



Universiteit
Leiden
The Netherlands

Word order, information structure and agreement in Teke-Kukuya Li, Z.

Citation

Li, Z. (2024, September 5). *Word order, information structure and agreement in Teke-Kukuya*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/4054947>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4054947>

Note: To cite this publication please use the final published version (if applicable).

CHAPTER 1

Introduction

This chapter first introduces the background and research questions of the thesis. Then I provide the geographic and demographic information of the Teke-Kukuya language, and I also introduce the people who speak the language. The methodology for fieldwork is discussed, and the conventions in the presentation of the data are also mentioned. The overview of each remaining chapter and the scope of the thesis are given in the last section.

1.1 Research questions

Earlier studies on the grammar of Bantu languages tend to focus on the noun class system and verbal morphology. In recent years, syntax has been a growing field in the studies of Bantu languages and the importance of information structure for the analysis of Bantu morphosyntax has been highlighted (van der Wal 2015; Downing and Marten 2019). Information structure (or information packaging) is the way in which speakers structure

2 *Word order, information structure and agreement in Teke-Kukuya*

the speech information they give so that the information can be more easily adapted into the current knowledge of the addressees. Different languages make use of various linguistic strategies for the sake of expressing what is given/known, what is new and/or what is contrasted. In many recently published grammars of Bantu languages, the expressions of information structure of the language have been described in separate chapters, e.g. Crane et al. (2011) and Guérois (2015). Information structure can affect many aspects of the morphosyntax of Bantu languages. The agglutinative verbal morphology such as subject/object marking and conjoint/disjoint alternation in Bantu is always on the boundary with syntax, which is often affected by information packaging. In addition to verbal inflection, word order variation is also commonly associated with discourse functions. In many Bantu languages, focused information and given information can (or tends to) occur in some particular positions in the clause or verb phrase, which is known as topic fronting and the use of dedicated focus positions. The studies on the interaction between syntax and information structure in Bantu have also influenced some theoretical debates on syntactic phenomena such as movement and agreement.

As for dedicated focus positions, the use of an immediate-after-verb (IAV) focus position has been reported for example in Aghem (Watters 1979), Bemba (Costa and Kula 2008), Makhuwa (van der Wal 2009), Matengo (Yoneda 2011) and Zulu (Buell 2006, 2009); and some formal analyses on the structural position of the IAV element have been proposed (Buell 2006, 2009; van der Wal 2006; Cheng and Downing 2012). A sentence-final focus position is attested in Kirundi (Sabimana 1986; Ndayiragije 1999) and Kinyarwanda (Gibson et al. 2017). Notably, there are also some West-Coastal Bantu languages such as Mbuun (Bostoen and Mundeke 2012), Kisikongo (De Kind 2014; De Kind et al. 2015) and Nsong (Koni Muluwa and Bostoen 2014) that have been claimed to have an immediate-before-verb (IBV) focus position. However, the exact formal and functional properties of the IBV focus position and its historical origin still merit more detailed research. The aim of this research is to concentrate on one understudied language in the West-Coastal area, Teke-Kukuya, which also displays this IBV focus position, investigating its functions and origin as well as providing a structural analysis of the IBV focus construction, in order to better understand this focus strategy as an areal feature in a broader sense.

This study is part of the Bantu Syntax and Information Structure (BaSIS) research project¹ which systematically studies the expression of information structure and its influence on nominal licensing in selected Bantu languages. Teke was chosen for the project and for my research, since in previous research the IBV focus position and an associated intriguing class 1 subject marking alternation are reported in some closely related West-Coastal Bantu languages such as Mbuun and Kisikongo, so at the start of the project we were curious about whether similar phenomena can be attested in Teke and how syntax and information structure interact in the presence of the IBV focus strategy. The existence of IBV focus in Teke and the class 1 subject agreement alternation were borne out at the very beginning of my first fieldwork period with elicitation on focus expressions such as interrogatives, and were investigated in detail throughout the rest of the fieldwork. Teke is also a yet insufficiently described Bantu language, so the project also aims at contributing to the description of the language.

Therefore, the goals in this thesis are first to provide a brief grammar description of the language, as well as to address the research questions mentioned above. The research focuses on the expression of information structure in Teke, paying particular attention to the word order variation and the use of the IBV focus strategy. In the major part of the thesis, I investigate the functions of the IBV focus construction, its diachronic development and its structural representation. I also provide a plausible explanation on the class 1 subject agreement alternation that conspires with IBV focus.

1.2 Teke-Kukuya: language and people

Teke is a group of Bantu languages which are spoken in Congo, DRC and Gabon, and is coded as B70 in Guthrie (1948)'s classification. Teke is a Kongo exonym for “the populations who live on the plateaus north of

¹My research was mainly supported by the China Scholarship Council and was also partly funded by the BaSIS project which was supported by NWO Vidi grant 276-78-001.

4 Word order, information structure and agreement in Teke-Kukuya

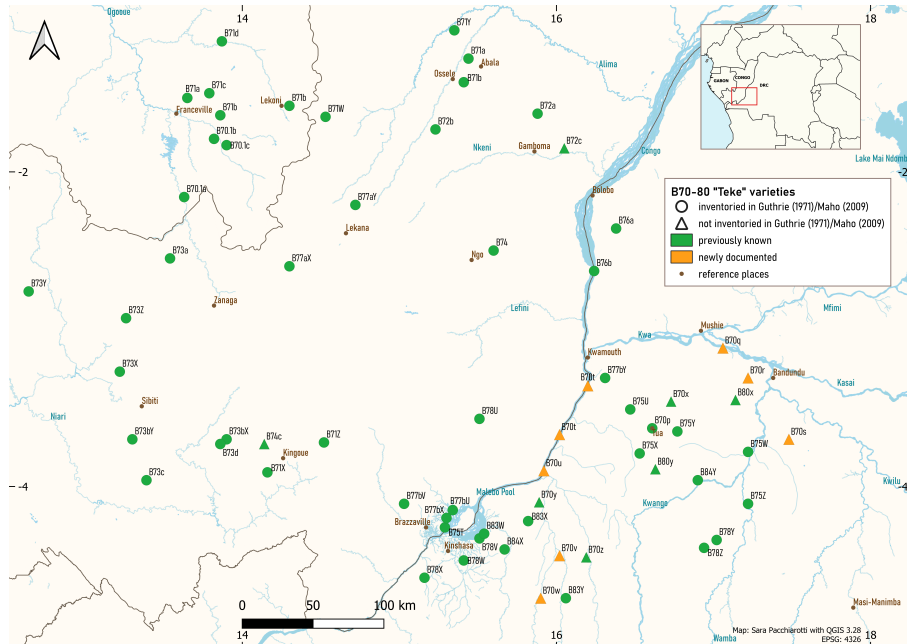


Figure 1.1: Map of the Plateaux region and location of Teke varieties (Kouarata et al. 2023)

Malebo Pool on both sides of the Congo River as far north as the mouth of Nkeni” (Vansina 1966: 102). Teke people usually name their subgroups after the plateau where they live (Pacchiarotti et al. 2019: 16). Since the internal classification of Teke and the labeling of its different varieties still meet a lot of confusion, here I do not list all the Teke varieties in the literature but I refer the readers to Pacchiarotti et al. (2019) and Kouarata et al. (2023) on the internal classification of Teke. A map of the geographic distribution of the known Teke varieties is illustrated in Figure 1.1 (Kouarata et al. 2023), which includes some newly documented varieties/dialects.

I chose the Kukuya variety which is coded as B77a (including B77aX and B77aY in the map above) for my research, because it shows less phonological reduction compared to other Teke varieties, and there has been a detailed study on its phonology (Paulian 1975) and some papers and notes by the same author on this language (Paulian 1997, 1998, 2001), which

allowed me to get familiarised with the language much more easily at the start of my research. During my fieldwork I also found that the speakers are quite aware of how Kukuya differs from other Teke languages on the grammatical features. In the rest of the thesis, I only use “Kukuya” to refer to the Teke-Kukuya language. In my description and analyses on the grammar of Kukuya, I don’t imply that the same holds for other Teke varieties but only make it explicit where I have solid data to prove.

The Kukuya language is mainly spoken around the Lékana district of the Plateau Department in Congo, to the east of the Lékéti river. The language was also referred to as Kukwa, Kikukuya, Koukouya, Kukwa, Kukwa or Küküia in some literature. Kukuya was reported to have about 38,800 native speakers according to Ethnologue (2000) (900,000 speakers of all Teke languages in Congo), but I don’t have a solid number of its current speakers. In Congo, French is the official language of the country, while Lingala and Kituba serve as the *lingua franca* in the northern and southern parts of the country respectively. According to my observation, each person that I made acquaintance with in Congo, no matter in the village or in urban areas, can speak at least three languages, namely French, Lingala or/and Kituba, and their native indigenous languages.

Kukuya is still used actively in daily life in the Lékana district, which is the primary communication language among its community members. It is currently considered as a language in vigorous status (6a*) according to Ethnologue (2024). However, there is an obvious tendency that the younger speakers are losing vocabulary and grammatical knowledge of the language. When I checked some words that are listed in Paulian’s (1975) work, younger speakers were less familiar with many of the words. In Brazzaville, young people are often reluctant to show that they are Teke speakers. Some speakers who have lived in Brazzaville for years also tend to mix the indigenous languages that they know, many words that I consulted with them turned out to be Lingala or Kikongo. So apparently there is risk for Kukuya to be used less and less in the future. The SIL Congo branch has made much effort on documenting the Congolese indigenous languages including Teke, and has worked on and published several orthography guide books, dictionaries, textbooks with exercises on various Teke varieties, but until now there has not been any textbook or grammar on Kukuya yet. For missionary

purposes, three Gospels of the New Testament were translated into Kukuya as “Mandaka Mabve ma Yezu Kristo” by Pierre Loubier and published in June 1979. The book is a good reference on what the language looked like a half century ago, but the transcription was rough and is not always reliable.

There were also some very detailed anthropological works in the 1970s-1980s on the history and society of the Kukuya people; to concentrate on the linguistic part of the thesis I don't present and discuss these here. I refer the readers to the studies of Bonnafé (1973, 1979, 1987).

1.3 Methodology and data

The data presented in this thesis are based on my two fieldwork periods in Lékana and Brazzaville in the Republic of Congo, from May until the end of August 2019, and from April until end of August 2021. The time I spent in Brazzaville and Lékana is almost equal. In Lékana I worked intensively with three local speakers and kept communicating with people who live in the village; in Brazzaville I worked with speakers who were all born in Lékana and were brought up there. My main language of communication in these sites was French. During these periods I collected almost 4000 sentences with grammaticality judgements and I transcribed 6 stories with the help of four main speakers who speak Kukuya as their first language and can speak French very fluently.

Gilbert Mbou was my main consultant in Lékana, with whom I worked for the most of the time in the summer of 2019 as well as in May 2021. He was born in 1967 and was brought up in Lékana where he spent all his previous years. He lives as a farmer and has a family with four children. Gabriel Ntsiebele who was born in the 1950s is his best friend in the village and they had known each other since their childhood. I worked with the two of them every evening after they finish their work in the field. Gilbert also often drove me out to see the village and visit the community members with his motorbike and often treated me with hearty dinners. Zacharie Ngouloubi (1949) and Alain Mbiambourou (1967) were the ones who I worked with

everyday in Brazzaville in July and August 2021. They were introduced to me by the spokesperson of the King of Teke. We worked in a café in the center of Brazzaville with beer and peanuts in the afternoons. They were both born and brought up in Lékana and have a very good sense of the Kukuya language.

My corpus consists of two different types of data: elicited and spontaneous data. For writing the grammar sketch part, I relied much on the elicited data. I checked different grammatical features of the language by asking the speakers to translate French sentences into Kukuya. With the help of the methodological guide “The BaSIS basics of information structure” (Van der Wal 2021) which was developed as part of the BaSIS project, I have collected both elicited and spontaneous data during the fieldwork, but since the marked information-structural expressions such as the use of the dedicated focus position are more obviously attested in elicited sentences, the examples presented in this thesis are largely extracted from the elicited data. I often get generalisations from the speakers’ judgement on the grammaticality and appropriateness of the elicited sentences when I intentionally change the word order and agreement patterns. I also included data from natural speech of different genres such as dialogues and storytelling, but I leave more elaborated investigation on the information structure of spontaneous speech for further research. Moreover, since the speakers are all literate, I also made some written tasks such as question-answer pairs, idioms and letters for them to write down, especially for Zacharie and Alain. All the data mentioned above are transcribed and stored in the OLD database which is accessed via the platform Dative, and will be archived in the Language Archive.

1.4 Outline of the thesis

The thesis is organised as two main parts. The first part is a grammar sketch of the Teke-Kukuya language (chapter 2), and the second part discusses word order variation and expressions of information structure of this language, with particular interests in a dedicated immediate-before-verb

8 *Word order, information structure and agreement in Teke-Kukuya*

(IBV) focus strategy in this language (chapter 3). I provide a diachronic account (chapter 4) as well as a synchronic analysis (chapter 5) of this IBV focus strategy, also discussing the agreement mechanism of the class 1 subject marking alternation that is associated with the syntax of IBV focus.

The grammar sketch in chapter 2 covers the basic grammatical properties of the Teke-Kukuya language, including its segmental phonology and prosodic system, noun classes and noun phrases, verbal morphology and TAM conjugations, as well as some syntactic issues such as complementisers and embedded clauses. The aim of this part is to provide a description on the structure of this language in theoretically neutral terms, for the readers to refer to while reading the thesis and to better understand the other chapters. It also allows readers who are interested in the typology of (West-Coastal) Bantu languages or the other varieties of Teke to use the data to conduct comparative research. Since there is no previous reference grammar of the Kukuya language, this part also contributes to the documentation of the language for the community.

The second part of the thesis consists of three chapters. Chapter 3 presents the influence of information structure on the word order of the language. Some notions of information structure are defined in this chapter. I describe different types of focus and topic expressions, showing how they deviate from the canonical SVO order. The main aim of this chapter is to illustrate the function and interpretation of the IBV focus position which is available for different kinds of argument and adjunct focus. In the rest of the chapter I also introduce different types of topical elements in the pre-verbal domain, functional passive constructions, and cleft constructions for focus expressions. This chapter also allows the readers who are interested in the expressions of information structure of (Bantu) languages to learn about the specific properties of Teke-Kukuya and the IBV focus position for comparative work.

In chapter 4, I investigate the diachronic origin of the IBV focus strategy. I compare the IBV focus construction with clefts, showing that they share many grammatical properties, both segmentally and tonally. I make the hypothesis that the IBV focus originates from a basic cleft, and it has been grammaticalised towards a monoclausal focus construction. I propose

a grammaticalisation pathway of the IBV focus strategy, and introduce its distribution as an areal innovation in some other West-Coastal Bantu languages. This chapter may also offer inspiration to the studies on the IBV focus position in other Bantu languages, especially with the solid tonal data on the connection between IBV focus and relative constructions, which was not taken into account in previous research.

Chapter 5 aims to provide analyses on the structural representation of the IBV focus construction and on the class 1 subject marking alternation. I first illustrate the structural representation of the canonical SVO order and look into the structural position of the IBV focused element. To account for the class 1 subject marking alternation, I start by discussing the subject agreement asymmetry in subject and non-subject relatives, and I propose that the class 1 subject marking allomorphy in IBV focus is connected to the agreement with a clause-internal ϕ P that is co-referential with an initial base-generated topic. This chapter contributes to a structural analysis of the IBV focus strategy in Bantu.

Chapter 6 concludes this thesis and highlights some further research questions. In the appendix I include one glossed and translated Kukuya texts, which is a dialogue between husband and wife.

