# Relative clause formation in the Bantu languages of South Africa

**Jochen Zeller**

*Linguistics Programme, Faculty of Human Sciences, University of KwaZulu-Natal,
King George V Avenue, Glenwood, Durban 4041, South Africa
e-mail: zeller@ukzn.ac.za, www.jzeller.de*

**Abstract:** This article discusses (verbal) relative clauses in the Bantu languages spoken in South Africa. The first part of the article offers a comparison of the relative clause formation strategies in Sotho, Tsonga, Nguni and Venda. An interesting difference between these language groups concerns the syntactic position and the agreement properties of the relative marker. Whereas the relative markers in Sotho, Tsonga and Venda are *clause-initial* elements, which express *agreement with the head noun*, the relative markers in the Nguni languages are relative concords, which are *prefixed to the verb* and *agree with the subject* of the relative clause. The second part of the article addresses this difference and shows that there is a historical relation between these two types of relative constructions. It is argued that earlier forms of Nguni employed relative markers similar to those used in present-day Sotho and Tsonga. In Nguni, these relative markers underwent a grammaticalisation process which turned them into relative concords. A detailed analysis of the syntactic conditions for, and the properties of, this grammaticalisation process leads to a hypothesis about the reasons why relative concords have developed in Nguni, but not (to the same extent) in Tsonga, Sotho and Venda.

## 1. Introduction

The purpose of this article is two-fold. First, it provides a discussion of the different strategies employed in (verbal) relative clause formation in the Southern Bantu languages spoken in South Africa. Relative clause constructions of many Bantu languages have been analysed in the literature (cf. e.g. Bokamba, 1976 for Dzamba; Poulos, 1982 for Zulu; Livingstone, 1996 for Luganda; Demuth & Harford, 1999; Harford & Demuth, 1999 for Sotho and Shona (and others); Ngonyani, 2001 for Swahili); however, there is no recent study that offers a systematic *comparison* of the relative clause formation strategies in Nguni, Sotho, Tsonga and Venda. Second, this article argues that the relative clause formation strategy of the Nguni languages, although synchronically different from the strategy used in Sotho and Tsonga, is diachronically related to the latter languages. I will demonstrate that the Nguni relative concord is the result of a grammaticalisation process in which a relative marker similar to those found in Sotho and Tsonga was reanalysed as a verbal prefix expressing relativisation. Although grammaticalisation and language change in African

languages have been studied extensively (see e.g. Givón, 1975; Hyman, 1975; Heine & Reh, 1984 (and references cited therein); Poulos, 1999; Güldemann, 1999) and although the possibility of a diachronic relation between the Nguni and the Sotho and Tsonga relative markers has been mentioned in passing by some authors (cf. e.g. Mischke, 1998; Poulos, 1999), I am not aware of any analysis in which the grammaticalisation process that derived the Nguni relative concords is discussed and spelled out in some detail.

In section 2, I give an overview of how (verbal) relative clauses are formed in Sotho, Tsonga, Nguni and Venda. I show that, whereas the relative markers used in Sotho, Tsonga and Venda are either *relative pronouns* or *relative complementisers*, the Nguni languages employ *relative concords*, morphologically complex elements affixed to the verb stem.

In section 3, I illustrate how relative concords in Nguni are derived historically from relative markers of the type employed in present-day Sotho and Tsonga. I investigate the syntactic configurations that gave rise to grammat-

icalisation and I provide synchronic evidence for this grammaticalisation process. Section 4 concludes the article with some speculations about why grammaticalisation took place in Nguni, but not in Sotho and Tsonga.

## 2. Relative clause formation in Sotho, Tsonga, Nguni and Venda

In the literature on Southern Bantu languages, a distinction is sometimes drawn between so-called 'direct' and 'indirect' relatives (cf. Doke, 1954). A direct relative is a relative clause construction in which the head noun corresponds to the grammatical subject of the relative clause (such as in the English examples in (1)). If the head noun corresponds to a constituent other than the subject (as in the examples in (2), where it corresponds to the object), then the construction is called an *indirect relative*:

(1) a **the dog [which bit you]**          (direct relative)
     b **the dog [that bit you]**            (direct relative)
(2) a **the dog [which you bit]**          (indirect relative)
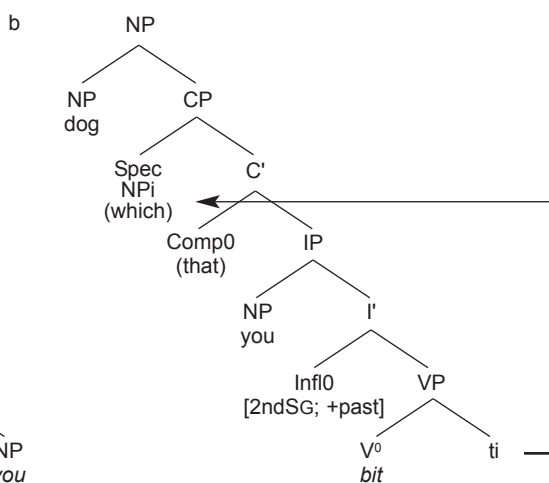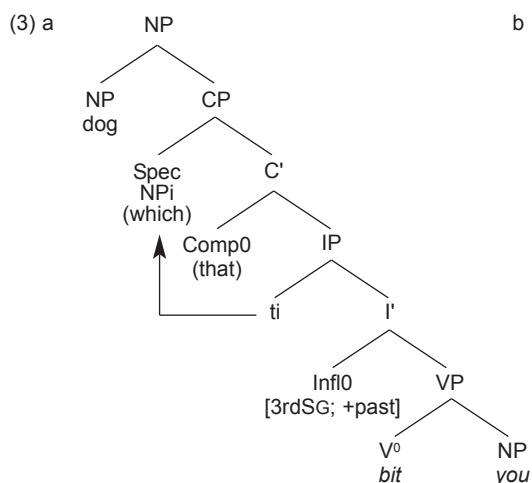     b **the dog [that you bit]**            (indirect relative)

In this section I discuss the relative clause formation strategies of Sotho, Tsonga, Venda and Nguni with respect to the formation of direct and indirect relatives. I use the syntactic framework developed in Chomsky (1986) as the theoretical basis of my discussion. In this framework, a sentence is taken to be the maximal projection of a complementiser (Comp$^0$), i.e. a CP. The head of the CP selects the maximal projection IP as its sister. IP is the phrasal projection of the inflectional features of the sentence (which are located in the head of IP, Infl$^0$).

The subject is located in SpecIP and agrees with the inflectional features in Infl$^0$. The VP is the sister of Infl$^0$.

I assume that relative clauses, like all sentences, are CPs (cf. Chomsky, 1986; 1995; Chomsky & Lasnik, 1993; Kayne, 1994). The CP corresponding to a restrictive relative clause is usually taken to be adjoined to the head noun-NP (but see Kayne, 1994 for a different analysis). The NPs in (1) and (2) can be represented by the tree diagrams in (3).

As shown in (3), the relative complementiser **that** is located in Comp$^0$; the relative pronoun is moved to SpecCP, leaving a trace in its base position. This base position corresponds to the syntactic function of the head noun. In the direct relative clause (3a), the relative pronoun moves from the subject position, leaving a trace in SpecIP. In an indirect relative clause, the relative pronoun would be moved from a different position, as in (3b), where the trace is in object position.

An English relative clause does not tolerate both an overt complementiser and an overt relative pronoun. This means that the phonological content of either of the two (or of both) must be deleted. Relative clauses with relative pronouns, such as (1a) and (2a), therefore have phonologically "zero" or "covert" complementisers, whereas complementiser-introduced relative clauses (such as (1b) or (2b)) have covert relative pronouns. The phonologically zero relative pronoun is sometimes referred to as an "empty operator" (cf. Chomsky & Lasnik, 1993; Alexiadou *et al.*, 2000). It fulfils the same func-

(3) a

NP
├── NP dog
└── CP
    ├── Spec NPi (which)
    └── C'
        ├── Comp0 (that)
        └── IP
            ├── ti
            └── I'
                ├── Infl0 [3rdSG; +past]
                └── VP
                    ├── V0 *bit*
                    └── NP *you*

b

NP
├── NP dog
└── CP
    ├── Spec NPi (which)
    └── C'
        ├── Comp0 (that)
        └── IP
            ├── NP you
            └── I'
                ├── Infl0 [2ndSG; +past]
                └── VP
                    ├── V0 *bit*
                    └── ti

tion as an overt relative pronoun, i.e. it creates a trace which is semantically interpreted as a variable, thereby allowing for the relative clause to be interpreted as a predicate (cf. Heim & Kratzer, 1998 for details). Some languages differ from English in allowing for both the relative pronoun and the relative complementiser to be overtly realised in relative clauses (see, for example, Bayer (1984) for Bavarian; Pollock (1992) for French; Zwart (2000) for various Dutch dialects).[1] In section 2.1 I show that relative complementisers and relative pronouns can also co-occur in certain Bantu languages.

## 2.1 The Sotho group (Northern and Southern Sotho; Tswana)

Consider the formation of relative clauses in the Sotho languages:[2]

The bracketed constituents in (4) – (9) are relative clauses that modify their respective head nouns. As the examples show, the initial element in a relative clause in Sotho is a *relative marker* (printed in italics) that agrees with the head noun with respect to class membership. Relative clauses in Sotho occur in the *participial mood*, which is specified as an inflectional feature of the clause in Infl⁰ and overtly represented by the participial form of the subject prefix located in this position (notice e.g. that the subject prefix of class 1 in Northern Sotho is **a**-, and not the prefix **u**- of the indicative). The verb of the relative clause is modified

with a *relative suffix* (**-ng** in Southern Sotho and Tswana; **-go** in Northern Sotho). I assume that the verb moves to Infl⁰ in order to combine with the subject prefix; otherwise, adverbs adjoined to VP would falsely be expected to severe the subject prefix from the verb.

The syntactic function of the head noun is always marked through a *pronominal clitic* inside the relative clause. In direct relative clauses, where the head noun corresponds to the subject, this function is fulfilled by the subject prefix. In (7) – (9), where the head noun corresponds to the object of the verb, an object clitic that agrees with the head noun is attached to the verb stem. The object clitic is obligatory; without it, the constructions are ungrammatical.

The clause-initial relative markers in Sotho are based on *demonstrative pronouns* (cf. Doke, 1954; Mischke, 1998; Harford & Demuth, 1999; and many others), as is revealed through a comparison of the relative markers in (7)–(9) and the italicised elements in (10):

(10) a  **setulo** *seo*
        chair7  DEM7
        'this chair'              (Southern Sotho)
     b  **badišana** *ba*
        herdboy2  DEM2
        'these herdboys'         (Northern Sotho)
     c  **monna** *yô*
        man1  DEM1
        'this man'                      (Tswana)

Direct relatives

(4) **ngwana**        **[ya**           **bala-ng**        **hantle]**
    child1a          REL1a+SP1a        read-RS           well
    'the/a child who studies well'                        (Southern Sotho)

(5) **motho**         **[yô**           **a**             **sepela-go]**
    person1          REL1              SP1               walk-RS
    'a person who walks'                  (Northern Sotho; Mischke, 1998: 110)

(6) **dikgômo**       **[tsê**          **di**            **fula-ng**       **fâle]**
    cattle8          REL8              SP8               graze-RS         there
    'the cattle which graze there'

Indirect relatives

(7) **setulo**        **[seo**          **basadi**        **ba-se-rek-ile-ng**            **kajeno]**
    chair7           REL7              woman2           SP2-OC7-buy-PERF-RS             today
    'the chair which the women bought today'                        (Southern Sotho)

(8) **badišana**      **[ba**           **mosetsana**     **a**             **ba**           **tseba-go]**
    herdboy2         REL2              girl1            SP1               OC2             know-RS
    'the herdboys who(m) the girl knows'          (Northern Sotho; Poulos & Louwrens, 1994: 107)

(9) **monna**         **[yô-batho**     **ba-mo-nyatsa-ng]**
    man1             REL1-person2      SP2-OC1-disrespect-RS
    'the man whom the people disrespect'                      (Tswana)

What is the syntactic status of these demonstratives when they are used as relative markers? (7) – (9) show that they *precede* the subject in indirect relatives. Given the structure of relative clauses in (3) above, there are therefore two possible structural analyses: the demonstrative/relative marker in Sotho could either be a relative *pronoun* like English **who** or **which** (and as such would be located in SpecCP) or a relative *complementiser* like **that** (and would be located in Comp⁰).

It is difficult to decide between these two possible analyses; as far as I can see, the morphosyntactic properties of the Sotho relative markers are compatible with either. Demuth and Harford (1999) and Harford and Demuth (1999) treat the Sotho relative markers as relative complementisers. Although their view seems to be motivated mainly by theory-internal considerations, it might also be prompted by the fact that the English complementiser **that** is historically derived from a demonstrative pronoun. However, if one accepts that conclusions about the status of the relative marker in Sotho can be drawn on the basis of a comparison with relative markers in other languages, then one should perhaps rather compare the Sotho relative clause construction to those of other *Bantu* languages. As it turns out, at least in some Bantu languages, demonstratives unequivocally function as relative pronouns.

As shown in Chumbow (1977), demonstratives are used as relative markers in Ngemba (Cameroon). Importantly, in this language, a relative clause-initial demonstrative can *co-occur* with an overt relative complementiser **bah** (11). The same co-occurrence of a demonstrative and a relative complementiser can be observed in Bafut (Cameroon) (12). Since **bah** and **mə̂** are relative complementisers and therefore in Comp⁰, the relative clause-initial demonstratives in (11) and (12) must be relative pronouns in SpecCP.[3] The relative clause-initial elements in Sotho are also demonstratives; there-

fore, an analysis that treats them as relative pronouns would allow for a uniform analysis of the use of demonstrative pronouns in relative clauses in Bantu languages.

However, although the data in (11) and (12) can be interpreted in favour of a relative pronoun-analysis, they do not conclusively rule out a relative complementiser treatment of the Sotho relative markers. The fact that many Bantu languages use demonstratives in relative clauses does not necessarily imply that these elements have synchronically the same syntactic status in these languages. Below I argue that the Nguni relative concords, which are prefixed to the verb of the relative clause, are *historically derived* from demonstrative pronouns. But if demonstratives could develop into relativising verbal prefixes in the Nguni languages, then it is also possible that they developed into relative complementisers in Sotho.

In the light of these considerations, I remain agnostic about the precise syntactic position of the relative marker in Sotho. For the analysis which I present in section 3, it is of no particular relevance whether the relative marker in Sotho is synchronically a relative pronoun or a relative complementiser. However, what is important to note is that the Sotho relative marker appears in front of the subject in indirect relative clauses and that it agrees with the head noun (13).

In direct relatives, the subject position is filled with the trace of the moved relative pronoun or empty operator (see (3a)). As a consequence, the relative marker (either in SpecCP or Comp⁰) and the subject prefix of the verb (in Infl⁰) are not separated by any phonological material, but are linearly adjacent, as is clearly illustrated by example (5) from Northern Sotho and example (6) from Tswana. However, Southern Sotho direct relatives differ in an interesting way from their counterparts. Although (7) clearly demonstrates that Southern Sotho uses a relative-clause initial relative marker in indirect relatives, (4) shows

(11) **nyung** **[wá**     **bah**     **a-keshung-ne**     **mung**     **wa**     **la]**     **a**     **kung**     **atsang**
     man     Dem/Rel     RComp     Sp+Tns-beat-Rs     child     Det     Def     Sp     enter     into.prison
     'The man who beat the child went to prison'             (Ngemba; Chumbow, 1977: 290)

(12) **mbú** **yá**     **[yiĩ**     **mə̂**     **kǐ**     **kwôá]**     **ywìinə̀mə́**
     dog     Det     Dem/Rel     RComp     Tns     die     resurrected
     'the dog which died was resurrected'                     (Bafut; Mutaka, 2000: 212)

(13) [Head Noun] [CP **relative marker** [IP subject [I' Sp-Oc-stem [VP ] ] ] ]

           agreement

that direct relatives are formed by means of a *single monosyllabic marker* that replaces the subject prefix (cf. Doke, 1954; Mischke, 1998). Mischke (1998) shows that this single relativiser is the result of a phonological process of coalescence between the relative marker and the subject prefix of the verb. For example, the demonstrative pronoun (= relative marker) of class 1a in Southern Sotho is **eo**; and the (participial) subject prefix of this class is **a**. The phonological merging of both elements derives the element **ya** used in (4). Similar processes derive the relative markers of the other noun classes (cf. Mischke, 1998). I return to this point in section 3.1.

## 2.2 Tsonga

Tsonga relative clause formation patterns in all relevant ways with Sotho. Consider the direct relative in (14) and the indirect relative in (15).

*Direct relative:*
(14)  **munhu        [*loyi*      a-famba-ka]**
        person1      REL1      SP1-travel-RS
        'a person who travels'        (Doke, 1954: 204)

*Indirect relative*
(15)  **buku  [*leyi* munhu a    yi    hlaya-ka]**
        book9  REL9 person1 SP1  OC9  read-RS
        'the book that the person is reading'

   As in Sotho, the verb in relative clauses in Tsonga is in the participial form and modified with a relative suffix (-**ka**). The head noun is represented inside the relative clause through a pronominal clitic (or the subject prefix) and through a relative marker; both elements agree in noun class with the head noun. The relative markers in Tsonga, like their Sotho counterparts, are identical to the demonstrative pronouns, (16):

(16)  **buku  *leyi***
        book9  DEM9
        'this book'                        (Tsonga)

   The position of the relative marker in front of the subject in (15) shows that these markers, like their Sotho counterparts, can be analysed as either relative complementisers or relative pronouns. Again, I leave open which of these two analyses is correct and whether the Tsonga relative marker is located in SpecCP or Comp⁰.

## 2.3 The Nguni group (Zulu, Swati, Xhosa, Southern Transvaal Ndebele)

The following examples illustrate the dominant relative clause formation strategy of the Nguni languages (an alternative strategy is discussed

in section 3.3):[4]

*Direct relatives*:
(17)  **indoda  [*e*-hleka    kakhulu]**
        man9    RC9-laugh  a.lot
        'the man who laughs a lot'        (Zulu)
(18)  **abafazi  [*aba*-lila-yo]**
        woman2  RC2-weep-RS
        'the women who are weeping'      (Xhosa)
(19)  **umfana  [*lo*-ti-senga-ko        tinkomo]**
        boy1    RC1-OC10-milk-RS  cow10
        'the boy who milks the cows'        (Swati)

*Indirect relatives*:
(20)  **incwadi  [isitshudeni *esi*-yi-funda-yo]**
        letter9  student7    RC7-OC9-read-RS
        'the letter that the student is reading'    (Zulu)
(21)  **indoda  [amakhwenkwe *a*-yi-bon-ile-yo]**
        man9    boy6                RC6-OC9-see-PERF-RS
        'the man whom the boys saw'        (Xhosa)
(22)  **umfati        [tintfombi    *leti*-m-elekelela-ko]**
        woman1      girl10        RC10-OC1-help-RS
        'the woman whom the girls help'        (Swati)

   There are obvious similarities between relative clauses in Nguni and those in Sotho and Tsonga. Relative clauses in Nguni are also in the participial mood. The verb is extended by means of a relative suffix (–**yo** in Zulu and Xhosa; –**ko** in Swati) that usually appears when the relative clause predicate is phrase-final (it can otherwise be omitted). As in Sotho and Tsonga, the syntactic role of the head noun in indirect relative clauses is represented by means of a pronominal clitic in the relative clause which agrees with the head noun. Furthermore, Nguni relative clauses also exhibit the occurrence of a relative marker which is printed in italics in (17) – (22).

   However, there is an important difference. The Nguni languages do not use overt relative pronouns or relative complementisers, which is clearly shown by the indirect relatives in (20) – (22). Here, the subject is the first element of the relative clause, and is followed by the relative marker. The relative markers in Nguni are what Doke (1954) calls *relative concords*. The relative concord is attached to the verb; it takes the place of the subject prefix. Consequently, it always agrees with the subject of the relative clause, regardless of whether we are dealing with a direct or an indirect relative clause. SpecCP is filled with an empty operator (a phonologically zero relative pronoun) like the one that occurs in English relative clauses introduced by **that** (23).

   The underlying morphological structure of relative concords in Nguni is complex. A relative

concord is the result of merging a *relative morpheme* (whose underlying form is **a-** in Zulu and Xhosa and **la-** in Swati) with a prefix which is identical in form with the subject prefix of the respective noun class. The overt phonological form of the relative morpheme is determined by a general rule of Vowel Raising that causes the vowel of **(l)a-** to assimilate in vowel height and backness to the high vowel of this following prefix (cf. e.g. Khumalo, 1992 for Zulu), deriving the two allomorphs **(l)o-** and **(l)e-**. Consequently, if the subject prefix of a noun class starts in a consonant, the relative concord has the form relative morpheme + subject prefix, cf. the examples in (24). However, if the subject prefix is a vowel, it is deleted.[5] In noun classes with vowel subject prefixes, the relative concord therefore only consists of the relative morpheme, cf. the examples in (25).

Nguni employs different relative concords for class 1/1a in direct and indirect relatives. As is shown in (25a), the relative concord of class 1/1a is **(l)o-** in direct relatives; however, in indirect relatives, it is **(l)a-**. I return to this fact in section 3.2 where I also address some further properties of the relative concords in Nguni.

Notice that in direct relatives, both the relative concord in Nguni and the relative marker in Sotho and Tsonga agree with the head noun.

However, this agreement is established through different structural configurations. As was shown in sections 2.1 and 2.2, the relative marker in Sotho and Tsonga appears in clause-initial position (in SpecCP or in Comp[0]) from where it directly agrees with the head noun. In contrast, the relative concord in Nguni is part of the verbal morphology in Infl[0] and therefore agrees with whatever element fills SpecIP in the relative clause. In direct relatives, this position is filled with the trace of the empty operator in SpecCP. The empty operator in turn agrees with the head noun. This means that the agreement between the head noun and the relative concord is established via an operator chain that links SpecCP and SpecIP.

The relative clause formation strategies in Nguni on the one hand and in Sotho and Tsonga on the other illustrate a striking difference between the grammars of these otherwise quite closely related language groups. In section 3 I try to show that the relative concord-strategy of Nguni has been derived historically from a relative clause formation strategy comparable to the one attested synchronically in Sotho and Tsonga. Before I turn to this analysis, let me take a look at relative clause formation in Venda.

(23)  [Head Noun] [$_{CP}$ Op [$_{IP}$ subject [$_{I'}$ **relative concord**-(O$_C$)-stem [$_{VP}$ ] ] ] ]

agreement

(24)                                       relative concord      <          relative morpheme + subject prefix
    a.   class 2:     **(l)a-ba-**    <    (l)a- + ba-
    b.   class 5:     **(l)e-li-**     <    (l)a- + li-
    c.   class 15:   **(l)o-ku-**    <    (l)a- + ku-
    d.   1stPers.Sg:  **(l)e-ngi-**   <    (l)a + ngi-
(25)                                       relative concord      <          relative morpheme + vowel subject prefix
    a.   class 1/1a:  **(l)o-**       <    (l)a- + u- (direct relatives only)
    b.   class 6:     **(l)a-**       <    (l)a- + a-
    c.   class 9:     **(l)e-**       <    (l)a- + i-
    d.   2ndPers.Sg: **(l)o-**       <    (l)a + u-

*Direct relatives*:
(26)  **nngwa**          **[dzi**                    **huvha-ho]**                                    (construction (i))
      dogs10            S$_P$10                    bark-R$_S$
      'the dogs which are barking'
(27)  **nngwa**          **[dzi**                    **no huvha]**                                    (construction (ii))
      dogs10            S$_P$10                    R$_F$ bark
      'the dogs which bark'
(28)  **nngwa**          **[*dzine***              **dza**           **huvha]**                    (construction (iii))
      dogs10            R$_{COMP}$10            S$_P$10          bark
      'the dogs which bark'

## 2.4 Venda

Venda differs from Sotho, Tsonga and Nguni in that there are three different ways of forming direct relative clauses, (26) – (28) (cf. Doke, 1954; Poulos, 1990).[6]

In construction (i), the verb in Venda, as in Sotho, Tsonga and Nguni, is modified with a relative suffix (–**ho**), (26). In construction (ii), there is no relative suffix; instead, the formative –**no**– appears between verb stem and subject prefix, (27). Construction (iii) employs a relative marker (in italics) which agrees with the head noun, (28). In contrast to their Sotho and Tsonga counterparts, the relative markers of construction (iii) bear no resemblance to the Venda demonstratives; below I show that these markers can be analysed as relative complementisers.

According to Doke (1954), the verb in the constructions (i) and (ii) in (26) and (27) is in the participial mood, whereas the verb in construction (iii) in (28) occurs in the so-called dependent mood, a verb form also found in various other subordinate constructions in Venda. Importantly, the form of the subject prefixes in these two moods is different in some noun classes; for example, the class 10 prefix is **dzi** in the participial mood, (26) and (27), but **dza** in the dependent mood, (28). Poulos (1990) labels the subject prefixes of the three relative constructions in Venda "relative concords". I think this terminology is slightly misleading, because it suggests a relation between the Nguni relative concords and the Venda prefixes. However, the Nguni relative concords are clearly different from the prefixes in (26) – (28); they are morphologically complex elements and they occur typically in relative constructions. In contrast, the subject prefixes in Venda

relative clauses are simplex and are also attested in other constructions which require the participial or dependent mood. It seems more appropriate to classify Venda with Sotho and Tsonga with regard to the subject prefixes and to maintain that of the Bantu languages spoken in South Africa, only the Nguni languages have genuine relative concords.

If a language has more than one relative construction, the choice between the different alternatives sometimes depends on whether the relative clause is used restrictively or non-restrictively. For example, non-restrictive relative clauses in English require an overt relative pronoun; they can never be formed with the relative complementiser **that** (cf. e.g. Alexiadou *et al.,* 2000). However, no such condition seems to determine the choice between the three relative constructions in Venda direct relatives; all three constructions can be used as non-restrictive relative clauses (29), (30) and (31).

However, this does not imply that constructions (i), (ii) and (iii) are synonymous. Interestingly, my Venda informant notes that whereas (26) is the preferred construction to describe dogs which are barking at the time of the utterance, (27) and (28) are better suited to describe a generic or habitual property of the dogs (i.e. they are the ones which bark, as opposed to those which bite). If this observation is in fact indicative of a systematic difference between the three ways of forming direct relative clauses in Venda, then it provides an interesting starting point for future research into the relation between the morphosyntactic properties of Venda relatives and their aspectual characteristics. However, since this is not the primary concern of this article, I do not discuss

| (29) | **Romeo** | **[a** | | **takalela-ho** | **u bika]** | **u** | **shuma** | **tshikoloni** | (construction (i)) |
|---|---|---|---|---|---|---|---|---|---|
| | Romeo1 | SP1 | | like-RS | to-cook | SP1 | work | at.school | |
| | 'Romeo, who likes to cook, works at the school' | | | | | | | | |
| (30) | **Romeo** | **[a** | **no** | **takalela** | **u bika]** | **u** | **shuma** | **tshikoloni** | (construction (ii)) |
| | Romeo1 | SP1 | RF | like | to-cook | SP1 | work | at.school | |
| | 'Romeo, who likes to cook, works at the school' | | | | | | | | |
| (31) | **Romeo** | **[*ane*** | **a** | **takalela** | **u bika]** | **u** | **shuma** | **tshikoloni** | (construction (iii)) |
| | Romeo1 | RCOMP1 | SP1 | like | to-cook | SP1 | work | at.school | |
| | 'Romeo, who likes to cook, works at the school' | | | | | | | | |

Indirect relative

| (32) | **munna** | **[*ane*** | **nngwa** | **dza** | **mu-pandamedza]** |
|---|---|---|---|---|---|
| | man1 | RCOMP1 | dog10 | SP10 | OC1-chase |
| | 'the man whom the dogs are chasing' | | | | |

this aspect of Venda any further.

Importantly, only construction (iii) can be used to form indirect relative clauses in Venda (cf. Doke, 1954; Poulos, 1990) (32).[7] As in Sotho, Tsonga, and Nguni, the grammatical function of the head noun in an indirect relative clause in Venda is indicated through a pronominal clitic which, like the relative marker, agrees with the head noun. (32) shows that the relative markers of construction (iii) precede the subject position of the relative clause. Doke (1954) classifies them as relative pronouns, in which case they would be located in SpecCP. However, the following observation suggests that the relative markers of construction (iii) in Venda are in fact relative complementisers in Comp$^0$. If an indirect relative clause in Venda is in the *perfect tense*, the relative marker takes on a different form. Compare the relative marker in (28) with the contracted form of this marker in (33) which illustrates that in Venda, relative clauses in the perfect tense require the choice of the short form of the relative marker. This requirement recalls the well-known relation between complementisers and tense, which is attested in many languages. For example, the selection of a particular complementiser in English depends on whether the clause is tensed or not (34). Given that the choice of the relative marker in construction (iii) in Venda is also sensitive to the tense specification of the relative clause, I conclude that these markers are relative complementisers and as such located in the Comp$^0$-position of the relative clause (35).

To conclude, Venda has a unique status among the South African Bantu languages with respect to many characteristics of relative clause formation. Only Venda has three differ-

ent strategies to form direct relatives, and only Venda employs relative complementisers whose overt forms depend on the tense of the relative clause. Venda also differs from Sotho, Tsonga and Nguni in that its relative markers are not related to the form of the demonstratives. Since this latter relation is particularly important for the following discussion of the development of the Nguni relative concord, I do not consider Venda relatives any further in the remainder of this article.

## 3. Relative markers in Nguni, Sotho and Tsonga: a historical relationship
In this section I suggest that at an earlier historical stage, the Nguni languages formed relative clauses by employing fully-fledged demonstrative pronouns as relative markers, which were located in relative clause-initial position and which agreed with their head nouns. (As in the case of Sotho and Tsonga, I leave open the question whether the early Nguni relative markers were relative pronouns in SpecCP or relative complementisers in Comp$^0$.) With respect to their syntactic position and their agreement properties, the early Nguni relative markers were hence different from the relative concords of modern Nguni and more like the relative markers of present-day Sotho and Tsonga. However, I assume, following Mischke (1998) and Poulos (1999), that the early Nguni relative markers, in contrast to their Sotho and Tsonga counterparts, underwent a grammaticalisation process in which they were reanalysed as part of the predicate's morphology in Infl$^0$, i.e. as relative concords (36).

In the discussion that follows I consider the conditions and properties of this grammaticalisation process in more detail.

(33)    **nngwa**        **[*dze*        **dza**        **huvha]**                        (construction (iii))
        dogs10          RCOMP10        SP10        bark
        'the dogs which barked'

(34)    a.    **I asked *for* him to drink his beer**
        b.    **I know *that* he drinks beer**

(35)    [Head Noun] [$_{CP}$ **relative complementiser** [$_{IP}$ subject [$_{I'}$ SP-OC-stem [$_{VP}$ ] ] ] ]
        └──────────────────────────┘
                agreement

(36)    a.    Input:      [Head Noun] [$_{CP}$ [**relative marker**] [$_{IP}$ subject [predicate] ]]
                            →        *Reanalysis*        →
        b.    Output:    [Head Noun] [$_{CP}$ [$_{IP}$ subject [**relative concord -**predicate] ]]

## 3.1 Syntactic contexts for reanalysis

Many, if not most, grammaticalisation processes are the result of reanalysis. In a typical reanalysis process, a structure α with two independent expressions X and Y is rebracketed as a structure β in which X is part of Y; sometimes, X and a part of Y thereby fuse into one morpheme which may then be associated with a new grammatical function (cf. Heine & Reh, 1984; Hopper & Traugott, 1993). For example, Heine and Reh (1984) discuss the development of a focus marker in Ewe (a Kwa language spoken in Togo and Ghana). This marker is the result of a grammaticalisation process in which the subject prefix of the verb immediately following a focussed NP in a copula clause was reanalysed as a focus-marking enclitic of the NP (37).

Importantly, reanalysis requires X and Y to be phonologically adjacent in α; only then can the linear string corresponding to α be interpreted as β. I have suggested that the grammaticalisation process which has derived the relative concords in Nguni is the result of reanalysis. This claim implies that there were syntactic contexts in early Nguni in which a relative marker in the left periphery of the relative clause and a predicate in Infl⁰ were in fact adjacent, such that the relative marker could have been reinterpreted as being part of the predicate. However, as is shown in (36a) earlier, the subject position intervenes between a relative marker in SpecCP or Comp⁰ and the predicate; an overt subject NP would hence block adjacency. In the following, I therefore discuss those Southern Bantu relative clause constructions in which the subject position is not filled with phonological material. In these contexts, a predicate in Infl⁰ would be phonologically adjacent to a relative marker located either in SpecCP or Comp⁰ of the relative clause.

### 3.1.1 Direct relative clauses

A verb is usually the first word of the IP in a *direct* relative clause, since here, SpecIP is only filled with the phonologically invisible trace of a subject relative pronoun or empty operator (cf. (3a) in section 2) (38). Direct relatives hence provided a context in which reanalysis could take place in Nguni.

In this regard, it is interesting to recall the properties of Southern Sotho direct relatives. As pointed out in section 2.1, although Southern Sotho uses clause-initial relative markers in indirect relatives, it employs a monosyllabic marker in direct relatives ((39) repeats example (4) from section 2.1).

As discussed above, the direct relative markers in Southern Sotho are the result of the phonological merging of the relativising demonstrative and the subject prefix (cf. Mischke, 1998). The same development seems under way in some Tsonga dialects (cf. Doke, 1954). According to Doke (1954), the marker **lw'a**- (**la**- according to my Tsonga informant) which precedes the verb stem in (40) is the result of coalescence of the demonstrative pronoun **lweyi** and the subject prefix **a**-.

I assume that the development of Nguni relative concords has been triggered by phonological merging processes similar to those observed in (39) and (40). Interestingly, Mischke (1998) analyses the single morphemes in Southern Sotho direct relatives as relative concords, i.e. as independent morphemes which replace the subject prefix and

| (37) | a. | Input α: | **[ [nye]** | **[é-vá] ]** | |
|---|---|---|---|---|---|
| | | | it.is.me | SP-come | |
| | b. | Output β: | **[ [nye-é]** | **[vá] ]** | |
| | | | 1stSG-FOC | come | (Ewe; Heine & Reh, 1984: 111) |

(38)         [CP [NP relative pronoun/operator]i [C' Comp⁰ [IP ti [I' SP-predicate [VP ] ] ] ] ]

| (39) | **ngwana** | **[ya** | **bala-ng** | **hantle]** | |
|---|---|---|---|---|---|
| | child1a | REL1a+SP1a | read-RS | well | |
| | 'the/a child who studies well' | | | | (Southern Sotho) |

| (40) | **mhunu** | **[lw'a-famba-ka]** | | |
|---|---|---|---|---|
| | person1 | REL1+SP1-travel-Rs | | |
| | 'a person who travels' | | | (Ronga (Tsonga dialect); Doke, 1954: 204) |

express agreement with the subject and rela-tivisation simultaneously. This means that the grammaticalisation process that derived rela-tive concords in Nguni has also partially taken place in Southern Sotho (and in the relevant Tsonga dialects). The main difference is that this process has been completed in Nguni, where relative concords are also used in indi-rect relatives, whereas Southern Sotho and Tsonga continue to use non-affixal, clause-ini-tial relative markers in these constructions. It is an interesting question why only Nguni extend-ed the relative concord-strategy to indirect rela-tives. I offer some speculations about a possi-ble answer in section 4.

### 3.1.2 pro-drop

The subject position may also remain phono-logically unrealised in indirect relatives. Southern Bantu languages are *pro*-drop lan-guages, i.e. the subject position need not be filled by an overt NP, in which case the subject prefix is interpreted as a pronoun.[8] Therefore, if the subject of a relative clause is pronominal, nothing intervenes between the predicate in Infl[0] and an element adjacent to IP(41).

(42) and (43) from Tswana and Tsonga illus-trate this situation. The clause-initial relative markers are adjacent to the pronominal subject prefixes (relative markers are in italics; subject prefixes are underlined). As is illustrated by (42) and (43), indirect relatives with pronominal sub-

jects create a context in which a relative mark-er immediately preceding the relative clause-IP is linearly adjacent to the inflected predicate of the relative clause.

### 3.1.3 Subject extraposition

Another context in which the subject position of an indirect relative is not phonologically realised is created by subject extraposition. As shown in (45b) and (46b), Southern Bantu lan-guages allow for right-dislocation of the subject NP (44). For many Nguni speakers, subject extraposition is preferred in relative clauses. This means that even in indirect relative claus-es with full subject NPs, the inflected predicate in Nguni frequently becomes the first overt ele-ment of IP. In early Nguni, this situation would have given rise to adjacency of the predicate and the clause-initial relative marker.

Having shown that the grammar of the Southern Bantu languages does indeed create contexts in both direct and indirect relatives in which relative markers in SpecCP or Comp[0] and the finite verb are adjacent, I now turn to synchronic evidence for the claim that the Nguni relative concord is derived from relative markers of the type used in present-day Sotho and Tsonga.

## 3.2 The relative concord and demon-strative pronouns

It has frequently been noted in the literature (cf.

(41)        [CP [relative marker] [IP **e** [I' SP-predicate [VP ] ] ] ]

(42)        **golô**          [*mô-ba-go-rata-ng*]
            place15          REL15-SP2-OC15-like-RS
            'a place which they like'                                                                       (Tswana)

(43)        **tihomu**        [*leti*        hi-ti-risa-ka]
            cattle10         REL10          1stPL-OC10-herd-RS
            'the cattle which we herd'                                                                      (Tsonga)

(44)        [CP [relative marker] [IP **t**i [I' SP-predicate ] ... (*subject*)i ] ]

(45) a.  **Indoda**        [amakhwenkwe          *a-yi-bon-ile-yo*]           **i-ya-fika**
         man9             boy6                  RC6-OC9-see-PERF-RS          SP9-Foc-arrives
     b.  **Indoda**        [*a-yi-bon-ile*       amakhwenkwe]                 **i-ya-fika**
         man9             RC6-OC9-see-PERF      boy6                         SP9-Foc-arrives
         'The man whom the boys saw is arriving.'                                                           (Xhosa)
(46) a.  **Setulo**  [*seo*    **basadi**          **ba-se-rek-ile-ng**         **kajeno]**     **se**    **robehile**
         chair7    REL7    woman2             SP2-OC7-buy-PERF-RS          today          SP7     be.broken
     b.  **Setulo**  [*seo*    **ba-se-rek-ile-ng**    **kajeno**                  ***basadi*]**   **se**    **robehile**
         chair7    REL7    SP2-OC7-buy-PERF-RS    today                      woman2         SP7     be.broken
         'The chair which the women bought today is broken.'                                                (Southern Sotho)

Doke, 1954; Hendrikse, 1975a; Poulos, 1982; Khumalo, 1992; and many others) that the form of the relative concords in Nguni closely resembles the form of the demonstrative pronouns of the so-called "1st position" (which expresses proximity). (47) – (49) illustrate this fact with a few examples.

Demonstratives, like relative concords, consist of an initial morpheme plus an affix which is identical to the subject prefix; this affix determines the phonological form of the initial morpheme as **(l)a-**, **(l)o-** or **(l)e-** (cf. Khumalo, 1992; Van der Spuy, 2001), and is deleted if it is a vowel. The differences between relative concords and demonstratives are almost negligible. In Zulu, the relative concord has lost the initial **I-** of the demonstrative. The same holds for all Xhosa relative concords, but the close relation between demonstratives and relative concords is reflected in this language by the fact that some demonstrative pronouns (e.g. **eli** of class 5) have also lost the lateral consonant (in some non-standard Xhosa dialects, the **I-** is dropped with demonstratives in all noun classes). In Swati, both demonstrative pronouns and relative concords have maintained the **I-**; relative concords and demonstratives are therefore identical in all classes.

I assume that the parallels illustrated by (47) – (49) reflect a diachronic relation between demonstratives and relative concords, and I suggest that the latter are historically derived from the former. Now recall from section 2 that in Sotho and Tsonga, demonstratives are used as relative markers in relative clause-initial position. The claim that the relative concords of present-day Nguni are derived from relative markers of the Sotho and Tsonga type now provides the missing link that connects these two observations. Relative concords and demonstratives in Nguni have identical structures because in early Nguni, like in present-day Sotho and Tsonga, demonstratives were used as fully-fledged, non-affixal relative markers – they were either relative pronouns in SpecCP or relative complementisers in Comp$^0$. Without a subject noun in SpecIP (a situation which, as was shown in section 3.1, frequently occurs in Bantu relatives), the relative marker/demonstrative pronoun would have been adjacent to the relative clause predicate and ultimately was reanalysed as being prefixed to it — it would have become a relative concord.

This historical relation between demonstrative pronouns and relative concords has left an interesting synchronic trace in Xhosa. According to Jordan (1967) and Hendrikse (1975a), if the head noun in a Xhosa relative is modified by a demonstrative, the relative concord cannot occur. Instead, the subject prefix must be chosen. Although according to my Xhosa informant, the co-occurrence of a demonstrative and a relative concord is not entirely ungrammatical, but only slightly marked, (50b), it is a crucial observation that the presence of a demonstrative on the head noun can trigger the occurrence of the subject prefix instead of the relative concord, (50c). Importantly, using the subject prefix in a relative clause is *only* possible if the head noun phrase

| (47) | Zulu: | | | class 2 | class 9 | class 5 |
|------|-------|---|---|---------|---------|---------|
| | | relative concord | | **aba-** | **e-** | **eli-** |
| | | demonstrative | | **laba** | **le** | **leli** |
| (48) | Xhosa: | | | class 2 | class 1/1a | class 5 |
| | | relative concord | | **aba-** | **o-** | **eli-** |
| | | demonstrative | | **aba** | **lo** | **eli** |
| (49) | Swati: | | | class 10 | class 1 | class 5 |
| | | relative concord | | **leti-** | **lo-** | **leli-** |
| | | demonstrative | | **leti** | **lo** | **leli** |

| (50) | a. | **umntwana** | **[o-lila-yo]** | (**the/a** child who is crying) |
|------|----|--------------|-----------------|---------------------------------|
| | | child1 | Rc1-cry-Rs | |
| | b. | **lo mntwana** | **[o-lila-yo]** | (**this** child who is crying) |
| | | Dem child1 | Rc1-cry-Rs | |
| | c. | **lo mntwana** | **[u-lila-yo]** | (**this** child who is crying) |
| | | Dem child1 | Sp1-cry-Rs | |
| | d. | **umntwana** | **[u-lila-yo]** | (**the/a** child who is crying) |
| | | child1 | Sp1-cry-Rs | (Xhosa) |

includes a demonstrative, (50d).

The contrast between (50c) and (50d) provides strong evidence that the ancestor of the Xhosa relative concord was itself a demonstrative pronoun. Given that demonstratives are frequently used as relative pronouns or relative complementisers in other Southern Bantu languages, this in turn can be taken as support for my claim that Nguni relative concords are historically derived from relative markers occupying the SpecCP or Comp[0]-position of the relative clause.[9]

Since the relative marker in early Nguni was a demonstrative pronoun, its form would have been based on the initial morpheme **(l)a**- and an element identical to the subject prefix of the respective noun class. I have suggested that relative concords were derived by merging the relative marker with the (participial) subject prefix in Infl[0]. This raises the interesting question of whether the subject prefix of the relative concord is a remnant of the subject prefix-part of the demonstrative pronoun, or if it is based on the participial subject prefix of the relative clause's predicate.

It is difficult to answer this question on the basis of the synchronic properties of the relative concord in present-day Nguni. Some observations suggest that the relative marker/demonstrative pronoun completely *replaced* the subject prefix of the verb (this is the position advocated by Doke, 1954). For example, the participial prefix of class 2 in Zulu is **be**-, but the relative concord is **aba**- (not **abe**-), indicating that the participial mood of the relative clause has no impact on the form of the relative concord. The same point can be made with respect to the relative concord of class 1/1a in Zulu direct relatives, which is **o**-, and therefore seems based on the form of the demonstrative

**lo** (and not on the participial subject prefix **e**-). However, there is also evidence that suggests that the relative concords are based on the subject prefixes. First, there are relative concords of the first and second person, for which there are no corresponding demonstrative pronouns. Hence, these markers must be derived by merging the relative morpheme with the subject prefix of the verb (e.g. 1stPers.Sg **(l)engi**- < (l)a- **+** ngi-). Second, morphological rules that apply to regular subject prefixes also apply to the class prefix of the relative concord. For example, the rule that deletes subject prefixes in front of the 'exclusive'- marker **se** in Zulu in certain contexts, (51a), also deletes the subject prefix-part of the relative concord in these contexts, (51b) (cf. Khumalo, 1992: 216f.).

Finally, recall that the relative concord of class 1/1a in Zulu and Xhosa indirect relatives is **a**-. This follows neither from the form of the demonstrative, which is **lo**, nor from the form of the participial subject prefix, which is **e**-. Moreover, the fact that the set of relative concords in Nguni includes two different markers for the same noun class is itself a puzzling property of the Nguni paradigm, which seems difficult to explain.

These complex morphological properties of the relative concord are less surprising, however, if the different contexts in which reanalysis took place in Nguni are taken into consideration. Since the relative marker located in SpecCP or Comp[0] and the subject prefix could end up in adjacent positions in both direct and indirect relatives, a number of morphophonological patterns could arise which were determined by the conditions in (52).

I assume that, due to the different combinatorial possibilities which arise from (52), the phonological merging process that combined

(51) a. **se-li-hamba**                         <        *li*-se-li-hamba
           ∅-Excl-Sp5-leave                            Sp5-Excl-Sp5-leave
           'it is already going'

      b. **ihhashi  [e-se-li-hamba]**  <    a-*li*-se-li-hamba
           horse5    Rm-∅-Excl-Sp5-leave      Rm-Sp5-Excl-Sp5-leave
           'the horse which is already going'

(52) a. The relative marker could have been based on a vowel subject prefix and therefore would have been monosyllabic, consisting only of the relative morpheme, or it could have been morphologically complex, consisting of the relative morpheme and the part that is identical to the subject prefix of the respective noun class.

      b. The noun class of the relative marker could have been different from the noun class of the participial-subject prefix of the verb (in indirect relatives), or both could have belonged to the same noun class.

relative markers and subject prefixes in early Nguni did not take place in a uniform way. I believe that synchronically, relative concords are complex Infl$^0$-elements whose parts may be affected by the same rules as subject prefixes. However, their morphophonological properties reflect different and partly idiosyncratic aspects of a historical process in which a morphologically complex relative marker and a participial subject prefix merged into one single element.

It should be mentioned that various researchers have tried to offer a *synchronic*, rather than a diachronic, explanation for the similarities between the morphological structure of demonstratives and relative concords. For example, since demonstratives and relative clauses typically establish *referentiality* of the noun with which they combine, Hendrikse (1975a) argues that the source of referentiality is the initial morpheme **(l)a**-, which forms part of the morphological structure of both demonstratives and relative concords.

Two points can be made with respect to this analysis. On the one hand, a synchronic approach such as Hendrikse's is certainly not incompatible with the diachronic analysis proposed here. On the contrary, it is well possible that a common synchronic function of the initial morpheme which occurs with the relative concord and the demonstrative is the result of a diachronic process that derived the former from the latter.

On the other hand, the idea that the same element is the source of referentiality in both demonstratives and relative clauses is not entirely unproblematic. For example, Schmitt (2000) also argues that demonstratives and relative clauses provide referentiality; however, she contends that the referential property of relative clauses is actually linked to the fact that they are specified for *tense*. In contrast, Schmitt shows that the referential property of demonstratives is due to their *deictic* function. Clearly, demonstratives are not specified for tense, and the deictic component of demonstratives is absent from relative clauses. If Schmitt's analysis is correct, then the initial morpheme **(l)a**-

cannot be the locus of referentiality in both demonstratives and relative concords.

In light of these comments, I conclude that the historical analysis offered here provides the most plausible explanation for the similarities between relative concords and demonstrative pronouns in Nguni. In the next section I present synchronic evidence for this analysis which to my knowledge has not been discussed in this context before.

### 3.3 "Strategy 2" in Zulu and Xhosa relatives

Consider the two relatives in (53), which show a relative construction in which the head noun corresponds to a *possessor* of the subject NP in the relative clause. The structure of the relative in (53a) is in accordance with the properties of relative clauses in Nguni discussed in section 2.3. The complex subject NP **umntwana wayo**, 'his child', is of class 1, and the relative concord agrees with this NP. The possessive marker combines with the pronominal clitic -**yo** of class 9 which expresses agreement with the head noun.[10]

However, according to Poulos (1982), (53a) is slightly marked. Poulos claims that the preferred alternative to form possessive relatives in Zulu is (53b), a construction that he labels "Strategy 2". In Strategy 2, there is no relative concord attached to the verb; the relative suffix may be omitted,[11] and the verb appears in the indicative mood, as illustrated by the focus marker -**ya**-, which occurs with intransitive verbs in the indicative. The most interesting aspect of Strategy 2, however, is the occurrence of an element which is identical to a relative concord, but which seems to be prefixed to the *initial noun* of the relative clause, which usually is the subject (but see (56)). The relative concord thereby seems to replace the initial vowel of this noun. Importantly, the "displaced" relative concord no longer agrees with the subject, but with the head noun (note that in (53b), we get class 9 **e**-, not class 1 **o**-).

Poulos (1982) argues that Strategy 2 is not restricted to relative constructions such as (53),

| (53) | a. | **inkosi** | **[umntwana** | **wa-yo** | **o-gula-yo]** |
|------|----|-----------|---------------|----------|----------------|
| | | chief9 | child1 | Poss1-Pc9 | Rc1-be.sick-Rs |
| | b. | **inkosi** | **[e-mntwana** | **wa-yo** | **u-ya-gula]** |
| | | chief9 | Rc9-child1 | Poss1-Pc9 | Sp1-Foc-be.sick |
| | | 'the chief whose child is ill' | | | (Zulu; Poulos, 1982: 111) |

where the head noun corresponds to a posses-sor. He claims that Strategy 2 is also possible with indirect relatives like (54), where the head noun corresponds to an object. (54a) shows the relative concord-strategy; (54b) is the Strategy 2-alternative.

The examples (55) – (56) from Pahl (1983) show that Strategy 2 is also possible in Xhosa. Notice that the relative marker in (56) is not attached to the relative clause-subject, but to an object noun. The object NP (which includes the possessor) has been preposed and there-fore is the initial element of the relative clause.

I assume that Strategy 2 in Zulu and Xhosa provides further synchronic evidence for the existence of a relative clause formation strate-gy in early Nguni which was similar to the strat-egy that is still synchronically used in Sotho and Tsonga. The crucial property of Strategy 2, namely the occurrence of a relative marker which *precedes* the initial noun of the relative clause and which *agrees with the head noun*, exactly resembles the situation in Sotho and Tsonga. Strategy 2 can hence be regarded as a reflex of the relative clause-formation strategy that I suggested has existed in early Nguni.

In this regard, it is relevant to mention that Strategy 2 is not accepted as a possible con-struction by all speakers. Many of my Zulu informants rejected Strategy 2 in relative claus-es; some of those speakers who accepted it still considered it "clumsy". It seems that Strategy 2 in Zulu is generally preferred if the sentence in which it occurs has the status of an idiomatic expression or proverb (cf. Zeller, 2003: 234). My Xhosa informant accepted Pahl's (1983) examples in (55) and (56), but pointed out that even these acceptable exam-ples somehow "sounded strange". This prob-lematic status of Strategy 2 might be because it

reflects properties of an earlier stage of Nguni. Strategy 2 might not be fully productive in mod-ern Zulu or Xhosa any longer, at least not in all areas. Its rare occurrences appear to be relics of an older form from which the relative con-cord-strategy of present-day Nguni has been derived. The fact that one finds dialectal and geographical variation with respect to the acceptability of this strategy is not surprising if we assume that in some speaker groups, the older form has just been preserved longer than in others. The occurrence of Strategy 2 in proverbs and fixed expressions is also not unexpected, because it is exactly this kind of linguistic data that often reflect older historical stages of a language.

Although the relative concords of Strategy 2 are similar to those found in present-day Sotho and Tsonga with respect to their agreement properties and their relative clause-initial posi-tion, there is also an obvious difference between these markers. Whereas the Sotho and Tsonga relative markers are phonologically independent words, the Nguni markers of Strategy 2 are *clitics*; i.e. phonologically depen-dent elements which need a host. This host is provided by the adjacent noun; therefore, the relative markers of Strategy 2 cliticise to this noun, creating the impression that they are "prefixed" to the noun stem. However, in con-trast to genuine word-level affixes, clitics do not select particular hosts; any adjacent word will do. This explains why the relative marker of Strategy 2 is not only found attached to sub-jects, but also attached to fronted objects.

In Zeller (2003), I analyse the relativising clitics of Strategy 2 as "phrasal affixes", follow-ing Anderson's (1992) theory of inflection. Anderson (1992) shows that inflectional affixes are not restricted to the domain of words. They

(54) a. **inja**         **[umfana**       *a***-yi-theng-ile-yo]**
         dog9     boy1            Rc1-Oc9-buy-Perf-Rs
     b. **inja**         **[e-mfana**       **u-yi-theng-ile]**                                         (Strategy 2)
         dog9     Rc9-boy1      Sp1-Oc9-buy-Perf
         'the dog which the boy bought'                                          (Zulu; Poulos, 1982: 171f.)

(55)   **umfazi**              **[o-ndoda**         **ya-khe**          **i-swelek-ile-yo]**
         woman1           Rc1-man9        Poss9-Pc1      Sp9-die-Perf-Rs
         'the woman whose man died'
(56)   **izithethi**            **[ezi-ntetho**       **za-zo**           **si-zi-phulaphule-yo**          **namhlanje]**
         speakers10     Rc10-talk10     Poss10-Pc10   1stPl-Oc10-listen.to-Rs   today
         'the speakers whose talks we listened to today'                      (Xhosa; Pahl, 1983: 217)

may also be attached to whole phrases, there-by expressing a particular function associated with these phrases. According to Anderson, inflectional phrasal affixes are typically clitics. For example, the English possessive clitic **'s** is a phrasal affix; it attaches to the whole preced-ing noun phrase by cliticising to its last word, thereby marking the whole phrase as a pos-sessor (57). Whereas the possessive clitic in English is a phrasal suffix, the Nguni relativising clitic of Strategy 2 is a phrasal prefix, because it precedes the whole relative clause IP. Its function is to express agreement between the relative clause and the head noun (see Zeller, 2003 for details of this analysis).

I assume that Strategy 2 represents an intermediate stage of the grammaticalisation process which turned demonstratives/relative markers into relative concords. Before the fully-fledged markers of early Nguni were reanalysed as relative concords, they turned into clitics — phrasal affixes which were attached to the relative clause IP. As such, the relative markers of Strategy 2 show a typical characteristic of grammaticalisation processes which Heine *et al.* (1991) call *overlapping*. When an element or structure develops from a stage A to a stage C, there is always an inter-mediate stage B in which the respective ele-ment exhibits properties of both stages A and C. The relative marker of Strategy 2 shows pre-cisely this behaviour. It still has the agreement properties of an early-Nguni relative marker (stage A), but it is phonologically bound to a host and therefore also reflects properties of the affixal relative concord-strategy of present-day Nguni (stage C).

I have left open the question whether the fully-fledged relative markers of Stage A in early Nguni were relative pronouns in SpecCP or rel-ative complementisers in Comp$^0$. However, I assume that the phrasal affixes of Stage B of the grammaticalisation process were unambigu-ously Comp$^0$-elements. This assumption follows from the fact that, in the syntactic framework that I have used in this article, inflectional fea-tures like agreement or tense, and some of the elements which express these features (such as auxiliaries, determiners, etc.), are usually asso-ciated with functional heads. For example, the English possessive marker **'s** discussed above is commonly treated as the functional head D$^0$ of a determiner phrase DP (which is the stan-dard representation of noun phrases since Abney (1987)) (58). Since possessor noun phrases are located in SpecDP, the D$^0$-position directly *follows* the possessor, and the clitic therefore attaches to the last word of this phrase. If the phrasal affixes of Strategy 2 are also analysed as functional heads, then it fol-lows that the syntactic position of these rela-tivising clitics must be a functional head which directly *precedes* the relative clause IP. This functional head is Comp$^0$. A relative clause con-struction like that in (55) can therefore be repre-sented syntactically as (59). As an element in Comp$^0$, the relative marker of Strategy 2 is an affixal relative complementiser which, due to this position, is adjacent to the relative clause IP and cliticises to the first word of this phrase.

(57)    I once knew [that guy you're talking about]'s brother in law.

(Anderson, 1992: 212)

(58)



(59)
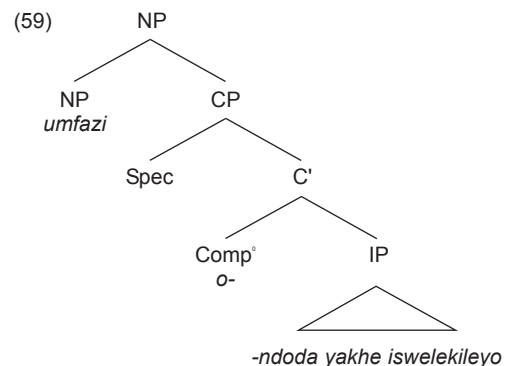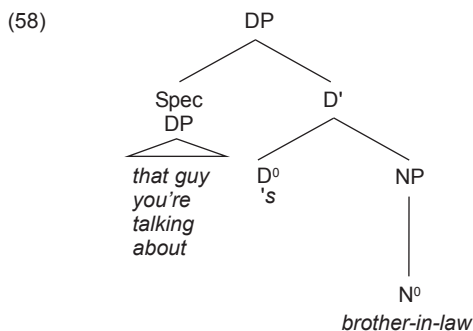
To summarise, early Nguni used fully-fledged demonstratives as relative markers, either as relative pronouns in SpecCP or as relative complementisers in Comp⁰. At some stage, these relative markers were phonologically weakened; they turned into affixal Comp⁰-elements and cliticised to the first word of the relative clause. In contexts where this first word was the relative clause predicate, the relativising clitic and the adjacent subject prefix would phonologically merge into a single inflectional prefix. As a consequence, the Nguni languages developed a full set of relative concords with the agreement properties of subject prefixes.

## 4. Conclusion

In this article I have provided a comparison of the relative clause formation strategies in the Southern Bantu languages of South Africa. I have argued that the relative concords found in Nguni are derived historically from clause-initial relative markers, suggesting that relative clauses in early Nguni were formed in the same way as in present-day Sotho and Tsonga. One obvious question that arises from this analysis is why relative concords have developed only in Nguni indirect relatives, and not in Sotho or Tsonga (although Southern Sotho and certain Tsonga dialects have undergone similar developments in direct relatives).

A possible answer might be provided by the clitic properties of the relative marker in Strategy 2. As was shown in section 3.3, this marker is phonologically bound to an adjacent host. If the assumptions made in this article are correct, then Strategy 2 reflects properties of an historical stage at which relative markers in early Nguni, *in contrast to those in Sotho and Tsonga*, turned into clitics. It is possible that this difference is responsible for the synchronic difference between these language groups: The clitic-status of the Nguni relative marker may have smoothed the progress of reanalysis and the development of affixal relative concords. It is well-known that clitics are more likely to develop into bound morphemes than independent words, given that clitics are already phonologically bound to a host. In losing their status as phonologically independent words, the demonstratives/relative markers of early Nguni were already one step closer to the relative concord-stage than the demonstratives/relative markers occurring in the same syntactic environment in Sotho and Tsonga. I believe that this idea warrants more detailed consideration; in Zeller (2004), the difference between Nguni relative concords on the one hand and Sotho relative markers on the other is further explored along these lines.

## 5. Notes

1. In contrast to modern English, Middle English also had constructions which featured both a relative pronoun and a relative complementiser: **This yongeste, *which that* wente unto the toun** (Chaucer, *Pardoner's Tale*, quoted from Faiß, 1977: 249).

2. The majority of the examples that I present in this article have been provided or checked by native speakers; otherwise, I have given the reference to the original example in the text. In the glosses, I mark the noun classes and agreement through numbers, according to Meinhof's (1906) numbering system of Proto-Bantu. Morphemes are glossed as follows: Cop = copula; Def = sentential definitiser; Dem = demonstrative pronoun; Det = determiner; Excl = exclusive tense; Foc = focus marker; Neg = negation; Oc = object clitic; Pc = pronominal clitic; Perf = perfect tense; Pl = plural; Poss = possessive marker; Rc = relative concord; Rcomp = relative complementiser; Rel = relative marker; Rf = relative formative; Rm = relative morpheme; Rs = relative suffix; Sg = singular; Sp = subject prefix; Tns = Tense.

3. Chumbow (1977) shows that the element **bah** in Ngemba is optional and that the demonstrative/relative pronoun can also occur on its own in relative clauses. This argues against an analysis (suggested by a reviewer) according to which the demonstrative and **bah** would form a morphologically complex complementiser in Comp⁰.

4. The Nguni data I present are from Zulu, Xhosa and Swati. Since Southern Transvaal Ndebele behaves like Zulu in all respects relevant for this article, I do not provide Ndebele examples.

5. If a high vowel subject prefix is followed by a vowel-commencing morpheme, the former is replaced by a glide, cf. Khumalo (1992).

6. Special thanks to Romeo Matumba for providing and discussing the Venda data.

7. A reviewer claims, contra Doke and Poulos, that constructions (i) and (ii) are also possible

with indirect relatives. (S)he quotes the examples in (1), where the head noun is used as an object in the main clause. However, according to my Venda informant, (1a) and (1b) are ungrammatical. He also rejected the indirect relatives in (2) (corresponding to the indirect relative in (32) in the text), which modify the subject of the sentence. Nevertheless, if the reviewer is correct, and (1) and (2) are indeed possible for some Venda speakers, then it has to be concluded that there is significant idiolectal variation among Venda speakers with respect to the acceptability of constructions (i) and (ii) in indirect relatives.

[8] A popular view in some syntactic theories is that the pronominal reading is not induced by the verb's inflectional morphology, but by *pro*, a pronominal with no phonetic content which is located in SpecIP (cf. Chomsky, 1982). The *pro*-drop properties of Xhosa are analysed along these lines in Visser (1986); arguments against the existence of *pro* in Zulu are provided by Van der Spuy (2001).

[9] The use of demonstratives as relative pronouns or relative complementisers is itself the result of a reanalysis process. The Xhosa data in (50) suggest that the demonstrative used in relative clauses originally was part of the head noun phrase. Since demonstratives can also follow their nouns, a structure was created in which a demonstrative was adjacent to a subsequent sentence which probably was added as an afterthought to the preceding sentence including the head noun. This structure was then reanalysed, with the demonstrative being treated as the initial element of a subordinate clause (cf. Heine & Reh (1984: 29, 108f.) on a similar development in Ewe).

[10] Since the head noun does not correspond to the subject of the relative clause (it corresponds to a possessor which is *part of* the subject noun phrase), the construction in (53) should be classified as an indirect relative. However, recall that the relative concord of class 1 or 1a in Zulu shows up in two different forms, depending on whether it occurs in a direct relative (the relative concord is **o**-) or in an indirect relative (the relative concord is **a**-). Curiously, and for reasons that are not clear, the relative concord in (53a) is **o**-. Therefore, constructions in which the head noun corresponds to a possessor of the subject of the relative clause are also called direct relatives, a misleading terminology, as pointed out by Hendrikse (1975b) and Poulos (1982, chapter 5).

[11] The data from Strategy 2 in Zulu are not uniform with respect to the occurrence of the relative suffix -**yo**. Whereas it does not seem to be licensed in the examples presented by Poulos (1982), the relative suffix appears in other examples I found in the literature. Furthermore, the examples in (55) and (56) below show that the relative suffix still occurs in Strategy 2 in Xhosa. These observations raise interesting questions that I cannot address here (but see Zeller (2003) for some discussion of the differences between Strategy 1 and Strategy 2 in Zulu).

(1)  a.  **Musidzana**  **u**    **bika**  **zwiliwa**   **[vhana**  **vha**   **zwi-funa-ho]**   (construction (i))
          girl         SP1    cook    food8        child2    SP2     OC8-like-RS
     b.  **Musidzana**  **u**    **bika**  **zwiliwa**   **[vhana**  **vha**   **no zwi-funa]**   (construction (ii))
          girl         SP1    cook    food8        child2    SP2     RF OC8-like
          'The girl cooks food which the children like'

(2)  a.  **\*Munna**  **[nngwa**  **dzi**   **mu-pandamedza-ho]**  **u**    **takalela**  **u bika**   (construction (i))
          man1       dog10    SP10    OC1-chase-RS          SP1    like        to-cook
     b.  **\*Munna**  **[nngwa**  **dzi**   **no mu-pandamedza]**  **u**    **takalela**  **u bika**   (construction (ii))
          man1       dog10    SP10    RF OC1-chase          SP1    like        to-cook
          intended: 'The man whom the dogs are chasing likes to cook'

# 6. References

**Abney S**. 1987. *The English Noun Phrase in Its Sentential Aspect*. PhD Thesis, Massachusetts Institute of Technology, Cambridge, USA.

**Alexiadou A, Law P, Meinunger A & Wilder C**. 2000. Introduction. In: Alexiadou A, Law P, Meinunger A & Wilder C (eds) *The Syntax of Relative Clauses*. Amsterdam / Philadelphia: John Benjamins. pp.1–51.

**Anderson S-R**. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.

**Bayer J**. 1984. COMP in Bavarian Syntax. *The Linguistic Review* **3**: 209–274.

**Bokamba EG**. 1976. *Question Formation in Some Bantu Languages*. PhD Thesis, Indiana University, Bloomington, USA.

**Chomsky N**. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge: MIT Press.

**Chomsky N**. 1986. *Barriers*. Cambridge: MIT Press.

**Chomsky N**. 1995. *The Minimalist Program*. Cambridge: MIT Press.

**Chomsky N & Lasnik H**. 1993. The theory of principles and parameters. In: Jacobs J, Von Stechow A, Sternefeld W & Vennemann T (eds) *Syntax: An International Handbook of Contemporary Research*. Berlin: De Gruyter.

**Chumbow B**. 1977. Relatives as determiners: A case from Ngemba. In: Kotey PFA & Der-Houssikian H (eds) *Language and Linguistic Problems in Africa* (Proceedings of the Seventh Conference on African Linguistics). Columbia: Hornbeam Press Incorporated. pp. 283–302.

**Demuth K & Harford C**. 1999. Verb raising and subject inversion in Bantu relatives. *Journal of African Languages and Linguistics* **20**: 41–61.

**Doke CM**. 1954. *The Southern Bantu Languages*. London / New York / Cape Town: Oxford University Press.

**Faiß K**. 1977. *Aspekte der englischen Sprachgeschichte*. Tübingen: Gunter Narr.

**Givón T**. 1975. Serial verbs and syntactic change: Niger-Congo. In: Charles N Li (ed) *Word Order and Word Order Change*. Austin: University of Texas Press. pp. 47–112.

**Güldemann T**. 1999b. Toward a grammaticalization and typological account of the *ka*-possessive in southern Nguni. *Journal of African Languages and Linguistics* **20**(2): 157–184.

**Harford C & Demuth K**. 1999. Prosody Outranks Syntax: An Optimality Approach to Subject Inversion in Bantu Relatives. *Linguistic Analysis* **29**: 1–2, 47–68.

**Heim I & Kratzer A**. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

**Heine B & Reh M**. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.

**Heine B, Claudi U & Hünnemeyer F**. 1991. From cognition to grammar — evidence from African languages. In: Closs Traugott E & Heine B (eds) *Approaches to Grammaticalization*. Amsterdam: John Benjamins. pp. 149–187.

**Hendrikse AP**. 1975a. A pre-theoretical analysis of the relative marker in Xhosa. Topics in Xhosa Relativization. (Communication No. 4, Department of African Languages). Grahamstown: Rhodes University.

**Hendrikse AP**. 1975b. Direct and indirect relative clauses. Topics in Xhosa Relativization. (Communication No. 4, Department of African Languages). Grahamstown: Rhodes University

**Hopper PJ & Closs Traugott E**. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.

**Hyman LM**. 1975. On the change from SOV to SVO: Evidence from Niger-Congo. In: Li CN (ed) *Word Order and Word Order Change*. Austin: University of Texas Press. pp. 113–147.

**Jordan AC**. 1967. The adjective in the Nguni languages of Southern Africa with special reference to Xhosa. *Journal of African Languages* **6**.2: 132–145.

**Kayne R**. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.

**Khumalo JSM**. 1992. The morphology of the direct relative in Zulu. In: Gowlett DF (ed) *African Linguistic Contributions*. Hatfield: Via Afrika Limited. pp. 210–226.

**Livingstone W**. 1996. *Relative Clauses in Luganda*. Cologne: Rüdiger Köppe.

**Meeussen AE**. 1971. Relative clauses in Bantu. *Studies in African Linguistics*, Suppl. 2: 3–10.

**Meinhof C**. 1906. *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin: Dietrich Reimer.

**Mischke G**. 1998. Southern Sotho verbal relative constructions. *South African Journal of African Languages* **18**(4): 106–111.

**Mutaka NN**. 2000. *An Introduction to African Linguistics*. Lincom Europa: München.

**Ngonyani D**. 2001. Evidence for head raising in Kiswahili relative clauses. *Studies in*

*African Linguistics* **30**(1): 59–73.

**Pahl HW**. 1983. *IsiXhosa*. Johannesburg: Educum Publishers.

**Pollock J-Y**. 1992. *Opérateurs neuls, dont, questions indirectes et théorie de la quantification*. Unpublished manuscript, Université de Rennes.

**Poulos G**. 1982. *Issues in Zulu Relativization*. PhD Thesis, Rhodes University, Grahamstown, South Africa. (Communication No. 11, Department of African Languages).

**Poulos G**. 1990. *A Linguistic Analysis of Venda*. Hatfield: Via Afrika.

**Poulos G**. 1999. Grammaticalisation in South-Eastern Bantu and the linguistic 'dynamics' underlying this process. *South African Journal of African Languages* **19**(3): 204–214.

**Poulos G & Louwrens LJ**. 1994. *A Linguistic Analysis of Northern Sotho*. Pretoria: Via Afrika.

**Schmitt C**. 2000. Some consequences of the complement analysis for relative clauses, demonstratives and the wrong adjectives. In: Alexiadou A, Law P, Meinunger A & Wilder C (eds) *The Syntax of Relative Clauses*. Amsterdam / Philadelphia: John Benjamins. pp. 309–348.

**Van der Spuy A**. 1993. Dislocated noun phrases in Nguni. *Lingua* **90**: 335–355.

**Van der Spuy A**. 2001. *Grammatical Structure and Zulu Morphology*. PhD Thesis, University of the Witwatersrand, Johannesburg, South Africa.

**Visser MW**. 1986. Cliticization and case theory in Xhosa. *South African Journal of African Languages* **6**(3): 129–137.

**Wald B**. 1970. Relativization in UmBundu. *Studies in African Linguistics* **1**(2): 131–256.

**Zeller J**. 2003. Word-level and phrase-level prefixes in Zulu. *Acta Linguistica Hungarica* **50**(1–2): 227–244.

**Zeller J**. 2004. *On the relation between noun prefixes and grammaticalisation in Nguni relative clauses.* Unpublished manuscript, University of KwaZulu-Natal, Durban, South Africa.

**Zwart JW**. 2000. A head raising analysis of relative clauses in Dutch. In: Alexiadou A, Law P, Meinunger A & Wilder C (eds) *The Syntax of Relative Clauses*. Amsterdam / Philadelphia: John Benjamins. pp. 349–385.