Linguistique et langues africaines

9(1) | 2023
Special issue: Bantu universals and variation

Bantu word order between discourse and syntactic relations

Elisabeth J. Kerr, Allen Asiimwe, Patrick Kanampiu, Zhen Li, Ernest Nshemezimana and Jenneke van der Wal

OpenEdition Journals

Electronic version
URL: https://journals.openedition.org/lla/9496
DOI: 10.4000/lla.9496
ISSN: 2822-7468

Publisher
Llacan UMR 8135 CNRS/Inalco

Electronic reference
Elisabeth J. Kerr, Allen Asiimwe, Patrick Kanampiu, Zhen Li, Ernest Nshemezimana and Jenneke van der Wal, "Bantu word order between discourse and syntactic relations", *Linguistique et langues africaines* [Online], 9(1) | 2023, Online since 30 June 2023, connection on 21 July 2023. URL: http://journals.openedition.org/lla/9496 ; DOI: https://doi.org/10.4000/lla.9496

Bantu word order between discourse and syntactic relations

Elisabeth J. Kerr
Leiden University

Allen Asiimwe
Makerere University

Patrick Kanampiu
University of Edinburgh

Zhen Li
Leiden University

Ernest Nshemezimana
University of Burundi

Jenneke van der Wal
Leiden University

Abstract

Discourse function has often been noticed to be a strong factor in conditioning Bantu word order. The importance of discourse function for determining the word order of Bantu languages is visible for example in locative inversion and dedicated focus positions. As a result of such phenomena, it has been proposed that Bantu word order is best captured by reference to discourse roles, e.g. Topic-VerbNontopic. Nevertheless, we typically see statements describing Bantu word order in relation to grammatical roles (e.g. "SVO"), and the notions "subject" and "object" remain core in analyses of Bantu. In this paper we present the result of a study reconsidering Bantu word order from a discourse-configurational perspective, asking how far we can get without reference to grammatical roles. We use a parametric approach to investigate this syntactic variation, presenting new discourse-oriented field data collected on 9 Bantu languages. We show how these parameters highlight variation within the family, with each language sitting at a different point on a continuum between grammatical role-oriented and discourse role-oriented. We therefore argue against a one-size-fits-all account of Bantu word order and advocate for approaches that include both grammatical and discourse roles.

Keywords

Bantu, discourse configurationality, parameters, syntax, word order

Résumé

La fonction discursive a souvent été identifiée comme un facteur important dans le conditionnement de l'ordre des mots en bantou. L'importance de la fonction discursive dans la détermination de l'ordre des mots des langues bantoues est visible par exemple dans l'inversion locative et les positions du focus. À la suite de tels phénomènes, il a été proposé que l'ordre des mots en bantou est mieux re-

présenté par les rôles discursifs, par ex. Topique-Verbe-Non-topique. Néanmoins, nous observons généralement des déclarations décrivant l'ordre des mots bantou par rapport aux rôles grammaticaux (par exemple « SVO »), et les notions de « sujet » et « objet » restent au cœur des analyses du bantou. Dans cet article, nous présentons les résultats d'une étude reconsidérant l'ordre des mots bantou dans une perspective des configurations discursives, en nous demandant jusqu'où nous pouvons aller sans référence aux rôles grammaticaux. Nous utilisons une approche paramétrique pour étudier cette variation syntaxique, en présentant de nouvelles données de terrain, axées sur le discours, recueillies sur 9 langues bantoues. Nous montrons comment ces paramètres mettent en évidence la variation au sein de la famille, chaque langue se situant à un point différent sur un continuum entre les rôles grammaticaux et les rôles discursifs. Nous nous opposons donc à une description unique de l'ordre des mots bantou et préconisons des approches qui incluent à la fois les rôles grammaticaux et discursifs.

Mots clés
Bantou, configurations discursives, ordre des mots, paramètres, syntaxe

1. Introduction

Discourse function has often been noticed to be a strong factor in conditioning Bantu word order (Van der Wal 2015a; Downing & Hyman 2016; Downing & Marten 2019). Core concepts of discourse function are topic, defined as what the sentence is about, and focus, most simply defined as new information added to the presupposition (Lambrecht 1994).[1] The importance of discourse function for determining the word order of Bantu languages is visible for example in locative inversion, which presents a topical element preverbally and a non-topical logical subject postverbally (1), and dedicated focus positions —immediate-before-verb (IBV) in (2a-b).

(1) Context: What has happened on the bridge?
A-ha *ru-tindo* *ha-a-raba=hó* *e-motoka* *ny-îngi.*
AUG-16 11-bridge 16SM-N.PST-pass=16 AUG-9.car 9-many
'On the bridge have passed many cars.' [Rukiga]

(2) a. Context: Who attacked the hunter?
Mu-bhii *kí-mbúlí* *kí-siim-i.*
1-hunter 7-lion 7SM-attack-PST
'[The lion]$_{FOC}$ attacked the hunter.'
b. Context: Who did the lion attack?
Ki-mbúli *mú-bhii* *kí-siim-i.*
7-lion 1-hunter 7SM-attack-PST
'The lion attacked [the hunter]$_{FOC}$' [Kukuya]

As a result of such data, it has been proposed that Bantu word order is best captured by reference to discourse roles, for example as Topic-Verb-Nontopic (see, e.g., Good 2010 for Naki; Yoneda 2011 for Matengo; and Morimoto 2000; Morimoto 2006 for Bantu discourse configurationality in general). Nevertheless, we typically see statements such as "The default order of sentence constituents across Bantu is S (Aux) VO (Adjuncts)" (Nurse & Philippson 2003: 9), and these traditional syntactic roles of "subject" and "object" continue to shape our descriptions and analyses. We thus find two factors

1. More specifically, we take focus to trigger a set of relevant alternatives (Rooth 1985; 1992), with information focus differing from exhaustive focus in that exhaustive focus excludes all other possible alternatives. We use question-answer congruence tests and modification by the exhaustive particle 'only' as tests for focus. While we also collected data on subtypes of focus, such as corrective and contrastive focus, we use the general term "focus" in this paper as evidence for discourse configurationality, leaving more detailed discussion to future work.

in the literature that are said to influence or determine word order in Bantu languages: discourse roles on the one hand, and grammatical roles on the other.

This discussion of Bantu fits into a larger debate about word order and types of languages: Li & Thompson (1976) proposed topic-prominent vs. subject-prominent languages, Hale (1983) suggested that languages can be non-configurational, and Mithun (1987) describes pragmatically based word order, already familiar from the Prague School (Sgall *et al.* 1973; Hajičová 1983; Sgall 1985) and later termed "discourse configurational" (É. Kiss 1995). Over the years the debate has shifted from asking "Which type does language L belong to?" to arguing against a simple binary classification, instead asking "To what extent is word order in language L determined by discourse roles and to what extent by grammatical roles?". For example, Öhl (2010: 231) notes about discourse configurationality vs. "relation configurationality" that "there is no real sharp division between these two classes. Almost every language can be said to have either property to at least some degree", and Morimoto (2009) similarly argues that many Bantu languages are in a transitional stage between topic-based and subject-based agreement systems, being neither fully topic-based nor subject-based.

Given this background, in this paper we want to investigate the relative influence of discourse prominence on word order, and how this varies across Bantu languages. As our starting point, we take a language's word order to be discourse-prominent if it shows a dedicated position for topic and/or focus —modifying É. Kiss's (1995: 6) definition of a discourse-configurational language.[2] In order to determine whether such dedicated positions for topic and focus exist, we established parameters for focus and parameters for topic, with various sub-questions. This is explained in further detail in Section 2 below. By concentrating on the discourse-configurational properties, we essentially ask ourselves how far we can get in describing Bantu word order without reference to grammatical roles. As a result, we can pinpoint more precisely where we *do* need them. We essentially explore the fact that "flexible" word order seen from the point of view of grammatical roles may instead turn out to be quite *in*flexible from the discourse-configurational perspective. Our findings show that different Bantu languages place varying levels of importance on discourse vs. grammatical roles, with no language fully dependent on one or the other. We discuss this continuum of variation in Section 4, arguing that models of Bantu word order must take (at least) both these factors into account, and that discourse roles are for some Bantu languages in fact more important than grammatical roles in determining word order.

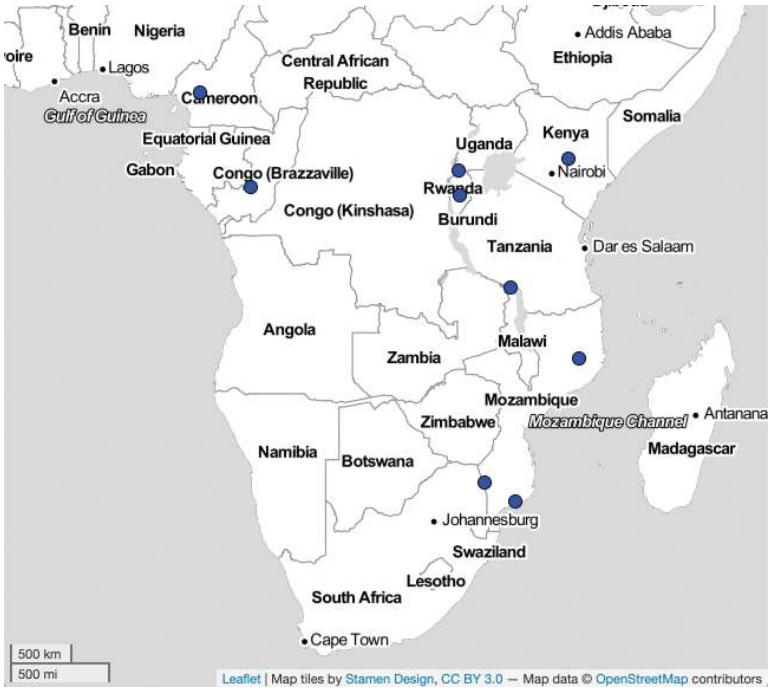
In Section 2 we introduce the languages of the sample, the parameters chosen to investigate the extent of influence of discourse and grammatical roles on word order, and the parameter values found. In Section 3 we make observations based on the parameter values, in Section 4 we reflect on the methodology, findings, and implications for theoretical models, and in Section 5 we conclude. As supplementary material, we provide all the data on which the parameter values were coded, together with a copy of the parameter list with further explanatory notes.

2. Methodology

In order to determine the influence of discourse roles on word order in a given language, detailed data on information structure are required. Unfortunately, such data are often not available from existing grammatical descriptions of Bantu languages, which frequently provide words or sentences in isolation from their discourse context. As part of the Bantu Syntax and Information Structure (BaSIS) project we therefore gathered new fieldwork data focusing on syntax and information structure using the project's own methodology,[3] which is based on (i) the Questionnaire on Information Structure (QUIS) (Skopeteas *et al.* 2006), (ii) the diagnostics in Van der Wal (2016), and (iii) collection of spontaneous speech of various genres. We conducted fieldwork studies on the following 9 languages: Tunen (A44, Cameroon), Teke-Kukuya (B77, Republic of Congo), Kîîtharaka (E54, Kenya), Kirundi

2. The definition of discourse configurationality given by É. Kiss (1995) specifically takes a structural position for topic and/or focus as a necessary characteristic. In Surányi's (2016) overview of discourse configurationality he makes a point of distinguishing discourse configurationality, based on structural positions, from the wider notion of discourse prominence. In the current paper we can only work with discourse prominence, as we do not have the test results for each language to establish whether a dedicated linear position also corresponds to a dedicated structural position.

3. The BaSIS project methodology is provided as part of the supplementary materials.



Co-ordinates from Glottolog (Hammarström *et al.* 2022); plotted in R using lingtypology package (Moroz 2017).

Figure 1 — Map of the 9 languages in our sample

(JD62, Burundi), Rukiga (JE14, Uganda), Kinyakyusa (M31, Tanzania), Makhuwa-Enahara (P31, Mozambique), Changana (S54, Mozambique) and Copi (S61, Mozambique).[4] The map in Figure 1 below shows the approximate location of these languages.

As we can see from Figure 1, a large geographical spread is represented, but there is a skew towards languages of the East. This is because we used a convenience sample —aside from the two languages of the North-West which were investigated by the PhD candidates on the project, the languages were chosen due to the location of the African-based collaborators and therefore the ability to conduct collaborative fieldwork. Moreover, the sample is too small in size to itself show any strong typological generalisations, and is therefore intended only as a starting point for investigating variation in Bantu languages. In Sections 3 and 4 we show how the use of these parameters can guide follow-up research questions for discourse and syntactic relations in other Bantu languages, which we hope will result in a more complete picture of the universals and variation in the family.

We drew up parameters investigating focus and topic prominence, with sub-parameters that serve as checkpoints to set the parameter, based on factors taken from the literature on discourse configurationality. This parametric set-up was inspired by previous parametric work on Bantu languages conducted under the SOAS, University of London project "Morphosyntactic variation in Bantu" (see e.g. Guérois *et al.* 2017). The parameters were originally set up to give a binary value of yes (Y) or no (N), although we allowed for intermediary coding where a binary coding was not possible. The parameters we developed evolved during the research project, as we discussed how well the parameters captured the variation after the initial coding and then revised the parameters accordingly, which helped resolve some of the unclear coding issues.

Once the parameter list was made and revised, we added some guidance on the purpose of the parameters. For example, notes are added below Question 1.3. ("Can the external and internal argument be focused in the same position?") to show how positive coding provides evidence for information structure (focus) being more important for word order than grammatical roles, while negative coding

4. Following Bantuist tradition, Guthrie classifications are given alongside language names and locations, following the updated classifications in Maho (2003; 2009). These classifications indicate geographical groupings.

1. Is there a dedicated focus position?

1.1. Do internal arguments need to be in a particular position when focused?

Definition: a linear position in which elements are focused (and possibly focus can be elsewhere in the clause).

Checkpoints:

a. Can the recipient and the theme be questioned in their canonical position?

[Example: if a focus-neutral ditransitive has V REC TH order, can we ask both "give who book?" and "give teacher what?"]

b. Can the recipient and the theme be an answer to a content question in their canonical positions?

[Example: if a focus neutral ditransitive has V REC TH order, can the questions "who did you give a hat?" and "what did you give Jane?" both be answered as "give Jane hat"?)

c. Can the recipient and the theme be modified by "only" in their canonical positions?

→Answering "yes" suggests that there is no dedicated focus position

1.2. Do adverbs need to be in a particular position when focused?

Checkpoints:

d. Can non-arguments be questioned in their canonical position?

[Example: if you normally say "they played football in the park", can you ask "they played football where?")

e. Can non-arguments be the answer to a content in their canonical position?

f. Can non-arguments that are modified by "only" occur in their canonical position?

→Answering "yes" suggests that there is no dedicated focus position

1.3. Can the external and internal argument be focused in the same position?

Checkpoints:

g. Can the external argument (EA) be questioned in the same position as internal arguments (IAs)?

[Example 1: IAs can be questioned in immediate postverbal position, and the EA can also be questioned postverbally in a VS subject inversion construction.

Example 2: IAs are questioned in immediate before verb position or in a cleft, and the EA can also be focused in those two ways]

h. Can the EA as the answer to a content question occupy the same position as the IA in the answer to a wh-question?

i. Can the EA modified by "only" occur in the same position(s) as the IA with "only"?

→Answering "yes" suggests that external and internal argument behave similarly

→Answering "no" suggests that there is a subject-object asymmetry for focus

2. Is there a dedicated topic position?

Definition: a linear position in which elements are topical (and possibly topics can be elsewhere in the clause). A topical subject is defined here as the subject in a categorical sentence with predicate focus (Lambrecht 1994).

Checkpoints (for otherwise unmarked clauses, i.e. no cleft):

a. Can a subject* interrogative phrase occur in the same position as a topical subject?

b. Can the subject as answer to a subject content question occur in the same position as a topical subject?

c. Can a subject modified by "only" occur in the same position as a topical subject?

d. Can a subject in a thetic sentence occur in the same position as a subject in a categorical sentence? (check for intransitive as well as transitive)

→Answering "yes" suggests that there is no dedicated topic position

e. Are topical arguments typically or preferably expressed in a preverbal position? (think of left-dislocation and locative inversion)

→Answering "yes" suggests that there is a preverbal topic position

* "Subject" here is understood as the single argument of an intransitive predicate, or the agent of a transitive predicate – not as the argument triggering subject marking.

Figure 2 — BaSIS parameter checklist

suggests that there is a subject/object asymmetry with respect to focus (and therefore gives evidence for grammatical roles conditioning word order). We also added definitions for what we mean by a dedicated focus or topic position, as there are two possible interpretations of what such a position is: (i) a linear position in which elements are focused/topical, or (ii) a linear position in which focused/topical elements must occur. We specified the former version, using a definition based on É. Kiss (1995) (see Section 3.2.1 for discussion). These notes were added to help prevent different people interpreting the terminology differently when coding the parameter values, as is always a risk when multiple authors conduct the survey. The resultant parameter list is given in Figure 2.

Having drawn up and revised the parameters, we each re-checked the coding for our individual languages, then paired up to check each other's parameter codings on the basis of the provided examples. There were no cases of direct disagreements in coding but there were cases in which the coding was less clear, as noted in the Appendix. The results for these parameter codings are given in Table 1; see the Appendix for the table with cross-referencing to all the supporting data and further notes where applicable in order to explain why a given parameter coding was decided upon. The colour coding in the table represents the extent to which a language shows evidence for grammatical role conditioning word order (indicated by yellow shading) and the extent to which discourse role conditions word order (indicated by blue shading). For each language we can therefore see a colour profile giving an approximate overview of the variation, and we can also read the codings horizontally to identify interesting areas for comparisons across languages.

Table 1 — Parameter coding for the BaSIS sample.

Parameter	Checkpoint	DR	GR	Tunen A44	Kukuya B77	Tharaka E54	Kirundi JD62	Rukiga JE14	Kinyakyusa M31	Makhuwa P31	Copi S61	Changana S53
1	a.	N	Y	N	N	Y	N	N	Y	N	Y	Y
	b.	N	Y	Y/N	Y	Y	N	Y	Y	N	Y	Y
	c.	N	Y	N	Y	Y	N	Y	Y	Y/N	Y	Y
	d.	N	Y	Y	Y/N	Y	Y/N	Y/N	Y	N	Y?	Y/N
	e.	N	Y	Y	Y	Y	Y	Y	Y	N	Y	-
	f.	N	Y	Y	Y	Y	Y	Y	Y	N	Y	-
	g.	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y
	h.	Y	N	N	Y	Y	Y	Y	Y	N	Y	Y?
	i.	Y	N	N	Y?	Y	Y	Y	Y	N	Y	Y
2	a.	N	Y	Y	Y*	N	N	N	Y?	N	N	N
	b.	N	Y	N	Y*	N	N	N	Y	N	N	N
	c.	N	Y	Y	Y*	N	N	N	Y	N	N?	N
	d.	N	Y	Y	Y	Y?	Y	Y	Y	N	-	-
	e.	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y

Blue shading: evidence for discourse roles (DR) conditioning word order; yellow shading: evidence for grammatical role (GR) conditioning word order; green shading: = mixed behaviour. A column for DR and GR is given to represent what the parameter settings would be for a purely DR/GR-conditioned language.

As can be seen in Table 1, there are still some cases for which no data was found (indicated by "-") or for which the parameter value is unclear (indicated by "?" for uncertainty or "Y/N" for mixed behaviour). The lack of data is to do with parameter checkpoints being added after the end of data collection, as we will discuss in Section 3. In other cases, the languages showed mixed behaviour, as discussed in the Appendix in the notes for each provided example. We discuss the use of the asterisk "*" for Kukuya in Section 3.2.1. We will evaluate our methodology further in Section 4, but first, we will consider in Section 3 what we can observe about Bantu word order based on the parameter values in Table 1.

3. Observations

The process of investigating the parameter values, comparing the results, and conducting further research led us to observations on three main topics: (i) the degree of crosslinguistic variation found, (ii) the evidence for and against focus prominence (Parameter 1), and (iii) the evidence for and against topic prominence (Parameter 2). We will go through these in turn.

3.1 Degree of crosslinguistic variation

As is clear from a first glance at Table 1, there is a lot of variation between the languages in our sample —no language patterns exactly the same as another with respect to these word order parameters. We should therefore avoid making claims of word order in "the Bantu languages" as if they were a monolithic entity. Such an observation also matches results from previous research which found variation between Bantu languages in terms of reliance on discourse vs. syntactic relations (e.g. Morimoto 2000). We expect that the study of additional Bantu languages would uncover further patterns of variation.

That being said, as this project only uses a sample of 9 different Bantu languages out of approximately 555 languages in the family (Hammarström 2019), further research is needed in order to make reliable generalisations. The study of more languages may for example show areal patterns conditioning some of the variation found, as we suggest for preverbal focus in Section 3.2.2. We may also find implicational relations between (sub-)parameters. For example, all languages had the same values for Parameters 1e and 1f —this may mean that a language in which a focused non-argument can appear in its canonical position as answer to a content question can always be modified by the exhaustive 'only' in that same position. If such implicational relations are proven to hold, there would be motivation to collapse those checkpoints into a single sub-parameter, as we could conclude that new information focus and exhaustive focus for non-arguments are always expressed the same when it concerns word order and only one answer would therefore be needed to check that sub-parameter in future field research.

Furthermore, although there is variation between languages, we can see that there is no language on the extreme, i.e. there is no language in our sample where word order is fully conditioned by grammatical role or fully conditioned by discourse role. This fits the expectation from prior research discussed in Section 1, namely that Bantu word order shows a mixture of grammatical and discourse roles. In this subsection we will consider the extent of the variation.

3.1.1 Most grammatical role-oriented languages

The languages which had the most parameter values indicating influence of grammatical role on word order were Tunen and Kinyakyusa. In Section 3.2.2 we discuss how the result for Kinyakyusa appears to be linked to a possible innovation of preverbal focus. In this section we discuss Tunen, which appears to be the most grammatical role-oriented language.

Tunen being at the extreme is not surprising given that Tunen has been previously noted to show unusual syntax for a Bantu language, displaying SOV word order as opposed to the SVO order described across other Bantu (Dugast 1971; Bearth 2003; Mous 1997; 2003; 2005). In our study it only displayed a few sub-parameters providing evidence for discourse role as the determining factor, with all other sub-parameters providing evidence for grammatical role. The parameter values reflect differences between external and internal arguments with respect to focus marking and the ability for non-topical elements to occur preverbally. For instance, focused external arguments must be clefted (3), while focused internal arguments may be in-situ (4a) or clefted (4b) (Parameter 1h).

(3) Context: Who shut the door?

a. *(*Á*) *Píél ána nikí kwiyí.*

á	Piélə	á-ná	nɛ-kí	kuiyí
FOC	1.Pierre	1SM.REL-PAST2	5-door	shut

'Pierre shut the door.' [Tunen]

b. **Píél á ána nikí kwiyí.*

(4) Context: What is the man holding?

a. *Məndɔ aná kalɔ́tɔ itíə́ ɔ məkat.*

mɔ-ndɔ	a-ná	kalɔ́tɔ	itíə́	ɔ	mɔ-kátá
1-person	1SM-PST2	7.carrot	hold	PREP	3-hand

'The man is holding a carrot in his hand.'

b. *Kalɔ́tɔ á məndɔ ána itíə́ ɔ məkat.*

kalɔ́tɔ	á	mɔ-ndɔ	á-ná	itíə́	ɔ	mɔ-kátá
7.carrot	FOC	1-person	1SM.REL-PST2	hold	PREP	3-hand

'The man is holding a carrot in his hand / It's a carrot that the man is holding in his hand.' [Tunen]

In addition, we can note that the structure of the cleft in (3) is different from (4b) in that the focus marker/copula *á* must precede the external argument but must follow the internal argument. Such a distinction also appears in fragment answers, where non-arguments pattern with internal arguments (see data in Appendix). These results support an account of Tunen using grammatical roles, with evidence for a subject/non-subject asymmetry.

A pertinent question considering Tunen's position in Cameroon close to the proposed Bantu homeland (see e.g. Bostoen 2018; Bostoen & Van de Velde 2019) is whether this word order is innovative or historical. Mous (1997; 2005) argues that the preverbal object position is innovative, arising from a previous stage in which SVO was the dominant word order. If we recast Bantu word order in discourse role terms, we lead to the question as to whether Tunen SOV derived from an earlier discourse-based word order rather than the earlier SVO order generally claimed in the literature on Proto-Bantu (see e.g. Heine & Reh 1984; Claudi 1993). Relevantly, Güldemann (2007) links the position of the object in Benue-Congo languages to information structure, arguing that the order of object and verb can at first vary dependent on the information structure status of the object, with a basic SVO word order co-occurring with a preverbal object position employed for extrafocal objects. From this situation, SOV word order can be grammaticalised, with the preverbal position being generalised across all objects regardless of information status. Reframing the historic word order as discourse-based instead of SVO leads to a possible hypothesis that Tunen SOV word order grammaticalised from an earlier stage in which discourse role was a more significant factor influencing the word order.

In order to investigate such a hypothesis into the diachrony of grammatical role-oriented word order, further research into the Northwestern region is needed, including non-Bantu Niger-Congo languages. The skew of our sample towards the Eastern and Southern Bantu languages limits our ability for detailed investigation. The only other Northwestern Bantu language in our sample, Kukuya, patterns quite differently to Tunen as it has developed an IBV focus position, and therefore shows evidence for discourse prominence conditioning its word order. The two languages do show overlap though, in that they both lack inversion constructions and a morphological passive. Interestingly, Kukuya's IBV focus position is also considered to be innovative, as we will discuss further in Section 3.2.

3.1.2 Most discourse role-oriented languages

Within our sample, Makhuwa patterns towards the discourse role-oriented end of the scale: it restricts the preverbal domain to topics, and focus is expressed in a dedicated position immediately after the verb (Section 3.2.1; Van der Wal 2009). However, 4 of the 14 checkpoints for Makhuwa presented (partial) evidence for grammatical role conditioning word order, as the subject cannot be focused in a postverbal position (see also Van der Wal 2009; 2012). Kirundi similarly has more checkpoints showing evidence for the significance of discourse roles, but non-topical subjects can appear preverbally.

In order to investigate whether there are Bantu languages which are fully discourse role-oriented, with no influence of grammatical role, we applied the parameter list in Figure 2 to a few other languages, using data available in the literature. We first investigated Zulu (S42, South Africa), which has been argued to have a discourse-leaning bias, with both a dedicated (immediate-after-verb/IAV) focus position and a preferential preverbal topic domain (Buell 2005; 2009, among others). For Zulu, 8 out of 9 of the checkpoints for which we could find relevant information showed evidence for discourse roles, indicating that Zulu syntax is indeed largely discourse-oriented and possibly more so than Makhuwa or Kirundi, although the lack of data for the other checkpoints means it is not possible to make a full comparison. The same reliance on discourse roles was identified for Matengo (N13, Tanzania; Yoneda 2011; Van der Wal 2015b), where 7 of the 8 checkpoints for which we found relevant data showed discourse influence. Covering the full range of checkpoints for these (and other) languages will show whether they are truly on the extreme end of the word order continuum and therefore test the extent to which Bantu languages' word order can be understood without reference to grammatical roles.

3.2 Parameter 1: Focus prominence

Our first parameter investigated the extent to which languages show evidence for focus conditioning their word order. There were four main points of interest that appeared from our study, namely (i) variation in dedicated focus positions, (ii) preverbal focus, (iii) thetics, and (iv) question words. We discuss these in turn.

3.2.1 Variation in dedicated focus positions

In our sample we observed that three languages have a dedicated focus position, but all are different: we found an immediate-after-verb (IAV) focus position in Makhuwa, an immediate-before-verb (IBV) focus position in Kukuya, and a sentence-final focus position in Kirundi. We illustrate each for recipient focus, where the question word appears in the focus position, as shown in (5)-(7).

(5) a. *Saárá onthumenlé* ***páni****, ekolár' íile?*
Sara o-n-thum-el-ale pani ekolari ile
1.Sara 1SM-1OM-buy-APPL-PFV.CJ 1.who 9.necklace 9.DEM.DIST
'Who did Sara buy the necklace for?'

b. **Saárá onthumenlé ekolár' íile* ***páni****?* [Makhuwa, IAV]

(6) a. *N-kaaká ma-désu* ***ná*** *ndé ká-wî?*
1-granny 6-bean who 1.PRO 1SM-give.PST
'To whom did grandmother give the beans?'

b. **N-kaaká á-wî ma-désu kukí* ***ná****?*
1-granny 1SM-give.PST 6-bean PREP who
Intended: 'To whom did grandmother give the beans?' [Kukuya, IBV]

(7) a. *U-ø-há-ye* *i-ki-jumbu* ***ndé****?*
2SG.SM-PRS.CJ-give-PFV AUG-7-sweet.potato who
'Who do you give a sweet potato?'
b. **Uhaye* ***ndé*** *ikijumbu?* [Kirundi, sentence-final]

The IAV focus strategy is well attested and described in many Bantu languages, such as Aghem (Watters 1979), Bemba (Costa & Kula 2008), Matengo (Yoneda 2011), Makhuwa (Van der Wal 2009), and Zulu (Buell 2009). The IBV focus position attested in Kukuya in our sample is relatively uncommon, with SVO languages stated to be more likely to have postverbal focus positions (Kidwai 1999; Horvath 1985). The IBV focus strategy found in Kukuya seems to be an areal feature as it is also encountered in languages in the Kwilu-Ngounie group, in the extended Kikongo cluster including some B80 languages (Hadermann 1996; Bostoen & Koni Muluwa 2021), first documented in Mbuun (Bostoen & Mundeke 2012), Kisikongo (De Kind 2014) and Nsong (Koni Muluwa & Bostoen 2014). Kukuya's IBV focus position also means that some checkpoints for Parameter 2 are interpreted differently, which we indicated by asterisks (*) after relevant values in Table 1. The availability of preverbal focus allows topical and focal subjects to both occur in the same linear position in these languages (preverbal), without excluding a dedicated focus position. Finally, the sentence-final focus position is to our knowledge much more rarely found in Bantu and has only been observed in Kirundi (Ndayiragije 1999; Nshemezimana 2016) and Kinyarwanda (Kimenyi 1980; Ngoboka & Zeller 2017).

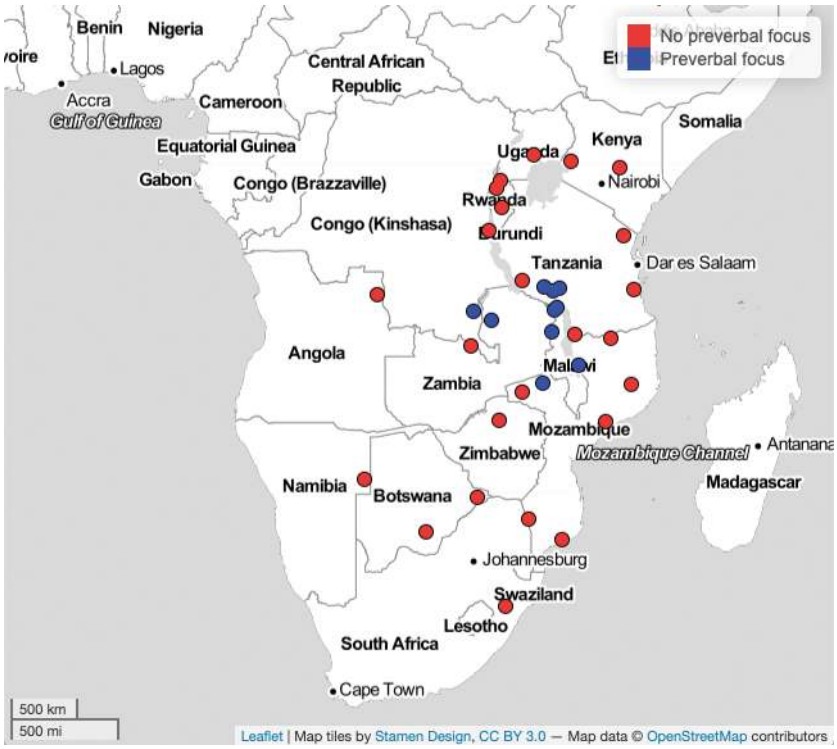
While these focus positions are supported by large amounts of data, there is evidence that they may not always be "dedicated" in the stricter sense of the term. For instance, we see in Kukuya that being focused in the IBV position is just a preference —in some circumstances focal arguments and adverbs can alternatively be placed in postverbal positions (see also Section 3.2.4). This reflects our earlier decision to use É. Kiss's (1995) definition of a focus position as a position in which elements are focused, rather than the stricter definition of a position elements *must* appear in, in order to be focused. In other words, our parameter codings do not show whether other word order patterns are also possible for focused constituents beyond the use of a particular focus position.

In summary, the IBV and sentence-final focus positions are found in a fairly restricted area (languages of Congo/DRC and Kinyarwanda and Kirundi, respectively), while IAV languages have been found throughout the area in which Bantu languages are spoken. It therefore seems most likely that a predecessor of Bantu featured a postverbal focus domain, even if this was not the only position for focus. We have not seen any clear correlation so far between the location of the focus position (IAV, IBV, final) and other parameter settings in the languages, but an interesting follow-up research question would be to investigate what in a language determines where a focus position develops (see also Gibson *et al.* 2017 for correlations with morphological focus marking).

3.2.2 Preverbal focus

While many eastern and southern Bantu languages are known to disallow preverbal focus (Buell 2005; Zerbian 2006; Van der Wal 2009; Yoneda 2011, among others), Kinyakyusa allows preverbal question words and preverbal answers, as shown in (8)-(9), and preverbal modification by 'only', as shown in (10)-(11) for Checkpoints 2a-2c.

(8) ***Juani*** *akuuliile ifilato?*
ju-ani a-ku-ul-il-ile i-fi-lato
1-who 1SM-2SG.OM-buy-APPL-PFV AUG-8-shoe
'Who bought you shoes?'



Co-ordinates from Glottolog (Hammarström *et al.* 2022); plotted in R using lingtypology package (Moroz 2017). The languages in blue pattern with Kinyakyusa in allowing preverbal focus, while those in red do not allow preverbal focus.

Figure 3 — Map showing further research into Parameter 2a on preverbal focus

(9) a. ***(*i)Fiki*** *fisatwiike?*
i-fi-ki fi-satuk-ile
AUG-8-what 8SM-fall-PFV
'What has fallen?'

b. ***Imbwa*** *jasatwike paasi.*
i-mbwa ji-a-satuk-ile pa-asi
AUG-9.dog 9SM-PST-fall-FV 16-down
'The dog fell down.' [Kinyakyusa]

(10) ***Beene abapuuti*** *batikubomba imbombo* (*abangi boosa bikubomba*).
ba-ene a-ba-puuti ba-ti-ku-bomb-a i-mbombo
2-only AUG-2-priest 2SM-NEG-PRS-work-FV AUG-9.job
a-ba-ngi ba-oosa bi-ku-bomb-a
AUG-2-other 2-all 2SM-PRS-work-FV
'Only the priests do not work. All the other people work.'

(11) ***Mweene Pita*** (*jo uju*) *alyendile mwigali*.
mu-eene Pita jo uju a-ali-end-ile mu-i-gali
1-only 1.Peter 1.PRO 1.DEM.PROX 1SM-PST-walk-PFV 18-5-car
'Only Peter (is the one who) came by car.' [Kinyakyusa]

This exceptional parameter setting for Kinyakyusa inspired us to explore Checkpoint 2a more broadly in eastern and southern Bantu languages. Plotting the settings for the various languages for which we found relevant data in the literature on a map revealed an interesting geographical spread, as shown in Figure 3. We see that preverbal focus patterns cluster around the region where Kinyakyusa is spoken, suggesting that preverbal focus may be an areal innovation. Such a hypothesis would be strengthened in further diachronic work investigating whether other phenomena in the languages of this region can be shown to be innovative (see Section 4.4 for some potential candidates).

It should be noted that in these languages the preverbal position is neither a dedicated focus position, nor the standard position for focus, which is still postverbal. This is where they differ from Kukuya (B77, Section 3.2.1), Mbuun (B87, Bostoen & Mundeke 2012), Nsong (B85d, Koni Muluwa & Bostoen 2014), and Kikongo (H16, De Kind 2014), which have been described as preferring or requiring focus in the position immediately before the verb (IBV). It should also be clarified that we are dealing here with the linear position of the preverbal element rather than its structural position. For the Parameters 2a-d we found that a topical subject can occur in the same linearly preverbal position as a focal or non-topical subject in Kukuya, but they may occupy different *structural* positions in a derivation. This is reflected in the fact that in Kukuya, multiple topical elements are allowed in the preverbal domain while only one focal element is allowed and thus must be placed in a unique structural position. It is only when there is no other preverbal element in the clause that the preverbal subject can be either a topic or a focus.

3.2.3. Thetics

We can turn now to discussion of thetic sentences, i.e. sentences in which all information is presented as new, without a topic-comment split (Sasse 1987; 1996; Erteschik-Shir 1997; Lambrecht 2000). While subjects are typically topical in a categorical sentence, in a thetic sentence, there is no topic constituent, and the subject is therefore typically marked as non-topical. If the preverbal domain in Bantu was reserved for topics, as suggested by Good (2010) for Naki and Yoneda (2011) for Matengo, then we predict that any non-topical constituent would be disallowed preverbally (i.e. in the same position as a topical subject). We have already investigated this above for preverbal focus constituents, and so can turn now to the detopicalised subject in thetic sentences.

Our investigation found interesting differences between the ability for preverbal foci and subjects in thetic sentences to appear preverbally. In Kîîtharaka, Kirundi, and Rukiga, preverbal focus is disallowed, as evidenced by the ungrammaticality of preverbal interrogatives, answers, and arguments modified by 'only' (Parameters 2a-c). Nevertheless, thetic subjects *are* allowed preverbally, as in (12)-(13) (see the Appendix for Kîîtharaka, which is slightly more complicated). Example (14) shows that Kinyakyusa also allows preverbal thetic subjects, as expected since preverbal focus is also allowed.

(12) Context: What is the matter?
Omugôngo nigunsháasha.
o-mu-gongo ni-gu-n-shaasha
AUG-3-back PROG-3SM-1SG.OM-hurt
'My back is hurting.' [Rukiga]

(13) Context: What is going on?
U-mu-ntu a-ø-ra-pfu-ye.
AUG-1-person 1SM-PRS-DJ-die-PFV
'Someone dies.' [Kirundi]

(14) Context: You are sitting in a house as a small group. Someone stares out through the window. Another person asks: Why do you stare through the window?
Ikyura kikolile ulubwele.
i-ki-ula ki-kol-ile u-lu-bwele
AUG-7-frog 7SM-catch-PFV AUG-11-fly
'A/The frog caught a fly!' [Kinyakyusa]

This means that Sub-parameter 2d ("Can a subject in a thetic sentence occur in the same position as a subject in a categorical sentence?") is set differently from Sub-parameters 2a-c ("Can a focused subject occur in the same position as a topical subject?") in these languages. As subjects of thetics are not topics, the preverbal domain should therefore not be characterised as "topical" in these three languages, but rather as "non-focal". Note that this does not mean that there cannot be a structural dedicated topic position preceding the underspecified structural subject position.

In addition to providing useful comparative evidence for preverbal focus, thetics are important for investigating topic prominence. Specifically, our Sub-parameter 2d reflects É. Kiss's (1995: 6) use of the behaviour of a subject in a thetic vs. categorical sentence as an indicator of topic prominence. We will discuss other sub-parameters related to topic prominence in Section 3.3.

3.2.4. Exceptions for question words

Although most diagnostic questions/checkpoints align for the focus-related Sub-parameters 1a-c ("Do internal arguments need to be in a particular position when focused?") and 1d-f ("Do adverbs need to be in a particular position when focused?"), there are some interesting exceptions regarding question words.

The first exception is that questioned internal arguments (theme, benefactive, goal) may behave differently from other focused internal arguments (Sub-parameters 1a vs. 1b-c). In Rukiga, a question word needs to be in the IAV position whereas other focused elements may also appear in other positions. For instance, the answer in (15c) to the content question posed in (15a) is not in the IAV position but is in the canonical position, whereas the *kí* must be in the IAV position (15a, b).

(15) a. *Waaha **kí** Jéini?*
wa-a-h-a ki Jeini?
2SG.SM-N.PST-give-FV what 1.Jane
'What have you given Jane?'
b. **Waaha Jéini **kí**?*
c. *Naaha Jéín' **énkofiira**.*
na-a-h-a Jeini e-n-kofiira
1SG.SM-N.PST-give-FV 1.Jane AUG-9-hat
'I have given a hat to Jane.' [Rukiga]

Note that, while in some languages like Basaá there are differences in behaviour between full noun phrases and pronouns (Hamlaoui & Makasso 2015), this flexibility in position applies equally to pronouns and full noun phrases in Rukiga.[5] Unlike Basaá, there is no direct counterpart to (15a, b) in Rukiga with a pronoun replacing *Jéini*, as pronouns (i) are only used emphatically, (ii) are more typically found in subject position, and (iii) require object marking on the verb. If a pronoun is used —which necessitates a slightly different discourse context in which the referent Jane is contrasted with someone else— the word order is the same as in (15a), suggesting that the status of pronoun vs. full noun phrase is not significant (16a, b).

5. We thank an anonymous reviewer for raising this point. The extra data in (16) are from the second author's judgements.

(16) a. *Wa-a-*(mu)-ha* ***kí*** (*we*)*?*
2SG.SM-N.PST-OM.1-give-FV what 1.pro
'What did you give *her*?'
b. **Waamuha we* ***kí****?* [Rukiga]

Exceptions for question words are also found in Kukuya, a language which has a dedicated preverbal focus position. A question word that functions as an argument is obligatorily placed in the IBV position as in (17a), while the answer and an element modified by the focus particle 'only' can be placed in either the IBV position or in its canonical postverbal position, as in (17b).

(17) a. *N-kaaká* ***kí-má*** *ká-wí* *mvá*?
1-granny 7-what 1SM.PST-give.PST 1.dog
'What did grandmother give to the dog?'
b. *Ndé* *á-wî* *ndé* ***mii-nkaaní****.*
1.PRO 1SM.PST-give.PST 1.PRO 4-paper
'She gave it some papers.' [Kukuya]

This shows that in Rukiga and Kukuya, (non-subject) question words are subject to different constraints than other focused constituents. Surprisingly, however, it is the distribution of question words that is more restricted than that of their answers, unlike what has been reported in the literature for other languages (see e.g. Aboh 2007 on focused vs. non-focused wh-words).

A second exception is questioned adverbs (Sub-parameter 1d). Kirundi has a dedicated sentence-final focus position for arguments, but questioned adverbs can be found either in a non-final postverbal position or in the sentence-final position, as in (18). While sentences such as (18b) may have a prosodic break before the object (if the object is right-dislocated), they are also grammatical without one, i.e. with the question word in non-final position (18c). All three questioned adverbs 'when', 'where' and 'how' have this freedom, but the answers to those questions must occur in the sentence-final focus position. Arguments and adjuncts hence behave differently when questioned in Kirundi.

(18) a. *Tu-shir-e* *i-n-huzu* ***he****?*
1PL.SM-put-SBJV AUG-10-cloth where
b. *Tu-zi-shir-e* ***he****(,)* *i-n-huzu?*
1PL.SM-10OM-put-SBJV where AUG-10-cloth
c. *Tu-shir-e* ***hé*** *i-n-huzu?*
1PL.SM-put-SBJV where AUG-10-cloth
'Where can we put the clothes?' [Kirundi]

Interestingly, our data also reveal that not all kinds of adverbs are subject to the same constraints on their linear position, even within the same language. For example, in Rukiga, the time adverb 'when' can either appear in the IAV position (19a) or its canonical postverbal position, as in (19b), while the manner adverb 'how' must be placed in the IAV position, as in (19c) and (19d).[6]

6. A reviewer asks whether prosodic weight could be an independent factor accounting for the difference between (19a, b) and (19c, d), given that the noun *purésidenti* has 5 syllables while *Kampála* only has 3. However, shorter words show the same word order restriction as longer words when used with the manner adverb 'how'. For example, if the word *muha* 'fox' is substituted for *purésidenti* 'president' in (19c-d), the judgements still hold: *Obutúmwa bukahika búta muha?* is grammatical (meaning 'How did the message reach the fox?'), while the order *Obutúmwa bukahika muha búta?* is ungrammatical (second author's own judgements).

(19) a. *O-ka-za* ***ryarí*** *Kampála?*
2SG.SM-F.PST-go when 23.Kampala
'When did you go to Kampala?'
b. *O-ka-za* *Kampalá* ***ryári****?*
2SG.SM-F.PST-go 23.Kampala when
'When did you go to Kampala?'
c. *O-bu-túmwa* *bu-ka-hik-a* ***bú-ta*** *purésidenti?*
AUG-4-message 14SM-F.PST-reach-FV 14-how 1.president
'How did the message reach the President?'
d. **Obutúmwa bukahika purésidenti* ***búta****?* [Rukiga]

These different constraints on sentential adverbs are also attested in Kukuya, in which some adverbs must occur in the IBV focus position, as in (20a), while adverbs like 'why' can appear either in the IBV position or in its canonical position, as shown in (20b).

(20) a. *Mwáana* ***munkí*** *ká-dzí* *ntsúi?*
1.child when 1SM-eat.PST 1.fish
'When did the child eat fish?'
b. *Baa-ntaba* *bwě* *báá-kwî* ***mu*** ***ki-má****?*
2-goat 2.all 2SM-die.PST 18.LOC 7-what
'Why did all the goats die?' (lit. 'For what did all the goats die?') [Kukuya]

Looking beyond our sample for further examples from the literature, Koni Muluwa & Bostoen (2014) report a similar difference for Nsong (B85d), where adverbials except time adverbials must be focused in their canonical postverbal position and cannot be focused in the IBV position. This contrasts with the languages in our sample, where the focus position is always available. We therefore do not find a clear-cut generalisation such that a certain type of adverbs are more restricted to a dedicated focus position, while others less so.

There are interesting lessons to be drawn from these findings. Methodologically, we learn that it is important to study question words independently of focus elements (i.e. as separate sub-parameters), and we see that multiple sentential adverbs should be tested during data collection. Typologically, we can propose a distinction between languages in which question words are always focused (i.e. in which question words behave the same as other focused elements) and languages in which question words differ from focused elements, and we also highlight the need for further study on which adverbs require a dedicated focus position. Theoretically, the stronger restrictions on question words vs. focused elements should be represented in the features and/or syntactic constraints proposed in the analysis of a given language. In a generative account assuming that focused wh-phrases have both a [wh] and a [focus]/[F] feature, the more restricted nature of wh-phrases with respect to focused phrases in certain languages can be captured by situating the variation on the [wh] feature and its licensing, independently of [focus]/[F].

3.3 Parameter 2: Topic prominence

Having considered the findings related to focus prominence, we can turn now to the second parameter, investigating topic prominence. Crosslinguistically, there are two general strategies for marking topics: the use of topic markers, and the use of syntactic topic constructions such as left dislocation (Gundel 1988). The use of topic markers is characteristic of topic-prominent languages, while syntactic topic constructions are more universally used to express topics (Li & Thompson 1976; Fuller & Gundel 1987; Gundel 1988). While not all languages in our sample show a restriction on topics to

occur in the preverbal domain, all languages do make use of word order to mark the relatively high accessibility/identifiability status of some referents, by left-peripheral expression, passives, and/or inversion constructions.

3.3.1 Universal tendency for preverbal topic

All the languages in our sample show a preference for preverbal topics (Parameter 2e), supporting Gundel's (1988: 231) generalisation that "every language has syntactic topic constructions in which an expression which refers to the topic of the sentence is adjoined to the left of a full sentence comment". The following examples from three sample languages illustrate this universal tendency, in which the topical element is placed in the leftmost position of the sentence.[7]

(21) Context: (You are in a kitchen and see a lot of food.) Why did they cook this food?
ɔ bɛ́ɛ(bɛ) bɛlábɔ́néá bɛ́ɛbɛ*, báná talɛ́áká ɛlɔ́áyɛ́ ɛŋganda yɛ Básɛka.*

ɔ	bɛ́ɛbɛ	bɛ-lábɔ́néá	bɛ́ɛbɛ	bá-ná	talɛ́á-aka
PREP	DEM.PROX.8	8-food	DEM.PROX.8	2SM-PST2	cook-DUR
ɛlɔ́áyɛ́	ɛ-ŋganda	yɛ	básɛka		
for	7-holiday	ASSOC.7	Easter		

'As for this food here, they cooked (it) for Easter.' [Tunen]

(22) Context: What has happened on the bridge?
Aha rutindo *haaraabah' émótoka.*

a-ha	ru-tindo	ha-a-raaba=ho	e-motoka
AUG-16	11-bridge	16.SM-N.PST-pass=16	AUG-9.car

'On the bridge have passed cars.' [Rukiga]

(23) Context: What about the children?
Twáána *ngá*(*tû*)*cereria mabuku*.

tû-ana	n-ka-tû-cer-îr-i-a	ma-buku
12-children	1SG.SM-FUT-12OM-find-APPL-IC-FV	8-books

'The children, I will find books for them.' [Kîîtharaka]

Looking deeper into the tendency for left-peripheral topics, we find evidence for a number of different strategies, including left dislocation and various types of so-called "subject creating constructions" such as inversion and passive constructions (Eyamba G. Bokamba 1976; Eyamba Georges Bokamba 1976; Bresnan & Mchombo 1987; Gundel 1988; Demuth & Harford 1999; Henderson 2011; Marten & Van der Wal 2014; Hamlaoui & Makasso 2015; Hamlaoui forthcoming, amongst others). The following subsections will illustrate these different means of expressing topics.

3.3.2 Left-peripheral topics

Topical arguments or adjuncts may be found in the left periphery of the sentence, where they function as the core referent that the rest of the sentence is anchored to. In addition to the word order change, there is often a prosodic break after the topic ("comma intonation"), though this is not necessary (see

7. A reviewer asks whether contrast is necessary for examples (21) and (23). Both examples are compatible with a contrastive topic but are also possible in a non-contrastive context. For example, the Kîîtharaka Example (23) would be felicitous in the following two contexts (judgements from the third author):
Context A. You are talking about teachers and children and are giving teachers money as rewards. Someone asks: What about the children? (contrastive)
Context B. You are leaving for many days. You have not said whether you will be bringing anything back for your children. Someone asks: What about the children? (non-contrastive)

e.g. Byarushengo, Hyman & Tenenbaum 1976 on Haya; Downing *et al.* 2004 on Chichewa; Downing 2011 on Bantu in general). Left-peripheral topics are most easily seen for objects and adjuncts, but the strategy is possible for subjects too. In general, two analyses are possible, one in which the topic is base-generated in the left periphery, and one in which it is moved to this position ("left dislocation"; Gundel 1988).

(24) ***Ya*** ***nzu*** *n-a-a-yi-gur-ye*.
9.DEM$_{vi}$ 9.house 1SG.SM-N.PST-DJ-9OM-buy-PFV
'That house, I bought it.' [Kirundi]

(25) ***Etthú tsáu tsootéénè****, oreéké wákúsheke*.
Etthú ts-áu ts-ootéénè o-r-eék-é wá-kúsh-ek-e.
10.things 10-POSS.2SG 10-all 2SG.SM-go-DUR-SBJV 2SG.SM.SUBS-carry-DUR-SJBV
'All your things, go and take them!' [Makhuwa]

(26) ***Méémbeé,*** *mbendeete márá ˈméérú.*
ma-embe n-eend-ete ma-ra ma-eru
6-mango 1SG.SM-want-STAT.PFV 6-DEM.DIST 6-ripe
'As for mangos, I like ripe ones.' [Kîîtharaka]

It is also possible for there to be two or more elements in the left periphery of the sentence, referring to multiple topics. In the examples below, the topical subject is in the leftmost position and is followed by an object which is also topical.

(27) Context: What did father do with the beans?
Tááta ***e-bi-hîmba*** *a-bi-teek-íre*.
1.father AUG-8-beans 1SM-8OM-cook-PFV
'Father, the beans, he cooked them.' [Rukiga]

(28) ***Maamá*** ***mu-nkáání*** *ku-ní* *ká-wéek-i?*
1.mother 3-letter 17-which 1SM.PST-send-PST
'Where did mother send the letter?' [Kukuya]

We therefore see left-peripheral topics across the languages in our sample as an example of how word order relates to discourse roles.

3.3.3 Subject-creating constructions

In some Bantu languages, the fronted constituents can function as the syntactic subject, controlling subject marking on the verb (a possible diagnostic for the notion "subject" in a grammatical role-oriented approach). This differs from the left-dislocation construction, where topical elements are just adjoined to the left of the sentence without change of the subject marking on the verb. Most Bantu languages make use of a passive construction; another type of subject-creating construction is subject inversion, whereby a locative, instrument, or theme is fronted to the preverbal domain and becomes the subject of the sentence. We will discuss these in turn.

Passives in Bantu are typically marked in the verbal morphology as a derivational suffix on the verb (Guthrie 1967; Schadeberg 2003; Guérois forthcoming), for example *-iy-* in Makhuwa in (29). Syntactically, the passive demotes the agent and promotes the patient to subject, but the passive has an important pragmatic function too: it promotes the patient to topic and detopicalises or focuses the agent (see also Demuth 1990; Hamlaoui & Makasso 2015; Van der Wal 2015b; Li 2020).

(29) Context: A conversation about *nsiro*, a traditional paste made from tree bark
(*Nsiro*) *Wáámóótt**í**ya erúttú yootééné...*
nsiro o-aa-oott-iy-a eruttu e-oteene…
3.nsiro 3SM-IPFV-smear-PASS-FV 9.body 9-all
'(Nsiro) It was smeared (on) the whole body.' [Makhuwa]

In languages that do not feature a morphological passive, word order may be used as a "functional passive", for example in Kukuya. When the patient occurs sentence-initially and the agent is in IBV position, the sentence is translated and interpreted as a passive (see Li 2020; also Bostoen & Mundeke 2011 for Mbuun; Van der Wal 2015b for Matengo).

(30) *Mwáana ma-láara bí-búru bíí-wî.*
1.child 6-oranges 8-parents 8SM.PST-give.PST
'The child was given the oranges by the parents.' [Kukuya]

Subject inversion constructions differ from the passive in that the agent cannot be omitted, and can be interpreted as focal. In instrument inversion and locative inversion, the instrument or locative is instead promoted to topic, occupying a preverbal position and determining agreement on the verb, as shown for Kirundi and Rukiga in (31) and (32), respectively.

(31) *I-karamu y-andik-ish-a John.*
AUG-9.pen 9SM-write-CAUS-FV 1.John
'It is John who writes with a pen.' [Kirundi]

(32) *A-ha ru-tindo ha-a-raaba=hó e-mótoka.*
AUG-16 11-bridge 16SM-N.PST-pass=16 AUG-9.car
'On the bridge have passed cars.' [Rukiga]

These constructions are used to manipulate the word order to fit the discourse preferences, providing further evidence of discourse roles conditioning word order. However, such constructions are not always possible, as inversion constructions may be restricted by the valency of the predicate (Bresnan & Mchombo 1987; Demuth & Mmusi 1997; Marten 2006; Khumalo 2010; Salzmann 2001; Marten & Van der Wal 2014; see also Marten 2014 and Hamlaoui forthcoming for overviews of theoretical approaches to subject inversion). In Rukiga, locative inversion is only possible with intransitive predicates (both unaccusative and unergative) and passivised predicates. For transitive predicates to participate in locative inversion, they are therefore first passivised, as in (33). This means that word order and passivisation do not always go hand in hand: the passive may merely function to demote the agent, and the patient may still be presented as the focus.

(33) *Omu rufûnjo hahingirwey' ébihîmba.*
o-mu ru-funjo ha-hing-ir-w-e=yo e-bi-himba
AUG-18 11-swamp 16SM-sprout.CAUS-PFV-PASS-FV=23 AUG-8-beans
'In the swamp were dug beans.' [Rukiga]

As a final note, the languages in our sample lacking both a passive and inversion constructions are interestingly both found in the Northwest: Tunen and Kukuya. While Kukuya can use the IBV strategy as a functional passive, as noted above and in Li (2020), Tunen uses a middle verbal prefix *bɛ́-* (Mous

2008), an active impersonal construction with the class 2 *bá-* subject marker (*ba*-passive), or a participle verb construction, with no change to word order.

3.4 Interim summary

Our research shows a rich amount of variation in Bantu languages when investigated for evidence of discourse prominence conditioning word order. Although no language is at the far end of the continuum of word order between grammatical and discourse roles, we do find a large range from Tunen, which is close to the grammatical role end, to Makhuwa and Kirundi close to the discourse role end. Furthermore, we found variation in the way in which discourse prominence is expressed, for example in the variation between IAV, IBV, and sentence-final focus positions, and the extent to which items such as question words, their answers, and sentential adverbs must be found in these positions.

Our parametric overview led to new observations on asymmetries between interrogatives and other focused elements, discovering a potential centre of innovation around Kinyakyusa allowing preverbal focus, and identifying via thetic subjects that the preverbal domain may be characterised as "topical" or "non-focal" in different languages. In terms of topic prominence, we found that all languages show a preference for initial preverbal topics, with most also employing passives and subject inversion constructions.

4. Discussion

Having shown the empirical variation highlighted in this study, we will reflect in this section on the methodology chosen to investigate discourse prominence, the theoretical modelling and implications of discourse prominence in Bantu, and the areas for further research.

4.1 Working with parameters

We chose to use a parametric approach in order to systematically investigate the influence of discourse roles across different languages. Parametric approaches aid comparison between languages by abstracting from the raw data to give an overview of the different patterns, which can be used to discover new generalisations (see e.g. Baker 2009). This is not to say that the raw data should subsequently be ignored; instead, the parameter values highlight where it is most interesting to go back to the raw data for deeper understanding, as we have done throughout this paper (for full data, see the Appendix). Furthermore, we believe that creating such parameters can serve as a useful guide for future fieldworkers interested in investigating the influence of discourse roles and grammatical roles on word order by providing a checklist of relevant points to address in data collection, which in turn will expand our ability for further typological studies on Bantu.

However, parametric overviews have been subject to various critiques (see for example Van de Velde this issue). They can be misleading if the parameters are not well-defined and/or not accurately coded. A common critique of parameters in the typological literature is that some language patterns are too complicated to distil into a binary value of "Y" or "N" (see e.g. Bickel 2015; Haspelmath 2018; Evans 2020), as we also found (see items coded Y/N in Table 1 and further discussion in the Appendix). We mitigated the risks of inaccurate coding by working iteratively as a team, checking each other's values and revising the parameter list where necessary until we were in agreement. We also developed checkpoints rather than single parameter settings for focus and topic prominence, to avoid the risk of collapsing too much variation into a binary "Y" or "N" value. This proved to be prescient, as the parameter values in Table 1 show variation between checkpoints within a single parameter. Future revisions of the parameters could consider further breaking down checkpoints that show mixed behaviour.

The critiques of parametric approaches relate to the broader point that typological research has a fundamental tension between a desire for generalisations and for understanding of particular language phenomena (Evans 2020). We hope to have shown in this paper that parameter codings should not be taken as results in and of themselves, but serve as a useful tool for guiding further investigation. Furthermore, with further revisions of this parametric approach to Bantu word order, we may be able to

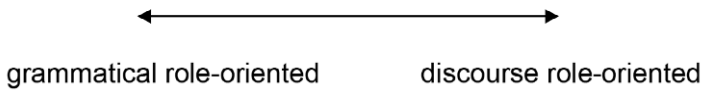


Figure 4 — Continuum of variation between grammatical role-oriented and discourse role-oriented word order

better capture the variation and investigate implicational relationships between the parameters (thus establishing parameter hierarchies), potentially also gaining new insight into the grammaticalisation mechanisms for such word order patterns. As it stands, however, our parameter codings should only be taken as generalisations over our data, rather than commitments to specific parameter "settings" in the minds of speakers of the languages.

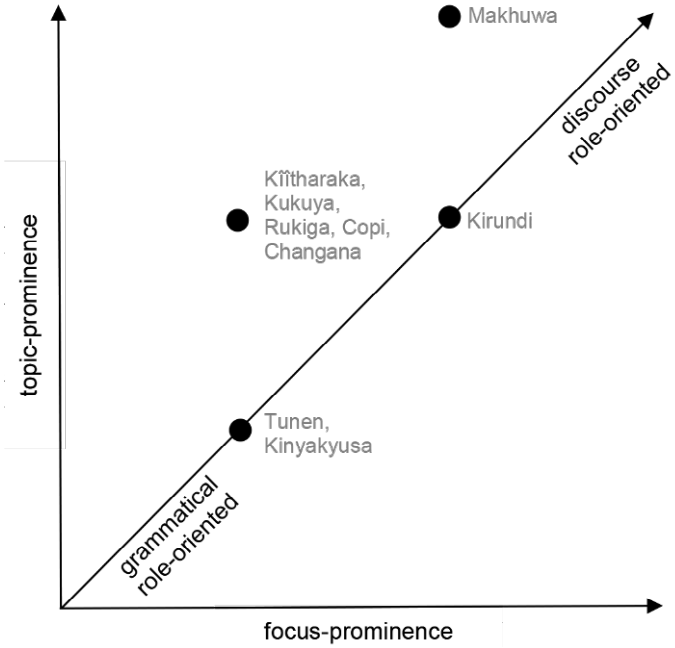
4.2 Continuum of variation — degree of discourse prominence

One way to conceptualise the variation we found is a continuum between word order being fully conditioned by grammatical roles on the one extreme and being fully conditioned by discourse roles on the other, as in Figure 4.

Surányi (2016) argues against such a simple continuum between grammatical role-oriented languages and discourse role-oriented languages ("configurational" and "discourse-configurational" in his terminology) on the grounds that other pragmatic factors such as animacy and evidentiality may result in deviations from grammatical role-oriented word order. There are two ways this critique can be taken on board. One option is to recast the continuum in Figure 4 as "configurational" vs. "pragmatic-configurational", with the latter being a more general category than discourse role-oriented. The second option is to propose multiple continua, i.e. multiple dimensions of variation. Our parameters allow us to easily draw up continua for both grammatical role-oriented vs. focus-oriented word order patterns (Parameter 1) and grammatical role-oriented vs. topic-oriented ones (Parameter 2). By tallying the number of grammatical role vs. discourse role codings for each group of sub-parameters, we arrive at the visual representation of the variation expressed in Figure 5.

Future work could investigate other dimensions of variation, e.g. between grammatical role and animacy or evidentiality, which could be plotted on separate axes resulting in further multidimensionality. As our study is limited to word order, Figure 5 only shows variation in that domain rather than being representative of the languages as a whole. We also note that such representations assume that each parameter checkpoint should be equally weighted, which introduces a degree of arbitrariness given that we expect the parameter checkpoints to evolve further on the basis of further study. The diagram in Figure 5 should therefore only be seen as a rough approximation of the variation found in the data, used to demonstrate how discourse prominence can be seen in terms of topic and focus prominence, with different Bantu languages behaving differently for these two properties.

Nevertheless, we can note some interesting points here. First, it is interesting that both Tunen and Kinyakyusa lean towards grammatical role determining word order, given that these languages are geographically quite distant. As mentioned above, the Northwestern Bantu languages are known to be typologically different from the other Bantu languages (see e.g. Bostoen & Van de Velde 2019), and Tunen is particularly unusual in showing dominant SOV word order, possibly as the only Bantu language and arguably as an innovation (Mous 1997; 2005; see Section 3.1.1). Kinyakyusa, on the other hand, might be part of a centre of innovation allowing preverbal focus, leading to it showing less evidence for focus prominence based on our parameters. Second, all languages show at least partial evidence for both topic and focus prominence, that is, if a language is topic-configurational to a certain degree, it is also focus-configurational and vice versa. Third, while languages may have the same number of Y vs. N parameter codings, they can still be very different. Kukuya, for example, falls in the middle with several other languages, even though it functions quite differently from them in having an IBV (not IAV) focus position: its overall profile shows that its word order is partly determined by grammatical roles and partly by information structure. The parameter codings are therefore only proxies to the variation.



The continuum between grammatical role-oriented and discourse role-oriented word order in Figure 4 is split into two sub-components: focus prominence (Parameter 1; x-axis) and topic prominence (Parameter 2; y-axis). Languages are plotted based on the values for the sub-parameters (1a-c, 1d-f, 1g-i, 2a-c, 2d-e) that showed evidence for grammatical vs. discourse role conditioning their word order.)

Figure 5 — A simple visualisation of the parameter codings in Table 1

We began this study by posing the research question "How far can we get in accounting for the word order of a language without reference to grammatical roles?". The continua of variation found show that we cannot fully describe the word order of Bantu languages in terms of information structural classifications such as "topic-nontopic" only, even in languages with discourse-driven constructions such as locative inversion and a dedicated focus position. However, the classical treatment of languages in terms of "SVO", "SOV" similarly fails to capture the full picture of word order in these languages, and grammatical role-oriented statements such as "Bantu languages are SVO" present a misleadingly homogenous view of Bantu word order. We therefore advocate a language-specific classification of the word order of each language which combines both approaches. For each individual language it needs to be determined whether grammatical roles or discourse roles capture most of its patterns, and what the exceptions are. Interestingly, these exceptions seem to always be motivated by the other force, for example how the passive in a "syntax-configurational" language like English adjusts word order for discourse-related reasons. Such an approach requires a fundamental rethinking of how we approach Bantu word order, as we are required to be able to flip the description to be discourse-oriented. While we are used to managing exceptions to the syntax, for example in saying that Italian is an SVO language with an extended left periphery for discourse roles (Rizzi 1997), we are yet to be trained in the opposite exceptions: "Makhuwa is a topic-V-focus language, with strict subject agreement".

4.3 Types of discourse prominence

In the debate on discourse prominence, a distinction is typically made between topic-prominent and focus-prominent languages (e.g. É. Kiss 1995; Surányi 2016). On the basis of our parameter values as visualised in Figure 5, we can tentatively conclude that Kirundi, Kukuya, and Makhuwa are focus-prominent (as they show a dedicated focus position). Disregarding Parameter 1d on thetic subjects, we could say all other languages are topic-prominent apart from Tunen and Kinyakyusa (which are neither topic- nor focus-prominent). However, the thetic data show us that the more correct generali-

sation is in fact a negative one: there is a position (typically preverbally for East-Bantu) that does *not* allow focus, and a position (typically postverbally for East-Bantu) that does *not* allow topics. Such negative characterisations of positions are more difficult to model theoretically but are more accurate descriptively.

We mentioned in Section 1 that we take the broad definition of discourse prominence, but it would be worthwhile investigating whether the linear positions can also be shown to be dedicated structural positions. Especially for Kukuya it is desirable to test two hypotheses: an analysis in which there is one position which can be occupied by constituents of different discourse roles, or a Cartographic analysis in which there is a dedicated topic position preceding a dedicated focus position (Rizzi 1997). This remains a topic for further research.

4.4 Areas for further research

There are various interesting avenues for further research highlighted in this study. Firstly, we observed above that there are different focus positions —IBV, IAV, and sentence-final. Further research could shed light on the question of what determines whether a language will grammaticalise one position or another.

Secondly, stronger crosslinguistic comparisons can be made with stronger crosslinguistic data. As this study aims at a shift in thinking from grammatical role-oriented to considering discourse relations, it required new data collection on the languages in question. We hope that the parameters and discussion in this paper can guide other researchers for future data collection that takes into account the (possible) influence of discourse relations on word order, in order to better understand the word order variations within the Bantu language family through future larger-scale typological study.

Thirdly, our investigation may also serve as the starting point for research on variation within Bantu and the influence of contact. One example is the presence/absence of preverbal focus, as discussed in Section 3.2.2, where we found evidence for an innovation of preverbal focus around the Kinyakyusa-speaking area. Interestingly, this same area around Guthrie's zone M was also found to be the centre of two waves of innovation for object marking, changing from non-doubling to doubling, and from multiple to single object markers (Van der Wal 2022; Marten *et al.* 2007 on Bemba and Chichewa as the centre). This suggests that the set of innovations may be linked. Another example of a potential areal phenomenon is inversion constructions —as Hamlaoui (2018) states (see also Hamlaoui & Makasso 2015), languages of the Northwestern region (including Tunen and Kukuya) typically have no inversion constructions. We saw above that Tunen has also been argued to have innovated an SOV word order pattern while Kukuya appears to have innovated an IBV focus position, showing another potential set of innovations that could be investigated in further comparative work.

Fourthly, when devising theoretical models of word order, we need to allow for varying influence of grammatical roles and discourse roles in the grammar, establishing on a language-individual basis which precise notions are active in a given language (see e.g. Stucky 1981; Bresnan & Mchombo 1987; Morimoto 2000; 2006; Hamlaoui & Makasso 2015 for discussion on this point in Phrase Structure, Lexical Functional Grammar, and Optimality Theory frameworks). We have investigated the broad notions of topic and (non-)focus here, but "contrast" or "unexpectedness" may also turn out to be relevant notions (see e.g. Bianchi *et al.* 2015; 2016; Cruschina 2019; 2021). Under the assumption that an acquirer only posits what is necessary and evidenced in the input (*contra* the Strong Uniformity Hypothesis of Chomsky 2001; see also Miyagawa 2009), the most appropriate theoretical model must allow for this variability, rather than proposing that all such notions are grammaticalised in all (Bantu) languages.

Finally, our study was limited to word order, which means that our evidence cannot by itself prove that information structure has a key role in regulating syntax in the studied languages. There are various other domains that should be investigated. For example, Gundel (1988) shows that in languages where topic-comment relations (i.e. information structure) control word order, there are other accompanying syntactic parameters, including a lack of dummy subjects, absence of restrictions on zero anaphora and absence or marginal presence of passivisation. Additionally, languages that have these parameters are also said to be topic-prominent, and are highly likely to have morphemes that mark

topics (such as Burmese, Japanese, Hua, Marathi, Tamil, *inter alia*, Gundel 1988: 216). In contrast, non-topic-prominent languages (which mostly have so-called free word order, such as Russian and Hebrew) show a closer association between subject and actor than topic, with word order relations the sole means of marking topics. Lastly, Morimoto (2006) discusses the influence of discourse roles on agreement. By investigating the variation in Bantu languages for such additional parameters, we can arrive at a fuller understanding of the extent of influence of discourse roles on Bantu syntax.

5. Conclusions

In summary, we have presented a parametric study of word order variation in a sample of 9 Bantu languages, asking the question of how far we can get in accounting for their word order without recourse to grammatical role relations such as "subject" and "object". We found that all languages showed influence of both grammatical role relations and discourse role relations, but there was a large amount of crosslinguistic variation, with some languages showing greater influence of discourse relations (e.g. Makhuwa, Kirundi) and some showing greater influence of grammatical relations (e.g. Tunen, Kinyakyusa). This means that the notions "subject" and "object" are still necessary, but other notions such as "topic" and "focus" are also necessary and for some languages in fact more insightful. When investigating the word order patterns, we observed that there is variation between focus positions (no dedicated position, IBV, IAV, and sentence-final), variation in whether question words pattern with other focused items, a crosslinguistic tendency for preverbal topics, and a possible areal innovation of preverbal focus in the region around where Kinyakyusa is spoken. We reflected on the methodology used, arguing that a parametric approach provides a template for collecting new data and helps provide an overview of the variation which can guide further research. Finally, we considered the implications of the variation found for theoretical models of Bantu word order, arguing that models must accommodate both discourse and grammatical roles, with different notions relevant for different languages. Further research could investigate the topics discussed in this paper in more detail, collect data on more Bantu languages to see whether there are further generalisations that can be made, and broaden the research into domains beyond word order, such as agreement.

Supplementary materials

Results table with supporting data and a copy of the parameters list with guidance notes (DOI: 10.17605/OSF.IO/PN2Q3).

Acknowledgements

This research was supported by an NWO Vidi grant 276-78-001 as part of the Bantu Syntax and Information Structure (BaSIS) project at Leiden University. We thank our speakers Bahati Laikon Mwakasege, Peter Mwasyika Mwaipyana, Yona Mwaipaja, Pamellah Geiga Birungi, Joel Tumusiime, Ronald Twesigomwe, Dennis Muriuki Katheru, Philip Murithi Nyamu, Onesmus Mugambi Kamwara, Constancia Zaida Mussavele, Arlindo João Nhanthumbo, Gomes David Chemane, Menrage Buananli, Joaquim Nazário, N'gamo Saida Aly, Pierre Molel, Étienne Ondjem, Jeanne Ong'omolaleba, Zacharie Ngouloubi, Gilbert Mbou, Alain Mbiambourou and Gabriel Ntsiebele for sharing their insights with us. We thank Helen Eaton, Hazel Gray, Heidrun Kröger and Julius Taji for checks on preverbal focus; Nelsa Nhantumbo for collaborative data collection on Cicopi, Aurélio Simango for collaborative data collection on Changana, and Simon Msovela for collaborative data collection on Kinyakyusa. We thank the audiences of ACAL51-52 and WOCAL10, especially the audience of the WOCAL10 Bantu Universals and Variation workshop, and the two reviewers of this paper for helpful feedback. Finally, we thank Irina Morozova and the LLA editorial board for proofreading and edits. As is customary, we alone remain responsible for the views expressed here and any errors made.

Abbreviations

1, 2, 3, …	Bantu noun class numbers (unless followed by SG or PL)
1/2SG, 1/2PL	1st/2nd person singular / plural
APPL	applicative
ASSOC	associative marker
AUG	augment (pre-prefix)
CAUS	causative
CJ	conjoint verb form
DEM	demonstrative
DIST	distal
DJ	disjoint verb form
DUR	durative
F.PST	far past
FOC	focus marker/copula
FUT	future tense
FV	final vowel (marker of aspect)
IC	inner causative
IPFV	imperfective
LOC	locative
NEG	negation
N.PST	non-past tense
OM	object marker
PASS	passive
PFV	perfective
POSS	possessive
PREP	preposition
PRO	pronoun
PROG	progressive
PROX	proximal
PRS	present
PST	past tense
PST2	2nd-degree past tense (hodiernal)
REL	relative
SM	subject marker
STAT	stative
SBJV	subjunctive
SUBS	subsective
VI	demonstrative series 6

References

Aboh, Enoch Oladé. 2007. Focused versus non-focused wh-phrases. In Enoch Oladé Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 287-314. Berlin: De Gruyter Mouton.

Baker, Mark. 2009. Formal generative typology. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 285-312. 1st ed. Oxford: Oxford University Press.

Bearth, Thomas. 2003. Syntax. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 121-142. 1st ed. New York: Routledge.

Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2016. Focus fronting, unexpectedness, and evaluative implicatures. *Semantics and Pragmatics* 9(3). 1-54.

Bianchi, Valentina, Giuliano Bocci, Silvio Cruschina, Enoch Oladé Aboh, Jeannette Schaeffer & Petra Sleeman. 2015. Focus fronting and its implicatures. In *Romance languages and linguistic theory: Selected papers from "Going Romance" Amsterdam 2013*, 1-20. Amsterdam: John Benjamins.

Bickel, Balthasar. 2015. Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 901-924. 2nd ed. Oxford: Oxford University Press.

Bokamba, Eyamba G. 1976. Authenticity and the choice of a national language: The case of Zaïre. *Présence africaine* (99-100). 104-142.

Bokamba, Eyamba Georges. 1976. *Question formation in some Bantu languages*. Bloomington, IN: Indiana University.

Bostoen, Koen. 2018. The Bantu expansion. *Oxford research encyclopedia of African history*.

Bostoen, Koen & Joseph Koni Muluwa. 2021. The immediate before verb focus position in West-Coastal Bantu: Some comparative data. Paper presented at the online workshop "Typological Perspectives on Focus Marking in African Languages", 27 May 2021, EHESS, Paris.

Bostoen, Koen & Leon Mundeke. 2011. Passiveness and inversion in Mbuun (Bantu B87, DRC). *Studies in Language* 35(1). 72-111.

Bostoen, Koen & Leon Mundeke. 2012. Subject marking, object-verb order and focus in Mbuun (Bantu, B87). *Southern African Linguistics and Applied Language Studies* 30(2). 139-154.

Bostoen, Koen & Mark L. O. Van de Velde. 2019. Introduction. In Mark L. O. Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 1-13. 2nd ed. London: Routledge.

Bresnan, Joan & Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63(4). 741-782.

Buell, Leston. 2005. *Issues in Zulu verbal morphosyntax*. Los Angeles, CA: University of California (PhD dissertation).

Buell, Leston. 2009. Evaluating the immediate postverbal position as a focus position in Zulu. In Masangu Matando, Fiona McLaughlin & Eric Potsdam (eds.), *Selected proceedings of the 38th Annual Conference on African Linguistics: Linguistic theory and African language documentation*, 166-172. Somerville, MA: Cascadilla Proceedings Project.

Byarushengo, Ernest Rugwa, Larry M. Hyman & Sarah Tenenbaum. 1976. Tone, accent, and assertion in Haya. In Larry M. Hyman (ed.), *Studies in Bantu tonology* (Southern California Occasional Papers in Linguistics 3), 181-205. Los Angeles, CA: University of Southern California.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1-52. Cambridge, MA: MIT Press.

Claudi, Ulrike. 1993. *Die Stellung von Verb und Objekt in Niger-Kongo-Sprachen: Ein Beitrag zur Rekonstruktion Historischer Syntax*. Cologne: Institut für Afrikanistik, Universität zu Köln.

Costa, João & Nancy C. Kula. 2008. Focus at the interface: Evidence from Romance and Bantu. In Cécile De Cat & Katherine Demuth (eds.), *The Bantu-Romance connection: A comparative investigation of verbal agreement, DPs, and information structure*, 293-322. Amsterdam: John Benjamins.

Cruschina, Silvio. 2019. Focus fronting in Spanish: Mirative implicature and information structure. *Probus* 31(1). 119-146.

Cruschina, Silvio. 2021. The greater the contrast, the greater the potential: On the effects of focus in syntax. *Glossa* 6(1). 1-30.

De Kind, Jasper. 2014. Pre-verbal focus in Kisikongo (H16a, Bantu). *ZAS Papers in Linguistics* 57. 95-122.

Demuth, Katherine. 1990. Subject, topic and the Sesotho passive. *Journal of Child Language* 17(1). 67-84.

Demuth, Katherine & Carolyn Harford. 1999. Verb raising and subject inversion in Bantu relatives. *Journal of African Languages and Linguistics* 20(1). 41-62.

Demuth, Katherine & Sheila Mmusi. 1997. Presentational focus and thematic structure in comparative Bantu. *Journal of African Languages and Linguistics* 18(1). 1-20.

Downing, Laura J. 2011. The prosody of "dislocation" in selected Bantu languages. *Lingua* 121(5). 772-786.

Downing, Laura J. & Larry M. Hyman. 2016. Information structure in Bantu. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 790-813. Oxford: Oxford University Press.

Downing, Laura J. & Lutz Marten. 2019. Clausal morphosyntax and information structure. In Mark L. O. Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 270-307. 2nd ed. London: Routledge.

Downing, Laura J., Al Mtenje & Bernd Pompino-Marschall. 2004. Prosody and information structure in Chichewa. *ZAS Papers in Linguistics* 37. 167-186.

Dugast, Idelette. 1971. *Grammaire du tùnen*. Paris: Klincksieck.

É. Kiss, Katalin É. 1995. *Discourse-configurational languages*. Oxford: Oxford University Press.

Erteschik-Shir, Nomi. 1997. *The dynamics of focus structure*. Cambridge: Cambridge University Press.

Evans, Nicholas. 2020. Introduction: Why the comparability problem is central in typology. *Linguistic Typology* 24(3). 417-425.

Fuller, Judith W. & Jeanette K. Gundel. 1987. Topic-prominence in interlanguage. *Language Learning* 37(1). 1-18.

Gibson, Hannah, Andriana Koumbarou, Lutz Marten & Jenneke van der Wal. 2017. Locating the Bantu conjoint/disjoint alternation in a typology of focus marking. In Jenneke van der Wal & Larry M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*, 61-99. Berlin: Mouton de Gruyter.

Good, Jeff. 2010. Topic and focus fields in Naki. In Ines Fiedler & Anne Schwarz (eds.), *The expression of information structure: A documentation of its diversity across Africa*, 35-68. Amsterdam: John Benjamins.

Guérois, Rozenn. Forthcoming. Passive constructions. In Lutz Marten, Nancy C. Kula, Ellen Hurst & Jochen Zeller (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.

Guérois, Rozenn, Hannah Gibson & Lutz Marten. 2017. *Parameters of Bantu morphosyntactic variation*. Draft master list. Alpha version, last amended 28 December 2017. Leverhulme project "Morphosyntactic variation in Bantu: Typology, contact and change". London: SOAS, University of London.

Güldemann, Tom. 2007. Preverbal objects and information structure in Benue-Congo. In Enoch Oladé Aboh, Katharina Hartmann & Malte Zimmermann (eds.), *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, 83-112. Berlin: De Gruyter Mouton.

Gundel, Jeanette K. 1988. Universals of topic-comment structure. In Michael Hammond, Edith A. Moravcsik & Jessica Wirth (eds.), *Studies in syntactic typology*, 209-242. Amsterdam: John Benjamins.

Guthrie, Malcolm. 1967. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. 4 vols. Farnborough: Gregg Press.

Hadermann, Pascale. 1996. Grammaticalisation de la structure infinitif + verbe conjugué dans quelques langues bantoues. *Studies in African Linguistics* 25(2). 155-169.

Hajičová, Eva. 1983. Topic and focus. *Theoretical Linguistics* 10(2). 268-286.

Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1(1). 5-47.

Hamlaoui, Fatima. 2018. Proto-Bantu word order. Paper presented at the International Conference on Reconstructing Proto-Bantu Grammar, 19-23 November 2018, Ghent University.

Hamlaoui, Fatima. Forthcoming. Inversion constructions. In Lutz Marten, Nancy C. Kula, Ellen Hurst & Jochen Zeller (eds.), *The Oxford guide to the Bantu languages*. Oxford: Oxford University Press.

Hamlaoui, Fatima & Emmanuel-Moselly Makasso. 2015. Focus marking and the unavailability of inversion structures in the Bantu language Bàsàá (A43). *Lingua* 154. 35-64.

Hammarström, Harald. 2019. An inventory of Bantu languages. In Mark L. O. Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 17-78. 2nd ed. London: Routledge.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2022. *Glottolog 4.7*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In Daniël Olmen, Tanja Mortelmans & Frank Brisard (eds.), *Aspects of linguistic variation*, 83-114. Berlin: De Gruyter Mouton.

Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.

Henderson, Brent. 2011. Agreement, locality, and OVS in Bantu. *Lingua* 121(5). 742-753.

Horvath, Julia. 1985. *Focus in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris.

Khumalo, Langa. 2010. Passive, locative inversion in Ndebele and the unaccusative hypothesis. *South African Journal of African Languages* 30(1). 22-34.

Kidwai, Ayesha. 1999. Word order and focus positions in Universal Grammar. In Georges Rebuschi & Laurice Tuller (eds.), *The grammar of focus*, 213-244. Amsterdam: John Benjamins.

Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda*. Berkeley, CA: University of California Press.

Koni Muluwa, Joseph & Koen Bostoen. 2014. The immediate before the verb focus position in Nsong (Bantu B85d, DR Congo): A corpus-based exploration. *ZAS Papers in Linguistics* 57. 123-135.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611-682.

Li, Charles N. & Sandra Thompson. 1976. Subject and topic: A new typology of language. In Charles N. Li (ed.), *Subject and topic*, 457-490. New York: Academic Press.

Li, Zhen. 2020. A note on the functional passives in Teke-Kukuya (Bantu B77, Congo). In Jenneke van der Wal, Heleen J. Smits, Sara Petrollino, Victoria Nyst & Maarten G. Kossmann (eds.),

Essays on African languages and linguistics: In honour of Maarten Mous, 267-288. Leiden: African Studies Centre Leiden.

Maho, Jouni Filip. 2003. A classification of the Bantu languages: An update of Guthrie's referential system. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 639-651. 1st ed. London: Routledge.

Maho, Jouni Filip. 2009. *NUGL online: The online version of the new updated Guthrie list, a referential classification of the Bantu languages*.

Marten, Lutz. 2006. Locative inversion in Otjiherero: More on morphosyntactic variation in Bantu. *ZAS Papers in Linguistics* 43. 97-122.

Marten, Lutz. 2014. The preverbal position(s) in Bantu inversion constructions: Theoretical and comparative considerations. *ZAS Papers in Linguistics* 57. 136-159.

Marten, Lutz, Nancy C. Kula & Nhlanhla Thwala. 2007. Parameters of morphosyntactic variation in Bantu. *Transactions of the Philological Society* 105(3). 253-338.

Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318-368.

Mithun, Marianne. 1987. Is basic word order universal? In Russell S. Tomlin (ed.), *Coherence and grounding in discourse: Outcome of a symposium, Eugene, Oregon, June 1984*, 281-328. Amsterdam: John Benjamins.

Miyagawa, Shigeru. 2009. *Why agree? Why move? Unifying agreement-based and discourse configurational languages*. Cambridge, MA: MIT Press.

Morimoto, Yukiko. 2000. *Discourse configurationality in Bantu morphosyntax*. Stanford, CA: Stanford University (PhD dissertation).

Morimoto, Yukiko. 2006. Agreement properties and word order in comparative Bantu. *ZAS Papers in Linguistics* 43. 161-187.

Morimoto, Yukiko. 2009. From topic to subject marking: Implications for a typology of subject marking. In Helen de Hoop & Peter de Swart (eds.), *Differential subject marking*, 199-221. Dordrecht: Springer.

Moroz, George. 2017. lingtypology: Easy mapping for Linguistic Typology.

Mous, Maarten. 1997. The position of the object in Tunen. In Rose-Marie Déchaine & Victor Manfredi (eds.), *Object positions in Benue-Kwa*, 123-137. The Hague: Holland Academic Graphics.

Mous, Maarten. 2003. Nen (A44). In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 283-306. 1st ed. London: Routledge.

Mous, Maarten. 2005. The innovative character of object-verb word order in Nen (Bantu A44, Cameroon). In Koen Bostoen & Jacky Maniacky (eds.), *Studies in African comparative linguistics with special focus on Bantu and Mande*, 411-424. Tervuren: Musée royal de l'Afrique centrale.

Mous, Maarten. 2008. Voice in Tunen: The so-called passive prefix *bé-*. *Studies in Slavic and General Linguistics* 33. 303-312.

Ndayiragije, Juvénal. 1999. Checking economy. *Linguistic Inquiry* 30(3). 399-444.

Ngoboka, Jean Paul & Jochen Zeller. 2017. The conjoint/disjoint alternation in Kinyarwanda. In Jenneke van der Wal & Larry M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*, 350-389. Berlin: Mouton de Gruyter.

Nshemezimana, Ernest. 2016. *Morphosyntaxe et structure informationnelle en Kirundi : focus et stratégies de focalisation*. Ghent: Ghent University (PhD dissertation).

Nurse, Derek & Gérard Philippson. 2003. Introduction. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 1-12. 1st ed. London: Routledge.

Öhl, Peter. 2010. Formal and functional constraints on constituent order and their universality. In Carsten Breul & Edward Göbbel (eds.), *Comparative and contrastive studies of information structure*, 231-275. Amsterdam: John Benjamins.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281-337. Dordrecht: Springer.

Rooth, Mats. 1985. *Association with focus*. Amherst, MA: University of Massachusetts (PhD dissertation).

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75-116.

Salzmann, Martin. 2001. *Theoretical approaches to locative inversion*. Zurich: University of Zurich (M.A. thesis).

Sasse, Hans-Jürgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25(3). 511-580.

Sasse, Hans-Jürgen. 1996. *Theticity*. (Arbeitspapiere des Instituts für Sprachwissenschaft der Universität zu Köln 27). Cologne: Universität zu Köln.

Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*, 143-163. 1st ed. London: Routledge.

Sgall, Petr (ed.). 1985. *Contributions to functional syntax, semantics and language comprehension*. Amsterdam: John Benjamins.

Sgall, Petr, Eva Hajičová & Eva Benešová. 1973. *Topic, focus and generative semantics*. Kronberg im Taunus: Scriptor.

Skopeteas, Stavros, Ines Fiedler, Sam Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on information structure (QUIS): Reference manual*. Potsdam: Universitätsverlag Potsdam.

Stucky, Susan U. 1981. Free word order languages, free constituent order languages, and the gray area in between. *Proceedings of the 11th Annual Meeting of the North Eastern Linguistic Society* 11(1). Article 25.

Surányi, Balázs. 2016. Discourse-configurationality. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 422-440. Oxford: Oxford University Press.

van der Wal, Jenneke. 2009. *Word order and information structure in Makhuwa-Enahara*. Leiden: Leiden University (PhD dissertation).

van der Wal, Jenneke. 2012. Subject agreement and the EPP in Bantu agreeing inversion. *Cambridge Occasional Papers in Linguistics* 6. 201-236.

van der Wal, Jenneke. 2015a. Bantu syntax. In *Oxford handbooks online*. Oxford: Oxford University Press.

van der Wal, Jenneke. 2015b. A note on the (non-existing) passive in Matengo. *Linguistique et langues africaines* 1. 81-98.

van der Wal, Jenneke. 2016. Diagnosing focus. *Studies in Language* 40(2). 259-301.

van der Wal, Jenneke. 2022. *A featural typology of Bantu agreement*. Oxford: Oxford University Press.

Watters, John Robert. 1979. Focus in Aghem. In Larry Hyman (ed.), *Aghem grammatical structure*, 137-197. Los Angeles, CA: University of Southern California.

Yoneda, Nobuko. 2011. Word order in Matengo (N13): Topicality and informational roles. *Lingua* 121(5). 754-771.

Zerbian, Sabine. 2006. *Expression of information structure in the Bantu language Northern Sotho*. Berlin: Humboldt-Universität (PhD dissertation).