

Morphosyntactic variation in Bantu: The case of Setswana*

Nancy C. Kula
University of Essex

Lutz Marten
SOAS University of London

Within the context of microvariation in Bantu, three processes are examined in Setswana – object marking, inversion constructions and diminutive marking. Setswana morpho-syntactic structures for these constructions present instances that distinguish Setswana from Eastern Bantu languages and yet also cases of similarity with patterns more commonly attested in Eastern rather than Southern Bantu languages. Using a comparative-parametric approach Setswana is shown to pattern with Southern Bantu in its default agreement and diminutive marking patterns while its flexibility in allowing multiple object markers with free ordering is more closely related to Eastern Bantu, suggesting complex historical patterns of contact.

Keywords: Setswana, Bantu languages, object marking, inversion constructions, diminutive marking

1. Introduction

Over the last few decades, research in Bantu morphosyntax has much increased and flourished. There have been several works on different linguistic topics related to Bantu languages, several studies adopting typological, formal, or comparative-historical approaches (see e.g. Downing and Marten 2019 for an overview), as well as studies focussing on individual Bantu languages, including Setswana (e.g. Chebanne, Creissels & Nkhwa 1997, Chebanne 2003, Otlogetswe & Chebanne 2018), which is also the focus of the current paper.

Bantu languages exhibit several broad typological similarities such as dominant SVO constituent order, the presence of complex noun class systems, and in many instances highly agglutinating verbal morphology. However, within this overall uniformity, there is a wide range of smaller-scale differences, resulting in numerous patterns of microvariation through historical processes of contact and change (see e.g. Bostoën et al. 2023, Bloom Ström et al. forthcoming, Marten et al. forthcoming).

In this paper, we are focussing on three such domains of micro-variation, namely object marking, inversion constructions, and diminutive marking. Against the background of cross-Bantu variation and adopting a broadly comparative-parametric approach, we will situate the specific morphosyntactic features of Tswana in these three areas. The results of the analysis provide a snapshot of morphosyntactic variation in Bantu and Tswana morphosyntactic structure.

* We are grateful to Denis Creissels for detailed comments and suggestions on an earlier draft of this paper, as well as to two anonymous reviewers and the editors of the special issue. Financial support from the Leverhulme Trust for Research Project Grant ‘Morphosyntactic variation in Bantu: contact and change’ (RPG-2014-208) is hereby gratefully acknowledged.

2. Object marking

Object marking is well-researched area in Bantu, and it is also related to cross-linguistic studies of pronominal, clitic, and agreement systems more widely. Bantu object marking has long been noted as being both complex and rich, and there are several dimensions of cross-linguistic variation (e.g., Polak 1986, Beaudoin-Lietz et al. 2004, Riedel 2009, Marten & Kula 2012, Zeller 2014, Marlo 2015, Taji 2017, van der Wal 2022): Morphologically, object markers may be restricted to one per inflected verb form in some languages, while in others more than one object marker is possible. There is also variation in terms of the position of object markers as occurring before or after the verb root, and with respect to any order restrictions on multiple object markers. Syntactically object markers sometimes fulfil anaphoric function, and so replace the lexical NP they refer to, or alternatively may fulfil (grammatical) agreement function, where both the object marker and a co-referential NP can (or sometimes have to) co-occur. A cross-linguistically unusual aspect of Bantu object marking – and of the notion of grammatical object more widely – is that pre-stem object marking can also extend to grammatical functions which are not normally seen as typical objects, such as locatives or instruments, often, but not always in conjunction with applicative morphology (Schadeberg 1995).

Against this background, Marten & Kula (2012) develop six variables or surface parameters to capture part of the variation found among Bantu languages, with a specific focus on Eastern and Southern Bantu languages. The parameters are summarised in (1):

- (1) Parameters of variation in Bantu object marking (following Marten & Kula 2012: 239)
1. Can the object marker and the lexical object NP co-occur?
 2. Is an object marker obligatory with particular object NPs?
 3. Are there locative object markers?
 4. Is object marking restricted to one object marker per verb?
 5. Is the order of multiple object markers free or fixed?
 6. Can either benefactive or theme objects be expressed by an object marker in double object constructions?
 7. Is an object marker required/optional/disallowed in object relatives?

In (1) we have modified and increased the parameters from Marten & Kula (2012) slightly by including a separate parameter 5 (relating to the order of multiple object markers). Marten & Kula (2012) discuss this as a sub-case of the number of object markers, but since the issue of order is relevant for Tswana, we have included it here as a separate case.

The first two parameters in (1) relate to the co-occurrence of object markers and co-referential object NPs in the same syntactic or prosodic domain (i.e. excluding dislocated or topicalised object NPs). In some Bantu languages, such as Swahili, object marker and object NP can, and in some cases must, co-occur. However, in Tswana, the two are in complementary distribution and the object marker and the object cannot co-occur in the same domain, and interact with the system of conjoint and disjoint verb forms:¹

¹ All data in this paper are taken from various sources in the literature, as indicated. We have provided consistent glosses as far as possible, but we have not changed the representation of the actual examples.

(2) kè rát-á Mphó
 SM1SG like-FV.CJ Mpho
 ‘I like Mpho’ (conjoint)
 (Tswana, Sengwato, McCormack 2006)

(3) kè à mó-rát-à
 SM1SG DJ OM1-like-FV.DJ
 ‘I like him’ (disjoint)
 (Tswana, Sengwato, McCormack 2006)

The example in (2) shows that when an object NP such as *Mphó* is present, no object marker is used with the verb. The verb form itself is in the conjoint form, which indicates that the following constituent is part of the comment. Examples like (2) are thus typically used when the object is not topical and is introduced as (part of) new information. On the other hand, when the object is old information and is referred to anaphorically, it can be expressed by an object marker as in (3). In this case, no overt object NP is possible, and the verb is in the disjoint form, indicating the verb is in the final position of the clause (except for adjuncts forming part of the comment). Any co-referential NP following the verb is marked prosodically and in terms of information structure as an ‘afterthought topic’, as in (4).

(4) kè à mó-rát-à Mphó
 SM1SG DJ OM1-like-FV.DJ Mpho
 ‘I like him, Mpho’ (disjoint)
 (Tswana, Sengwato, McCormack 2006)

In terms of the parameters in (1), then, for Tswana for both the first and the second parameter, the answer is ‘no’ – the object marker and the lexical NP cannot co-occur in the same domain, and object markers are never obligatory with a particular NP.

The third parameter in (1) is related to locative object markers, and is simply concerned with whether they exist in a given language. While in some Bantu languages, there are no (pre-verbal) locative object markers, in others one or more are found. In part, this is related to the presence of locative noun classes in the language, and to what extent the main Proto-Bantu locative classes (classes 16-18) are reflected in the system. In Tswana, the locative system has been much restructured. An innovative locative suffix *-(i)ng* is found in the nominal domain, while in the verbal domain only one locative agreement prefix *gò-* is found, corresponding to the reconstructed Proto-Bantu locative class 17. In addition, there are three different locative adverbs *fá-*, *gó-*, and *mó-*, corresponding to noun classes 16, 17, and 18 and semantically typically encoding proximate, distal and interior locations, respectively (Cole 1955: 341). These adverbs have developed from locative demonstratives and can also function as locative prepositions (Creissels 2011, forthcoming). For locative object marking this means that there is only one locative marker, namely *gò-* as seen in (5).

Unfortunately, this means that there are some inconsistencies in data representation, and that not all examples have tone marking of a distinction between all seven vowels of Tswana.

- (5) ke a gò itse
 SM1SG DJ OM17 know
 ‘I know it (there, the place)’
 (Marten et al. 2007: 330)

As the example in (5) shows, locative object marking is possible in Tswana, making use of the class 17 locative object marker *gò-*.

The next two parameters are concerned with the number and order of object markers. While some Bantu languages only have one object marker slot per verb (e.g. Chichewa), others allow more than one (pre-verbal) object marker. Tswana belongs to this second group as (7) shows, although Cole (1955: 431), from where the examples are taken, notes that structures with three objects like (6) and (7) are ‘rare’.

- (6) ke-tla-kwal-êl-êl-a ngwana ba-tsadi lo-kwalô
 SM1SG-FUT-write-APPL-APPL-FV 1.child 2-parents 11-letter
 ‘I will write a letter to the parents on behalf of the child’ (Cole 1955: 431/2)
- (7) ke-tla-lo-ba-mo-kwal-êl-êl-a
 SM1SG-FUT-OM11-OM2-OM1-write-APPL-APPL-FV
 ‘I will write it (a letter) to them (the parents) on behalf of him/her (the child)’ (Cole 1955: 431/2)

The three post-verbal object NPs in (6), some of which are introduced by the applicative morphemes in the verb form, are expressed by pre-verbal object markers in (7). The example thus shows the possibility of having more than one object marker in an inflected verb, at least in terms of morphosyntactic restrictions. It would be interesting to investigate further the semantic and pragmatic constraints on the use of multiple object markers in Tswana, and other Bantu languages.

Another relevant aspect of Tswana grammar is the possibility to change the order of object markers while maintaining the same reading. Of course, the question only arises for languages that allow multiple object markers in the first place. However, among those, it appears that most languages have a fixed order of object markers (with typically the first-person singular object marker being closest to the verb root, cf. Marlo 2015). In Tswana in contrast, at least in some dialects, such as Segkatla, it seems possible to have two orders:

- (8) ke mo e ape-ets-e
 SM1 OM1 OM9 cook-APPL-PERF
 ‘I cooked him/her it’
 (Tswana, Segkatla, Marten et al. 2007: 330)
- (9) ke e mo ape-ets-e
 SM1 OM9 OM1 cook-APPL-PERF
 ‘I cooked him/her it’
 (Tswana, Segkatla, Marten et al. 2007: 330)

- (15) ?dikwalo tse ke bone-ng ts-one
 10-books REL10 SM1SG.PAST see-REL 10-DEM
 ‘The books which I saw *those* ...’
 (Tswana, Segkatla, Marten et al. 2007: 332)

From a cross-Bantu perspective, some languages are like Tswana in requiring a resumptive pronoun, in others a resumptive pronoun is possible, but not required, while others do not allow the use of a resumptive pronoun at all.

Having discussed the object marking properties of Tswana concerning our seven parameters, we can now consider Tswana (final column of Table 1) in the context of a small set of comparative Bantu languages, following Marten & Kula (2012): Chichewa, siSwati, Otjiherero, Lozi, Bemba, Chaga, Ha, and Swahili. A summary of this comparison is provided in Table 1.

Table 1: Variation in object marking in 9 Bantu languages (Chichewa, siSwati, Otjiherero, Lozi, Bemba, Chaga, Ha, Swahili, Tswana)

	Chew	Swati	Her	Lozi	Bemb	Chag	Ha	Swah	Tswa
1 OM-objNP	*	*	*	✓	✓	*	✓	✓	*
2 OM required	*	*	*	*	*	✓	?	✓	*
3 OM loc	✓	*	✓	*	✓	✓	✓	✓	✓
4 Multi OM	*	*	*	*	(✓)	✓	✓	*	✓
5 Free order	*	*	*	*	*	*	*	*	✓
6 Sym OM	*	✓	✓	✓	*	✓	?	*	✓
7a Res oblig	✓	✓	*	*	*	?	*	*	✓
7b Res opt	*	*	*	*	*	?	✓	✓	*
7c No Res	*	*	✓	✓	✓	?	*	*	*
Summary	2	2	3	3	4	4	4	4	5

In Table 1, ‘yes’ values (marked with a tick ‘P’) are shaded, and in the ‘Summary’ row we have counted all the ‘yes’ values for the given language. This might be seen as an interesting matrix for comparing the similarities between the different languages. However, we do not mean to imply that there is a hierarchy or a cline along which all the properties can be related. The first two parameters are related primarily to the difference between (grammatical) agreement marker and (incorporated) pronoun, so maybe ‘yes’ values can be seen as indicating more grammaticalized object markers – and so Tswana object markers would be comparatively less grammaticalized. The next four

parameters – locative object markers, number and order of object markers, and symmetrical object marking – could be thought of as more or less restricted object marking systems, where ‘yes’ values indicate more complex or liberal systems. With four ‘yes’ values Tswana is the most ‘liberal’ system under this interpretation. Finally, parameter 7 is split into three possible systems, which, as noted above, probably reflect relative clause syntax more directly than object marking systems. We might note in passing, though, that in this little mini-typology, all three languages that require a resumptive object marker (‘yes’ in 7a) also have ‘no’ in parameters 1 and 2, and thus have fewer grammaticalized object markers (cf. Henderson 2006 for more discussion of this point).

Having said that, looking at the different distribution of values in the summary row, it is interesting to note that Tswana patterns with the Eastern Bantu languages Bemba, Chaga, Ha, and Swahili, all of which have four or five ‘yes’ values. In contrast, the Southern Bantu languages in the sample, Swati, Herero, and Lozi are at the lower end of the pattern, with two or three ‘yes’ values. Particularly striking is the contrast with Lozi, which is often seen as a Sotho-Tswana language (e.g. Gowlett 1989), as well as with siSwati which is often seen as closely related. Of course, the typological space we have explored here is small – based on only seven parameters – but it is nevertheless tempting to draw a tentative conclusion that morphosyntactic similarity does not necessarily correspond to genetic relationships. This in turn raises interesting questions for language contact and change, and the roles of phonology and the lexicon in this on the one hand, and morphosyntactic structure on the other.

3. Inversion constructions

Our next empirical domain is the area of inversion constructions. This is a construction type that has been widely attested in Bantu languages, and which has attracted considerable scholarly attention (e.g. Bresnan & Kanerva 1989, Harford 1990, Demuth & Mmusi 1997, Rugemalira 2004, Marten 2006, Marten & van der Wal 2014, Diercks 2010, 2011, Zeller 2013, Guérois 2014, Zeller & Ngoboka 2018). There is considerable variation in the details of the construction between different Bantu languages, and it thus presents a good area for comparative study.

The main hallmark of the construction is that the logical subject or agent of the clause is found in the non-canonical post-verbal position in an otherwise SVO language and that the verb either shows some sort of default agreement or that a non-agentive NP or PP, such as a locative phrase, is found in preverbal position. Pragmatically, the construction typically functions to express presentational focus. Against this uniformity, three dimensions of variation are usually identified: 1) The status of the pre-verbal phrase as either grammatical subject or topic, and the associated agreement relation between the pre-verbal phrase and the verb as either grammatical agreement or default agreement; 2) thematic restrictions on the predicate type which licenses inversion (e.g., unaccusative, transitive, etc.); and 3) thematic restrictions on the pre-verbal phrase (e.g., locative, instrument, etc.).

Due to the seminal work of Demuth & Mmusi (1997) on locative inversion, inversion constructions in Tswana have been well documented, and the examples in (15) provide relevant evidence:

- (16) a. Mó kéreké-ng gó ópélá ba-sádi.
 18LOC 9.church-LOC SM15/17 sing 2-woman
 ‘In the church there are women singing.’ (Lit.: ‘There sing women.’)
 (Creissels 2011: 42)
- b. Fá-se-tlharé-ng gó-émé ba-símané.
 16-7-tree-LOC SM17-stand.PRF 2-boys
 ‘By the tree stand the boys.’ (Demuth & Mmusi 1997: 4)
- c. Mó-le-fátshé-ng gó-fúla di-kgomo.
 18-5-country-LOC SM17-graze 10-cattle
 ‘In the country are grazing the cattle.’ (Demuth & Mmusi 1997: 4)

Following work by Creissels (2011, forthcoming), locative and other inversion constructions in Tswana are typically analysed as ‘default agreement’ construction. The pre-verbal, locative phrases are analysed as prepositional phrases, which pragmatically function as (background) topics, and which are headed by one of three locative prepositions (which have historically developed from locative adverbs). Agreement marking on the verb – the subject marker *gó-* – is a default or expletive marker, and not in a grammatical anaphoric agreement relation with the preceding locative phrase. Part of the evidence for this analysis is the absence of class agreement: The locative phrases in (16) are marked by class 16 *fá-* and class 18 *mó-* locative prepositions, but the verbal agreement is the default class 17. Furthermore, default agreement can also be used for expletive constructions as in (17):

- (17) Gó tsámá-flé Mphó.
 SM17 go-PRF.CJ Mpho
 ‘There has gone Mpho.’ (Creissels 1996: 113)

In terms of the first dimension of variation identified above, Tswana thus belongs to the group of languages where pre-verbal phrases are topics, and agreement marking on the verb is default or expletive agreement.

In terms of the second dimension of variation, which relates to thematic constraints on the predicates that can take part in locative inversion, recent work has shown that verbs of different argument structures show different capacities for locative inversion (Bresnan & Kanerva 1989, Harford 1990, Demuth & Mmusi 1997, Marten 2006, Khumalo 2010, Marten & van der Wal 2014, Diercks 2011), providing a fertile area in which to investigate morphosyntactic variation. With respect to this dimension, Tswana belongs to the languages where there are comparatively few restrictions, as shown in Table 2, based on a small sample of eight Bantu languages.

Table 2: Locative inversion variation with respect to predicate type in eight Bantu languages, with Tswana highlighted (*italics: default agreement inversion*) (after Marten & van der Wal 2014)

	Cilubà	Chichewa/ Chaga	Shona	Sesotho/ Tswana	Ojibherero	Ndebele
Copula 'be'	✓	✓	✓	✓	✓	✓
Unaccusative active	*	✓	✓	✓	✓	✓
Transitive passive	*	✓	✓	✓	✓	✓
Unaccusative passive	*	*	✓	✓	✓	✓
Unergative passive	*	*	✓	✓	✓	✓
Unergative active	*	*	*	✓	✓	✓
Transitive active	*	*	*	*	✓	✓

As can be seen in Table 2, there is a wide variety of restrictions on predicate types in locative inversion. For example, Bresnan & Kanerva (1989: 16) show that Chichewa allows locative inversion with the copula, with unaccusative (active) verbs, and with passivized transitive verbs, but not with passivized unaccusatives, unergatives or active transitives. Languages like Cilubà are even more restrictive, and locative inversion is only possible with copulas. By contrast Demuth & Mmusi (1997) show that Setswana allows locative inversion with both unergative active and unaccusative active verbs, as shown respectively in (16a, b) and (16b, 17), above, as well as with passivized unaccusatives (18) passivised unergatives (20), and passivized transitives (21). Only active transitives do not take part in the alternation in Tswana (22).

- (18) Unaccusative passive (Demuth & Mmusi 1997: 12)
 Gó-fith-ilwe
 SM17-arrive-PRF.PASS
 'There has been arrived.'
- (19) Unergative passive (Demuth & Mmusi 1997: 13)
 Gó-á-leng-w-a
 SM17-PRS-plough-PASS-FV
 'There is being ploughed.'
- (20) Transitive passive (Demuth & Mmusi 1997: 12)
 Gó-bólá-ílwé di-phológóló dí-lé di-ńtsi
 SM17-kill-PRF.PASS 10-animals 10-COP 10-many
 'There have been killed many animals.'

- (21) Transitive active (Demuth & Mmusi 1997: 14)
- a. *Gó-ét-el-a ba-símané kokó
 SM17-visit-APPL-FV 2-boys 1a.grandmother
 ‘There are boys visiting the grandmother.’
- b. *Gó-kwál-él-a kokó lo-kwálo
 SM17-write-APPL-FV 1a.grandmother 5-letter
 ‘There is writing the grandmother a letter.’

Moving further to the right of Table 2, we see that only Otjiherero and Ndebele have fewer restrictions on predicates in locative inversion than Tswana: In these two languages, also active transitive verbs can take part in locative inversion.

Looking at the distribution in Table 2, two interesting further generalisations can be drawn. First, there is a hierarchy of predicate types, ranging from copula to active transitive predicates, and variation appears to follow strictly incremental increases along the hierarchy. In other words, if a language allows locative inversion with unergatives, it will also allow it with unaccusatives, and further, if a language allows locative inversion with transitives, it will allow it with both unaccusatives and unergatives. Correspondingly, there is no language in which locative inversion is possible with, for example, copulas, unergative passives, and active transitives. This means that locative inversion is more likely to occur with predicates on the left of the hierarchy. All languages in the sample in Table 2 allow locative inversion with copulas, but only two allow locative inversion with active transitive predicates. As Bresnan & Kanerva (1989), Demuth & Mmusi (1997), Diercks (2011), and many others have noted, locative inversion seems to be related to the absence of an external argument, and only a few languages allow locative inversion when an external argument is present. Second, it seems that there is no correlation between restrictions on predicate type and the nature of the syntactic/agreement relation. In Table 2, languages that have default agreement and syntactic topics, rather than subjects, in locative inversion, like Tswana, are marked in italics – so these are Sotho, Tswana, and Ndebele in the sample. While all three of these languages have comparatively few restrictions on predicate types, they are not alone in this: Like Ndebele, Otjiherero allows active transitives but has agreeing formal locative inversion. What might be said, based on this small sample, is that southern Bantu languages such as Shona, Sesotho, Tswana, Otjiherero, and Ndebele have fewer restrictions on predicate types than Eastern and Central Bantu languages such as Cilubà, Chichewa, and Chaga, but of course more comparative data is needed to confirm this.

The final dimension of variation noted for inversion constructions relates to (thematic) constraints on the pre-verbal phrase, or in other words, the availability of different inversion construction types in different languages. In a comprehensive comparative study, Marten & van der Wal (2014) distinguish between seven different types of inversion in Bantu: Formal locative inversion, semantic locative inversion, instrument inversion, patient inversion, complement inversion, and the two expletive constructions, agreeing inversion and default agreement inversion. Marten & van der Wal’s (2014) study included some 47 Bantu languages to show the range of variation encountered. In Table 3, we have reproduced results from the study relevant to southern Bantu languages.

Table 3: Inversion constructions in 10 southern Bantu languages, with Tswana highlighted (after Marten & van der Wal 2014)

		Formal Locative Inversion	Semantic Locative Inversion	Patient Inversion	Instrument Inversion	Complement Inversion	Default Agreement Inversion	Agreeing Inversion
R31	Otjiherero	✓	*	*			✓	*
S10	Shona	✓					✓	
S31	Tswana	*		*	*		✓	*
S32	Sesotho	*	*	*	*		✓	*
S33	N. Sotho	*	*	*	*		✓	*
S41	Xhosa		✓				✓	
S42	Zulu	*	✓	*	✓		✓	*
S43	Swati	*	✓		✓		✓	
S44	Ndebele	*	✓	*	✓	*	✓	*
S53	Tsonga/ Changana	*		*			✓	

As can be seen in Table 3, the ten languages in the sample fall into a few distinct groups (although keeping in mind that we do not have all relevant data for all languages). First, all languages have default agreement constructions, thus classifying the group as a whole. Recall that we have analysed inversion constructions in Tswana as default agreement constructions, with optionally pre-verbal PP locative topics, and so not as formal locative inversion. Accordingly, in Table 3 Tswana only has a ‘yes’ (P) value for ‘Default Agreement’. This is true for all Sotho-Tswana languages of the sample, neatly identifying the genetic grouping. Similarly, the four Nguni languages of the sample – isiXhosa, isiZulu, siSwati, and isiNdebele – pattern alike, have semantic locative inversion and instrument inversion in addition to default agreement constructions. We do not have enough data for Tsonga/Changana to see how the language(s) fit into the typology, although the data we have would be consistent with either the Sotho-Tswana pattern or the Nguni pattern. Finally, Otjiherero and Shona differ from the rest of the sample in that both have formal locative inversion. In contrast to our tentative conclusion about morphosyntactic variation in object marking in Section 2, where we noted that patterns of variation do not correspond to genetic sub-groupings, in the case of inversion constructions, the fit between morphosyntactic variation and sub-groupings seems to be much better. It would seem that morphosyntactic variation may arise independently of genetic classification in some cases, while in other cases, variation and sub-family can go hand in hand.

4. Diminutive marking

The final domain of morphosyntactic variation we discuss in this paper is diminutive marking. Many Bantu languages employ noun classes to mark diminutive, where often nouns that belong lexically to any class can be shifted into a class or classes with diminutive meanings. This is a process sometimes called secondary classification, and several noun classes have been reconstructed for Proto-Bantu as indicating diminutive meaning. The most common diminutive class is class 12 **ka-* for singular diminutives with class 13 **tu-* as the corresponding diminutive plural (Meeussen 1967, Maho 1999). In addition, classes 7 **ki-*, 8 **βi-*, 19 **pi-*, and 20 **yu-* have also been reconstructed as including diminutive (and indeed in some cases augmentative) meaning. The use of these noun classes for diminutives can still be seen in synchronic grammars, however, there are also several other strategies that different Bantu languages have developed to mark diminutives including the use of other noun classes, compounding, reduplication, and suffixation (Gibson et al. 2017).

Many of the southern Bantu languages, Tswana included, have developed a diminutive suffix *-ana*. This form is considered to be grammaticalised from the word for ‘child’, which is *mwana* in many Bantu languages, and reconstructed as **-jánà* ‘child’ in Proto-Bantu (e.g. Poulos 1999). It has also been argued that this suffixal form of the diminutive may be a result of contact with Khoisan languages which use a head-final structure and a diminutive based on the word for *child* (see e.g. Engelbrecht 1925, Güldemann 1999). In many cases, the suffix has completely replaced the older diminutive system based on noun class prefixation, making noun class use no longer an option for marking diminutives.

Southern Bantu languages pattern in two ways concerning the use *-ana* as the diminutive marker. One group of languages, including e.g., Ronga, Tsonga, and Venda, still use diminutive noun classes in addition to the derivational suffix *-ana* as the examples in (22) to (24) show. Venda, in particular, is interesting as one of the noun classes involved has a *ku-* prefix (cf. 24b), which Poulos (1986, 1990) analyses as a class 20 prefix.

- (22) Ronga (S54, Bachetti 2006: 63-64)
- | | | | |
|----|-------------------|-------------------|----------------------------|
| a. | yi-ndlu | ‘house’ | (class 9) |
| | xi-yi-ndlw-ana | ‘small house’ | (class 7 + class 9 + -ana) |
| | swi-yi-ndlw-ana | ‘small houses’ | (class 8 + class 9 + -ana) |
| b. | mu-lungu/va-lungu | ‘European(s)’ | (class 1/2) |
| | xi-lungw-ana | ‘small European’ | (class 7 + -ana) |
| | swi-lungw-tana | ‘small Europeans’ | (class 8 + -ana) |
- (23) Tsonga (S53, Poulos 1999: 206)
- | | | | |
|--|-------------|------------------|------------------|
| | muti | ‘village’ | (class 3) |
| | xi-mut-ana | ‘small village’ | (class 7 + -ana) |
| | swi-mut-ana | ‘small villages’ | (class 8 + -ana) |
- Venda (S53, Poulos 1986: 289, 1990: 38)
- | | | | |
|----|------------|---------------|------------|
| a. | lu-fhanga | ‘knife’ | (class 11) |
| | tshi-panga | ‘small knife’ | (class 7) |

b.	thavha	‘mountain’	(class 9)
	thavh-ana	‘small mountain’	(class 9 + <i>-ana</i>)
	ku-thavha	‘small mountain’	(class 20)
	ku-thavh-ana	‘very small mountain’	(class 20 + <i>-ana</i>)

By contrast, Tswana along with Northern Sotho and Nguni languages such as Zulu, only uses the innovative suffixes *-ana* or *-nyana*, reflecting a complete loss of using noun classes for diminutives.

(24) Tswana (S31, Cole 1955: 105-109, Creissels 1999)

a.	tau	‘lion’	(class 9)
	taw-ana	‘young lion’	(class 9)
b.	molapo	‘river’	(class 3)
	molatsw-ana	‘stream’	(class 3)
c.	thaba	‘mountain’	(class 9)
	thaba-nyana	‘small mountain’	(class 9)
d.	mosima	‘boy’	(class 1)
	mosima-nyana	‘small boy’	(class 1) ²

Similar to when noun classes are used to mark diminutives, the two diminutive suffixes *-ana* and *-nyana* can be added to a range of nouns (and adjectives) to express simple diminution in size or to encode diminution in quantity when used with a plural noun. Other uses of the diminutive include individuation and (pejorative or ameliorative) connotation. Phonologically, as the examples above show, the suffixation of *-ana* can trigger phonological changes in the final consonant of the root, e.g., /p/ becoming /tsw/ as in (25b).

In a wider comparative context, Gibson et al. (2017) show that the use of noun classes is the main means of marking diminutives in Bantu, in particular in the centre of the Bantu-speaking area, while both further to the south and further to the north, other means of expression are often found, including reduplication and compounds based on **-jána* ‘child’, in addition to the suffix *-ana* found in southern Bantu. Table 4 summarises the results of the study by Gibson et al. (2017) for 13 southern Bantu languages.

² We are grateful to an anonymous reviewer for this example.

Table 4: Diminutive marking in 13 southern Bantu languages, with Tswana highlighted (after Gibson et al. 2017)

Language	12/ 13	12	13	12/ 14	12/ 8	12/ 19	5/ 13	19/ 13	19	7/ 8	-jáná + N	N + -jáná	Red
Umbundu R11	✓												
Kwanyama R21				✓									
Herero R30				✓									
Yeyi R41	✓												
Matuumbi P13	✓												
Yao P20	✓												
Cuwabo P34											✓		
Venda S21										✓		✓	
Tswana S31												✓	
N. Sotho S32												✓	
Sesotho S33												✓	
Zulu S42												✓	
Tsonga S53										✓		✓	

As Table 4 shows all Southern Bantu languages (of Zone S) make use of the diminutive suffix *-ana* (or a similar form like *-nyana*) mostly as the sole diminutive marker. Only Venda and Tsonga use noun classes in addition to the suffix. In contrast, the use of classes 12/13 or 12/14 is the most prominent strategy for expressing diminutive meaning in the remaining languages of the sample. Only Cuwabo appears to be an exception to this generalisation, as diminutives are expressed by a compound based on **-jáná* as the first member, so not as a suffix.

The study of diminutives provides another illustration of the complex morphosyntactic microvariation observed in Bantu languages. The situation in Southern Bantu provides a good example of a change in progress, where the old system of using noun classes for expressing diminutives is being replaced by an innovative system based on the grammaticalized diminutive suffix *-ana*. Tswana, together with several other Zone S languages, has travelled furthest on this development path, and *-ana* and *-nyana* are the only means to form diminutives in the language.

5. Conclusions

In this paper we have discussed patterns of morphosyntactic variation in Bantu languages, focussing on the Southern Bantu language Tswana. We have highlighted three areas of variation: Object marking, inversion constructions and diminutive marking. For each of these areas we have provided the cross-linguistic context, and then situated the specific structures found in Tswana within this context.

Results of the investigation have shown that Tswana often patterns with other southern African Bantu languages, and that the group can be characterised by a number of shared features – such as

the presence of default agreement constructions, or diminutive marking with the suffix *-ana*. However, we have also seen that morphosyntactic variation is sometimes independent of genetic classification, for example concerning object marking, Tswana patterned more closely with Central and Eastern Bantu languages such as Bemba, Chaga, Ha, and Swahili, as opposed to Southern Bantu languages such as Swati, Herero, and Lozi. The study thus gives food for thought about the historical development, transmission, and contact effects relevant to different parts of language – e.g., lexicon and morphosyntax – and how they interact typologically and diachronically.

Based on the findings presented in this paper, further research can easily be conceived – more languages could be added to our three case studies, and many more aspects of morphosyntactic variation could be investigated. Another fruitful avenue specifically concerning Tswana might be a closer exploration of morphosyntactic differences between different dialects of the language, following the lead of a range of syntactic dialect atlas projects underway recently.

Finally, the paper has provided evidence of the impressive progress in Bantu linguistics over the last few decades. While description and analysis of individual languages have historically fed into comparative studies, we are now in a situation where in addition, and conversely, comparative studies can provide insights for our understanding of individual languages by providing detailed accounts of variation, in which specific languages can be contextualised and situated.

References

- Bachetti, Claudio. 2006. *Gramática de lingua Ronga*. No 2. Paulinas Editorial.
- Batibo, Herman M. 1999. A lexicostatistical survey of the Setswana dialects spoken in Botswana. *South African Journal of African Languages* 19(1). 2-11.
- Beaudoin-Lietz Christa, Derek Nurse & Sharon Rose. 2004. Pronominal object marking in Bantu. In Akinbiyi Akinlabi & Oluseys Adesola (eds.) *Proceedings of the World Congress of African Linguistics, New Brunswick 2003*, 175-188. Cologne: Köppe.
- Bloom Ström, Eva-Marie, Hannah Gibson, Rozenn Guérois & Lutz Marten (eds.) Forthcoming. *Morphosyntactic variation in Bantu: Current approaches*. Oxford: OUP.
- Bostoen, Koen, Gilles-Maurice de Schryver, Rozenn Guérois & Sara Pacchiarotti (eds.) 2023. *On Reconstructing Proto-Bantu Grammar*. Berlin: Language Science Press.
- Bresnan, Joan & Jonni M. Kanerva. 1989. Locative Inversion in Chichewa - a Case-Study of Factorization in Grammar. *Linguistic Inquiry* 20(1). 1-50.
- Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2). 147-185.
- Chebanne, Andy M. 2003. Reviving the Sotho-Tswana languages: Some considerations for the Union linguistic development of Sotho-Tswana variants. In A. Chebanne, M. Jokweni, L. Mokitimi, S. Ngubane (eds.) *Unifying Southern African languages: Harmonization and Standardization*, 75-86. Cape Town: CASAS.
- Chebanne, Andy M., Denis Creissels & H. W. Nkhwa. 1997. *Tonal Morphology of the Setswana Verb*. München: Lincom Europa.
- Cole, Desmond. T. 1955. *An introduction to Tswana grammar*, 9th impression. Cape Town: Longman.
- Creissels, Denis. 1996. Conjunctive and disjunctive verb forms in Setswana. *South African Journal of African Languages* 16(4). 109-115.
- Creissels, Denis. 2005. A typology of subject and object markers in African languages. In Voeltz, Erhard Friedrich Karl (ed.), *Studies in African linguistic typology*, 43-70. Amsterdam; Philadelphia: John Benjamins.

- Creissels, Denis. 2011. Tswana locatives and their status in the inversion construction. *Africana Linguistica* 17. 33-52.
- Creissels, Denis. 2017. The conjoint/disjunct distinction in the tonal morphology of Tswana. In van der Wal, Jenneke & Hyman, Larry M. (eds.), *The conjoint/disjunct alternation in Bantu*, 200-238. Berlin, New York: De Gruyter.
- Creissels, Denis. 2020. Grammaticalization in Tswana. In Walter Bisang & Andrej Malchukov (eds). *Grammaticalization scenarios. Areal patterns and cross-linguistic variation*, 769-801. Berlin: De Gruyter.
- Creissels, Denis. Forthcoming. The grammaticalization path DEICTIC SPATIAL ADVERB > ADPOSITION. To appear in Cuadernos de Lingüística de El Colegio de México.
- Demuth, Katherine & Sheila Mmusi. 1997. Presentational focus and thematic structure in comparative Bantu. *Journal of African Languages and Linguistics* 18. 1-19.
- Diercks, Michael. 2010. Agreement with subjects in Lubukusu. Georgetown: Georgetown University.
- Diercks, Michael. 2011. The morphosyntax of Lubukusu locative inversion and the parameterization of Agree. *Lingua* 121. 702-720.
- Downing, Laura & Lutz Marten. 2019. Clausal morphosyntax and information structure. In Mark van de Velde, Koen Bostoen, Derek Nurse and Gérard Philippson (eds.) *The Bantu Languages*, 270-307. 2nd ed. London and New York: Routledge.
- Engelbrecht, J. 1925. Suffixbildung in den südafrikanischen Bantusprachen. *Mitteilungen des Seminars für Orientalische Sprachen* 28, III. Abteilung, 86-131.
- Gibson, Hannah, Guérois, Rozenn & Lutz Marten. 2017. Patterns and developments in the marking of diminutives in Bantu. *Nordic Journal of African Studies* 26(4). 344-383.
- Gowlett, Derek F. 1989. The parentage and development of Lozi. *Journal of African Languages and Linguistics* 11. 127-149.
- Guérois, Rozenn. 2014. Locative inversion in Cuwabo. *ZAS Papers in Linguistics* 57. 49-71.
- Güldemann, Tom. 1999. Head-initial meets head-final: nominal suffixes in eastern and southern Bantu from a historical perspective. *Studies in African Linguistics* 28. 48-91.
- Halme Riikka. 2004. *A Tonal Grammar of Kwanyama*. Cologne: Köppe.
- Harford, Carolyn. 1990. Locative inversion of Chishona. In John P. Hutchison and Victor Manfredi (eds.) *Current Approaches to African Linguistics*, vol. 7, 137-144. Berlin, Boston: De Gruyter Mouton.
- Henderson, Brent M. 2006. The syntax and typology of Bantu relative clauses. Urbana-Champaign: University of Illinois.
- Khumalo, Langa. 2010. Passive, locative inversion in Ndebele and the unaccusative hypothesis. *South African Journal of African Languages* 30(1). 22-34.
- Maho, Jouni F. 1999. A comparative study of Bantu noun classes. Göteborg: Acta Universitatis Gothoburgensis.
- Marlo, Michael R. 2015. On the number of object markers in Bantu languages. *Journal of African Languages and Linguistics* 36. 1-65.
- Marten, Lutz. 2006. Locative Inversion in Otjherero: More on Morphosyntactic Variation in Bantu. *ZAS Papers in Linguistics* 43. 97-122.
- Marten, Lutz & Nancy Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies* 30. 237-253.
- Marten, Lutz & Jenneke van der Wal. 2014. A typology of Bantu subject inversion. *Linguistic Variation* 14(2). 318-368.
- Marten, Lutz & Nancy Kula & Nhlanhla Thwala. 2007. Parameters of Morphosyntactic Variation in Bantu. *Transactions of the Philological Society* 105(3). 253-338.
- Marten, Lutz, Nancy Kula, Jochen Zeller & Ellen Hurst (eds.) Forthcoming. *The Oxford Guide to the Bantu Languages*. Oxford: OUP.

- McCormack, Anna. 2006. A further look at conjunctive and disjunctive forms in Setswana. *ZAS Papers in Linguistics* 43. 123-141
- Meeussen, Achille Emile. 1967. Bantu grammatical reconstructions. *Africana Linguistica*. 3(1). 79–121.
- Otlogetswe, Thapelo & Andy M. Chebanne. 2018. Setswana. In T. Kamusella & F. Ndhlovu (eds.) *The Social and Political History of Southern Africa's Languages*. London: Palgrave Macmillan.
- Polak Louise. 1986. Les infixes ('préfixes objets') du bantou et leur reconstruction. *Africana Linguistica* 10. 365-422.
- Poulos, George. 1986. Instances of semantic bleaching in South-Eastern Bantu. In Gerrit J. Dimmendaal (ed.) *Current Approaches to African Linguistics* (vol. 3), 281-296. Dordrecht: Foris.
- Poulos, George. 1990. *A linguistic Analysis of Venda*. Pretoria: Via Afrika Limited.
- Poulos, George. 1999. Grammaticalisation in South-Eastern Bantu and the linguistic 'dynamics' underlying this process. *South African Journal of African Languages* 19. 204-14.
- Riedel, Kristina. 2009. The Syntax of Object Marking in Samba. A comparative Bantu perspective. Utrecht: LOT.
- Rugemalira, Josephat M. 1991. What is a symmetrical language? Multiple object constructions in Bantu. *Proceedings of the annual meeting of the Berkeley Linguistics Society* 17. 200-209.
- Rugemalira, Josephat M. 2004. Locative arguments in Bantu. In Akinbiyi Akinlabi & Oluseys Adesola (eds.) *Proceedings of the World Congress of African Linguistics*, New Brunswick 2003. Cologne: Köppe, 285-296.
- Schadeberg, Thilo C. 1995. Object diagnostics in Bantu. In E. 'Nolue Emenanjo and Ozo-Mekuri Ndimele (eds.). *Issues in African Languages and Linguistics: Essays in Honour of Kay Williamson*. Aba: National Institute for Nigerian Languages, 173-80.
- Taji, Julius. 2017. Object marking strategies in Chiyao. In Shegeki Kaji (ed.) *Proceedings of the 8th World Congress of African Linguistics*. Tokyo: ILCAA, 227-242.
- Van de Velde, Mark, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.) 2019. *The Bantu Languages*. 2nd ed. London and New York: Routledge.
- Van der Wal, Jenneke. 2022. *A Featural Typology of Bantu Agreement*. Oxford: Oxford University Press.
- Zeller, Jochen. 2013. Locative inversion in Bantu and predication. *Linguistics* 51(6). 1107-1146.
- Zeller, Jochen. 2014. Three types of object marking in Bantu. *Linguistische Berichte* 239. 347-367.
- Zeller, Jochen & J. Paul Ngoboka. 2018. Agreement with locatives in Kinyarwanda: a comparative analysis. *Journal of African Languages and Linguistics* 39. 65-106.

Nancy C. Kula
 Department of Language and Linguistics
 University of Essex
ncula@essex.ac.uk

Lutz Marten
 Department of Linguistics
 SOAS University of London
lm5@soas.ac.uk