Chapter 5

Clitic doubling in Albanian dialects from the perspective of functional transparency

Veton Matoshi

LMU Munich

The relevant literature reports differences in the use of clitic doubling in Albanian dialects. Quantitative corpus studies show that all dialects spoken outside of the Republic of Albania show a more frequent use of clitic doubling. The data of this corpus prove that the less restrictive use of clitic doubling is not accompanied by increasing transparency of its usage. In contrast to Standard Albanian, where the usage of clitic doubling is not optional and can almost without exception be explained by topic and focus marking, in the peripheral Albanian dialects outside of the Republic of Albania numerous exceptions from the general tendency can be detected. In order to explain these exceptions, a wide variety of factors must be taken into account and, in certain contexts, point to the optional use of clitic doubling. From a descriptive point of view, these exceptions suggest an increasing degree of functional opaqueness.

1 Introduction

Most studies on differential object marking and clitic doubling, including this one, follow a simple key question: What are the properties of objects that trigger clitic doubling? In order to approach this question from a broader perspective, it is useful to outline general aspects of typologically grounded theories on agreement and transitivity. Typological studies question the purely syntactic notion of transitivity which posits the existence of an object as a sufficient indicator

of transitivity. This dichotomous view (intransitive vs. transitive) can be supplemented by a semantic approach to transitivity, which states that clauses exhibit a certain degree of transitivity depending on the pragmatic and semantic properties of the agent and the patient. The degree of transitivity can be measured on the basis of certain properties of the core arguments and the semantics of the verb. Table 1 summarizes the basic idea of prototypical transitivity according to Hopper & Thompson (1980).

Table 1: Parameters of transitivity according to Hopper & Thompson (1980: 252).

| Factor | Degree of transitivity | |
	High	Low
a. Participants	2 or more participants, A and O	1 participant
b. Kinesis	action	non-action
c. Aspect	telic	atelic
d. Punctuality	punctual	non-punctual
e. Volitionality	volitional	non-volitional
f. Affirmation	affirmative	negative
g. Mode	realis	irrealis
h. Agency	A high in potency	A low in potency
i. Affectedness of O	O totally affected	O not affected
j. Individuation of O	O highly individuated	O non-individuated

Prototypical transitive clauses exhibit an object which shows a high degree of individuation and affectedness due to the action executed by the agent which is prototypically a deliberately acting agent. For instance, definite and highly referential objects (criterion J) are analyzed as being a more typical component of transitive clauses than indefinite and less referential objects. Hopper & Thompson (1980: 259) go so far as to posit – more like an "extreme restatement" – that "an indefinite O is not really an O at all, but is a subordinate part of a compound of which the verb stem is the head (i.e., it is incorporated into the verb)".

Instances of "special" object marking can be considered, in a broad sense, as morphosyntactic realizations of a higher or lower degree of transitivity, among them:

- clitic doubling = high transitivity

- differential object marking = high transitivity

- object omission (anti-causative, incorporation) = low transitivity

Another way of viewing transitivity is the notion of distinctness of participants or the "maximally distinguished argument hypothesis" (Næss 2007) which states, in simple terms, that the two main participants (subject and object) must be semantically maximally distinguishable from each other. Lenz (1920: 52) was one of the first to assume a distinguishing function of differential object marking. He assumed that in Spanish the direct object bears the preposition *a* only if it is logically possible to perceive it as the subject of the clause.

The introduction so far has shown that clitic doubling can be viewed in the wider context of argument alignment and, what is more important, as a realization of differential object marking (Kallulli 2016). Bossong (1991: 151) regarded the emergence of differential object marking as a form of "grammemic replacement" of eroded case systems with the important difference that case systems, such as the Latin case system, are "a petrified grammatical category whereas the more recent DOM [differential object marking] systems are living ones". As a consequence, the former "are used mechanically and without exception" and are "meaningless", whereas the usage of differential object marking is not pervasive and is dependent on different factors, mostly regarding the defining properties of the object and the action denoted by the verb.

One major factor which can be associated with the emergence of a new case or agreement system is topicality. Givón (1976: 152) introduces a universal hierarchy of topicality that ranks the core arguments regarding their likelihood to be topics as follows: agent > dative > accusative. This hierarchy serves as a starting point for Givón's theory on the emergence of object (and subject) agreement which basically states that an overuse of the topic-shift construction will eventually lead to the grammaticalization of object agreement, including clitic doubling in this context, cf. Table 2.

Table 2: Grammaticalization of object agreement via topic-shift constructions according to Givón (1976: 157).

topic shift ("marked")	afterthought-topic construction ("semi-marked")	neutral ("demarked")
The man, I saw him.	*I saw him, the man.*	*I saw-him, the man.*

Very often topic and focus are viewed as being in a complementary distribution (see for example Buchholz & Fiedler 1987: 538, Kallulli 2000: 218 for Alba-

nian), which is reminiscent of what is traditionally called topic and comment. It should not go unmentioned that major caveats have been expressed concerning such a definition of topic and focus (see for example Leafgren 2002: 30). The absence of a widely accepted definition of the terms *topic* and *focus* and their importance for clitic doubling in Albanian require a general definition of the terms as they are used in this article. The term *topic* is used according to the notion of *aboutness*-topic, thus the "topic of a sentence is the thing which the proposition expressed by the sentence is about" (Lambrecht 1994: 118). Focus, on the other hand, "refers to significant emphasis on a particular element within the context of the information conveyed in a particular clause" (Leafgren 2002: 23–24). It becomes clear that these definitions call into question the view of a complementary distribution of topic and focus. They do not, however, refute the view that in the majority of cases topic and focus do not overlap, as it is usually the part of a clause which conveys the new information that also bears a significant emphasis and consequently stands out. This gradual tendency can be depicted using a modified version of the topic acceptability scale of Lambrecht (1994: 165) in Table 3.

Table 3: Topic acceptability scale according to Lambrecht (1994: 165) (the grey part is a modification by the author).

active	most acceptable topic	least acceptable focus
accessible		
unused	↓	↑
brand-new anchored		
brand-new unanchored	least acceptable topic	most acceptable focus

The theoretical framework outlined above allows the general statement that objects which incorporate the typical properties of a subject, such as [+animate/human, +definite, +specific, +given, +topic, −focus] are prone to some sort of differential marking. This is, simply put, the basic idea behind the theory of grammaticalization of differential object marking along the referentiality hierarchy, cf. Figure 1.

The theory implies that at the end of the grammaticalization process all objects, irrespective of their pragmatic-semantic quality, will undergo clitic doubling or, as Bossong (1991: 152) puts it, "such a differential system may ultimately become non-differential again".

Most marked for objects →

Human Pronoun

Human PN   Animate Pronoun

Human Definite   Animate PN   Inanimate Pronoun

Human Specific   Animate Definite   Inanimate PN

Human Non-Specific Animate Specific Inanimate Definite

Animate Non-Specific Inanimate Specific

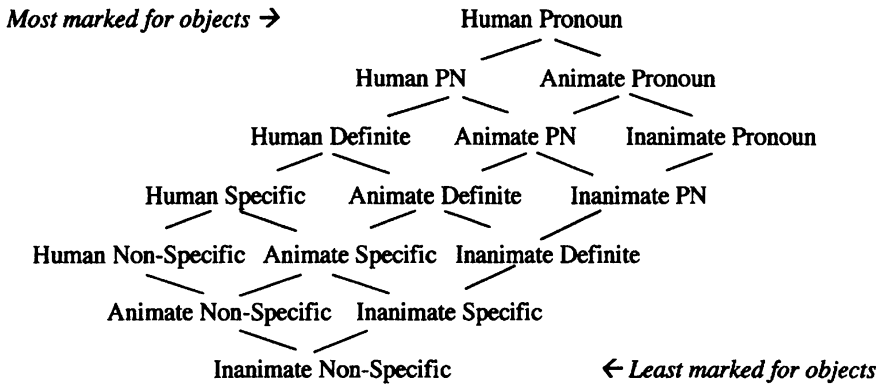Inanimate Non-Specific   ← Least marked for objects

Figure 1: Grammaticalization of differential object marking along the referentiality hierarchy (Aissen 2003: 459).

An intuitive conclusion would be to associate an increasing grammaticalization of differential object marking with a loss of restrictive usage rules, more freedom of use and consequently, in the broadest sense, a language change for the better. Note, however, that Figure 1 viewed in isolation, suggests well-defined intermediate stages, which does not do justice to the fact that grammaticalization is a continuous process with several transitional stages which, in turn, may allow a certain extent of optionality. Optional differential object marking is evidenced for many languages, such as Sinhalese, Romanian, and Yiddish (Aissen 2003). Such cases cannot be explained on the basis of well-defined factors or rules and can be considered, at least from a descriptive point view, as functionally less motivated and therefore opaque. The leading question will be: Does a higher degree of grammaticalization and, therefore, a less restricted usage of clitic doubling consequently lead to a higher degree of functional transparency? Or does an increased frequency of usage lead to the disappearance of a stable system and therefore result in a higher degree of optionality and functional opaqueness in certain contexts? This question will be pursued on the basis of a quantitative and qualitative corpus analysis of clitic doubling in Albanian dialects.

2 Clitic doubling in Albanian

2.1 Syntax of clitic doubling in Albanian

Albanian, as a member of the so-called *Balkan Sprachbund*, shares some structural similarities with surrounding languages which cannot be attributed to ge-

nealogical relationships, but rather to intensive language contact. It is important to note that these similarities go beyond lexical borrowings or phonological approximations and extend to the level of morphosyntax (Friedman 2006), one case being clitic doubling. In the case of Albanian and other Balkan languages, clitic doubling constitutes an additional marking of the direct or indirect object with depronominal clitics; cf. (1), where the clitic *e* is coreferential to *djalin* 'the boy'. Note: In the translations the accusative clitic in the 3rd person is always indicated by a cl and its absence with a ∅.

(1) Northwest Gheg (Montenegro; 42.43041, 19.25936)[1]

    Plaka   *e*   *ndali*   *djalin*   *me têtjë*
    old_woman.nom.def cl.acc.3sg stop.aor.3sg boy.acc.def inf stay.ptcp
    me   *tê*
    with her
    'The old woman cl stopped the boy so that he would stay with her.'

In contrast to other Balkan languages, the inventory of pronominal clitics in Albanian is rather simple since they show no gender distinction and dative and accusative clitics in the 1st and 2nd person coincide in their form, cf. Table 4.

Table 4: Paradigm of pronominal clitics in Standard Albanian. Some dialects may show slightly different paradigms.

	Dative	Accusative
1sg		*më*
1pl		*na*
2sg		*të*
2pl		*ju*
3sg	*i*	*e*
3pl	*u*	*i*

Clitics appear mostly as proclitics, i.e. preceding either the inflected verb, compare (1), or the main verb in an infinite verbal construction, compare (2).

---

[1]Most language examples have glosses. If an example exceeds a certain length, only the translation is given. The information given for each example are: Albanian dialect, country, geographical coordinates (longitude and latitude). If not otherwise stated, all examples are taken from a dialect corpus which was compiled for this study that is described in §3.1. §3.1 gives an overview of modern Albanian dialects.

(2) Northeast Gheg (Serbia; 42.30917, 21.64986)

Ridvani   don   me ja   falë
Ridvan.NOM.DEF want.PRS.3SG INF CL.DAT.3SG.CL.ACC.3SG give.PTCP

lojën   edhe e   hup
game.ACC.DEF and CL.ACC.3SG lose.prs.3sg

'Ridvan wants to let him win and loses the game.' (Lit. 'Ridvan wants to CL give him <u>the game</u> and loses it.')

Clitics can be attached to the verb stem in combination with imperative verb forms, cf. (3).

(3) Northwest Gheg (Montenegro; 42.43041, 19.25936)

Hê,   ktu tek jam,   po çil-ma-ni
INTERJ here at be.PRS.1SG so open-CL.DAT.1SG.CL.ACC.3SG-PRS.2PL

derën
door.ACC.DEF

'Hey, here I am, so open-CL <u>the door</u> for me.'

2.2 Use of clitic doubling

One major function of the clitics is the phoric resumption of discourse referents; cf. (4), where the *të shoqen dhe të dy ushtarët* 'his wife and both soldiers' are referenced in the subsequent sentence with a clitic. This usage of pronominal clitics will not be addressed in this paper.

(4) South Gheg (Albania; 41.90111, 20.0475)

Sa   vajtën,   mreti   e   njofti   të
once go.AOR.3PL king.NOM.DEF CL.ACC.3SG recognize.AOR.3SG ART

shoqen   dhe të dy ushtarët.   I   pyti   se
wife.ACC.DEF and ART two soldier.ACC.PL.DEF CL.ACC.3PL ask.AOR.3SG that

si   ish   puna
how be.IPF.3SG matter.NOM.DEF

'The moment they came, the king CL recognized <u>his wife and both soldiers</u>. He CL(=them) asked what their story was.'

Without exaggeration it can be said that almost any research on clitic doubling in Albanian focuses on 3rd person accusative clitic doubling. This interest in 3rd person accusative clitic doubling is due to the simple fact that all dative objects and 1st/2nd person accusative objects are unexceptionally doubled (Buchholz &

Fiedler 1987: 444). On the other hand, grammarians have struggled to determine the usage of 3rd person accusative clitic doubling (hereafter clitic doubling if not stated otherwise) with a great amount of certainty (Buchholz 1977: 3–4).

Clitic doubling can only occur if the associated nominal or pronominal phrase has the syntactic status of direct object. Nominal phrases in the accusative which do not qualify as such cannot be associated with a coreferential clitic. Thus, bare nominal indefinites are never doubled (Kallulli 2016). These cases can typically be found in recurrent light verb constructions, cf. (5), but less frequent constructions fall under this category also, cf. (6).

(5)   South Tosk (Albania; 40.2115, 19.68922)
      Finoku   as   hante   sa   duhej,   as
      Finok.NOM.DEF neither eat.IPF.3SG as_much_as need.IPF.PASS.3SG neither
      <u>bënte</u>   <u>muhabet</u>   si përpara
      do.IPF.3SG conversation as before
      'Neither did Finoku eat as much as he should, nor did he ∅ <u>have conversations</u> like before.'

(6)   North Tosk (Albania; 40.52797, 20.97003)
      Plaka   <u>gjen</u>   <u>ustallarë</u> dhe fillojnë
      old_woman.NOM.DEF find.PRS.3SG mason.PL and begin.PRS.3PL
      ndërtimin
      constructing.ACC.DEF
      'The old woman ∅ finds masons and they start to build.'

According to the notion of transitivity, as outlined in §1, the status of the noun phrases in (5) and (6) as object becomes questionable in favor of another analysis, namely considering them as components of a verbal incorporation.

Once the status as a genuine direct object is identified, the use of clitic doubling becomes contingent on information structure. Buchholz (1977: 192) and later Buchholz & Fiedler (1987: 440) cite the "primary accent" (*Primärakzent*), i.e. focus, as an important factor, stating that clitic doubling is applied if the object is topicalized[2] and not within the focus domain, which, in general, is in compliance with the descriptions we find in other standard grammars (Agalliu et al. 2002) and in recent works on clitic doubling in Albanian (Kallulli 2016). This systematicity is well depicted in (7): when the key to room forty is mentioned for the first time in an object position, it is not doubled. In the subsequent clause, when the same referent, the key to room forty, occurs again in the object position and is topicalized, it undergoes clitic doubling.

---

[2]They use the terms *theme* and *rheme*.

(7) North Tosk (Albania; 40.52797, 20.97003)

Ky djali, pas një kohe, si u rrit, i kërkon të jëmës çelsat e pallatit. Dhe ajo i thotë se: - Çfarë të duhen ty çelsat?- Unë i dua, të di se çfarë ka në pallat. E ëma i jep tridhjetë e nëndë çelsa dhe i hap dhomat, por dhoma e dyzetë nuk i hapet me ato çelsa. Djalit i mbeti merak që e ëma nuk i dha <u>çelsin e dhomës dyzetë</u>_1 dhe, mbas lutjesh që i bëri i biri, ajo ia jep <u>çelsin</u>_2.

'The boy, after some time, grew up and asks his mother for the keys to the palace. And she says to him: "What do you need them for?" "I want them to know what is in the palace". The mother gives him thirty-nine keys and he opens the rooms, but room forty does not open with those keys. The boy was worried that his mother did not ∅ give him <u>the key to room forty</u>_1 and, after begging her, she CL gives him <u>the key</u>_2.'

Although very rare, clitic doubling is also possible in combination with indefinite nominal objects if the object is topicalized, cf. (8).

(8) North Tosk (Albania; 40.60275, 19.61929)

<u>Një</u> <u>vajzë</u>	<u>e</u>	<u>njëj fshatari,</u>	*tridhjetë e*	*ca*	*vjeçe,*
a girl.NOM	ART	a peasant.GEN	thirty and	some	years_old

<u>e</u>	*kërkoi*	*mbreti*	*ta*	*merrte,*
CL.ACC.3SG	WANT.AOR.3SG	king.NOM.DEF	SBJV.CL.ACC.3SG	take.IPF.3SG

se	*ishte*	*e*	*bukur.*
because	be.IPF.3SG	ART	beautiful

'It was <u>a peasant girl</u>, thirty and a few years old, that the king CL wanted to have because she was beautiful.' (Lit. '<u>A girl of a peasant</u>, thirty and a few years old, the king CL wanted to have because she was beautiful.')

While the rules outlined above suffice to explain most cases of clitic doubling in Standard Albanian, there are also cases which deserve special attention. These are the quantifiers *të gjithë* and *të tërë* 'all' that are invariably clitic doubled (Kallulli 2016), cf. (9), and interrogative pronouns or wh-elements, which never occur in clitic doubling constructions (Kallulli 2000: 220), cf. (10) and (11).

(9) South Gheg (Albania; 41.42836, 19.66541)

por pyet	*nonën*	*ti, qi*	<u>i</u>	*ka*
but ask.IMP.2SG	grandmother.ACC.DEF	you who	CL.ACC.3PL	have.PRS.3SG

provu	<u>të gjitha</u>
experience.PTCP	everything

'but ask the grandmother who CL has experienced <u>everything</u>.'

(10) South Gheg (Albania; 40.96166, 20.21805)
<u>Çfarë</u> kërkon   ti   nga unë
what want.prs.2sg you from I
'<u>What</u> do you ∅ want from me?'

(11) North Tosk (Albania; 40.52797, 20.97003)
Vajza   Zili i   thoshin   kësaj   në shtëpinë
girl.NOM.DEF Zili CL.DAT.3SG say.IPF.3PL DEM.DAT.3SG in house.ACC.DEF
e   saj nuk pranonte   <u>asnjeri</u>
ART her NEG accept.IPF.3SG no_one
'The girl – she was called Zili – did not ∅ receive <u>anyone</u> in her house.'

Cases such as (9) and (10) cannot be viewed in complete isolation from the general dependency of clitic doubling on topic and focus. According to Kallulli (2016) doubling clitics trigger givenness of their associates and since 'all'-quantifiers are always non-novel they trigger clitic doubling. Likewise, the incompatibility of novel accusative interrogative pronouns and clitic doubling can be attributed to the simple fact that they are [−given, −topic] and mostly also [+focus] and therefore prohibit clitic doubling. This is corroborated by the example (12). While the pronoun *disa* 'some' may be classified under the category indefinite pronoun, it differs from *çfarë* 'what' in (10) and *asnjeri* 'no one' in (11) in as much as it has an antecedent and is therefore [+given, +topic, −focus] so that clitic doubling is required.

(12) South Tosk (Albania; 40.2115, 19.68922)
Ajo Lubi kush e   di   sa   njerëz ka
that Lubi who CL.ACC.3SG know.PRS.3SG how_many people have.PRS.3SG
mbytur e   i   ka   vdekur si hakë për ujë,
drown.PTCP and CL.ACC.3PL have.PRS.3SG kill.PTCP as payment for water
po <u>disa</u> nuk <u>i</u>   ka   ngrënë.
but some NEG CL.ACC.3PL have.PRS.3SG eat.PTCP
'That Lubi, who knows how many people it has drowned and it has killed as a payment for water, but <u>some</u> (of them) it CL has not eaten.'

In summary, the description given so far attests to an already existing highly redundant use of clitic doubling in all Albanian varieties, making its use restricted only in 3rd person accusative.

2.3 Clitic doubling in Albanian dialects from an areal perspective

Varieties of one language very often undergo grammaticalization processes at different speeds and to a different degree. For Albanian, as a Balkan language with multiple contact languages, variations within dialects require analysis of those variations in the wider context of areal linguistics. This approach is corroborated by Friedman (2008: 36), who states: "Of particular importance is the fact that the phenomenon [clitic doubling] shows varying degrees of encoding (as pragmatic or grammatical devices) on the basis of areal rather than genealogical relations." Assenova (2002: 110), cited in Friedman (2008: 40), gives the following overview of the conditions for clitic doubling, which applies, by and large, to all Balkan languages:

- the object is most often marked with a definite article

- the object is more often pre-verbal than post-verbal

- clitic doubling is especially common when the object is a personal pronoun

- indirect objects are more redoubled than direct objects

- objects that are not definite are not reduplicated

Concerning the frequency and degree of grammaticalization of clitic doubling in Balkan languages, i.e. their standard varieties, the ranking in Table 5 applies. The overview is by no means exhaustive, but it suffices to point out the rough areal tendency in the Balkans to further grammaticalize clitic doubling from the East to the West, whereby the highest degree of grammaticalization has been reached at the intersection of Central Gheg, Western Macedonian and Northern Aromanian (Curtis 2012, Lopašov 1978). In this region clitic doubling is almost exclusively contingent on definiteness and consequently occurs very frequently. Friedman (2006: 662) points out the fact that in "Macedonian […], unlike in the other Balkan languages, it can even occur (facultatively) with indefinite indeterminate pronouns such as *nikoj* 'nobody'". These cases, as sporadic as they may be, indicate an ongoing disengagement of clitic doubling from obvious triggers, like definiteness or givenness.

Many authors have stated the existence of such usage differences of clitic doubling between Albanian dialects which, furthermore, tend to match the general areal tendencies in the Balkans. For example, Curtis (2012: 310) states:

Table 5: Conditions for clitic doubling in Balkan languages (Kallulli & Tasmowski 2008: 9–10).

Macedonian	all definite direct objects and all indirect objects
Albanian	all IOs, DOs instantiated by first and second person pronouns, and all non-focal/non-rhematic DO DPs
Romanian	all full personal and definite pronouns, preverbal indirect objects and not [−specific] DPs, postverbal direct object DPs that are not [−specific] and are introduced by *pe*, and postverbal indirect object DPs which are not [−specific] and/or [−human] Goal
Greek	no obligatory context, except with *olos* 'all'
Bulgarian	all objects that are interpreted as Experiencers and objects of *ima, njama* 'there is (not)'

> [...] dialects of Albanian and Aromanian in contact with Greek do not show the same tendencies as those further to the north, namely that object reduplication is used for contrast and topicalization rather than as a strict grammatical obligation.

Regarding specifically the Albanian variety spoken in Kosova, Pani (2006: 70) remarks the following, using "double accusative" in the sense of clitic doubling and "Kosovar" as a term for the Albanian variety spoken in Kosova:

> Also in constructions with a double accusative object there are remarkable differences between Kosovar and Albanian. Speakers of Kosovar do not use proclitic pronouns in the same way as the Albanians do. On the other hand, constructions with double accusative object are used in Kosova in contexts where in Albania simple accusative objects are used.

Përnaska (2012) goes so far as to contend that Albanian, in general, shows the tendency to generalize clitic doubling, although this tendency is most evident in Northeast Gheg, especially the varieties spoken in Kosova.

The evidence which has been brought up so far allows for the assumption that the disparity is particularly striking between Albanian dialects spoken in today's Republic of Albania and those spoken in Kosova (Pani 2006, Përnaska 2012) and West Macedonia (Friedman 2008). Despite this evidence, down to the present, there have been no large-scale corpus-based analyses on the functional variation of clitic doubling between Albanian dialects.

3 Clitic doubling in Albanian dialects

3.1 Overview of Albanian dialects and the corpus

Figure 2 shows the territory in question and the Albanian dialects which are spoken today. The dialect map in Figure 2 is based on classifications as we find them in the current literature on Albanian dialectology, such as Gjinari & Shkurtaj (2000). On top of this classification, we will consider national borders, which leads to the division of some dialect areas into separate subareas, cf. Table 6, Version A.

Table 6: Number of transitive clauses for each subdialect (dialect samples of Albanian dialects spoken in Greece are missing). NEG: Northeast Gheg, NWG: Northwest Gheg, CG: Central Gheg, SG: South Gheg, NT: North Tosk, ST: Lab/South Tosk, CT: Cham Tosk

	Albania	Kosova	Serbia	Macedonia	Montenegro	Total
Version A						
NEG	261	304	80	90		735
NWG	405	91			366	862
CG	391			260		651
SG	538					538
NT	360			224		584
ST	176					176
CT	44					44
Total	2175	395	80	574	366	3590
Version B						
NEG	261	474				735
NWG	405	91			366	862
CG	391			260		651
SG	538					538
NT	360			224		584
ST	220					220
Total	2175	565		484	366	3590

A great many dialect descriptions[3] provide transcriptions of spoken language material in the form of monologues, narratives, riddles, etc., which lend themselves to quantitative corpus analyses. Similarly, the Albanian folkloristic literature[4] offers an extensive amount of narratives written in the dialect of the region they originate from. On the basis of samples from these two sources a small dialect corpus was compiled consisting of approximately 67.000 tokens and 3.590 transitive clauses and covering 69 places in the territory where Albanian is spoken, cf. Figure 3.

The available data is unequally distributed across the whole area in question, cf. Table 6, Version A. As can be clearly seen, very little data is available for some of the subareas which makes them not suitable for quantitative analysis. Since the corpus is relatively small overall for a quantitative analysis, the subareas Albania_CT and Albania_ST are grouped together into Albania_ST and also the subareas Macedonia_NEG, Se_NEG and Kosova_NEG are grouped together into Kosova_NEG, resulting in Table 6, Version B; this second classification is used for the subsequent quantitative analyses. This approach is supported by the fact that the subareas Albania_CT, Albania_ST and Macedonia_NEG, Se_NEG and Kosova_NEG show a comparable usage frequency of clitic doubling, despite the fact that it is obviously incorrect from a geopolitical point of view to make this grouping.

Not all available material could be included in the corpus for quantitative analyses. In the work at hand additional dialectal material was consulted for illustrative purposes if the corpus lacked language examples. Furthermore, the corpus contains exclusively dialect material. Standard Albanian data are missing, resulting in a lack of comparative data which would have been useful to assess to which degree the respective varieties deviate from the Standard variety on a quantitative scale. Thus, the data on South and North Tosk spoken in Albania (i.e. the subareas Albania_NT and Albania_ST in Table 6, Version B) will be employed instead. This approach is legitimate in so far as the Albanian standard variety is based for the most part on the South Albanian dialect so that no considerable differences vis-à-vis the Standard variety are to be expected.

---

[3]The dialectological sources which were consulted to compile the corpus: Ahmetaj (1989), Basha (1984), Beci (1974, 1982), Frano Luli (1975), Gecaj (2005), Gjinari & Shkurtaj (2000), Gosturani (1975, 1982), Keshi (2005), Lafe (1964), Mulaku (2005), Shefqet Hoxha (1975), Shkurtaj (1967, 1974, 1975, 1982), Topalli (1974).

[4]The folkloristic sources which were consulted to compile the corpus: Panajoti et al. (1988), Çetta (1982).

Figure 2: Dialects of modern Albanian (CC-BY SA ArnoldPlaton, https://commons.wikimedia.org/wiki/File:Albanian_dialects.svg)

Figure 3: Places covered by the dialect corpus.

3.2 Clitic doubling from the perspective of functional transparency

The fastest and most intuitive approach to assessing the degree of grammaticalization of clitic doubling is to measure its frequency in general. Such an approach requires that all objects be extracted, irrespective of the semantic and morphological properties, and are checked as to whether they undergo clitic doubling or not. Bare indefinites as well as the 'all'-quantifiers *gjithë* and *tërë* were not included in the analysis since no variation could be detected vis-à-vis Standard Albanian (cf. §2.2).

Figure 4 shows the ratio of doubled to non-doubled objects for each dialect area, irrespective of the countries they are spoken in, as described in Table 6, Version B.
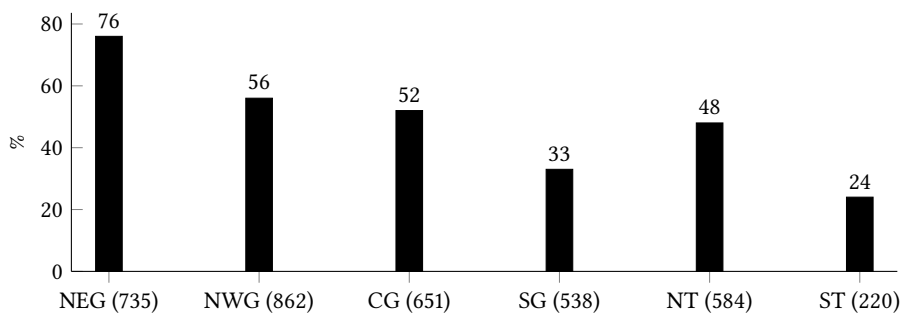

Figure 4: Frequency of clitic doubling in Albanian dialects. The bars display relative frequencies of clitic doubling in each dialect area. The number in brackets next to the abbreviation for the dialect area shows the absolute total number of cases of accusative objects for that specific area.

In general, the frequency pattern in Figure 4 corroborates the existence of an increasing frequency of clitic doubling towards the North of the Albanian speaking territory (Curtis 2012: 310). This picture would be immaculate if it was not for the high frequency of clitic doubling in North Tosk (52%), which comes as a surprise and does not correspond to the overall areal tendency. A tentative explanation can be found when the same frequency analysis is applied to the areas as defined in Table 6, Version B, i.e., considering also the countries the respective dialects are spoken in (cf. Figure 5).

After including the second classification parameter, i.e. national borders, the areal development of clitic doubling is shown from a completely different perspective: it suggests that the frequency of clitic doubling is to a lesser extent dependent on the dialectal affiliation and/or the geographical area (North vs.
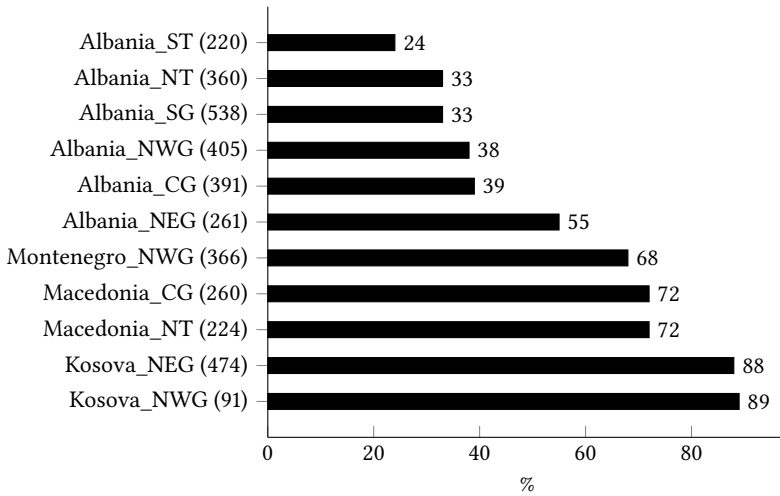
Figure 5: Frequency of clitic doubling in Albanian dialects considering the countries they are spoken in. The bars display relative frequencies of clitic doubling in each dialect area. The number in brackets next to the abbreviation for the dialect area shows the absolute total number of cases of accusative objects for that specific area.

South). Of central importance appears to be the division into center and periphery which, more or less, corresponds to the national separation into (a) dialects spoken within the Republic of Albania and (b) dialects spoken in countries of former Yugoslavia (Kosova, Macedonia, Montenegro, Serbia). While North Albanian dialects within Albania, especially Northeast Gheg, also show a tendency to make more use of clitic doubling, an actual abundance of clitic doubled objects is noticeable only for the dialects spoken outside of Albania. This becomes most obvious when one compares the areas in Albania and Macedonia where North Tosk dialects are spoken (Albania: 33% vs. Macedonia: 72%). However, it must be emphasized that the rather small amount of data and sparse coverage of the speaking territory requires further studies on the basis of larger amounts of data and focusing especially on the areas along today's national borders in order to assess the influence of these borders.

Such apparent differences in frequency usually go along with functional differences and require more elaborate analyses, ideally capturing aspects of information structure. However, annotating corpora on the level of information structures has proven to be a cumbersome and error-prone task. Instead, I will draw indirect conclusions on the basis of robust morphological and lexical criteria. According to the introductory remarks in §1, we arrive at the trivial, yet

helpful for the purpose at hand, conclusion that there is a strong tendency for arguments which are given to be also morphologically definite, topicalized and not focused, while newly introduced arguments are usually morphologically indefinite, not topicalized and within the focus domain. What is more, we can state that subjects are better candidates to be topic than objects, which subsequently leads to the conclusion that on a quantitative scale objects commonly tend not to be doubled, which has proven to be correct for most dialects spoken in Albania and especially for Tosk, cf. Figure 4. Lastly, I contend that pronouns outrank nouns within the referentiality hierarchy, i.e. indefinite pronouns are rated lower than indefinite nouns (for example: *whom* vs. *a man*) and definite pronouns are rated higher than definite nouns (*him* vs. *the man*).

This correlation between definiteness and topic acceptability allows for the sketching of clitic doubling frequency patterns along a morpho-lexical definiteness hierarchy, cf. Table 7. The classification made here is based on purely morphological and lexical criteria and does not claim to be complete. Thus, for example, the category "indefinite pronouns" also includes interrogative pronouns. Nevertheless, an additional division of the four classes would lead to an unnecessary complexification and not serve the purpose of the hierarchy, which is to describe general areal tendencies regarding the usage of clitic doubling on a quantitative scale and to compare these findings with the descriptions we find in the current literature on Albanian and other Balkan languages.

Table 7: Morpho-lexical definiteness hierarchy.

1. definite pronouns (prop_def)	
2. definite nouns (np_def)	↑ Increasing frequency of clitic doubling
3. indefinite nouns (np_indef)	
4. indefinite pronouns (prop_indef)	

The underlying assumption is that clitic doubling would increase along this hierarchy, which is confirmed by the results of the corpus analysis for the regions Albania_NT and Albania_ST (≈ Standard Albanian), cf. Figure 6.

As the figure shows, clitic doubling mostly occurs in association with definite objects. Despite this strong tendency, definiteness cannot be judged as the decisive factor for clitic doubling, otherwise a higher percentage of doubled definite objects would be expected. Table 8 displays the usage pattern of clitic doubling for each of the subareas as described in Table 6, Version B, along the same morpho-syntactic definiteness hierarchy.

Table 8: Frequency of clitic doubling along a morpho-lexical definiteness hierarchy in Albanian dialects, where n_{cd} is the number of cases with clitic doubling and n_{ao} is the total number of accusative objects.

| Dialect regions	prop_indef		np_indef		np_def		prop_def	
	n_{cd}	n_{ao}	n_{cd}	n_{ao}	n_{cd}	n_{ao}	n_{cd}	n_{ao}
Albania_CG	1 (6%)	16	12 (13%)	93	137 (51%)	269	4 (31%)	13
Albania_NEG	1 (8%)	13	23 (55%)	42	112 (57%)	197	8 (89%)	9
Albania_NT	0 (0%)	24	5 (5%)	93	100 (44%)	226	14 (82%)	17
Albania_NWG	1 (3%)	32	14 (16%)	87	131 (48%)	275	7 (64%)	11
Albania_SG	0 (0%)	34	9 (6%)	152	150 (47%)	322	20 (67%)	30
Albania_ST	1 (5%)	31	1 (2%)	58	48 (38%)	128	3 (23%)	13
Kosova_NEG	2 (5%)	38	85 (90%)	94	321 (97%)	332	10 (100%)	10
Kosova_NWG	0 (0%)	9	7 (100%)	7	69 (99%)	70	5 (100%)	5
Macedonia_CG	0 (0%)	22	21 (35%)	60	161 (93%)	174	4 (100%)	4
Macedonia_NT	1 (5%)	20	3 (9%)	33	142 (92%)	154	15 (88%)	17
Montenegro_NWG	1 (4%)	28	67 (69%)	97	173 (75%)	230	9 (82%)	11
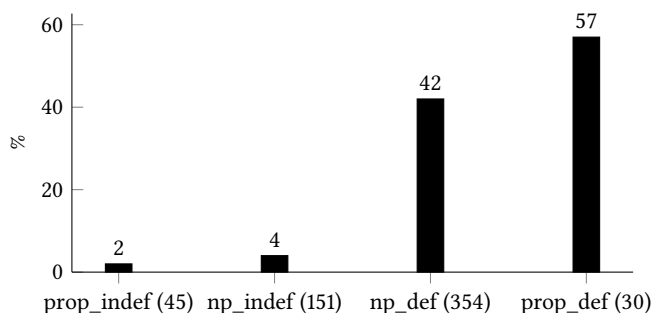
Figure 6: Frequency of clitic doubling in the regions Albania_NT and Albania_ST along the morpho-lexical definiteness hierarchy. The bars display relative frequencies of clitic doubling. The number in brackets shows the absolute total number of cases of accusative objects.

Of particular interest remain the areas outside the Republic of Albania. The patterns allow for positing grammaticalization along the morpho-syntactic definiteness scale: first, clitic doubling will become almost obligatory with definite pronominal objects, a process that is already in progress for some varieties in North Albania. Then definite nominal objects will follow as the main trigger of clitic doubling while still excluding most indefinite nominals, a stage that has been reached especially in Southwest Macedonia. Subsequently, instances of doubled indefinite nominal objects will become more common, as is the case in Northwest Macedonia. Eventually, the majority of indefinite nominals will undergo clitic doubling in almost any context, a stage that has been reached in Montenegro and, even more so, in Kosova. Novel indefinite pronominal objects, however, remain very resistant to clitic doubling.

It is important to stress that Table 8 displays a continuous grammaticalization path and that exceptions from the general tendencies do occur in all of the areas in question. In the following, the Albanian speaking territories of Montenegro, Macedonia and Kosova will be subject to a cursory examination from a functional perspective. The leading question will be whether the higher degree of grammaticalization and frequency of clitic doubling in these areas can be viewed as a higher degree of functional transparency.

The entire region of Montenegro appears to be most opaque regarding the functional motivation of clitic doubling. On the one hand, the dense occurrence of clitic doubling shows a clear development towards a loss of any restrictions whatsoever, a pattern that comes close to the situation we find in Kosova. This is, however, inconsistent with the relatively high number of exceptions: while most indefinite and definite nominals are doubled, a relatively large number of them are not. See, for instance, example (13), which presents the beginning of a story:

The first object (*i djalë* 'a son') is not doubled, which is less surprising as it has the feature matrix [−given, −definite, −topic, +focus] and corresponds to the pattern of Standard Albanian. Note, however, that in the subsequent part of the introduction all other instances of an object undergo clitic doubling independent of whether they are [+given/−given], [+def/−def], [+topic/−topic], [+focus/−focus].

(13)   Northwest Gheg (Montenegro; 42.44626, 19.4564)
        Â' kenë i nanë e ka pa' <u>i djalë</u>. Ajo â' kenë fukara e <u>e</u> ka çuo <u>djalin</u> rrogtar
        për me mujë me jetuo. Djali â' kenë qiros edhe ka tejë tu 'i zotni rrogtar. 'I
        ditë prej ditsh djali â' ba me do fmi e kanë luoj bashkë. Njari prej asi fmísh e
        ka pa' <u>i kapicë</u> ën krye. Qirosi i ka lakmuo kapicës edhe prej atyt shkon fill
        tu nana e vet e i thotë: - Me <u>ma</u> ble '<u>i kapicë</u>! Nana i përgjégj: - <u>Ta</u> ka pa'
        lanë baba '<u>i kapicë</u>, por po ta gjêj.

        'Once upon a time there was a mother and she ∅ had <u>a son</u>. She was poor
        and CL sent <u>her son</u> as a day laborer so that they would survive. The boy
        was bald and stayed with a lord as a day laborer. One day the boy
        gathered with some children and they played together. One of the
        children CL had <u>a hat</u> on his head. The bald boy also wanted such a hat
        and went from there to his mother and said: "Buy CL me <u>a hat</u>!". The
        mother answers him: "Your father CL left you <u>a hat</u>, I will find it for you".'

Especially in need of explanation are undoubled definite nouns. Later on, in the same story, a magic flute is introduced and doubled when it is mentioned the first time, cf. (14). The flute plays an important role in the subsequent part of the story (as it empowers the boy to summon a large army whenever he likes) and, therefore, is mentioned several times in the storyline. Nevertheless, in example (15) when a reference is made to the "story behind the flute" (*punën e fellit*) the respective object is not doubled.

(14)   Northwest Gheg (Montenegro; 42.44626, 19.4564)
        I ditë prej ditsh bâhet me çobana e <u>ja</u> sheh njaj çobanit <u>fellin</u> tye i ra. Prap
        shkon e i thotë nanës: - Nanë, me <u>ma</u> ble '<u>i fyell</u>! Nana i përgjégj: - Qiroso,
        Qiroso, ma prune punën gusht! <u>Ta</u> ka pa lanë baba '<u>i dreq fellit</u>, por ruoju,
        se ke me ja pa sherrin! Çohet nana, ja gjê e <u>ja</u> nep <u>fellin</u>.

        'One day he gathers with some shepherds and CL sees <u>the flute</u> of a
        shepherd and him playing with it. Again he goes to his mother and says
        to her: "Mother, buy CL me <u>a flute</u>". "My boy, you really put me in a
        difficult position! Your father CL left you <u>a damn flute</u>, but be careful
        because it will cause you many more problems!" She gets up, finds it for
        him and CL gives him <u>the flute</u>.'

(15)   Northwest Gheg (Montenegro; 42.44626, 19.4564)

- Si â' kjo punë e ç' fuqi ke ti qi e ke gjidh kët ushtri? Qirosi, si budallë, i kalzon e i thotë: - Tash me da' û', tanë kta periherë i tres mos m' u duktë ma. Bija e mretit e bvét: - Si muç ti m' e ba ata? Ky i kuvet krejtsisht <u>punën e fellit</u>.

'"What's going on and what power do you have that you have this whole army?" The boy, stupid as he was, tells her: "Now, if I want, I could make the whole army disappear, as if it had never been there". The king's daughter asks him: "How can you do that?" He ∅ tells her in great detail <u>the story behind the flute</u>.'

The status on the level of information structure of the undoubled objects in examples (13) and (15) allow us to draw on factors such as focus or givenness in order to explain the absence of the clitic. Such a solution is not satisfactory in as much as it fails to provide a holistic explanation of the usage of clitics in this variety. The scattered instances of clitic omission rather suggest that in the area in question clitic doubling is the default case in association with indefinite nominal and definite nominal and pronominal objects and may be left out facultatively for contrastive or focussing purposes.

Moving on to the region of West Macedonia, the overall pattern appears to be more transparent. Table 8 shows that irrespective of the dialectological classification, i.e. Gheg vs. Tosk variety, in West Macedonia the tendency to almost exclusively double definite objects prevails. This tendency is somewhat stronger in the Southwest area than in the Northwest area where clitic doubling of indefinite objects is more common. In addition, cases with an undoubled definite object are recorded as well. This leads to the intricate question of how to account for these cases which clearly deviate from the overall pattern. Additional factors, alongside definiteness, such as specificity, focus or contrast, have been brought up in the current literature on clitic doubling in Macedonian, the main contact language (cf. for example Friedman 2008) which shows similar tendencies to double exclusively definite objects. While some cases, such as (16), clearly suggest contrast, alongside focus, as an important co-factor which prohibits clitic doubling in combination with definite objects, others are harder to account for, such as (17): the object *një plakë* 'an old woman' is doubled despite being [−given, −definite, −topic, +focus]. A possible explanation could be found if the referent is viewed as being [+specific]. Drawing on specificity is also corroborated by example (18) which is comparable to (17): the first mentioning of *ni qakë* 'an old woman' [−given, −definite, −topic, +focus] is not doubled and, therefore, adheres to the general areal tendency. However, the second instance of the same phrase

ni qakë 'an old woman' undergoes clitic doubling despite having the same features [−given, −definite, −topic, +focus]. One possible explanation would be to view the first instance as [−specific], as it does not make any reference to a specific person in the story, and the second as [+specific] since it refers to a specific person in the story.

(16) Central Gheg (Macedonia; 42.03503, 20.9161)
Gje   kërkojshe   <u>shatkën</u>,   sot   kërkojshe   <u>shatokin</u>
yesterday want.ipf.2sg duck.acc.def today want.ipf.2sg drake.acc.def
'Yesterday you ∅ asked for <u>the duck</u>, today you ∅ asked for <u>the drake</u>.'

(17) North Tosk (Macedonia; 41.09028, 21.01325)
oxha ja vuri synë kësaj // po kjo mazalla s' e deshte // si t' ja bëj ?! <u>e</u> thiri
<u>një plakë</u> t' ia ndreqë punën // plaka tha : unë do ta regulloj që të marë //
plaka mori një perusti dhe e vuri tersëne // çupa i thotë : moj plakë / nuku
viet ashtu perustija / po ndryshe viet // po nuku di // zbrit e më trego //
'the Imam had an eye on her [the girl] // but she didn't want anything from him // what should he do now ?! he cl called <u>an old woman</u> to settle the matter for him // the old woman said: I'll make sure she takes you // the old woman took a tripod and puts it upside down // the girl says : you there / you don't put a tripod up like this // I don't know // come down and show me //'

(18) Central Gheg (Macedonia; 41.23853, 20.64221)
Hajduti i porosojti : Çitni <u>ni qakë</u> t' lipe vjom devje n' at mahallë ku u vdir
devja , se ai deven e ka therrë ! E qaka, kur t' dale ne ta boe derën pagës me
vjom! Mbasandoj ju shkoni bastojsni ! <u>E</u> çesin <u>ni qakë</u> . Qaka i merr shpojt
me rent.
'The thief told them, "∅ Send <u>an old woman</u> to beg for camel fat in the area where the camel disappeared, because he [the other thief] slaughtered the camel already! And the woman, when she goes out of the house again, let her mark the front door with camel fat! Then you can search the house!" They cl send <u>an old woman</u> out. The old woman takes the houses one by one.'

While the argumentation above seems coherent, one cannot get around the fact that such explanations always bear a certain amount of vagueness and posit a certain degree of arbitrariness and optionality. In any case, drawing on specificity cannot explain the lack of clitic doubling in (19), as there is no reason why the referent *një çupë* 'a daughter', the main figure in the story, should not be regarded as being [+specific].

(19)   North Tosk (Macedonia; 40.8876, 21.31207)
        qënka   një plak   edhe një plakë   / kanë   edhe një
        be.PRS.ADM.3SG a old_woman and a old_man have.PRS.3PL also a
        çupë   //
        daughter.ACC
        'once upon a time there was an old woman and an old man / they ∅ had
        <u>also a daughter</u>'

It should be emphasized that the entire Albanian-speaking territory of Macedonia poses another difficulty, its dialectological heterogeneity. A more fragmented division of the area into a northern, central and southern part would do more justice to the increasing frequency of clitic doubling towards the North where the varieties are actually (transitional to) Northeast Gheg and feature a similarly high frequency of clitic doubling as in the Albanian varieties of Kosova. It is also Northwest Macedonia where we find a case with a doubled indefinite pronoun *kërkân* 'no one', cf. (20). Note that Friedman (2006) proved the existence of such cases in Slavic dialects of Northern Macedonia, as well.

(20)   Northeast Gheg (Macedonia; 42.16205, 21.61688)
        ama kërkân   s' un p' e   zanë   tu vedhë
        but no_one.ACC NEG can PROG CL.ACC.3SG catch.PRS.3PL GER steal.PTCP
        'but he couldn't CL catch anyone stealing'

The varieties in Kosova, both Northeast and Northwest Gheg, show a higher degree of homogeneity regarding the usage pattern of clitic doubling which is subject to barely any restrictions whatsoever. While the overall picture is reminiscent of the findings on Albanian dialects in the Montenegrin area, the usage of clitic doubling is more pervasive and fewer exceptions occur. Most objects without clitic doubling fall under the category of indefinite and interrogative pronouns, and only sporadic instances of undoubled nominals can be found. Example (21) shows the interrogativ pronoun *kâ* 'whom' that is not doubled.

(21)   Northwest Gheg (Kosova; 42.73953, 20.05389)
        Kâ   po   pret   ktu?
        whom PROG wait.PRS.2SG here
        '<u>Whom</u> are you ∅ awaiting here?'

Note, however, that in view of example (20) from North Macedonia, one should actually expect clitic doubling in example (21). In the analysis, no distinction was made between interrogative pronouns and indefinite pronouns as they mostly do

not differ in their form in Albanian and share similar features regarding their degree of definiteness and givenness. In this context, it must be stated that example (20) is marked as it contains a genuine indefinite pronoun in a left-dislocation, in contrast to example (21) which contains an interrogative pronoun, whose initial position is predefined and not marked in any way. Left-dislocated objects are notoriously associated with topicalization and thus predisposed for clitic doubling (Friedman 2008: 40). In order to draw direct comparisons other examples of left-dislocated indefinite pronouns are required. Unfortunately, these cases are very rare in the corpus used so far for quantitative comparisons. Thus, additional material from the folkloristic literature from the region of Drenica in central Kosova (hereafter called Drenica corpus) was searched manually. The search unveiled other cases of left-dislocated indefinite pronouns; clitic doubling occurred in two of them, cf. (22) and (23). Nevertheless, not all instances of left-dislocation indefinite pronouns are doubled, cf. (24).

(22) Northeast Gheg (Kosova; 42.52, 20.85527 (Drenica corpus))
      Pos   teje <u>qeterkan</u>   për nafakë nuk <u>e</u>   due.
      except_for you no_one.ACC for fate   NEG CL.ACC.3SG want.PRS.1SG
      '<u>No one</u> else I CL want as my husband except for you.'

(23) Northeast Gheg (Kosova; 42.52, 20.85527 (Drenica corpus))
      Une <u>qetërkân</u>   për nafakë nuk <u>e</u>   due   vetëm ktâ
      I   no_one.ACC for fate   NEG CL.ACC.3SG want.PRS.1SG only   him
      '<u>No one</u> else I CL want as my husband, only him.'

(24) Northwest Gheg (Kosova; 42.73953, 20.05389)
      <u>Tjetër gjâ</u>   s'   folën   motrat.
      other thing.ACC NEG speak.AOR.3PL sister.NOM.PL.DEF
      '<u>something else</u> the sisters didn't ∅ speak'

The fact that all three cases (20), (22) and (23) with clitic doubled indefinite pronouns are [+human] deserves attention and raises the question whether the additional semantic factor [+human/−human] must be considered. Against this background the following excerpts from newspaper articles extracted from the Albanian National Corpus[5] may be classified as evidence for an ongoing or already completed grammaticalization process towards an obligatory clitic doubling of indefinite (including interrogative) pronouns which have the semantic feature [+human]:

---

[5]http://web-corpora.net/AlbanianCorpus/search/

(25) Kosova sot[6]

> *Ja   kë   e   falënderon   e   kë   e   shan*
> here whom CL.ACC.3SG thank.PRS.3SG and whom CL.ACC.3SG scold.PRS.3SG
> *M. Z.*
> M. Z.
> 'One can tell <u>whom</u> M. Z. CL thanks and <u>whom</u> he CL scolds.'

The sentences above appeared in newspapers from Kosova. Newspapers from Kosova contain many cases of clitic doubling with the interrogative pronouns *kë* 'whom'. This is remarkable in the sense that the same standard variety is used as the official language in Kosova, Albania and in the Albanian communities of Macedonia. The fact that the authors of the articles were not aware of this redundant use of clitic doubling may serve as an indicator for how far the grammaticalization process has gone in these varieties. Nevertheless, further studies with the aid of larger corpora and/or experiments are required to furnish proof.

The last desideratum that remains is to explain the cases with undoubled indefinite and definite nominals in the Albanian varieties of Kosova, which do not match the overall pattern at all. Their scattered occurrence allows for several different tentative explanations, all of them implying that clitic doubling serves an expressive function, leaving its omission in some cases at the discretion of the speaker. A similar conclusion by drawing on optionality was also made for the region of Montenegro to explain the rather arbitrary cases of clitic omission. One such explanation can be found in Buchholz (1977: 192–193), who contends that the omission of clitic doubling in combination with non-novel definite objects emphasizes a change in the situation. Indeed, some examples with undoubled definite objects describe a turning point, such as (26), where the word order V O S is also marked. Example (27) is "marked" as well, in as much as the respective clause without clitic doubling is embedded in a chain of successive short clauses and therefore conveys a certain degree of dynamicity. Other cases, such as (28), speak in favor of viewing the associated object not as definite, despite its definite affix, but as a component of an incorporation into the verb, since the meaning of the noun-verb construction is not compositional. Example (29), on the other hand, remains opaque since the object is not only [+given,+topic,+definite,−focus] it is also left-dislocated and therefore expected to undergo clitic doubling.

---

[6]https://www.kosova-sot.info/lajme/19663/ndihmohen-femijet-e-drenices-ja-ke-e-falenderon-e-ke-e-shan-milaim-zeka/

(26) Northeast Gheg (Drenica corpus;)
Kap *shtagën* *baba* *e* *shko* *tu begu*
grab.PRS.3SG stick.ACC.DEF father.NOM.DEF and go.PRS.3SG to bey.NOM.DEF
'The father ∅ takes the <u>stick</u> and goes to the bey.'

(27) Northeast Gheg (Kosova; 42.38441, 20.4285)
Çohet */ e* *kish* *pas* *marr*
get_up.PRS.PASS.3SG CL.ACC.3SG have.PRS.3SG have.PTCP take.PTCP
etja *// niset* *për shpi* *// lyp*
thurst.NOM.DEF set.off.PRS.PASS.3SG for home look_for.PRS.3SG
<u>*kofen*</u> *e* <u>*bunarit*</u> *me pi* *uj* *//*
bucket.ACC.DEF ART well.GEN.DEF INF drink.PTCP water
kofa *kërkun*
bucket.NOM.DEF nowhere
'he stands up / he was thirsty // he heads for home // he ∅ seeks the <u>well bucket</u> to drink water // the bucket was nowhere'

(28) Northeast Gheg (Kosova; 42.38441, 20.4285)
Muer <u>*dyjênë*</u> *n' sy* *Baba Hasani për* *m' u*
take.AOR.3SG world.ACC.DEF in eye.ACC Baba Hasan SUPIN INF PASS
largue *pej shpije.*
leave.PTCP from house.ABL
'Baba Hasan set off to leave the house.' (Lit. 'Baba Hasan ∅ took <u>the world</u> in his eye to leave the house')

(29) Northwest Gheg (Kosova; 42.73953, 20.05389)
Mreti *u* *kallxoi* *: − Qeso* <u>*héshë*</u> *veç*
king.NOM.DEF CL.DAT.3PL tell.AOR.3SG this_kind_of food.ABL only
grueja *m'* *goditke!*
wife.NOM.DEF CL.DAT.1SG make.IPF.3SG
'The king told them: "<u>This kind of food</u> only my wife ∅ made for me".'

4 Concluding remarks

Although peripheral Albanian dialects outside the Republic of Albania feature a higher frequency of clitic doubling than most Albanian dialects spoken in the Republic of Albania, the relatively high number of opaque usage cases does not allow us to view the redundant use of clitic doubling as evidence of a higher degree of functional transparency and, therefore, as a language change for the

better. In case of Standard Albanian, most instances of clitic doubling in main clauses allowed for an explanation by defining whether the respective object was [+focus, –topic] or not. While notions, such as *topic* or *focus*, prove to be intricate at first glance, once detected they can be applied to explain almost any instance of clitic doubling in Standard Albanian. In the case of the dialects outside Albania, the data allowed for the sketching of an interim and coarse usage pattern along a morpho-lexical definiteness hierarchy which seems neat *prima facie* and suggests more easily detectable factors, to wit *definiteness*.

However, it is all the more difficult to explain cases which do not adhere to this usage pattern. In particular need of explanation are cases with undoubled definite objects. In contrast to Standard Albanian, one must avail oneself of additional factors, such as *contrast*, cf. (16), *specificity*, cf. (17), (18), or even *animacy*, cf. (22), (23), to explain the exceptional cases. What is even more important, the explanations offered in the work at hand are tentative and far from sufficient and do not offer a holistic description. The rather scattered and unsystematic cases of clitic omission in combination with indefinite nominals and definite nominals, especially in Albanian varieties spoken in Montenegro and Kosova, and the difficulties of pinning down clear factors suggest a certain degree of optionality, which is not given in Standard Albanian and presumably in the majority of varieties spoken in the Republic of Albania. Thus, from a descriptive point of view, with the claim of finding regularities with a great amount of certainty, optionality can be regarded as an indicator of a higher degree of functional opaqueness.

Abbreviations

1	first person		GER	gerund
2	second person		IMP	imperative
3	third person		INF	infinitive
ABL	ablative		IPF	imperfect
ACC	accusative		NEG	negation, negative
AOR	aorist		NEG	Northeast Gheg
ART	linking article		NOM	nominative
CG	Central Gheg		NT	North Tosk
CL	clitic		NWG	Northwest Gheg
CT	Cham Tosk		PASS	passive
DAT	dative		PL	plural
DEF	definite		PROG	progressive
GEN	genitive		PRS	present

PTCP	participle	ST	Lab/South Tosk
SG	singular	SUPIN	supin.
SG	South Gheg		

References

Agalliu, Fatmir, Shaban Demiraj & Mahir Domi. 2002. *Gramatika e gjuhës shqipe*. Mahir Domi (ed.). Vol. 1. Tiranë: Akademia e Shkencave e Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë.

Ahmetaj, Mehmet. 1989. E folmja shqipe e Plavës dhe e Gucisë. *Studime gjuhësore* 2. 223–342.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483.

Assenova, Petja. 2002. *Balkansko ezikoznanie: Osnovni problemi na balkanskija ezikov săjuz*. Sofia: Faber.

Basha, Naim. 1984. *E folmja e qytetit të Dibrës: Punim magjistrature*.

Beci, Bahri. 1974. E folmja e Luznisë së Dibrës. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 2*, 223–281. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Beci, Bahri. 1982. E folmja e Mirditës. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 4*, 26–143. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Bossong, Georg. 1991. Differential object marking in Romance and beyond. *New analyses in Romance linguistics: Selected papers from the XVIII Linguistic Symposium on Romance Languages, Urbana-Champaign, April.* 143–170.

Buchholz, Oda. 1977. *Zur Verdoppelung der Objekte im Albanischen*. Berlin: Akademie der Wissenschaften der DDR.

Buchholz, Oda & Wilfried Fiedler. 1987. *Albanische Grammatik*. Leipzig: Verlag Enzyklopädie.

Çetta, Anton. 1982. *Përralla*. Vol. 2. Prishtinë: Instituti Albanologjik i Prishtinës.

Curtis, Matthew C. 2012. *Slavic-Albanian language contact, convergence, and co-existence*. The Ohio State University. (Doctoral dissertation).

Frano Luli. 1975. E folmja e Postrripës. *Studime filologjike* 1. 79–99.

Friedman, Victor A. 2006. The Balkans as a linguistic area. In Keith Brown (ed.), *Elsevier encyclopedia of language and linguistics*, 657–672. Oxford: Elsevier.

Friedman, Victor A. 2008. Balkan object reduplication in areal and dialectological perspective. In Dalina Kallulli & Liliane Tasmowski (eds.), *Clitic doubling in the Balkan languages*, 35–63. Amsterdam: Benjamins.

Gecaj, Murat. 2005. Vështrim rreth morfologjisë të së folmes së Krasniqes. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 7*, 221–257. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 151–188. New York (etc.): Academic Press.

Gjinari, Jorgji & Gjovalin Shkurtaj. 2000. *Dialektologjia*. Reprint. Tiranë: Shtëpia Botuese e Librit Universitar.

Gosturani, Xheladin. 1975. E folmja e Hasit. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 3*, 211–269. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Gosturani, Xheladin. 1982. E folmja e çeremit dhe e Valbonës (rrethi i Tropojës). In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 4*, 279–344. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.

Kallulli, Dalina. 2000. Direct object clitic doubling in Albanian and Greek. In Maria-Luisa Rivero & Angela Ralli (eds.), *Comparative syntax of the Balkan languages*, 209–248. Oxford: Oxford University Press.

Kallulli, Dalina. 2016. Clitic doubling as differential object marking. *Rivista di Grammatica Generativa: Research in Generative Grammar (RGG)* 38. 161–171.

Kallulli, Dalina & Liliane Tasmowski. 2008. Introduction: Clitic doubling, core syntax and the interfaces. In Dalina Kallulli & Liliane Tasmowski (eds.), *Clitic doubling in the Balkan languages* (Linguistik Aktuell/Linguistics Today 130), 1–32. Amsterdam: John Benjamins.

Keshi, Skënder. 2005. Disa dukuri për të folmen e Mbiudhës së Maqellarës në Dibër. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 7*, 158–220. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Lafe, Emil. 1964. E folmja e Nikaj-Mërturit. *Studime filologjike* 3. 107–148.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Leafgren, John. 2002. *Degrees of explicitness: Information structure and the packaging of Bulgarian subjects and objects*. Amsterdam: John Benjamins.

Lenz, Rodolfo. 1920. *La oración y sus partes.* Madrid: Centro de Estudios Históricos.

Lopašov, Jurij Aleksandrovič. 1978. *Mestoimennye povtory dopolnenija v balkanskich jazykach.* Leningrad: Nauka.

Mulaku, Latif. 2005. *E folmja e Shalës së Bajgorës.* Prishtinë: Instituti Albanologjik i Prishtinës.

Næss, Åshild. 2007. *Prototypical transitivity* (Typological studies in language). Amsterdam: Benjamins.

Panajoti, Jorgo, Agron Xhagolli & Albert Doja (eds.). 1988. *Folklor shqiptar Ser. 1. Proza popullore; 6. vël. 6.* Tiranë: Akademia e Shkencave e RPSSH, Instituti i Kulturës Popullore.

Pani, Pandeli. 2006. Some differences between varieties of Albanian with special reference to Kosovo. *International Journal of the Sociology of Language* 178. 55–73.

Përnaska, Remzi. 2012. Bashkësia e të ashtuquajturave trajta të shkurtëra të përemrave vetorë dhe zgjedhimi nënanor-kundranor (subjektor-objektor) sui generis në gjuhën e sotme shqipe. In *Shqipja dhe gjuhët e Ballkanit: Konferencë shkencore e mbajtur më 10 – 11 nëntor 2011 në Prishtinë.* 397–410. Prishtinë.

Shefqet Hoxha. 1975. Vëzhgime mbi të folmen e vilë-e-kalisit të Lumës. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 3*, 130–210. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Shkurtaj, Gjovalin. 1967. E folmja e Kastratit. *Studime filologjike* 3. 47–66.

Shkurtaj, Gjovalin. 1974. E folmja e Hotit. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 2*, 344–428. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Shkurtaj, Gjovalin. 1975. E folmja e Kelmendit. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 3*, 5–129. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Shkurtaj, Gjovalin. 1982. E folmja e Ranxave të Mbishkodrës. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 4*, 144–278. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.

Topalli, Xh. 1974. Disa vërejtje për të folmen e krahinës së pukës. In Mahir Domi, Jorgji Gjinari & Gjovalin Shkurtaj (eds.), *Dialektologjia shqiptare 2*, 282–343. Tiranë: Akademia e Shkencave e RPS të Shqipërisë, Instituti i Gjuhësisë dhe i Letërsisë, Sektori i Gramatikës dhe i Dialektologjisë.