

THE ROLE OF CLITIC DOUBLING IN THE INFORMATION STRUCTURE OF THE ALBANIAN LANGUAGE

Qëndresa JAKUPI

University of Prishtina “Hasan Prishtina”, Kosovo

E-mail: gendresa.jakupi@uni-pr.edu

ABSTRACT: This paper attempts to argue whether the repetition of the object or clitic doubling as a syntactic mechanism affects the encoding of the information structure in the Albanian language. In three previous studies by Kallulli (2000, 2008); Kapia (2010), it is reported that direct object clitic doubling is not an optional phenomenon in the Albanian language, that direct object clitic doubling systematically produces and encodes the information structure, and that clitic doubled objects are interpreted as topics. Thus, the object and the clitic have the status of topic in the sentence, as doubling occurs when the clitic refers to the already present discourse referent. That is, the NP is previously introduced in the discourse, thus becoming a specific referent available to be later selected via the doubling mechanism. In addition, according to them, when the object bears the topic feature, the direct object doubling is obligatory. Unlike Kallulli (2000, 2008) and Kapia (2010, 2012), Buchholz & Fiedler (1987) draw different conclusions, according to which the direct object clitic doubling in the Albanian language may be optional if the word order is canonical (SVO). This paper, on the other hand, aims at proving whether the object clitic doubling in our corpus marks the information structure by giving the feature [+ topic] or [-focus].

KEYWORDS: information structure, clitic doubling, topic, focus, direct object, *Koha ditore* newspaper, *Një tregim Prishtine* movie

1. Introduction

Clitic doubling means the doubling (or “repetition”) of an **object** or verbal argument by a clitic pronoun, which argument “is associated” with the clitic within the same propositional structure or within the same clause (CP). The role of the “associate” within the verbal argument is realized by a pronoun, a non-pronoun referential expression, or a clause (CP). “When the verbal argument or the associate is a pronoun or a DP, the clitic bears the same phi-features (ϕ) as the pronoun or the DP, whereas when the argument is a CP, the clitic is a single element in the accusative” (Kallulli and Tasmowski, 2008, p. 1). To argue that the clitic bears the same phi-features of the associate, we analyze the following examples:

1. *Ajo i takoi ata.* “She met them.”
2. *Ajo e këndoi këngën.* “She sang the song.”
3. *Ajo e dinte që nuk ishte në shtëpi.* “She knew that you were not home.”

In such cases, *i* and *e* are functioning as clitics, whereas *ata* “they” (pro), *këngën* “the song” (DP), *që nuk ishte në shtëpi* “that you were not home” (CP) are in the position of a verbal argument functioning as direct objects. This grammatical phenomenon of direct object clitic doubling in the Albanian language is supposedly related to the presence or absence of the feature [+/- topic], through which the direct object clitic doubling is given or not given the status of the topic. Therefore, according to Kallulli (1995, 1999, 2000, 2008), clitic doubling of direct objects is never optional. Conversely, Buchholz & Fiedler (1987) draw conclusions different from the above, according to which the direct object clitic doubling in the Albanian language can be optional if the word order is canonical (SVO). Thus, according to them, repeating the direct objects via clitics not always marks the information structure in the Albanian language.

1.1. Literature review

In modern linguistics, clitic doubling is defined as “the co-occurrence of a direct or indirect object and a co-referential clitic pronoun attached to the verb” (Janse, 2008, p. 165). According to Gabriel and Rinke (2010) clitic doubling is a phenomenon that occurs when the clitic pronoun and the corresponding object XP appear in the same clause, in which case the object clitic functions as an argument marker, and the object (XP) satisfies the thematic and case conditions within the vP. The direct object doubling has been treated not only as a morpho-syntactic phenomenon, but also as an element that marks the information structure (Gabriel and Rinke, 2010; Kallulli, 2000; Fridman, 2008; Kapia, 2008). According to Gabriel and Rinke (2010, p. 58–59), based on the information conveyed by the clitics, two types of object clitic doubling are to be distinguished: clitic

doubling and clitic right dislocation; the latter is interpreted as topic, and the former is interpreted as focus. Holton, Mackridge & Philippaki-Warbuton (1997), on the other hand, describe clitic doubling as “a mechanism that removes the object from the comment (focus) of the sentence and makes it part of the given information (topic). In terms of generating feature phrases, clitic doubled NPs (DPs) are marked as [-Focus] as opposed to clitic undoubled NPs (DPs), which are interpreted as [+Focus]. Regarding the approach of clitic doubling within the information structure in the Albanian language, two studies have been conducted by Kallulli (2002; 2008) and Kapia (2010). Kallulli points out that direct object doubling via third-person clitic pronouns in Albanian (and modern Greek) produces information structure in a systematic way, in that doubled DPs are unambiguously interpreted as topics. In fact, Kallulli’s (2000 and 2008) research assume that (third person) clitics together with the direct object which they repeat form the topic in a clause (with some exceptions related to verb features). Therefore, to interpret these DPs as topics, they have to necessarily be clitic doubled. This stands despite the variations in word order. That is, the clitic doubling of direct objects (third person singular and plural) is not an optional phenomenon. Likewise, according to Buchholz & Fiedler (1987), direct object clitic doubling occurs only when the noun phrase in the accusative case bears the known information (i.e., lies within the topic domain), whereas when the clitic is absent, the direct object bears new information (i.e., lies inside the focus domain). However, direct object clitic doubling in the Albanian language can be optional and not obligatory when the word order is canonical (SVO), for example, *Agim-i po Ø/e vështron hënë-n* “Agim is looking at the moon”; whereas indirect object clitic doubling is never optional, for example, *Shfaqj-a *Ø/i pëlqeu Agimit* “Agim liked the play”.

1.2. Research questions

The questions to be answered in the present study are as follows:

- a. Does clitic doubling of the object (XP) affect the encoding of the information structure?
- b. Is object clitic doubling in Albanian language always interpreted as topic, as has been considered so far?
- c. Is direct object clitic doubling optional or obligatory in conveying information units, topic or focus, that is, does clitic doubling affect the occurrence of the object inside the domains of new or given information.

2. Method

To conduct this research, the following methods have been employed: method of direct observation, which aims at investigating the empirical data when clitic doubling produces the information structure; method of analyzing the interpretation of these data to see if clitic doubling always produces the information structure; comparative method, especially between written and spoken language, to see whether in the presented data the realization of topic via NP/DP clitic doubling is more frequent in written or in spoken Albanian.

2.1.1. Data collection and analysis

To observe the role of object clitic doubling in the information structure, two corpora have been selected: the newspaper *Koha ditore* and the feature film *Një tregim Prishtine*. These types of texts have been selected to analyze the role of clitic doubling in the information structure of the news, “which deviates from rules and norms, representing an event extra ordinary from the everyday life” (Whitaker et al, 2004; as cited in Rugova, 2016, p. 97), but also to analyze the role of the clitic doubling mechanism in relation to the information structure of the film language, which contains **exophoric references**, where a clitic pronoun refers to a unit outside the text (Rugova, Sejdiu - Rugova, 2015). Therefore, in order to analyze the correlation between direct object clitic doubling and the encoding of the information structure, we have analyzed the context (both in the newspaper *Koha ditore* and in the movie *Një tregim Prishtine*) where the DPs appear functioning as direct objects. In such cases it has been observed if DPs functioning as objects are entities already activated from the previous discourse and if these DPs have been repeated via clitics, that is, if the clitic doubling is related to the realization of the given information or topic.

3. Results and discussion

It was said that direct object clitic doubling is not an optional phenomenon in the Albanian language. It was also said that clitic doubling produces and encodes the information structure in a systematic

way, in that doubled objects are unambiguously interpreted as topics. According to this definition, doubling occurs when referring to a referent already present in the discourse. A necessary condition for doubling is for the DP to be already present, thus becoming a specific referent available to be later selected via the doubling mechanism (Kallulli, 2000, 2008; Kapia, 2010). Conversely, based on an unsystematic observation made earlier, we have noticed that the direct object is not always clitic doubled, even when the referent has been already introduced in the text. Therefore, the object is not repeated even when the DP is known to the hearer/receiver, i.e., a topic, having already been introduced earlier in the text. To see how accurate this theory is, a pragmatic research has been conducted, analyzing the anaphoric and cataphoric relations of the referent and the element to which it refers, which also indicates whether the referent has already been introduced earlier in the discourse. Thus, the context in which the direct objects tend to be or not be doubled by clitics is also analyzed, as the syntactic relations alone between the clitic and the object are not sufficient to prove or disprove the above arguments. Therefore, the main element for analysis is the context in which the reference between the clitic and the direct object appears. The reference can be anaphoric or cataphoric. It is *anaphoric* when the referential element enters the discourse after the unit to which it refers (Rugova, Sejdiu-Rugova, 2015). It is *cataphoric*, however, when the referential unit precedes the referenced unit. Therefore, an anaphoric clitic is a morpheme morphologically related to a constituent mentioned earlier; it refers to the same and completes the position of a DP in the clause, for example: *Librin e lexoi dje* “He read the book yesterday”. *Librin* “the book” is the referenced unit and *e* is the referential unit. They are also morphologically related to each other. A cataphoric clitic, on the other hand, is a morpheme morphologically related to a constituent mentioned later in the discourse; it refers to the same and completes the position of a DP in the clause, for example: *E lexoi librin* “He/she read the book”. Anaphoric/cataphoric relations occur only if: a) one or more clitics carry the role of a verbal argument; b) a semantic reference head (which may be a phrase, sentence, discourse, or context) is set for each clitic (Cresti, 2009).

3.1. Analysis

- a. Examples from the newspaper *Koha ditore*: *Context*: Korisha village is known for the self-organization of inhabitants, this practice will be applied even during the quarantine period, imposed due to the coronavirus pandemic (overline).

[1] **Fshati Korishë Ø vazhdon vetorganizimin edhe gjatë karantinës** (headline)

[_{TOP} *Fshati Korishë*] Ø_(cl) [_{FOC} *vazhdon*] [_{TOP} *vetorganizimin edhe gjatë karantinës*]

“Korisha village continues self-organization even during quarantine”

In the given example the NP functioning as direct object (*vetorganizimin* “self-organization”) is not doubled by the clitic (*e*) although as a phenomenon self-organization has been previously introduced in the text by overline and does not constitute new information for the reader, hence it lies within the borders of the topic. Therefore, although known to the reader, so topic, it is not clitic doubled.

Context: In its campaign Lëvizja Vetëvendosje promised to depoliticize public enterprises... Albin Kurti said that these boards are temporary and that they will deal with the fight against crime and corruption from within.

[2] **Kurti kritikohet se Ø ka shkelur premtimin për depolitizimin e bordeve** (headline)

Kurti kritikohet/ se [_{TOP} *pro*] Ø_(cl) [_{FOC} *ka shkelur*] [_{TOP} *premtimin për depolitizimin e bordeve*]

“Kurti criticized for violating the promise to depoliticize boards”

Gjashtë muaj më vonë, qeveria në detyrë, e udhëhequr nga Albin Kurti Ø ka shkarkuar përbërjen e tri bordeve të ndërmarrjeve publike për t’i zëvendësuar ato me disa persona të afërt me këtë subjekt politik (nëntitull).

[_{TOP} *Gjashtë muaj më vonë, qeveria në detyrë, e udhëhequr nga Albin Kurti*] Ø_(cl) [_{FOC} *ka shkarkuar*] [_{TOP} *përbërjen e tri bordeve të ndërmarrjeve publike për t’i zëvendësuar ato me disa persona të afërt me këtë subjekt politik*]

“Six months later, the incumbent government led by Albin Kurti dismissed the composition of three boards of public enterprises to replace them with certain persons close to this political party (subheading).”

In the given example the NP functioning as direct object (*premtimin* “promise”) is not doubled by the clitic (*e*), although as a concept it is already known to the public, as it has been previously introduced in the text. It is not repeated via the short form, even though the direct object (*premtimin* “promise”) together with the determiners enter the topic domain and constitutes given information. Likewise, in the subheading *Albin Kurti Ø ka shkarkuar përbërjen e tri bordeve të ndërmarrjeve publike* “Albin Kurti dismissed the composition of three boards of public enterprises”, the noun phrase (*përbërjen e tri bordeve të ndërmarrjeve publike* “the composition of three boards of public enterprises”) is not repeated via the short form even though it is given information for the reader by use of hyronims which connect the headline and subheading: *depolitizimi i bordeve ~ përbërje e tri bordeve* “depoliticization of boards ~ composition of three boards” (for hyronims see Rugova, Sejdiu - Rugova, 2015).

Context: The picture shows four facilities which have been identified by the Municipality of Peja to serve as quarantine.

[3] **Komuna e Pejës Ø ka identifikuar katër objekte për karantinë** (headline) p. 4;

[_{TOP} Komuna e Pejës] Ø(_{cl}) [_{FOC} ka identifikuar] [_{TOP} katër objekte për karantinë]

“The Municipality of Peja has identified four facilities for quarantine”

Komuna e Pejës Ø ka identifikuar katër objekte publike, të cilat ndonëse kanë destinime të tjera, mund të shfrytëzohen edhe për vendosjen e presonave të prekur nga koronavirusi (subheading).

[_{TOP} Komuna e Pejës] Ø(_{cl}) [_{FOC} ka identifikuar] [_{TOP} katër objekte publike, të cilat ndonëse kanë destinime të tjera], [_{FOC} mund të shfrytëzohen edhe për vendosjen e presonave të prekur nga koronavirusi].

“The Municipality of Peja has identified four public facilities, which although have other purposes, can also be used to accommodate persons affected by the coronavirus.”

In the picture inside the text are given four facilities selected for quarantine, thus, the noun phrase *katër objekte publike* “four public facilities” functioning as direct object (plural) is already known to the reader and, according to research done so far (Kallulli, 2000, 2008; Kapia, 2010), to double the clitic and the information structure, this noun phrase should have been doubled by the short form of the third person singular pronoun *i*, to have the status of the topic. However, even without being marked with the clitic *i*, this noun phrase “four public facilities” carries the status of the topic, as it is not entirely new information for the reader. *Context:* During the pandemic period, imports and exports in Kosovo decreased, causing a reduction of the trade deficit to 161.2 million. Overline: *Blokimi i aktiviteteve ekonomike dhe rënia e konsumit ulin importet dhe eksportet.* “Blocking economic activity and falling consumption reduce imports and exports.”

[4] **Pandemia Ø ul deficitin tregtar për 37 për qind** p. 2; 25 May; no. 8206

[_{TOP} Pandemia] [_{FOC} ul] [_{TOP} deficitin tregtar për 37 për qind]

“Pandemic reduces trade deficit by 37 percent”

As for this given example, based on the assumptions made by Kallulli (2000, 2008) the DP, as it is not clitic doubled, will belong to the new information. However, based on the context, it is in the over line that the NP (*deficitin tregtar* “trade deficit”) is introduced to the reader via collocation (*importet dhe eksportet~deficit tregtar* “imports and exports~trade deficit”), thus, we categorize it inside the domain of known information (topic), although it is not repeated via the short form of the personal pronoun (*e*). In the two examples below, we will see a NP in function of the direct object that appears in almost the same context and in the first case [5] is part of the focus domain, while in the second case [5a] it is topic.

Context: Vermont Senator, Bernie Sanders, who less than four months ago led the race for the Democratic nomination to challenge Republican President Donald Trump ... Sanders reconsiders his run for the White House after deep losses to Joe Biden.

[5] **Sanders Ø rivlerëson fushatën pas humbjeve të thella ndaj Bidenit** (news headline) p. 11; 19 March; no.8139

[_{TOP} Sanders] Ø [_{FOC} rivlerëson fushatën pas humbjeve të thella ndaj Bidenit]

“Sanders reconsiders the campaign after deep losses to Biden”

[5a] Bernie Sanders është duke *e* rishikuar garën e tij për Shtëpinë e Bardhë, pas humbjeve të thella kundruall Joe Biden... (article entry p. 11).

[_{TOP} Bernie Sanders] [_{FOC} është duke *e*_(cl) rishikuar] [_{TOP} garën e tij për Shtëpinë e Bardhë, pas humbjeve të thella kundruall Joe Biden]

“Bernie Sanders is reconsidering his run for the White House, after deep losses to Joe Biden”

In this example [5], which is also the news headline, the NP *-fushatën* “the campaign” is not clitic doubled; yet, based on the context, the NP is new information, as it is not given earlier in the overline. Conversely, in example [5a], the NP as a direct object, repeated via the clitic with which it is in anaphoric relation, is old information, because it is understood from the title via collocation (*fushatë ~ garë* “campaign ~ run”).

b. Examples from the feature film *Një tregim Prishtine*:

Context: A family of three (father, mother, and daughter) are in the car. In the background is heard a song by Ledri: *Ni got për ty* “a glass for you”. All three sing the chorus:

[1] // [^] *E pi ni got për ty* // [^] *Masi ty e kohën s’ muj mi kthy* //

// [_{TOP} *pro*] *e* [_(cl)] [_{FOC} *pi ni got për ty*] // [_{TOP} *Masi ty e kohën*] [_{FOC} *s’ muj mi kthy*]

“I drink a glass for you” “since I can’t return you or the time”

In the given example the NP in function of the direct object (*ni got* “a glass”) is doubled via the clitic (*e*), although, if taken out of context, the NP *ni got* “a glass” does not convey old information, but the transitive verb itself (*pi* “drink”) necessarily requires an object that conveys an expected information for the hearer. On the other hand, in the clause *masi ty e kohën s’ muj mi kthy* “since I can’t return you or the time”, the coordinated NP (you or the time) functions as old information, as pronoun *ty* “you” is mentioned in the preceding clause and likewise the clitic *i* is presented as a cataphora.

Context: Being in the car, singing at the peak of their happiness, the wife advises the husband to be careful:

[2] // [^] *O Ben kçyr-e rrugën* //

[_{TOP} *Ben*] [_{FOC} *kçyr-e*_(cl)] [_{TOP} *rrugën*]

“O Ben watch the road”

In the given example the NP functioning as a direct object *rrugën* “the road” (*kçyr-e rrugën* / “watch the road”) is doubled via the clitic (*e*) in the position of an enclitic, infixed in the verb. The NP *rrugën* “the road” conveys old information, as it is implied that when driving the car, one must be careful on the road, and not something else.

Context: So, the family is in the car, singing and, at the peak of their happiness, the husband asks the woman for a kiss on the cheek:

[3] // [^] *Ama ni façe* //

[_{TOP} *pro*] [_{FOC} *ni façe*]

“Gimme your cheek”

Formally it does not seem that there is any clitic referring to the NP *façe* “cheek”, which conveys new information, therefore we assume that the clitic *e* + the other clitic *më* for the first person are merged with the verb *jap* “give”, giving us the form *ama* “gimme”. Although we have a clitic present here, the NP *façe* to which it refers is part of the focus domain because it was not mentioned before in the discourse.

Context: Ben and his wife are discussing and Ben speaks ironically, he wants to quarrel with her, and she tells him:

[4] // [^] *Ben mos ja nis* // (zënkës)

[_{TOP} *Ben*] [_{FOC} *mos ja* _(cl) *nis* (zënkës)]

“Ben don’t start [the fight]”

In example [4] we have an idiomatic expression (*mos ja nis/a ja nise* a “don’t start”) with both doubling pronouns in dative and accusative (i+a). In such cases, it is assumed that the clitic doubling of the object is necessary, as Friedman (2008) states, the clitic doubling of a NP in an idiomatic construction is obligatory and has nothing to do with the information structure, since, according to him, the clitic doubling of the object here appears to be in part a morphologized or lexicalizing signal. In our example, the indirect

object, even though it is being doubled with two clitics, is still lying in the domain of focus, so the whole idiomatic expression (mos ja nis/a ja nise a “don’t start”) is interpreted as new information, since through this information we understand that she (his wife) is not ready to start fighting. Thus, doubling clitic of object occurs even when the NP is inside the domain of new information.

4. Conclusion

Clitic doubling does not serve as a pragmatic tool in the information structure and is not always used as a tool to mark direct objects as topics. Based on the data from this research, it follows that direct object clitic doubling is not always related to the marking of one of the two notions of information structure, topic and focus. In twenty out of thirty examples extracted from the newspaper it is noticed that direct objects are not repeated via short forms (clitics), even though it can be said that they enter the topic domain, in that they are considered as given information, that is, they have, in some form, been mentioned in the prior discourse. Of the above twenty examples extracted from the movie, we have noticed that in nine of them the direct object is not clitic doubled when it has the feature [+ topic]. In the studies by Kallulli (2000, 2008) and Kapia (2010, 2012) there is a certain tendency on the relation between the clitic doubling and the information structure, to normalize the use of clitic in cases where the NP, functioning as a direct object, is recognized by the prior discourse and it is topic.

REFERENCES:

- Brisard, F. Östman, J.-O. Verschueren, J. (Eds.) (2009)** Grammar, Meaning and Pragmatics: Handbook of Pragmatics Highlights 5. Cambridge: Cambridge University Press.
- Buchholz, O., Fiedler, W. (1987)** Albanische grammatik. Verlag Enzyklopädie, 443 p.
- Cresti, E. (2009)** Clitics and anaphoric relations in informational patterning: A corpus-driven research on spontaneous spoken Italian (C-ORAL-ROM). In L. Mereu (Ed.), *Information Structure and its Interfaces* (pp. 169-202). Berlin, New York: De Gruyter Mouton.
- Friedman, V. A. (2008)** Balkan object reduplication in areal and dialectological perspective. // *Clitic doubling in the Balkan languages* (Ed. by Dalina Kallulli, Liliane Tasmowski), John Benjamins, 35-63.
- Gabriel, C., Rinke, E. (2010)** Information packaging and the rise of clitic doubling in the history of Spanish. // *Diachronic studies on information structure. Language acquisition and change*, 63-86 p.
- Holton, D., Mackridge, P., Philippaki-Warbuton. (1997)** Greek: A comprehensive grammar of the modern language. London: Routledge, 432 p.
- Janse, M. (2008)** Clitic doubling from ancient to Asia Minor Greek. // *Clitic doubling in the Balkan languages*, (Ed. by Dalina Kallulli, Liliane Tasmowski), John Benjamins, 165 p.
- Kallulli, D., Tasmowski, L. (2008)** Clitic doubling in the Balkan languages (Vol. 130). John Benjamins Publishing, 1 p. – 227 p.
- Kallulli, D. (2000)** Direct object clitic doubling in Albanian and Greek. // *Clitic phenomena in European languages*, 209-248 p.
- Kapia, E. (2010)** The role of syntax and pragmatics in the structure and acquisition of clitic doubling in Albanian. Boston University, 82 p.
- Kapia, E. (2012)** Clitic doubling and information structure in Albanian. University of Sussex, 904 p.
- Rugova, B., Rugova, L. (2015)** Hyrje në gramatikën e tekstit të gjuhës shqipe. Prishtinë: Trembelat, 139 p.
- Rugova, B. (2016)** Riçke rome, itaçi e zakucati. Prishtinë: Trembelat, 97 p.
- Sejdiu – Rugova, L. (2019)** Strukturë dhe funksion studime në sintaksë të shqipes. Prishtinë: Trembelat.