Bantu word order between discourse and syntactic relations

Allen Asiimwe (Makerere University), Patrick Kanampiu (University of Edinburgh), Elisabeth Kerr (LUCL), Zhen Li (LUCL), Amani Lusekelo (University of Dar es Salaam), Ernest Nshemezimana (University of Burundi) & Jenneke van der Wal (LUCL)[1]

1. Introduction

Discourse function has often been noticed to be a strong factor in conditioning Bantu word order (Downing & Hyman 2015, Downing & Marten 2019). The importance of discourse function in 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 – immediately before the verb (IBV) in (2):

(1)   (What has happened on the bridge?)
        A-ha-ru-tindó   ha-a-rabá=hó   e-mótóka   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)   (Who attacked the hunter?)
        Mu-bhií   kí-mbúlí   kí-siim-i.
        1-hunter 7-lion   7SM-attack-PST
        '[The lion]FOC attacked the hunter.'   [Teke-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, 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.

These two positions on Bantu word order fit in the 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 and Hajičová 1973, Hajičová 1983, Sgall 1984), later termed 'discourse configurational (see 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?'. Öhl (2010: 251) 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 (2006) 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.

Considering the described word order flexibility for Bantu languages as demonstrated by phenomena such as locative inversion and focus positions,[2] we want to determine where the Bantu languages are on this continuum between discourse roles and grammatical roles in their effect on word order, and therefore take seriously the significance of information structure. We thus switch to a discourse-configurational approach, asking How far can we get in describing Bantu word order without reference to syntactic roles? This reframing of research on word order allows us to to pinpoint precisely where we do need syntactic relations. The results show Bantu-internal microvariation, with some languages heavily discourse-based and others more affected by grammatical role, highlighting the need for more detailed investigation of the grammars of Bantu languages, and arguing against a one-size-fits-all account of word order.

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. Such data is unfortunately often not available from the existing grammatical descriptions of Bantu language. As part of the BaSIS project we therefore gathered new fieldwork data focusing on syntax and information structure, using our own project methodology alongside the Questionnaire on Information Struture (Skopeteas et al., 2006) and the elicitation of natural speech. This talk presents results from our fieldwork studies on 8 of the languages in the project, spoken across sub-Saharan Africa: Tunen (A44, Cameroon), Teke-Kukuya (B77, Republic of Congo), Kîîtharaka (E54, Kenya), Kirundi (JD62, Rwanda), Rukiga (JE14, Uganda), Kinyakyusa (M31, Tanzania), Makhuwa (P31, Mozambique), and Copi (S61, Mozambique).[3] The geographical distribution of these languages is shown in the map in Figure 1 below.

---

[2] Note that the word order is only described as 'flexible' from the point of view of grammatical roles, not that of discourse roles!

[3] Guthrie classifications are given alongside language names and locations, from Maho (2009).
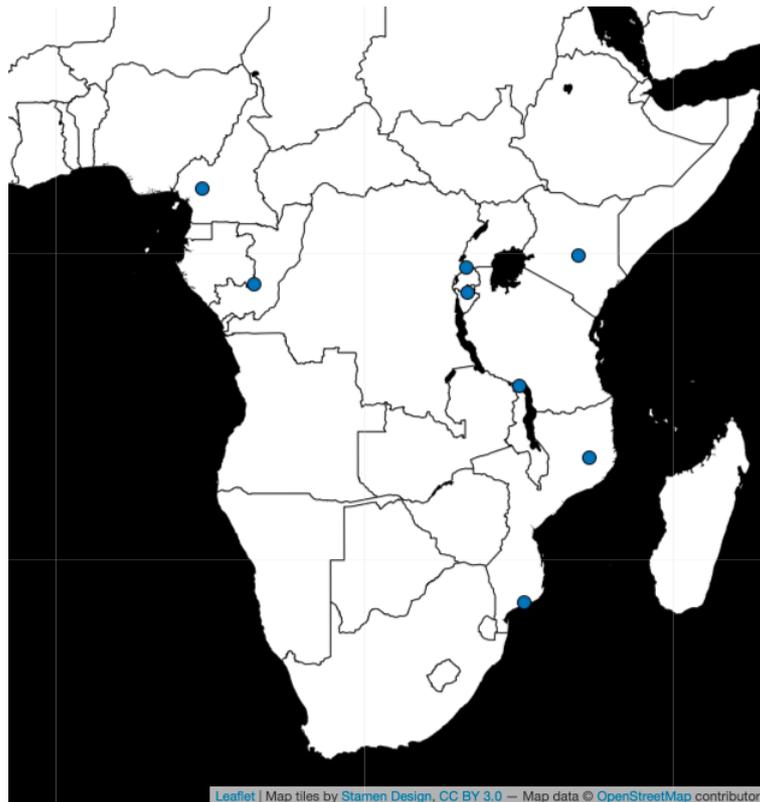
Figure 1. Map of languages in the BaSIS sample (co-ordinates from Glottolog).

For each of these BaSIS languages, we checked three factors of discourse-configurationality in word order (leaving aside asymmetries in other domains, such as agreement), with subquestions that serve as checkpoints to set the parameter (inspired by the SOAS project 'Morphosyntactic variation in Bantu'). For focus, we considered three diagnostics: wh words (which are inherently focused), answers to wh questions (simple/information focus), and modification by the focus-sensitive particle 'only' (exhaustive focus).

BaSIS word order parameters:

1. **Is there a dedicated focus position?**
    Checkpoints:
    a. Can the recipient and the theme be questioned in their canonical position?
    b. Can the recipient and the theme be an answer to a wh question in their canonical positions?
    c. Can the recipient and the theme be modified by 'only' in their canonical positions?
    d. Can adverbs be questioned in their canonical position?

2. **Is the preverbal domain reserved for topics?**
    Checkpoints (for otherwise unmarked clauses, i.e. no cleft):
    a. Can preverbal arguments be questioned?
    b. Can preverbal arguments be the answer to a wh question?
    c. Can preverbal arguments be modified by 'only'?

d. Can thetic sentences be expressed as SV(O)? (where S is not an always-available referent)?

e. Are topical arguments typically or preferrably expressed preverbally?

3. **Is there symmetry between syntactic roles relating to word order?**
   Checkpoints:
   a. Can the subject be questioned in the same position as the object?
   b. Can the subject as the answer to a subject wh question occupy the same position as the object in the answer to an object wh question?
   c. Can the subject be modified by 'only' in the same position(s) as the object?

The answers for each language are given in Table 1 below. The colour coding indicates whether the answer provides evidence for the influence of grammatical roles (yellow) or discourse roles (blue) on word order. A column is added to show what the parameters would be for a language that entirely relies on grammatical roles (GR) or discourse roles (DR).

Parameter	Checkpoint	DR	GR		Tunen	Kukuya	Tharaka	Kirundi	Rukiga	Kinyakyusa	Makhuwa	Copi
1	a	N	Y		-	N	Y	N	N	Y	N	Y
	b	N	Y		-	N	Y	N	Y	Y	N	Y
	c	N	Y		Y/-	Y	Y	N	Y	Y	N	Y
	d	N	Y		Y	Y/N	Y	Y/N	Y/N	Y	N	Y
2	a	N	Y		Y	Y*	N	N	N	Y?	N	N
	b	N	Y		Y	Y*	N	N	N	Y?	N	N
	c	N	Y		Y	Y*	N	N	N	Y	N	Y/N?
	d	N	Y		Y	Y	Y	Y	Y	N	N	N
	e	Y	N		Y	Y	Y	Y	Y	Y	Y	Y
3	a	Y	N		N	Y	Y	Y	Y	Y	N	Y
	b	Y	N		N	Y	Y	Y	Y	Y	N	Y
	c	Y	N		N	Y	Y	Y	Y	Y	N	Y

Table 1. Parameter settings for word order in the BaSIS languages.

Key: blue shading = evidence for discourse roles (DR) conditioning word order, yellow shading = evidence for grammatical role (GR) conditioning word order (a column for DR and GR is giving to make clear what the parameter settings would be for a purely DR-conditioned or a purely GR-conditioned language).

*Yes, but only in IBV position, and therefore still evidence for IS-conditioned word order.

3. Observations

From the table we can make at least the following observations.

3.1. Cross-linguistic variation

There is a lot of variation between the languages in our sample and indeed cross-Bantu. We have to stop making claims about word order in 'the Bantu languages' as if they were a monolithic entity. As this project only uses a sample of 10 different Bantu languages out of approximately 555 languages in the family (Hammarström 2019), further research is needed in order to make reliable generalisations.

3.2. Universal tendency for preverbal topics

Answers to parameter 2e were all Y - all languages show a preference for topics in the preverbal domain (see Gundel's 1988 generalisation d: "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."

3.3. Variation in focus positions

There are three languages with a dedicated focus position, but all are different: immediately before the verb in Kukuya, immediately after the verb in Makhuwa, and sentence-final in Kirundi.
Why? Link to grammaticalization/other parts of these languages' grammars?

- Kukuya's IBV focus position also means that parameter 2 is interpreted differently: checkpoints 2abc are 'yes' but coloured blue, as they aren't arguments for the influence of grammatical roles. Similarly for 2d, the fact that the subject can be preverbal in a thetic sentence cannot be interpreted as evidence for one or the other type of roles, as the subject may be in the focus position.
  - Kukuya's IBV focus position is marked, based on typological work that states that SVO languages are more likely to have postverbal focus positions (Kidwai, 1999; Horvath 1986)
  - The existence of IBV focus positions in the family may mean that parameter 2 needs to be tweaker to exclude focus positions

3.4. Variation within parameters

The checkpoints align for parameter 3 (within a language, all answers are Y or all answers are N), but not for parameters 1 and 2. This can point to the need to rethink the parameters and checkpoints, but some interesting exceptions can be understood in the following way:

- 1a for Rukiga shows that wh questions may behave differently from other focused NPs. A wh needs to be in the IAV position whereas other focus elements may appear in other positions . For instance, the answer (in b) to the wh question posed in (a) is not in the IAV but its canonical position.

(3) a.   Waaha kí Jéini?
       Wa-a-ha   ki   Jeini?
       2SG.SBJ-N.PST-give-FV   what   1.Jane
       'What have you given Jane?'

b.   Naaha Jéín' énkofiira.
     Na-a-h-a   Jeini   e-n-kofiira
     1SG.SBJ-N.PST-give-FV   1.Jane AUG-9-hat
     'I have given a hat to Jane.'   [Rukiga]

- 1d for Kukuya, Kirundi, and Rukiga shows that not all adverbs behave the same way. Interestingly, it seems that the set of interrogative adverbs being flexible in position is not the same across the languages (i.e., *how*, *when*, and *where* are each specified separately in each language). For example, the adverb 'when' in Rukiga is not restricted to the IAV position (4) but 'how' is (if necessary achieved by left-dislocating an intervening object):

(4) a.   Okaza Kampalá ryári?
          o-ka-za   Kampala   ryari
          2SG.SM-F.PST-go 23.Kampala when
          'When did you go to Kampala?'

     b.   Okaza ryarí Kampála?
          O-ka-z-a   ryari   kampala
          2SG.SM-F.PST-go when   23.Kampala
          'When did you go to Kampala?'   [Rukiga]

(5) a.   Obutúmwa bukahika búta purésidenti?
          O-bu-tumwa   bu-ka-hik-a   bu-ta   puresidenti?
          AUG-4-message 14-F.PST-reach-FV   14-how   1.president
           'How did the message reach the President?'

     b.   *Obutúmwa bukahika purésidenti búta?

     c.   Purésidenti óbutúmwa bukamuhika búta?
          Puresidenti   o-bu-tumwa   bu-ka-mu-hik-a   bu-ta
          1.Puresidenti   AUG-4-message 14-F.PST-1OM-reach-FV   14-how
          'As for the president, how did the message reach him?'   [Rukiga]

- 2d for Kîîtharaka, Kirundi and Rukiga shows that non-topical transitive subjects are allowed preverbally, while focused arguments are banned. This suggests that the preverbal domain in these languages is not restricted to topics, but can be characterised as 'non-focal'.
- Kinyakyusa is exceptional for an eastern Bantu language in allowing preverbal focus. This appears to be an areal innovation:
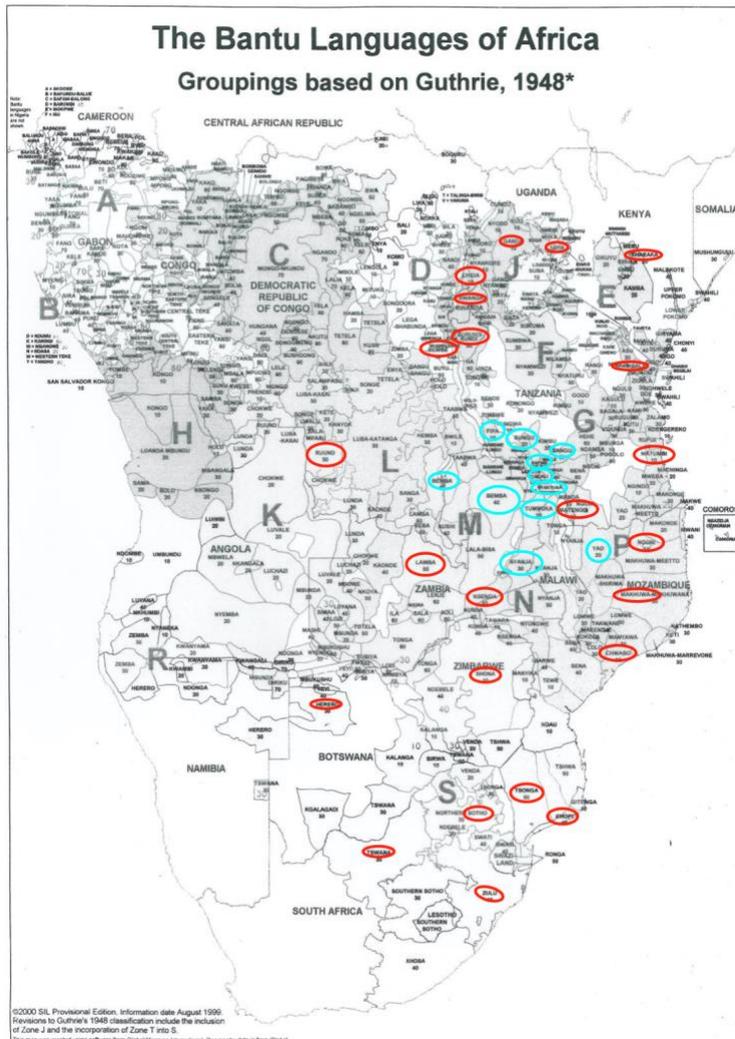
Figure 1 – Preverbal focus: blue = allowed, red = ban

3.5. Tunen as grammatical role-oriented

Tunen is exceptional in showing hardly any influence of information structure on word order, with only parameter 2e showing evidence for the influence of discourse role (as in all the languages; section 3.2) and all other parameter values showing evidence for grammatical role.[4]

For example, the results for parameter 3 for Tunen indicate that Tunen has word order asymmetries conditioned by grammatical role. Evidence for such an asymmetry is shown in the fragment answers below. Subject fragments must be preceded by the copula/focus marker á (ex6), but object fragments cannot be preceded by á (ex7). Non-arguments pattern with objects, i.e. are not preceded by á (ex8).

(6) Context: Which politician died?
   *(á) Píɛ̀lə̀
   /á Piɛlə/

---

[4] Parameter 1a/1b/1c are not filled in due to insufficient data; COVID-19 has unfortunately delayed follow-up research here.

FOC Pierre
'Pierre.'   [EO 270]

(7) Context: What is the man holding?
     (*á) kàlɔ́t
     /á kalɔtɔ/
     FOC 7.carrot
     'Une carrotte.' / 'A carrot.'   [PM 1266, 1267]

(8) Context: From the tree ?
     ìndínə́ yɛ̀ bɔ̀lɛ̀, ɛ́ɛ̀
     /ɛ-ndinə yɛ bɔ-lɛa ɛɛ/
     7-foot ASSOC.7 14-tree yes
     'Au pied de l'arbre, oui.' / 'At the foot of the tree, yes.'   [EO 591]

The difference in word order of the focussed argument and the focus/copular marker á show a subject/non-subject asymmetry, meaning that grammatical role has more effect on Tunen word order than information structural status (the fragment answers in (1) and (2) are both information focus contexts).

This exceptional lack of influence of information structure could be explained due to the position of Tunen at the Northwest of the Bantu family. It has been argued by Mous (1997) that the position of the object in Tunen is an innovation, with the word order having changed from the SVO found in other Bantu to SOV in Tunen. The innovative nature of Tunen's word order could explain why it patterns differently from the other languages in the BaSIS sample. Furthermore, the lack of inversion constructions in Tunen may be an areal pattern for languages of the North-Western region (see Hamlaoui 2018, Hamaloui & Makasso 2015).

4. Discussion

Working with parameters to compare languages has methodological pros and cons. Pros of such an approach is that it aids comparison between languages in abstracting from the raw data and giving an overview of the different patterns.[5] The parameter list also guides researchers in how to investigate the influence of discourse roles and grammatical roles on word order, which can help ensure that future fieldworkers pay attention to discourse roles when working on a language. Cons of such parametric overviews are that they can be misleading if the parameters are not well-defined and accurately coded, and that some language patterns may be too complicated to distil into a binary value of "Y"or "N".

In order to mitigate against these cons, we developed the parameters as a group and revised them if they were not clear or if a language did not fit them. By using checkpoints rather than single parameter settings, we avoided the risk of collapsing too much variation into a binary "Y" or "N" value, and we allowed a non-binary "Y/N" value in cases which showed mixed behaviour. The use of checkpoints turned out to be sensible, as parameters 1 and 2 both showed a mix of Y's and N's for many languages. As shown in section 3 above, these mixed results led us to better understanding of the different languages, for example in the different

---

[5] This does not mean that the raw data should be ignored – see the appendix for the full list of data on which the parameter values were based.

behaviour of wh-items and non-wh-items in Rukiga. When there is no clear language-internal explanation for mixed results, there is evidence that that parameter should be split into two different parameters in future work.

Problématique of the parametric approach:
- Criticisms of this kind of parametric approach: Bickel (2015), Evans (2020); cf Bickel (2010), Haspelmath (2018)
- Alternative approaches: Round & Corbett (2020), Bickel (2010, 2015)
- "[Typology] introduces a permanent tension between generalisation and particularisation in typological research" (Evans 2020:418)
- "Unlike description of languages, comparison is not a goal in itself. It always serves some other goal, such as learning about human language in general or answering questions about the historical origin and development of languages." (Haspelmath 2018:83)
- "There are many things that can usefully be compared across languages but each language also has highly idiosyncratic features that cannot be readily compared." (Haspelmath 2018:94)
- "General category fallacy
  We do not learn anything about particular languages merely by observing that category A in language 1 is similar to category B in language 2 or by putting both into the same general category C" (Haspelmath 2018:84)
- "on the present view of comparative linguistics, what we compare is not language systems (which are incommensurable) but "the phe- nomena of languages" (Haspelmath 2018:109)
- "Bickel (2010), in his influential deconstruction of grammatical relations typology, essentially proposes moving the problem down from grammatical relations to more oper- ationalisable and hopefully cross-linguistically comparable elements – e.g. replacing hard-to- compare concepts like 'subject', which ostensibly assemble a class of linked behaviours in their morphology, syntax, semantic and information-structure properties, with behaviour on more specific levels such as agreement, case assignment, quantifier scope, switch-reference and so on." (Evans 2020:422)

Examples of generative approach to parameters:
- Examples of generative approaches to typology: Biberauer et al
  o Notion of parameter hierarchies. For us, too early to say whether there are interdependencies within the BaSIS parameters
- Builds on earlier work in structuralism, Greenbergian typology
- McCloskey 1997, Baker 2010, Homberg 2010 generative typology

As expected based on prior research by Morimoto (2006, 2008) and Öhl (2010), we found that no language's word order was at the extreme of fully-conditioned by discourse roles, nor fully conditioned by grammatical roles. Instead, languages lie at different points of a continuum. This means that it is not enough to describe the word order of Bantu languages in terms of information structural classifications such as "topic-nontopic", even in languages with constructions such as locative inversion and focus positions. However, the classical treatment

of languages in terms of "SVO", "SOV" and so on similarly fails to capture the full picture of word order in these languages. We therefore advocate a language-specific classification of the word order of each language.

Figure 1. Diagram representing how the languages in the BaSIS sample are at different positions on a continuum between grammatical roles and discourse roles influencing the word order. Note that no language is found at the extremes.

Although we can clearly see variation between heavily grammatical role-oriented languages and heavily discourse role-oriented languages, the examples in Figure 1 are left abstract (represented by "x") rather than using language names, as the exact position on the continuum is influenced by the weighting of the parameters – it would only be an assumption to suggest our current parameters should be evenly weighted – and the current project only investigated word order. Further work should include other domains, such as agreement (cf Morimoto 2006, 2008) in order for a more complete picture to emerge.

From a typological perspective, the fact that all languages had a "Y" for parameter 2e supports Gundel (1988)'s generalization d "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." This result therefore seems to be indicative of word order patterns that generalise outside the Bantu family.

This type of research 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 shown above. Interestingly, the centre of the Bantu languages (around 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 to appear), see also Marten et al. (2007) on Bemba and Chichewa as the centre. Another example is inversion constructions – as Hamlaoui (2018) states (see also Hamlaoui & Makasso 2015), languages of the North-Western region (including Tunen and Kukuya) typically have no inversion constructions.

5. Conclusion

A very obvious conclusion is that our systematic overview confirms the interesting microvariation within the Bantu languages: no two of our languages show the same profile.

A second conclusion is that information structure is indeed a significant influence in the grammar of many of these languages. In order to fully understand what determines word order, then, the context needs to be taken into account for every individual language. The methodology proposed gives a framework for this, and may inspire research on discourse-configurational languages beyond Bantu.

This also means that theoretical models need to capture the variation and the influence of both information structure and grammatical role in the grammar. There are two necessary questions that need to be answered on an individual basis, which may then inspire a more encompassing theory:

A. Which precise notions are active in a given language? For Makhuwa, both 'topic' and 'focus' are needed to capture word order: the preverbal domain is reserved for topics, and there is a dedicated position for focus. For Rukiga, perhaps only 'focus' is necessary, describing the preverbal domain as 'non-focus'. Or perhaps different notions are needed (contrast, unexpectedness, …).

B. What type of features correspond to these notions? Semantic or syntactic features? A or A-bar features? The theoretical status of information structural features may be different from language to language: not in all languages are these A-bar features active in the left periphery (as is typically assumed in cartographic analyses), but they may be active in the A-syntax too, influencing licensing and agreement.

Naturally, this is only a first step in a work-in-progress, and there are plenty of areas of future work:

- As word order has been claimed to be intimately linked to agreement, the discourse-configurationality of agreement is a next step, for example determining whether the subject marker is perhaps better captured as a topic marker (Morimoto 2006).
- Apart from agrement, other morphological marking of information structure is also relevant in determining the discourse-configurationality of a language (Gibson et al. 2017).
- What can explain the different focus positions as found in Kukuya (IBV), Makhuwa (IAV) and Kirundi (final)?

Abbreviations

Bare numerals (1, 2, 3, etc.) refer to Bantu noun class; when followed by SG/PL they indicate person.

ASSOC	associative marker
AUG	augment (pre-prefix)
FOC	focus marker/copula
N.PST	near-past tense
PST	past tense
SM	subject marker

Acknowledgements

References

Bickel, Balthasar 2015. Distributional typology: Statistical inquiries into the dynamics of linguistic diversity. In: Heine, Bernd; Narrog, Heiko. *The Oxford handbook of linguistic analysis*, 2nd edition. Oxford: Oxford University Press, pp. 901 - 923.

Bickel, Balthasar. 2010. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 399–444. Oxford: Oxford University Press.

Downing, Laura J., and Larry M. Hyman. (2015). Information structure in Bantu. In Féry, Caroline & Shinichiro Ishihara. *The Oxford handbook of information structure*. Oxford: Oxford University Press.

Downing, Laura, and Lutz Marten. 2019. Clausal morphosyntax and information structure. In Van de Velde, Mark et al. (eds), *The Bantu Languages, second edition*, 270-307. London: Routledge.

É. Kiss, Katalin, ed. 1995. *Discourse Configurational Languages*. New York: Oxford University Press.

Evans, Nicholas. 2020. Introduction: Why the comparability problem is central in typology. Linguistic Typology 24(3). 417–425. DOI: https://doi.org/10.1515/lingty-2020-2055.

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

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

Hamlaoui, Fatima. 2018. Proto Bantu word order. Presentation at BGR2, Ghent.

Hamlaoui, Fatima, and 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 Van de Velde, M. et al. (eds.), *The Bantu Languages, second edition*, 17-78. London: Routledge.

Haspelmath, Martin. (2018). How comparative concepts and descriptive linguistic categories are different. In Van Olmen D., Mortelmans T., & Brisard F. (Eds.), Aspects of Linguistic Variation (pp. 83-114). Berlin; Boston: De Gruyter. Online file: http://www.jstor.org/stable/j.ctvbkjwxf.6.

Li, Charles N., and Sandra Thompson. 1976. Subject and Topic: A New Typology of Language. In Charles N. Li (ed.), *Subject and Topic*, 457-90. New York: Academic Press.

Maho, Jouni F. 2009. NUGL Online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages. *Online file: https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf* (version dated 04 June 2009).

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

Morimoto, Yukiko. 2000. Discourse Configurationality in Bantu Morphosyntax. Stanford University: Doctoral Dissertation.

Morimoto, Yukiko. 2006. Agreement properties and word order in comparative Bantu. *ZAS Papers in Linguistics* 43. 161-188.

Nurse, Derek and Gérard Philippson. 2003. Introduction. In Nurse, D. and G. Philippson (eds.), *The Bantu Languages*, 1-12. London: Routledge.

Öhl, Peter. 2010. Formal and Functional Constraints on Constituent Order and Their Universality. In Carsten Breul (ed.), *Comparative and Contrastive Studies of Information Structure*, 231-275. Amsterdam: John Benjamins.

Round, Erich R. and Corbett, Greville G. 2020. "Comparability and measurement in typological science: The bright future for linguistics" *Linguistic Typology* 24 (3), pp. 489-525, DOI: https://doi.org/10.1515/lingty-2020-2060

Skopeteas, Stavros, Fiedler, Ines, Hellmuth, Sam, Schwarz, Anne, Stoel, Ruben, Fanselow, Gisbert, Féry, Caroline and Manfred Krifka. 2006. *Questionnaire on Information Structure: Reference Manual.* ISIS Volume 3. Potsdam: Universitätsverlag Potsdam.

Van der Wal, Jenneke. To appear. A featural typology of Bantu agreement. OUP.

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

Appendix

RStudio BaSIS languages map:

Map created using *lingtypology* package in R, which takes language coordinates from Glottolog. In Glottology, Chiga is a dialect of Runyakore-Rukiga, Copi as a dialect of Chopi, and Nyakyusa as a dialect of Nyakusa-Ngonde.

```
map.feature(languages=c("Tunen", "Tharaka", "Teke-Kukuya", "Makhuwa", "Chiga", "Chopi",
"Rundi", "Nyakyusa-Ngonde"), tile = "Stamen.Toner")
```