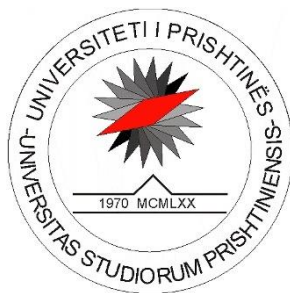


UNIVERSITETI I PRISHTINËS “HASAN PRISHTINA”
FAKULETI I FILOLOGJISË



MA QËNDRESA JAKUPI

**STRUKTURA E INFORMACIONIT
BRENDA FJALISË: TOPIKA DHE FOKUSI
NË GJUHËN SHQIPE**

PUNIM I DOKTORATËS

Prishtinë, 2023

UNIVERSITY OF PRISHTINA "HASAN PRISHTINA"

FACULTY OF PHILOLOGY

MA QËNDRESA JAKUPI

**INFORMATION STRUCTURE WITHIN THE
SENTENCE: THE TOPIC AND THE FOCUS
IN THE ALBANIAN LANGUAGE**

DOCTORAL THESIS

Prishtina, 2023

UNIVERSITETI I PRISHTINËS “HASAN PRISHTINA”

FAKULTETI I FILOLOGJISË

MA QËNDRESA JAKUPI

**STRUKTURA E INFORMACIONIT BRENDA
FJALISË: TOPIKA DHE FOKUSI NË GJUHËN
SHQIPE**

PUNIM I DOKTORATËS

Mentorja:

Prof. dr. Lindita Sejdiu-Rugova

Prishtinë, 2023

REZYME

Qëllimi i këtij studimi është ta hetojë mënyrën se si kodifikohet struktura e informacionit në gjuhën shqipe brenda fjalisë. Struktura e informacionit në literaturë nënkuptohet si një ndarje midis fokusit dhe topikës së një fjalie. Sipas Halliday, “çdo tekst [. . .] është i organizuar në ato që mund të quhen njësi informimi [. . .] dhe kjo nuk është përcaktuar [. . .] sipas strukturës përbërëse. Përkundrazi, mund të thuhet se shpërndarja e informacionit specifikon një strukturë të veçuar në një plan tjetër” (1967: 200).

Prandaj, për ta hulumtuar mënyrën e kodifikimit të strukturës së informacionit në gjuhën shqipe këtë punim të doktoratës e kemi ndarë në gjashtë kapituj, duke përfshirë hyrjen dhe përfundimet:

Në kapitullin e HYRJES (0) shtjellohet së pari koncepti i strukturës së informacionit, objekti i studimit, hipotezat e studimit si dhe metodat dhe detyrat. Qysh në fillim të këtij kapitulli është synuar të definohet koncepti i strukturës së informacionit brenda kornizave të gjuhësisë, duke i shpjeguar edhe ndërfaqet e strukturës së informacionit me fushat e tjera të gjuhësisë, si: sintaksa, semantika, prozodia dhe pragmatika, për të treguar se përdoren strategji të përziera që të strukturohet informacioni brenda një fjalie dhe struktura e informacionit si koncept nuk mund të funksionojë si e vetme pa ndërvepruar me domenet e lartpërmenduar në disa pika të caktuara. Tutje, pas definimit të konceptit, është synuar të përkufizohet objekti i studimit përmes të cilit provohet të shpjegohet se si informacioni është i organizuar brenda një fjalie të shqipes, dhe në veçanti se si sintaksa, semantika, prozodia dhe struktura e informacionit ndërveprojnë në kodifikimin e informacionit në gjuhën shqipe. Me ç’rast këto supozime nxisin hipoteza se shqipja përdor strategji të ndryshme, për ta kodifikuar strukturën e informacionit, të cilat do të shihen në vazhdim. Të gjitha këto strategji që përdoren për të mundësuar dhe hetuar veprimin topikë/fokus është përdorur një metodologji heterogjene, për shembull: për të hetuar topikën në një fjali janë përdorur analiza sintaksore dhe semantike, njësoj edhe për të hetuar fokusin etj.

Brenda boshtit të *kapitullit të parë* hulumtohet hollësisht nocioni i temës/topikës në gjuhën shqipe. Së pari, ofrohet një studim i shkurtër historik i cili do të tregojë disa nga përkufizimet e temës/topikës në literaturën e gjuhësisë. Pastaj, do të shqyrtohen karakterizimet e ndryshme dhe motivime që qëndrojnë pas tyre. Me ç'rast shqyrtohen propozime që lidhen me nocione të tilla, si: karakteristika tipike të topikës, rrethimi sintaktik, referencat diskursive, dija paraprake në diskurs dhe konteksti, si dhe ato që imponojnë disa kufizime sintaktike të tilla. Po ashtu, mënyra e realizimit të të sintagmës së topikës në gjuhën shqipe.

Në *kapitullin e dytë* një nocioni tjetër që trajtohet brenda boshtit të këtij disertacioni është nocioni i fokusit/remës. Për të cilin do të shqyrtohet pozicioni i tij në fjali, si dhe me çfarë mjete markohet në gjuhën shqipe. Po ashtu, do t'u kushtohet rëndësi disave prej llojeve të fokusit që janë propozuar në literaturë (Lambrecht 1994). Detyra e identifikimit të fokusit të një fjalie mund të duket më e lehtë sesa e identifikimit të një topike, pasi në krahasim me topikën, fokusi nuk markohet vetëm fonologjikisht përmes theksit siç është pretenduar deri më tani. Pra, jo të gjithë elementet që e përçojnë fokusin theksohen, disa elemente që marrin theks nuk e përcjellin domosdoshmërisht fokusin. Madje, elementet që fokusohen mund të jenë elemente që përcjellin informacione jo krejt të reja për dëgjuesin dhe se disa sintagma të fokusit mund të paraqiten në pozicionin fillestar të fjalisë. Prandaj, do të bëjmë analizë semantike në kuadrin gjenerativist përmes llojeve të pyetjeve për të gjetur fokusin.

Megjithëse, topika dhe fokusi janë trajtuar si nocione kryesore të strukturës së informacionit, prapëseprapë këto nocione nuk i përthekojnë studimet mbi të, sepse ky nuk është përdorimi i vetëm i mundshëm i termit informacion, pasi ai gjithashtu mund ta përshkruajë mënyrën e përgjithshme në të cilën strukturohet informacioni i çfarëdo lloji qoftë. E, po ashtu, struktura e informacionit markohet gjuhësisht ndryshe në gjuhë të ndryshme. Siç supozon edhe Vallduví (1996: 460), paketimi i informacionit është një strukturim i fjalive me mjete sintaksore, prozodike ose morfologjike që lind nga nevoja, për t'i përmbushur kërkesat komunikuese të një konteksti ose diskursi të veçantë.

Me këto nënkuptojmë se struktura e informacionit realizohet nga një bashkëveprim i shumëllojshëm i mjeteve apo strategjive gjuhësore. Faktorët semantikë, pragmatikë, sintaksorë dhe prozodikë ndërthuren në pasqyrimin e paketimit të informacionit, i cili nga ana tjetër siguron

depërtim, jo vetëm, në ndërfaqet e këtyre përbërësve, por gjithashtu, në mënyrën sesi gjuhët e ndryshme e kodifikojnë informacionin e tillë (Gürer 2020: 3).

Prandja, dy nocionet që u përmendën më parë, topikë dhe fokus, nuk janë elementet e vetme që e përbëjnë boshtin e këtij studimi, por ky bosht plotësohet edhe nga mekanizma të tjerë gjuhësorë, me të cilët supozohet se realizohet struktura e informacionit në gjuhën shqipe. Ndër mekanizmat gjuhësorë, që mendohet se luajnë rol në kodifikimin dhe strukturimin e informacionit në gjuhën shqipe, janë: mjetet prozodike që hamendësohet se e markojnë fokusin (FocP), duke e dalluar atë nga topika (TopP); tjetër mekanizëm është dyfishimi i kundrinave nëpërmes klitikëve përemërorë, për të cilin pretendohet se e kodifikon topikën ose informacionin e dhënë më parë në diskursin paraprak; një strategji tjetër që besohet se përdor shqipja për ta realizuar strukturën e informacionit është konfigurimi variacional i fjalive, meqë është një gjuhë deri diku me rend të lirë.

Prandaj, brenda kapitullit të dytë provohet nëse shqipja i përdor **mjetet prozodike**, për ta kodifikuar strukturën e informacionit, pra nëse në gjuhën shqipe markimi i fokusit bëhet nëpërmjet elementeve mbisegmentare përfshirë frekuencën themelore (F0), intensitetin dhe gjatësinë e tingullit të fjalës së dhënë. Dhe nëse fjala e dhënë e përmban tonin H*+ L, do të merret si fokus i fjalisë, pra kjo fjalë (**F-marked**) e bart theksin sipas teorisë autosegmentale metrike të intonacionit.

Në kapitullin e tretë synohet të vërtetohet nëse dyfishimi i kundrinave përmes përemrave klitikë e markon topikën, pasi nga disa studime (Kallulli 2000, 2008; Kapia 2010) që janë bërë më parë, është thënë se dyfishimi me klitikë i kundrinave të drejta nuk është një fenomen fakultativ në gjuhën shqipe. Si dhe është supozuar se dyfishimi i klitikut e prodhon dhe kodon strukturën e informacionit në mënyrë sistematike, në mënyrë që kundrinat e dyfishuara interpretohen qartësisht si topikë.

Tutje, një mekanizëm tjetër sinatksor që do të trajtohet në kapitullin e katërt ka të bëjë me konfigurimet variacionale të fjalive në raport me strukturën e informacionit se sa ndikon në kodifikimin e kësaj të fundit. Shqipja llogaritet të jetë gjuhë me rend të lirë, ndërsa rendi kanonik i fjalëve konsiderohet të jetë ai SVO. Nga një hulumtim i bërë nga Sejdiu-Rugova dhe Simitçiu (2018) del se shqipja *tregon një fleksibilitet të madh në renditjen e përbërësve të fjalisë. Për shkak të këtij fakti, ajo është përmendur shpesh si një gjuhë "me rend të lirë", me rendin SVO të*

përbërësve që paraqitet si themelor, por edhe jo i detyrueshëm. Pra, zbatimi i disa veprimeve të lëvizjes opsionale mund të japë edhe renditje të ndryshme fjalësh (**SVO, OSV, OVS, VSO, VOS, SV, VS**), që tregon se shqipja është një gjuhë “scrambling”. Dukuria “scrambling”, e cila së pari u përcaktua si rend i lirë i fjalëve, shihet si rezultat i operacioneve lëvizëse krejtësisht opsionale nga disa gjuhëtarë (si: Chomsky, Rizzi, Bošković), dhe si rezultat i veprimeve të caktuara të lëvizjes së detyrueshme nga disa të tjerë (si: Güreer; Şener). Ndërsa në gjuhën shqipe ndryshimet e mundshme të renditjes së fjalëve shpesh janë konsideruar si konfigurim variacional me qëllim për ta ndryshuar ose rritur peshën e informacionit (Sejdiu-Rugova dhe Simitçiu 2018; Xhaferi 2005).

Prandaj, këtu do të argumentohet nëse të gjitha operacionet e lëvizjes, që realizohen ndryshe nga struktura e përgjithshme e fjalive në shqip, drejtohen nga tipare të bazuara në diskurs; dhe nëse ka vend për ndonjë operacion që e ndryshon rendin linear të përbërësve në mënyrë të rastësishme dhe/ose pa përfshirjen e një tipari diskursiv. Pra, nëse ndërrimet e rendit të fjalëve përcaktohen plotësisht nga motive diskursive-pragmatike nëpërmjet të cilave kodifikohet struktura e informacionit.

Për t’i analizuar dhe hulumtuar të gjitha këto që i thamë më lart janë përzgjedhur tre korpusë: gazeta **Koha**; emisioni **Pressing** (30 maj 2022); filmi i metrazhit të gjatë **Një tregim Prishtine**; dhe komentimi i ndeshjes së futbollit **Serbi – Shqipëri** (14.10.2014 Beograd). Këto lloje tekstesh janë përzgjedhur me qëllim, pasi janë ndërtuar mbi stile të ndryshme që kanë funksione të ndryshme, e poashtu dallojnë për nga këndvështrimi folës, i cili së bashku me stilin mund të llogariten si dy faktorë të jashtëm, që luajnë rol në kodifikimin e strukturës së informacionit.

Gjatë hulumtimit, po ashtu, përdoren metoda deskriptive, statistikore (në kapitullin e tretë), analitike, krahasimtare dhe e vëzhgimit të drejtpërdrejtë.

Fjalët kyçe: strukturë informacioni, topikë, fokus, TopP, FocP, markimi-F, dyfishimi klitik, konfigurim variacional i fjalëve në fjali

SUMMARY

The purpose of this study is to investigate the way in which the information structure is codified in the Albanian language within the sentence. Information structure in literature is understood as a division between the focus and the topic of a sentence. According to Halliday, “every text [. . .] is organized into what may be called information units [. . .] and this is not defined [. . .] according to the constituent structure. Rather, it can be said that the distribution of information specifies a separate structure on another plane” (1967: 200).

Therefore, in order to investigate the way of codification of the information structure in the Albanian language, this doctoral thesis has been divided into six chapters, including the Introduction and the Conclusion:

In the INTRODUCTION chapter (0) the concept of the structure of information, the object of the study, the hypotheses of the study, as well as the methods and tasks are elaborated first. From the very beginning of this chapter, there is an intention to define the concept of information structure within the framework of linguistics, while explaining the interfaces of information structure with other fields of linguistics, such as: syntax, semantics, prosody and pragmatics, in order to show that mixed strategies are used to structure information within a sentence, and information structure as a concept cannot function alone without interacting with the aforementioned domains at certain points. Further, after defining the concept, it is intended to define the object of the study through which it is tried to explain how information is organized within an Albanian sentence, and in particular how syntax, semantics, prosody and structure of information interact in the codification of information in the Albanian language. Consequently, these assumptions lead to the hypothesis that Albanian uses different strategies to codify the structure of information, which will be seen in the following chapters. All these strategies that are used to enable and investigate the topic/focus action denote that a heterogeneous methodology is used, for instance: in order to investigate the topic in a sentence, a syntactic and semantic analysis is used, as well as to investigate the focus, and so on.

Within the axis of the first chapter, the notion of topic/topic in the Albanian language is explored in detail. First, a brief historical survey is provided which will show some of the definitions of theme/topic in the linguistics literature. Then, the various characterizations and motivations behind them will be examined. In this case, propositions related to such notions as: typical characteristics of the topic, syntactic context, discursive references, prior knowledge in discourse and context, as well as those that impose some such syntactic constraints, are examined. Also, the way of realizing the syntagm of the topic in the Albanian language.

In the second chapter, another notion that is treated within the axis of this dissertation is the notion of focus/rheme. Thus, its position in the sentence will be examined, as well as by what means it is marked in the Albanian language. Also, attention will be paid to some of the types of focus that have been proposed in the literature (Lambrecht 1994). The task of identifying the focus of a sentence may seem easier than identifying a topic, since compared to topics, focus is not marked phonologically only through stress. So not all elements that carry focus are highlighted, some elements that get focus don't necessarily carry focus. In fact, focus elements may be elements that convey information not entirely new to the listener, and some focus phrases may appear in sentence-initial position. Therefore, we will do semantic analysis in the generativist framework through question types to find the focus.

Although the topic and focus have been treated as the main notions of information structure, these notions still do not underline studies on it, because this is not the only possible use of the term information, as it can also describe the general way in which information of any kind is structured. Also, the structure of information is linguistically marked differently in different languages. As Vallduví (1996: 460) assumes, information packaging is a structuring of sentences with syntactic, prosodic or morphological means that arises out of necessity, to meet the communicative requirements of a particular context or discourse.

By these we mean that the information structure is realized by a diverse interaction of language tools or strategies. Semantic, pragmatic, syntactic and prosodic factors intertwine in reflecting the packaging of information, which in turn provides insight, not only into the interfaces of these components, but also into how different languages codify such information (Gürer 2020: 3).

Therefore, the two notions that were mentioned before, topic and focus, are not the only elements that make up the axis of this study, but this axis is also complemented by other linguistic mechanisms, with which it is assumed that the structure of information in the Albanian language is realized. Among the linguistic mechanisms, which are thought to play a role in the codification and structuring of information in the Albanian language, are: prosodic tools that are assumed to mark the focus (FocP), distinguishing it from the topic (TopP); another mechanism is the doubling of opposites through pronominal clitics, which is claimed to codify the topic or information previously given in the preceding discourse; another strategy that is believed to be used by Albanian to realize the structure of information has been the variational configuration of sentences, since it is a language with somewhat free order.

Therefore, within the second chapter, it is proven if Albanian uses prosodic tools to codify the structure of information, that is, if in the Albanian language focus marking is done through suprasegmental elements including the fundamental frequency (F0), intensity and length of the sound of the given word. And if the given word contains the tone H*+ L, it will be taken as the focus of the sentence, so this word (F-marked) carries the accent according to the metrical autosegmental theory of intonation.

In the third chapter, the aim is to prove whether the doubling of adverbs through clitic pronouns marks the topic, since in several studies (Kallulli 2000, 2008; Kapia 2010) that have been done before, it has been claimed that the doubling of adverbs with clitics is not an optional phenomenon in the Albanian language. Also, it has been assumed that clitic doubling produces and encodes information structure in a systematic way, so that doubled opposites are clearly interpreted as topical.

Further, another syntactic mechanism that will be treated in the fourth chapter has to do with the variational configurations of the sentences in relation to the structure of the information as it affects the codification of the latter. Albanian is considered to be a language with free order, while the canonical word order is considered to be SVO. From a research done by Sejdiu-Rugova and Simitçiu (2018) it appears that Albanian shows a great flexibility in ordering the components of the sentence. Because of this fact, it has often been referred to as a “free-order” language, with the SVO order of the constituents appearing as fundamental but also not mandatory. So, the

implementation of some optional movement actions can also give different word orders (**SVO, OSV, OVS, VSO, VOS, SV, VS**) which shows that Albanian is a "scrambling" language. The "scrambling" phenomenon, which was first defined as a free word order, is seen as the result of completely optional movement operations by some linguists (such as: Chomsky, Rizzi, Bošković) and as the result of certain compulsory movement actions from some others (like: Gürer; Şener). While in the Albanian language possible changes in the order of words are often considered as variational configuration in order to change or increase the weight of information (Sejdiu-Rugova and Simitçiu 2018; Xhaferi 2005).

Therefore, here it will be argued whether all movement operations, which are realized differently from the general structure of sentences in Albanian, are governed by discourse-based features; and whether there is room for any operation that changes the linear order of the constituents randomly and/or without the inclusion of a discursive feature. So, if word order changes are fully determined by discursive-pragmatic motives through which the structure of information is codified.

To analyze and research all that we said above, three corpora have been selected: *Koha newspaper*; the show **Pressing** (May 30, 2022); the feature film *A Story of Pristina*; and the commentary of the football match *Serbia - Albania* (14.10.2014 Belgrade). These types of texts have been selected on purpose, as they are built on different styles that have different functions, and they also differ from the speaker's point of view, which together with style can be considered as two external factors that play a role in the codification of information structure.

During the research, descriptive, statistical (in the third chapter), analytical, comparative and direct observation methods are also used.

Key words: information structure, topic, focus, TopP, FocP, F-marking, clitic doubling, variational configuration of words in the sentence