buzuku text) and dialectally in contemporary Albanian, there is also a form that is sometimes referred to as a locative case (so Newmark, Hubbard, and Prifti 1982), e.g. *malt* from *mal* ‘mountain’ with the preposition *në* ‘in, on’, thus *në malt* ‘in/on the-mountain’ (where accusative *malin* is found with *në* in other dialects and in the standard language now). This case is referred to as “instrumental” in Matzinger and Schumacher (this handbook, 2.1.).

2.4. Prepositions

In addition, Albanian has prepositions that govern nominals in different case forms and signal various adjunct and oblique grammatical relations within the clause. From an Indo-European standpoint, these are not all that remarkable, as all modern Indo-European languages have adpositions of some sort, even though older stages of some of them show adverbial elements (especially Vedic Sanskrit and Homeric Greek) that do not govern object nominals per se, suggesting that PIE may not have had any adpositions. If so, then the occurrence of prepositions in Albanian is an innovation away from PIE syntax but it is one that all the Indo-European languages took part in, a “drift”-like phenomenon.

Two Albanian prepositions, *nga* ‘from, by’ and *tek* ‘at/to the location of’, show the trait – unusual both from an Indo-European perspective and more generally cross-linguistically – of governing nouns in the nominative case. In the case of *tek*, this trait is explainable via its etymology, since this preposition apparently compresses within it traces of PIE correlative syntax, being originally ‘there where NOMINATIVE is’ (so Mann 1932: 72; see also Hamp *apud* Joseph and Maynard 2000), with the *t-* of *tek* reflecting the PIE *to- demonstrative and the *-k* the relative stem *k^w- (and with suppression of the copula, as is usual for PIE). The etymology of *nga* is more obscure, but one might expect a similar sort of explanation for its nominative “object”.

3. Verbal syntax

Several features of the Albanian verb qualify as noteworthy from the point of view of historical syntax, including the internal syntax of how certain verbal constructs are composed. Thus mention is made here of the way in which PIE preverbs are realized in Albanian, the formation of various multi-word periphrastic tenses, and the uses of the non-active (mediopassive) voice. Note too that the discussion of weak object pronouns below in 4.2. treats an aspect of Albanian verbal syntax in that the co-occurrence of such pronouns with full objects can be taken as a means of expressing transitivity and thus registering a verb’s argument structure.

3.1. Preverbs

Like all other Indo-European subgroups, Albanian shows the accretion onto a verbal root of prefixal elements generally referred to as “preverbs” that once (in PIE) were independent adverbial modifiers within the clause or verb phrase, as in dialectal *des* ‘die’ versus standard *vdes* ‘die’. In this regard, therefore, Albanian participated in the same “drift” as other Indo-European languages involving these original adverbials (see 2.4. on prepositions for another aspect of drift involving these elements). Many of these have traceable Indo-European pedigrees (e.g., regarding the form of *v-*, compare the Sanskrit preverb *ava*, and for the function of *v-*, compare Ancient Greek θνήσκω / ἀποθνήσκω ‘die’).

For the most part, just one preverb occurs on a verb at a time, so that in this way Albanian is unlike Indo-Iranian, Greek, Celtic, and Balto-Slavic. However, there are a few forms with multiple preverbs, at least from a diachronic perspective, since it is unclear that they could be so identified synchronically due to the degree of fusion between preverbs and verbal root. For instance, the stem *hëngër-*, which forms the suppletive past tense to the present *ha-* ‘eat’, derives from a sequence of multiple preverbs attached to a root: **Ho-en-g^wrō-*, where **Ho* corresponds to the initial element in Greek ὁ-κέλλω ‘I run (a ship) aground’ and **en* to Greek ἐν- as in ἐν-τρέπω ‘I turn in, and **g^wrō-* is the verbal root seen in Greek βι-βρώ-σκω ‘I eat’, Latin *vorō* ‘I devour’, etc. A similar phenomenon is seen with some preverbs and the PIE verbal past tense prefixal marker, the so-called “augment”, otherwise not overtly observable in Albanian. In particular, the verb *marr* ‘take’ is from a preverb **me* plus the root and nasal-present formation seen in Greek ἄρνημαι ‘I gain’, with the *-rr-* reflecting **-rn-*; to explain the vocalism in the past stem, *mor-*, one can posit **me* with the augment **e*, and just the root (with no nasal outside of the present system), with a fused (contracted) **mē* yielding Albanian *mo-*. The “interior” positioning of the augment parallels its placement with respect to preverbs in Greek and Sanskrit and thus may reflect an old feature, even if the univerbation took place at the level of the individual branches of Indo-European.

3.2. Periphrastic formations

Two-word syntactic combinations that fill paradigmatic slots, so-called “periphrastic” formations, are a key feature of Albanian morphosyntax. The future tense, the perfect system forms, and the modal category known as the “admirative” – indicating (among other modalities) a speaker’s surprise at some unexpected aspect of an event or situation – all now involve, or historically did involve, periphrasis, as does the expression of progressive aspect. In addition, various nonfinite formations are multiword periphrases based on the Albanian participle.

3.2.1. Future tense

There is a major dialectological split within Albanian between a periphrastic future based on ‘have’, found in Geg dialects, and one based on ‘want’, found in Tosk dialects (though the dialect distribution is somewhat more complicated). The Geg future uses an infinitive (marked by a prefixal element *me*) introduced by an inflected ‘have’ auxiliary, whereas the Tosk future uses a finite subjunctive, introduced by a fixed invariant form *do*, the third person singular form of ‘want’ (but with its volitional meaning depleted):

- (4) a. (Tosk)
- | | | | | |
|-----------------------------|-----------|--------------|---|---------------------|
| <i>do</i> | <i>të</i> | <i>shkoj</i> | / | <i>do të shkosh</i> |
| want/3SG | SUBJ | go/1SG | | go/2SG |
| ‘I will go’ / ‘you will go’ | | | | |

b. (Geg)

kam me shkue / ke me shkue
 have/1SG INF go/PPL have/2SG
 ‘I will go’ / ‘you will go’

Both formations represent innovations away from the PIE monolectal (synthetic) future, and both must be considered in the context of the Balkan Sprachbund. The ‘want’-based future, especially with an invariant future marker involved, is found in Greek, Aromanian, Daco-Romanian, Macedonian, Bulgarian, and Romani, whereas a ‘have’-based future is found in Macedonian and Bulgarian (where the distribution is grammatically determined, with ‘have’ found mainly in negative forms, and ‘want’ elsewhere) as well as in Daco-Romanian (competing with the ‘want’ future, with some nuanced meaning differences) and some dialects of Aromanian (in negated forms, probably calqued on Macedonian). The exact source of the Albanian futures may well thus lie in contact with one (or more) of those languages, though Vulgar Latin, an important contact language for prehistoric Albanian in the Balkans, may have played a role (note the ‘have’ futures of modern Romance languages, for instance, and there are future-like uses of *volō* ‘I want’ in late Latin). Moreover, given the existence of parallels outside of Indo-European to both types of future formation, independent emergence of each within Albanian cannot be discounted. But the periphrastic composition of each type historically is clear.

3.2.2. Perfect system

Replacing the synthetic perfect of PIE, Albanian developed a periphrastic perfect, with the verb ‘have’ as an auxiliary for active forms and ‘be’ as an auxiliary for non-active forms; in each case, the main verb is expressed as a participle. Examples are given in (5):

- (5) a. *kam larë / kemi larë*
 have/1SG wash/PPL have/1PL
 ‘I have washed’ / ‘we have washed’
- b. *jam larë / jemi larë*
 be/1SG wash/PPL be/1PL
 ‘I have been washed’ / ‘we have been washed’

A full set of forms is possible, covering all verbal categories of tense and mood; for instance, a pluperfect active and perfect subjunctive active are given in (6a), and an optative perfect non-active in (6b):

- (6) a. *kishim larë* ‘we had washed’
të kemi larë ‘that we have washed’
- b. *qofsha larë*
 be/OPT.1SG wash/PPL
 ‘may I have been washed’

The innovation of an analytic periphrastic perfect is found in the later stages of several branches of Indo-European (compare English and German, for instance, and Romance), so in that sense, here again, Albanian is taking part in a development that may be associated with another characteristic Indo-European drift, in this case towards analytic structures. At the same time, periphrastic perfects with ‘have’ are found in most of the languages of the Balkan Sprachbund (in Macedonian, for instance, such a formation has developed and has come to occupy a different niche in the verbal system from that of the inherited Slavic ‘be’-based perfect), with the pluperfect being a key point of convergence among the languages (it was the point of entry for the whole ‘have’-based perfect of Modern Greek, for instance).

3.2.3. Admirative

Although the use of the admirative is connected with pragmatics and discourse factors, its form clearly reflects an origin in a syntactic combination akin to a perfect formation, consisting of a truncated participle with a postposed inflected form of ‘have’ fused to the participle. There are admirative forms in all tenses and moods, active and non-active; (7) has a sampling (see 3.3. on the non-active formation in [7c]) with glosses that are inadequate as they are not in a suitable discourse context:

- (7) a. *paskam* ‘I might have’ (cf. participle *pasur* ‘had’)
 genke ‘are you really?’ (cf. participle *genë* ‘been’)
 b. *paskam larë* ‘I might have washed’ (PERF.ADM)
 c. *u lakam* ‘I might wash myself, I might be washed’ (cf. participle *larë*)

Although built with native Albanian material, the admirative is clearly an innovation, constituting a category that could not have been a part of the PIE verbal system (inasmuch as it is absent from every ancient Indo-European language). It shows affinities with similar categories in Macedonian and Bulgarian that were built on their perfect formations; in the emergence of this category, all of these languages may have been influenced by Turkish, a language with an inherited category marking a speaker’s epistemic stance towards a narrated event.

3.2.4. Nonfinite formations

Albanian inherited a participle, generally ending in *-r* in Tosk, reflecting a PIE **-no-* suffix, that, like analogously formed participles in Sanskrit (and cf. Hittite **-nt-* participles), generally has a passive value when formed from transitives and an active value when formed from intransitives, e.g. *shkruar* ‘(having been) written’, *shëtitur* ‘(having) strolled’. From this participle, a variety of periphrastic nonfinite formations are made, all innovative, *vis-à-vis* PIE, in form and to a large extent in function; following the terminology of Newmark, Hubbard, and Prifti (1982: 64–65), these are given in (8):

- (8) a. Privative: *pa larë* ‘without washing’
 b. Gerundive: *duke larë* ‘while washing’
 c. Infinitive: *për të larë* ‘(in order) to wash’
 d. Absolutive: *me të larë* ‘upon washing’

Note too that Geg has an infinitive formed with *me* and a shortened form of the participle, e.g. *me punue* ‘to work’, as opposed to the widespread Tosk *për të punuar*. Still, there are traces of a Geg-like infinitive with *me* in some dialects of Tosk.

The composition of these formations is fairly clear and suggests a relatively recent development; most of the relevant formative elements occur otherwise as prepositions with nominal objects (see 2.4.) – cf. *pa* ‘without’, *për* ‘for’, *me* ‘with’. It is likely moreover that the *të* of the Tosk infinitive is a nominalizing element (perhaps to be identified with the *nyje* particle) that combines with participles; cf. *të dhënat* ‘data’, from the participle of ‘give’ (see also 4.3.). In that regard, inasmuch as infinitives in other Indo-European languages typically are formed from deverbal nouns, and the **-no-* suffix of the participle that figures in the Albanian infinitival formation also occurs in the Germanic infinitive (cf. Gothic *bairan* ‘to bear’, from **bheronom*) and forms a deverbal derivative in Sanskrit (Ved. *bhāraṇam* ‘[an act of] bearing’), the Albanian infinitive may be a replacement for a PIE infinitival prototype rather than a wholly innovated category and formation. Further, if the occasional *me* formations in some Tosk dialects are taken seriously as relics, and not as borrowings from Geg, that proto-Albanian infinitive may well have been of the Geg type.

3.2.5. Progressive aspect

One further periphrasis with grammatical value is seen in the two ways in which the indication of progressive aspect in the present and past can be realized. The marker *po* can occur with present and imperfect tense forms, as in (9ab).

- (9) a. *Ç’ po bën tani?*
 what PO do/2SG now
 ‘What are you doing now?’
 b. *Po të vështroja*
 PO you/ACC watch/1SG.IMPF
 ‘I was watching you’
 c. *Jam duke të vështruar*
 be/1SG PROG you/ACC watch/PPL
 ‘I am watching you’

The second type seen (9c), being built on a relatively new nonfinite formation, most likely itself represents a recent development, but the type with *po* is surely an old feature of Albanian, as it is found in both major dialects, even if innovative from the standpoint of PIE. Newmark, Hubbard, and Prifti (1982: 36) identify this verbal *po* with the “emphatic particle” *po* meaning ‘yes, indeed, exactly so!’, though perhaps in a different way;

Hamp (*apud* Joseph 2011) has derived *po* from **pēst*, a combination of an asseverative particle **pe* (cf. Latin *quip-pe* ‘why so?; of course’ [from **k^wid-pe*]) with a 3sg injunctive form of **H₁es-* ‘be’, so that *po* is etymologically ‘[it is] just [now] so’, and this “just-now” meaning is the basis for the emergence of a temporal progressive sense for *po*. Interestingly, this usage has no counterpart in any of the other Balkan languages.

3.3. Non-active voice

Albanian has a categorial distinction between active and non-active voice, where the non-active corresponds to what is also called “middle” or “mediopassive”. There is a distinct set of endings added to a special stem in the present non-active system (taking in the present, imperfect, and future tenses and the subjunctive mood), and in other forms (taking in the past tense, the optative, admirative, and imperative moods, and nonfinite forms) the non-active is formed from the combination of active forms with a voice marker *u*, that is generally a prefix (but postposed in the imperative). Returning to the theme of 3.2.3., there is a periphrastic non-active in the perfect, consisting of ‘be’ plus the participle. Some examples of all of these formations are given in (10):

- (10) *laj* ‘I wash’ / *lahem* ‘I am washed, I wash myself’
lava ‘I washed’ / *u lava* ‘I was washed, I washed myself’
lafsha ‘may I wash’ / *u lafsha* ‘may I be washed, may I wash myself’
për të larë ‘(in order) to wash’ / *për t’u larë* ‘(in order) to be washed; to wash myself’
kam larë ‘I have washed’ / *jam larë* ‘I have been washed, I have washed myself’

As the glosses in (10) indicate, the uses of non-active forms include passive and reflexive meanings; in plurals, a reciprocal sense is possible too, e.g. *lahemi* ‘we wash each other’. Some verbs are deponent, occurring only in the non-active, even if their meaning is active, e.g. *kollem* ‘I cough’. In addition, there is an impersonal use of the third person non-active forms, most often negated, to indicate a generalized activity, even with intransitives, e.g. *s’shkohet* ‘there’s no going’ (cf. *shkon* ‘it goes’).

These uses are familiar and widespread across Indo-European (cf. the Greek and Sanskrit middle voice), and thus they surely continue PIE uses of non-active. From the standpoint of form, it is noteworthy that Albanian is one of the two modern Indo-European languages, along with Greek, that has an inherited distinct monolectal (synthetic) verbal form for the non-active. For Albanian, though, the synthetic form is restricted to the present system and related forms; in the aorist (and other categories, especially the nonfinite forms) one encounters the analytic formation, as in (10), employing the particle *u*, which derives (in a somewhat complicated way) from the PIE reflexive element **swe*.

4. Word order

The order of elements in the Albanian clause is typically subject – verb – object, when full nominals are involved as subject and object. Still, case-marking and the use of weak

object pronouns to register arguments on the verb (see 5.2.) allow for greater freedom of order for the major constituents of a clause, in some instances associated with pragmatic factors such as topicality. Moreover, it is quite common in discourse for full nominals to be replaced by pro-forms, in particular weak object pronouns for direct and indirect objects, and “zero” (the absence of an overt form altogether) for the subject. The freedom of constituent order in Albanian parallels what is found in other Indo-European languages with similar morphological cues for identifying arguments.

5. Sentential syntax

In the area of clausal syntax, there are three main phenomena to consider: negation, weak object pronoun (“clitic”) behavior, and complementation.

5.1. Negation

Albanian has a distinction between what may be called “modal” and “nonmodal” negation, roughly equivalent to nonindicative versus indicative negation. Thus, as in (11), the present, imperfect, aorist, perfect, and future tenses are negated with *s'* or *nuk*, whereas imperative, subjunctive, and optative forms (as well as nonfinite formations), as in (12), are negated with *mos*:

- (11) a. *Unë nuk e njoh* ‘I do not know him’ (PRES)
- b. *S’ke para* ‘You do not have money’ (PRES)
- c. *Nuk do të vijë atje* ‘He won’t come here’ (FUT)
- d. *Nuk lexonte* ‘He was not reading’ (IMPF)
- e. *S’lexuam një libër* ‘We did not read a book’ (AOR)
- f. *S’e kanë parë* ‘They haven’t seen him’ (PERF)
- (12) a. *Përpiqet të mos qeshë* ‘He tries not to laugh’ (SUBJ)
- b. *Mos shko në Tiranë* ‘Don’t go to Tirana’ (IMPV)
- c. *Mos vdeksh kurrë* ‘May you never die’ (OPT)
- d. *Erdha në Tiranë për të mos u mërzitur* ‘I came to Tirana (in order) not to be bored’ (NONFINITE)

This differential usage of *nuk/s'* and *mos* continues an old distinction, one that is inherited from Proto-Indo-European. Greek, Armenian, and Indo-Iranian show essentially this same distinction: Ancient Greek – οὐ versus μή (Modern Greek *den* [from Ancient Greek οὐδέν, built with the οὐ negator] versus *mi*); Armenian – *oč'* versus *mi*; Sanskrit/Avestan – *na* versus *mā*. μή/*mi*/*mā* negate modal forms and οὐ/*oč'*/*na* negate indicatives. They reflect a PIE distinction of indicative **ne* versus modal **mē* (Greek οὐ and Armenian *oč'* indirectly so, being from a truncation of **ne H₂oyu k^{wid}* ‘not ever at-all’ [Cowgill 1960], to which Albanian *as-* ‘no-’ [as in ‘no one’ or ‘nothing’ or ‘nowhere’] may belong, just as *s'* represents a truncation of **né k^{wid}*, with the same extension as in *mos*, from **mē k^{wid}*).

At the same time, *mos* shows some innovative uses that in part go beyond negation, and, interestingly, are shared by Greek and in some instances other Balkan languages. One such use is in dubitative questions, as in (13):

- (13) *mos e njihni atë?*
mos him know/2PL him/ACC
 ‘Do you perhaps know him?’

which is matched functionally by such questions with μή in Ancient Greek, and, as in (14), *mi* in Modern Greek, but is found in no other Indo-European languages:

- (14) *Min ides to peđi?*
mi saw/2SG the child
 ‘Did you perhaps see the child?’

Thus, this may well be an early Greek innovation that was borrowed (calqued) into Albanian, but still represents a new usage that entered Albanian post-PIE.

Another such innovation with *mos* is an independent use as a one-word prohibitive utterance (15a), also found in Modern Greek (15b) and Romani (15c), but interestingly, not in Ancient Greek nor in any other Indo-European language:

- (15) a. *Mos!* ‘Don’t’
 b. *Mi!* ‘Don’t!’
 c. *Ma be, Ismet!* ‘Hey you, Ismet, don’t [*ma*]!’

Given the absence of this usage from Ancient Greek, it quite possibly reflects an Albanian innovation that spread into Modern Greek (and Romani).

Both the question use and the independent prohibition use of *mos* may reflect extensions within Albanian of simple prohibitive **mē*, inasmuch as the usage in (15) is clearly related to the expression of verbal prohibitions (possibly, therefore, through elision of a now-only-implicit verb), and the uses in (13)/(14) are associated with weak negation of a modal type. However, given the chronological and geographical distribution of clear parallels in Indo-European outside of Albanian, they seem to represent innovations affecting Albanian that took place on Balkan soil, whether emanating from Albanian itself or finding their way into Albanian from some other Balkan language.

5.2. Clitics

Another important aspect of Albanian clausal syntax is the occurrence of weak (so-called “clitic”) forms of personal pronouns, e.g. accusative/dative *mē* ‘me’ (versus “strong” *mua*), accusative/dative *e* ‘him, her’ (versus strong *atë*), or dative *u* ‘to them’ (versus strong *atyre*). The presence of such forms in the grammar of Albanian is surely a reflex of a PIE strong/weak distinction, given that similar alternations are found in Greek, Hittite, Indo-Iranian (especially Vedic and Avestan), Old Church Slavonic, and Old Irish, among other languages, and to some extent, the forms of the weak pronouns match up

well (*m-* in first person singular, *t-* in second person singular, *n-* in first person plural, etc.).

The positioning of the weak pronouns, however, is probably not old but shows affinity with the innovative positioning of parallel elements in Greek (innovative *vis-à-vis* Ancient and Medieval Greek, cf. Pappas 2003) and Macedonian (innovative *vis-à-vis* South Slavic, as a comparison with Old Church Slavonic and Bulgarian shows) and is thus probably tied in some way to contact among these languages. In Albanian, the weak pronouns precede all verb forms, though with imperatives they may show postpositive placement:

- (16) a. *Unë nuk e njoh*
I/NOM NEG him know
'I do not know him'
- b. *Më njihni mirë*
me/ACC know/2PL well
'You know me well'
- c. *Mund të ju ndihmojmë*
can SUBJ you help/1PL
'We can help you'
- d. *Pa e pare ikën*
without him see/PPL left/3PL
'Without seeing him, they left'
- e. *C' ju paska ngjarë*
what you have/ADM.3SG happen/PPL
'What on earth happened to you?'
- f. *Na shkruaj! Shkruaj na!*
us/DAT write/IMPV
'Write to us!'

Assuming some sort of "Wackernagel" placement of weak pronouns for PIE, that is, in second position within their governing unit (phrase or clause), as proposed by Wackernagel (1892), the Albanian placement shows two innovations: it is verb-centered (always adjacent to the verb), rather than positioned relative to some element in the clause or phrase, and it involves (nearly) constant pre-positioning (proclisis) of the weak form. The postpositive (enclitic) placement in the imperative could, however, reflect an inherited trait, since imperatives typically would be initial within their clause (as the lone verb with the subject suppressed), and thus an enclitic element would actually be in second position.

One striking fact about the placement of weak pronouns is their positioning in the imperative plural, where the pronoun can be interior to the person/number marker, thus an apparent "endoclititic" (a word-internally positioned clitic):

- (17) *Shkruamëni* 'Write to me!' (vs. *Më shkruani* 'idem').

Admittedly, this placement may say more about the nature of the 2PL ending *-ni* than about the pronoun, since *-ni* shows other signs of having a “freer” status than that of other person/number endings. In particular, it can occur as a “plural” marker with a number of interjections, adverbials, particles, and even greetings, forms that would not ordinarily be thought of as being compatible with a verbal plural ending; a few such cases are given in (18), from Newmark, Hubbard, and Prifti (1982: 324):

- (18) a. *Mosni* ‘Don’t (you all)!’ (cf. *Mos!* ‘Don’t’ [15a])
 b. *Forcani* ‘Heave ho (you all)!’ (cf. *forca* ‘heave ho!’ [from *forca* ‘powers’])
 c. *Mirëmëngjesni* ‘Good morning (you all)!’ (cf. *Mirëmëngjes* ‘good morning’)

These suggest that *-ni* may have once had greater freedom than an ending like *-(j)më* for first person plural, and if so, then diachronically *shkrua-më-ni* might reflect a later accretion of a once-independent “ending” onto an imperative form with a postpositive weak pronoun object.

Interestingly, there is a parallel in Modern Greek to this seemingly unusual pronoun placement in imperatives; in Thessalian Greek (see Tzartanos 1909), one finds intercalated *-m-* for a first person singular object between the root and the second plural imperative ending, with a few verbs, e.g. *do-m-ti* ‘(You all) give me!’ (literally: “give-me-PL!”). The *shkrua-më-ni* placement, therefore, may represent a contact-induced innovation in Albanian, though it is as likely that Greek borrowed this construction from Albanian (specifically, from Arvanitika, the Tosk Albanian dialect spoken in Greece for the past 600 years or so, with a heavy concentration of speakers in central Greece), and independent innovation cannot be ruled out.

A further innovative aspect of Albanian syntax involving weak pronouns is that they can co-occur with full object forms, either strong forms of pronouns or full noun phrases, as in (19):

- (19) a. *E pashë Gjonin* ‘I saw John’ (literally: “him I-saw the-John”)
 b. *Të pamë ty* ‘We saw you’ (literally: “you we-saw you”)
 c. *I dha Gjonit një libër* ‘He gave John a book’ (literally: “to-him he-gave to-the-John a book”)

This “clitic doubling” (also called “object reduplication”) has a largely pragmatic function, having to do with information flow, topicality, focus, and the like (see Friedman 2008). However, in certain contexts, it has a purely grammatical (i.e. syntactic) function, occurring obligatorily when co-indexing a dative case-marked indirect object, so that (19c) without the cross-indexing *I* doubling the object, is ungrammatical (**Dha Gjonit një libër*).

Clitic doubling occurs throughout the Balkan languages. In Greek, it is entirely pragmatically linked, whereas in Macedonian, a grammatical use parallel to that in Albanian is found, with obligatory doubling of full indirect objects. Given the distribution of this phenomenon and its relatively late appearance in Greek (i.e., it is not part of Ancient Greek syntax) and in other Indo-European languages (e.g. in Spanish, but not in Latin), clitic doubling seems to be a Balkan innovation that has entered Albanian. Most likely its emergence is to be tied to the need for communicative clarity, as expressed through

redundancy and emphasis, in cross-language interactions among speakers of different languages with less-than-native command of the other language.

One final use of a subset of the weak pronouns that is noteworthy is for marking the speaker's emotional involvement in the action expressed in the clause. There is no doubling, since there is no formal object, and the only forms possible here are the first and second person singular and first person plural weak dative pronouns. The function is essentially that of the so-called "dative of interest" or "ethical dative", and as such, given that there are parallels for this usage in non-Balkan Indo-European languages (e.g. Latin and Germanic), it most likely represents an archaism in Albanian grammar. Some of the more complicated combinatory possibilities that result, as with the 1SG + 2SG + 3PL in (20) (from Newmark, Hubbard, and Prifti 1982: 27), indicating here both speaker and hearer involvement and a doubled object, may well be an Albanian innovation.

- (20) *ai fiku ... më t' i bënte*
 this-fig/NOM me/DAT you/DAT them/ACC made
ato kokrrat na ... sa një ftua
 these bits/ACC behold as-big-as a quince
 'That fig tree produced ("for me for you them") figs, wow ... the size of a quince'.

5.3. Complementation

One further striking feature of Albanian syntax that aligns it with Balkan Sprachbund languages and differentiates it from most other Indo-European languages is the preponderance of subordinate clauses with finite verbs – most typically subjunctives marked with *të* – inflected for person and number. This finite complementation means, from a structural standpoint, that all verbs in a sentence are fully "specified" as to person and number and in some instances, tense. This is a feature which links Albanian to the Balkan Sprachbund, as it is found, to varying degrees throughout the region, most thoroughly in Greek and Macedonian, and fairly intensely in Bulgarian, Aromanian, and Daco-Romanian. Presumably, therefore, this phenomenon is not all that old in Albanian, and dates to the period of intense contact with other Balkan languages in the Middle Ages (especially the Ottoman period). Like clitic doubling (5.2.), the use of finite complements instead of infinitives may have been a function of a desire on the part of speakers for clarity of communication via redundancy in a multi-lingual contact situation. (See Joseph 1983 on this Balkan trait, and Chapter 4 on Albanian specifically.)

The extensive use of finite complementation is actually more a feature of the Tosk dialect of Albanian (and thus of the standard language, which is generally based on Tosk) than of the Geg dialect. As noted in 3.2.4., Geg has an infinitive, consisting of the marker *me* with the participle, and it is used in complementation in contexts in which Tosk uses a finite complement. Some Tosk examples of finite complements, governed by verbs, adjectives, and nouns, are given in (21), and some Geg examples of infinitival subordination, governed by verbs, nouns, and a subordinating conjunction, are seen in (22).

- (21) a. *dua të shkruaj*
want/1SG SUBJ write/1SG
'I want to write' (literally: "I want that I write")
- b. *do të vazhdojmë të ulemi atje*
FUT continue/1PL SUBJ sit/1PL here
'We will continue to sit here' (literally: "We will continue that we sit here"; note that the future is formed by placing the particle *do* before the present subjunctive)
- c. *mund të shkoni në Tiranë*
can SUBJ go/2PL to Tirana
'You (all) can go to Tirane' (literally: "It-can that you go to Tirana")
- d. *është e vështirë të qeshin*
be/3SG difficult SUBJ laugh/3PL
'It is difficult for them to laugh' (literally: "It is difficult that they laugh")
- e. *propozimi të shkojmë në Shqipëri*
proposal-the SUBJ go/1PL to Albania
'the proposal for us to go to Albania' (literally: "the proposal that we go to Albania")
- f. *vijmë në shkollën që të mësojmë*
come/1PL to school in-order SUBJ learn/1PL
'We come to school in order to learn' (literally: "We come to school in order that we learn")
- (22) a. *dëshiroj me të pa*
desire/1SG INF you/ACC see/PPL
'I desire to see you'
- b. *puna me e shue këtë politike*
task-the INF it wipe-out/PPL this-policy/ACC
'the task of wiping out this policy'
- c. *shkoi përjashta me mësoje filozofi*
went/3SG abroad INF study/PPL philosophy
'He went abroad (in order) to study philosophy'
- d. *sado me u kujdesue ...*
despite INF REFL worry/PPL
'Despite (his) worrying,...'

Interpreting these facts historically is even further complicated by the fact that Tosk also has an infinitive, as seen above in 3.2.4., with the form *për të* + Participle. The infinitive in Tosk has rather limited uses, mainly occurring in the expression of purpose, though it can be used in complementation, as in (23).

- (23) a. *ata folja shërben për të emërtuar një veprim*
these verbs serve/3PL INF designate/PPL an action
'These verbs serve to designate an action'

- b. *është e vështirë për të thënë*
be/3SG difficult INF say/PPL
'It is difficult to say'
- c. *propozimi për të vënë në funksionin artilerinë*
proposal-the INF put/PPL in function artillery
'the proposal to put artillery into place'

The *për të* + participle formation has the appearance of being a relatively recent creation. Importantly, a formation that is somewhat similar, but at the same time different in a significant way, is found in Old Albanian. In the Buzuku text, *për të* occurs with a true nominalized element, clearly so since it shows marking for definiteness and case, e.g. *për të lutunit* 'for the prayer' (with definite dative case marking on the participle *lutun*-from *lus* 'invoke'). Moreover, non-active voice marking as illustrated in 3.3., (10), seems not to occur with these early *për të* formations (and is not allowed in the ostensibly parallel Arvanitika formation). The passage from a nominal formation to a verbal one, capable of marking voice distinctions, is thus an innovation that took place within historically documented Albanian.

6. References

- Cowgill, Warren
1960 Greek *ou* and Armenian *oč'*. *Language* 36: 347–350.
- Friedman, Victor A.
2008 Balkan object reduplication in areal and dialectological perspective. In: Dalina Kallulli and Liliane Tasmowski (eds.), *Clitic Doubling in the Balkan Languages*. Amsterdam: Benjamins, 35–63.
- Hamp, Eric P.
1982 The Oldest Albanian Syntagma. *Balkansko ezikoznanie* 25: 77–79.
- Joseph, Brian D.
1983 *The Synchrony and Diachrony of the Balkan Infinitive: A Study in Areal, General, and Historical Linguistics*. Cambridge: Cambridge University Press.
- Joseph, Brian D.
1989 Η ερμηνεία μερικών βορείων τύπων της προστακτικής κατά τη σημερινή μορφολογική θεωρία [The Interpretation of Several Northern Forms of the Imperative According to Current Morphological Theory]. *Ελληνική Διαλεκτολογία* 1: 21–26.
- Joseph, Brian D. and Kelly Maynard
2000 Hamp Lectures on the Albanian Language, Ohio State University 11/29–12/4, 1999. *Indo-European Studies Bulletin (University of California at Los Angeles)* 9(1): 25–27.
- Mann, Stuart
1932 *A Short Albanian Grammar with vocabularies, and selected passages for reading*. London: David Nutt.
- Newmark, Leonard, Philip Hubbard, and Peter Prifti
1982 *Standard Albanian. A reference grammar for students*. Stanford: Stanford University Press.
- Pappas, Panayiotis
2003 *Variation and Morphosyntactic Change in Greek: From Clitics to Affixes*. New York: Palgrave MacMillan.

Tzartanos, Achilleus

1909 Περὶ τοῦ συγχρόνου Θεσσαλικῆς διαλέκτου [On the modern Thessalian dialect]. Athens: P. A. Petrakou.

Wackernagel, Jacob

1892 Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1: 333–436.

Brian D. Joseph, Columbus, OH (USA)

98. The lexicon of Albanian

- | | |
|-------------------------|---------------------------------------|
| 1. Introduction | 5. Word-formation |
| 2. Inherited vocabulary | 6. Future perspectives and desiderata |
| 3. Loan-words | 7. References |
| 4. Specific vocabulary | |

1. Introduction

1.1. Fundamentals of Albanian lexicology

Apart from some scarce evidence (proper names, single words, and single phrases), the literary documentation of Albanian begins with theological texts in the mid-16th century CE. Mostly translations of Latin originals, these texts were written by representatives of the Catholic clergy. The first author known to us is Gjon Buzuku (“Missal” of 1555), a priest from the Geg dialect area. Literature in the Tosk dialect begins with the work of the Italo-Albanian priest Lekë Matrënga (*Dottrina cristiana* of 1592). After the Ottoman conquest of the Balkans – some one hundred years before the aforementioned first written Albanian records – Albania became an integral part of the Ottoman empire for 500 years. As a consequence of this long-lasting Ottoman rule Albanian literary production came to a standstill. The theological documents of the 16th and 17th centuries thus constitute the Old Albanian literature. It was only in the second half of the 19th century that Albanian literature outside and – to a smaller extent – inside Albania began anew. This fact was caused by the so-called National Awakening (*Rilindja* [*kombëtare*]), the struggle of the Albanians against Ottoman domination leading to the independence of the emerging Albanian state in 1912. With respect to literature, this comprises the very fruitful period of the great Albanian classical writers culminating in the works of Father Gjergj Fishta (1871–1940). During the time of the Rilindja, the Albanian lexicon increased impressively because of the great number of neologisms and calques created by the Rilindja writers (see Buchholz and Fiedler 1979). Albania’s best known contemporary author is Ismajl Kadare (born 1936).

<https://doi.org/10.1515/9783110542431-019>