

Questions in Northern Sotho^{*}

Sabine Zerbian

ZAS, Berlin

This article gives an overview of the marking of polar and constituent questions in Northern Sotho, a Bantu language of South Africa. It thereby provides a contribution to the typological investigation of sentence types in the world's languages. As will be shown, Northern Sotho follows cross-linguistic tendencies in marking interrogative sentences: It uses intonation as main indicator in polar questions and question words as main indicator in constituent questions. Nevertheless, it also shows interesting language-specific variation, e.g. with respect to the location of raised intonation in polar questions, the presence of two pragmatically distinct question particles in polar questions, or a split in the formation of constituent questions based on the grammatical function of the questioned constituent.

1 Introduction

Northern Sotho (Sesotho sa Leboa, also known as Sepedi after its standardized dialect) is one of the eleven official languages of the Republic of South Africa. It is spoken in the northern provinces of South Africa by approximately 4,208,980 speakers (Statistics South Africa 2004). According to Guthrie's (1967-71) classification it belongs to group S30. It is mutually intelligible with the other languages in this group, namely Tswana and Southern Sotho (also referred to as Sesotho). Like other Bantu languages Northern Sotho is a tone language that uses changes in fundamental frequency to indicate lexical and grammatical differences in meaning. Syntactically, it displays SVO basic word order. Morphologically, Bantu languages are famous for their rich noun class agreement system (for grammars of Northern Sotho see e.g. Ziervogel et al. 1969, Louwrens 1991, Poulos & Louwrens 1994, Louwrens et al. 1995).

* Thanks to my language informant Lister Maphuti Matloga for providing the data and Laura J. Downing, Andreas Haida, Hans-Martin Gärtner, and Lutz Marten for feedback on earlier versions of this paper.

As interrogatives in Northern Sotho are formed on the basis of declarative sentences, basic sentence structure in Northern Sotho declaratives is exemplified in (1).¹

- (1) Declarative sentences in Northern Sotho
- a. Mo-nna o ngwal-ela ba-sadi lehono.²
 CL1-man SC1 write-APPL CL2-woman today
 ‘The man writes to (the) women today.’
- b. Ba-sadi ba ngwal-ela mo-nna lehono.
 CL2-woman SC2 write-APPL CL1-man today
 ‘The women write to (the/a) man today.’

The sentences in (1) display the basic word order SVO. The subject in Northern Sotho declaratives (and more generally in Bantu languages) is marked unambiguously through agreement with the verb. In agreement the noun is related to the verb via an agreement marker that agrees with the noun in its noun class (in example (1a) this is class 1, in (1b) it is class 2). The object follows the verb. In Northern Sotho, there is no agreement between the object and the verb in basic SVO structures, as displayed in (1).

On the basis of the declarative sentence additional and diverging features will be presented that distinguish declarative sentences from interrogatives. Section 2 discusses interrogative clause formation in Northern Sotho, addressing the formation of polar questions and their prosodic and morphosyntactic aspects in section 2.1, and the formation of constituent questions in section 2.2. Though Northern Sotho uniformly uses question words in constituent questions, there is a striking asymmetry in the distribution of the question words. Therefore, non-subject and subject constituent questions are dealt with separately, in sections 2.2.1 and 2.2.2 respectively. Section 2.2.3 briefly mentions alternative strategies of constituent question formation for both subjects and non-subjects whereas

¹ The examples are given in Northern Sotho orthography which is disjunctive, which means that a.o. concordial markers are written separately. For an overview of Southern Bantu writing systems see Louwrens (1991).

The abbreviations used in the glosses are the following:

1SG	first person singular	PASS	passive
2SG	second person singular	POSS	possessive pronoun
PRES	present tense-a-	PST	past tense
APPL	applicative	PROPNAM	proper name
CL	noun class	QP	question particle
COP	copula	REL	relative tense
LOC	locative	RPRN	relative pronoun
OC	object concord	SC	subject concord

section 2.2.4 discusses the observed asymmetry in more detail. Section 3 provides a summary of the findings.

2 Interrogatives in Northern Sotho

Cross-linguistic research on sentence types (e.g. König & Siemund, to appear) investigates grammatical distinctions that can be correlated with a certain illocutionary force. Interrogatives are typically used for eliciting information and asking questions. Apart from this functional criterion, a definition of sentence types should also be based on formal criteria in that the formal properties characterizing a sentence type set up a system of alternative choices that are mutually exclusive.

Polar and constituent questions in Northern Sotho share the general illocutionary force of interrogatives. Moreover, they also find their place in a cross-linguistic definition of formal properties of interrogatives, as they display cross-linguistically prominent patterns both in the formation of polar as well as in the formation of constituent questions. These as well as interesting language-specific variation will be presented in the following sections. The presentation concentrates on interrogatives as main clauses only. Further research has to show if what holds for non-embedded interrogatives is also true for embedded structures.

2.1 Polar questions

Polar questions (also termed *yes/no-questions*) are interrogative sentences whose answer is expected to provide the truth-value of the corresponding declarative. In Northern Sotho they are parallel to corresponding declarative sentences in that they show basic word order and identical verbal morphology. Polar questions differ from declaratives in prosody and optionally in lexical means (particles). Among the prosodic and morphological means used to mark polar questions, prosody has to be regarded the crucial factor. Therefore, the presentation will start with prosodic marking in section 2.1.1 and turn to the (optional) morphological marking in section 2.1.2.

2.1.1 Prosodic marking of polar questions

In order to understand how polar questions are distinguished from declaratives with respect to prosody one first has to discuss the relevant prosodic properties at sentence level in Northern Sotho.

Vowel length is not contrastive in Northern Sotho. However, the penultimate syllable of a word in isolation or of the last word in a declarative

sentence is systematically lengthened (Doke 1954, Lombard 1979). Penultimate lengthening is common in many Bantu languages. Polar questions lack this length on the penultimate syllable so that the penultimate syllable of the question utterance does not show significantly different length than other syllables in the utterance (Lombard 1979). Additionally, the polar question as a whole is pronounced at a higher overall pitch (Ziervogel et al. 1969, Poulos & Louwrens 1994). The tonal structure of the utterance, however, remains unaltered in Northern Sotho. These observations are illustrated in the example in (2). The lengthened penultimate syllable is indicated by a colon following it. The upright arrow indicates raising of the overall pitch level.³

(2) Polar questions in Northern Sotho⁴

- a. Q: ↑O dumediša di-kgarebe?
 SC2SG greet CL10-lady
 ‘Are you greeting the ladies?’
 A: Ee, ke dumediša di-kgare:be.
 Yes SC1SG greet CL10-lady
 ‘Yes, I am greeting the ladies.’
- b. Q: ↑Mo-kgalabje ó dúmédíšá ngáka?
 CL1-old man SC1 greet CL9.doctor
 ‘Is the old man greeting the doctor?’
 A: Ee, mo-kgalabje ó dúmédíšá ngá:ka.
 Yes CL1-old man SC1 greet CL9.doctor
 ‘Yes, the old man is greeting the doctor.’

The transcription indicates that the question is pronounced at a higher overall pitch, and the absence of the colon indicates that lengthening does not occur. A close comparison of the tone marking of the questions and answers in (2a) and (2b) shows that there are no changes in the tone structure. To give an impression of the raised pitch register, measurements from two examples are given in (3) in order to illustrate the difference in the fundamental frequency in yes/no-questions and their corresponding answers.

³ Following common practice, accents mark high tones and underlying high tones are underlined. For details on the tonological processes in Northern Sotho, see Zerbian (2006a, b).

⁴ Unless stated otherwise, all examples are from my own research on Northern Sotho.

(3) Average F⁰ in yes/no-questions

	Mokgalabje	ó dúmédíšá ngá(:)ka (cf.2b)
F ⁰ Question	232 Hz	253 Hz
F ⁰ Answer	167 Hz	181 Hz
	Ngaka	é tlhókómélá mókgalá(:)bje
	The doctor	is looking after the old man.
F ⁰ Question	211 Hz	224 Hz
F ⁰ Answer	167 Hz	178 Hz

In the row named ‘F⁰ Question’ pitch peaks are given for the subject (*mokgalabje, ngaka*) and the following VP. As the subject is low-toned in both cases, the pitch on the subject is lower than the pitch on the following VP, which starts with a high-toned subject marker. As high tones spread to the right in Northern Sotho (see Lombard 1976, Zerbian 2006b) the whole VP is high-toned. In the row named ‘F⁰ Answer’ pitch measurements of the corresponding answers are given. Again, as the subject is low-toned in both cases the pitch on the subject is lower than on the following VP. However, comparing the pitch level of questions to the pitch level of answers, it becomes clear that the overall pitch of questions is well above the overall pitch of answers.

Jones et al. (2001 a,b) verify the informal observations made for the prosody of polar questions in Bantu languages of Southern Africa in a detailed phonetic study on Xhosa, investigating both the acoustic and perceptual aspect of prosodic marking in polar questions. Xhosa is a Southern African Bantu language also spoken in South Africa, belonging to the Nguni group (S40 in Guthrie’s (1967-71) classification). It is an areal neighbour to the Sotho languages but not a member of the same language sub-group. However, impressionistically the same observations have been made for Nguni as for Sotho (see references in Jones et al. 2001). The results of the phonetic studies show that acoustically, polar questions (called *queclaratives* in their study) differ from declaratives in duration specifically on the penultimate vowel, pitch on the penultimate vowel and overall raised pitch. Perceptually, it was found that duration and pitch on the penultimate syllable are highly significant prosodic features in disambiguating polar questions from declaratives. However, listeners make decisions about the kind of sentence type as early as the first syllable. That the pitch on the first syllable plays an important role in disambiguation reinforces the role of register raising as a cue to interrogativity.

The findings for Northern Sotho and Xhosa are not surprising cross-linguistically. Raised pitch is the most common feature in polarity questions

(Ulan 1978, Cruttenden 1997) and reflects uncertainty on the part of the speaker. Syntactic and morphological means play a more subordinate role.

Although many of the world's languages are reported to have a raise in pitch in polar questions, there is some variation as to where in the utterance the raise in pitch occurs. According to Ulan's sample a higher pitch towards the end of the contour seems to be the prominent pattern, with fewer languages that mark polar questions with a higher pitch towards the beginning of the utterance (Siemund 2001). In contrast, Northern Sotho has high pitch extending over the whole utterance. This is not the case in all Bantu languages. Data from Jita suggest that there is interesting language-specific variation to be found with respect to the domain in which the register raising applies. In Jita, a Bantu language spoken in Tanzania, the domain of register raising is determined by the prosodic structure of the utterance. All low-toned sequences are raised altogether, while in sequences containing one or more high tones the raising only affects the sequence up to the last high tone (Downing 1996:160).

Establishing the prosodic features of lack of vowel length and raised overall register as grammatical markers of polar questions not only shows by which formal means questions are differentiated from statements but also addresses another issue in tone languages. It is still largely uninvestigated if and how intonation is used in tone languages, the reason being that in tone languages change in pitch, the acoustic correlate for intonation, is already used for conveying lexical or grammatical meaning. Whereas some authors (e.g. Cruttenden 1997) claim that intonation only plays a limited role in tone languages to convey linguistic meaning, Yip (2002) collects examples from tone languages that show the use of pitch to convey pragmatic functions such as sentence type and focus. The observations from Northern Sotho can be interpreted exactly in this way. As intonation is used to differentiate between polar questions and statements, it fulfils a grammatical function.

2.1.2 Morphological marking of polar questions

Whereas the prosodic marking of polar questions is obligatory, grammar books on Northern Sotho state that polar questions can also be optionally marked by the particles *na/naa* or *a/afa* (Ziervogel et al. 1969, Poulos & Louwrens 1994). In this view, Northern Sotho supports cross-linguistic tendencies in using the two most common strategies for the marking of polar questions (Siemund 2001, König & Siemund, to appear), namely prosody and interrogative particles. According to Ziervogel et al. (1969), the two particles are not different in their meaning but only in their distribution. Although this view on the meaning of the question particles will be rejected later in this section, the distributional asymmetries are still valid. Whereas *na/naa* can occur not only in sentence-

initial position, sentence-finally, in both positions simultaneously as well as clause-second, *a/afa* only occurs in sentence-initial position.

Respective examples are given in (4). Example (4a) shows *na/naa* in sentence-initial position, (4b) in sentence-final position, (4c) in both positions simultaneously, and (4d) shows the question particle even in sentence-medial position. All examples are grammatical.

(4) Use of question particles in Northern Sotho

- a. Naa Shwahlane o a šoma?
 QP PROPNAME SC1 PRES work
 ‘Does Shwahlane work?’
- b. Shwahlane o a šoma na/naa?
- c. Naa Shwahlane o a šoma naa?
- d. Shwahlane na o a šoma?

(Poulos & Louwrens 1994: 374)

Example (5a) shows the occurrence of *a/afa* in sentence-initial position, example (5b) shows the particle in sentence-final-position. The asterisc marks example (5b) as ungrammatical.

(5)

- a. A(fa) puku ya gagwe ke ye botse (na)?
 QP CL9.book CL9 POSSPR2 COP CL9 beautiful (QP)
 ‘Is your book a beautiful one?’
- b. * O a n-tseba a(fa)?
 SC2SG PRES OC1SG-know QP
 ‘Do you know me?’

(Ziervogel et al. 1969: 119)

The use of either of the two variants of the interrogative particle (either *na* or *naa*; either *a* or *afa*) does not invoke any important differences with regard to the pragmatic function they fulfill. However, *naa* appears to be slightly more ‘inciting than’ *na*, and *afa* occurs more frequently than *a* (Louwrens 1987: 122).

The variety found in Northern Sotho in the positioning of the question particle *na/naa* is interesting. Although cross-linguistically question particles in polar questions appear in a variety of positions, their position seems to be relatively fixed within languages. They are mostly found sentence-finally or, though rarer, sentence-initially (Siemund 2001). The occurrence of *na/naa* following the subject, as in (4d), resembles languages like Russian or Khoisan languages in which the question particle cliticizes onto the first constituent of the clause (Siemund 2001, citing Comrie 1984 for Russian; Hagman 1977 for Khoekhoe).

Also the use of two different question particles in polar questions is a language-specific peculiarity in cross-linguistic comparison. Prinsloo (1985) and Louwrens (1987) investigate the use of the two question particles *na/naa* and *a/afa* in more detailed studies and reach the conclusion that there is a pragmatic difference between the two. Whereas *na/naa* is used to ask a standard polar question which could be answered in the affirmative or negative, *a/afa* is used for rhetorical questions where no answer is expected. Rhetorical questions therefore show a mismatch of form and function. The rhetorical use of *a/afa* is illustrated from two text excerpts in (6), which are reproduced from Louwrens (1987) and Poulos & Louwrens (1994) respectively:

(6) Rhetorical questions in Northern Sotho

- a. Kae, goreng, go dir-el-eng? Afa o šetše o
 where why SC17 do-APPL-what QP SC2SG remain.PST SC2SG
 lebetše gore go šetše ganyane gore o di
 forget.PST that SC17 remain.PST little that SC2.SG CL10
 tsen-e kgwedi-ng ya go feta?
 enter-PST CL9.moon CL9.POSS CL15 pass
 ‘Where, why, what happened? Have you already forgotten that
 you almost lost your life last month?’
- b. Mokalabi: Ke rek-ile ngwako kua Bulawayo mo-tse-ng
 SC1SG buy-PST CL9.house DEM PROPNAME CL3-village-LOC
 wa Morningside.
 POSS PROPNAME
 Mokalabi: ‘I bought a house in Bulawayo in Morningside.’
 Mmatlala: (Ka makalo) Eng? Afa ke go kwa gabotse?
 (with surprise) what QP SC1SG OC2SG hear well
 Mmatlala: (astonished) ‘What? Do I hear you right?’

Excerpt (6a) describes a mother’s reaction when being informed by her son who is a detective that he investigates a robbery. Only a month ago her son almost lost his life while tracking a gang of murderers. She expresses her disapproval as in (6a). Excerpt (6b) shows a mini-dialogue in order to illustrate another use of a polar question as a rhetorical question.

Having a closer look at the distribution of *na/naa* and its combinatoric properties with both constituent questions and *a/afa*, Prinsloo (1985) and Louwrens (1987) conclude that also *na/naa* has a pragmatic meaning: In contrast to *a/afa*, *na/naa* indicates that an answer is expected. To support their

conclusion they provide data that show that *na/naa* but not *a/afa* can co-occur with constituent questions, as shown in (7).⁵

(7) Constituent questions with question particles (Louwrens 1987)

- a. Na/ naa o sepetše le mang?
 QP SC2SG walk.PST with who
 ‘With whom did you go?’
- b. *Afa o sepetše le mang?
- c. *A o sepetše le mang?

The example in (7a) illustrates the co-occurrence of *na/naa* and a question word, which is in this case *mang* occurring in sentence-final position. Example (7b and c) show the ungrammaticality of *a/afa* and question word occurring in the same sentence.

Although *na/naa* may co-occur with another question word in the same interrogative structure while *a/afa* may not, *a/afa* and *na/naa* can also be found co-occurring in the same structure. Examples are given in (8).

(8) Combination of question particles (Louwrens 1987)

- a. A o tseba gore se-kgalabjane se-la se
 QP SC2SG know that CL7-old.man CL7-DEM SC7
 di tsen-e na?
 CL10 enter-PST QP
 ‘Do you realize that the poor old man is dead?’
- b. Afa o a n-kwa na?
 QP SC2SG PRES OC1.SG-hear QP
 ‘Do you hear me?’

In the example in (8a), the variant *a* in sentence-initial position co-occurs with *na* in sentence-final position. In (8b), *afa* co-occurs with *na*. As is suggested from the discussion above, context is crucial for the interpretation of these examples. Louwrens (1987) summarizes that questions as in (8) are asked in situations where the questioner already knows the answer but for some or other pragmatic reason wants the addressee nonetheless to answer the question. He gives two contexts for (8a) and (8b) respectively: Imagine a situation where a gangster killed an old man and is questioned in court by an attorney. The attorney is well aware of the fact that the old man did pass away but wants the accused to admit before court that he too is aware of the fact. By commencing

⁵ It is an open question if the particle *a/afa* can appear with rhetorically used constituent questions, like ‘Wer ist schon perfekt?’ (‘Who’s perfect?’) in German.

the question in (8a) with *a*, he indicates that he knows the answer to the question. By ending the question with *na*, however, he demands to be answered by the accused. For (8b), Louwrens (1987) proposes the context where an annoyed father is scolding his son for something the latter has done. In their argument the father might reach a point where he asks the question in (8b). Again, by starting with *afa* the speaker indicates that he knows the answer to the question, by finishing with *na* he nevertheless demands an answer from the hearer.

Contrary to the view maintained by grammar books and in accordance with Louwrens (1987) neither *a/afa* nor *na/naa* can be considered morphological markers of polar questions in Northern Sotho. The examples have shown that *na/naa* does not unambiguously mark polar questions, as it can also appear with constituent questions. *A/afa*, on the other hand, cannot occur with every polar question but only with those that are used as rhetorical questions. Therefore, polar questions are solely marked by prosody whereas the use of question particles can differentiate the pragmatic meaning of the question further.

2.2 *Constituent questions*

Constituent questions in Northern Sotho follow the cross-linguistic prominent pattern that intonation is much less important than it is in polar questions. In contrast to polar questions the penultimate syllable of a constituent question is lengthened (Lombard 1979) and no raise of the overall pitch register can be observed.

Instead of prosody, the use of interrogative words (also termed question words, *wh*-words) clearly indicates the sentence type of the utterance. (For the use of the question particle *na/naa* with constituent questions, see 2.1.2) Northern Sotho does not employ any other grammatical means to mark structures as constituent questions. Interestingly, there is a split with respect to the position of the interrogative word and thereby the formation of constituent questions in Northern Sotho. The split is obligatory across grammatical functions. Subjects are questioned differently from non-subjects in syntactic respect (subject/object-asymmetry). However, what both structures of constituent questions have in common is that question words are used to indicate the questioned element. The syntactic position of non-subject and subject question words will be illustrated in section 2.2.1 and 2.2.2, whereas section 2.2.3 lists more marked ways of question formation in Northern Sotho. Section 2.2.4 explores the subject/object-asymmetry further.

2.2.1 Non-subject constituent questions

With respect to the syntactic distribution of the *wh*-word, Northern Sotho shows a split according to the grammatical function the question word fulfills. This section deals with the syntactic position of the question word if it represents non-subjects.

Objects and adverbials are questioned in situ, i.e. the position of the interrogative word in a question corresponds to its syntactic position in basic word order in a declarative sentence. This is shown for objects in the examples in (9).

(9) Object questions in Northern Sotho

- a. Mo-kgalabje o nyaka ma:ng?
 CL1-old man SC1 look.for who
 ‘Who is the old man looking for?’
- b. Mo-kgalabje o jwala e:ng?
 CL1-old man SC1 plant what
 ‘What is the old man planting?’
- c. O bona mang ka me:hla?
 SC2SG see who always
 ‘Who do you always see?’
- d. Mo-gkalabje o nyaka ngaka e-fe:ng?
 CL1-old man SC1 look.for CL9.doctor CL9-which
 ‘Which doctor is the old man looking for?’

The examples in (9a, b) show that the *wh*-pronoun for objects is differentiated into *mang* for [+human] and *eng* for [-human]. Example (9c) further shows that the *wh*-word appears postverbally, in verb-adjacent position. As Northern Sotho is a head-initial language, a modifier will be questioned following its head, as shown in (9d).

However, in double object constructions where there is more than one object present, the *wh*-word does not necessarily appear verb-adjacent. It will appear in its in situ position which is determined by semantic factors such as thematic hierarchy or person hierarchy (see Hyman & Duranti 1982). This is shown in (10).

(10) Postverbal constituent order in Northern Sotho

- a. O fa mo-kgalabje se-hla:re.
 SC1 give CL1-old man CL7-medicine
 ‘He gives the old man medicine.’
- b. O fa mo-kgalabje e:ng?
 SC1 give CL1-old man what
 ‘What does he give the old man?’
- c. * O fa eng mo-kgalabje?
 SC1 give what CL1-old man
- d. O fa mang sehla:re?
 SC1 give who CL7-medicine
 ‘Who does he bring the medicine to?’
- e. * O fa sehla:re mang?
 SC1 give CL7-medicine who

The example in (10a) illustrates the basic word order in double object constructions in Northern Sotho. The beneficiary precedes the patient (or the animate object precedes the inanimate object). When asking for the patient, as in (10b), the question word for [-human] *eng* has to appear in the position corresponding to the patient in declarative word order. This is in sentence-final position. Putting the question word in verb-adjacent position leads to ungrammaticality, as shown in (10c). The same applies in a parallel fashion when questioning the beneficiary. The question word for [+human] *mang* has to occur in the position corresponding to the beneficiary in declarative sentence word order. As shown in (10d), this is the verb-adjacent position in this case. Placing the question word in sentence-final position, as in (10e), leads to ungrammaticality.

The same generalization made for objects applies to the questioning of local, temporal, and modal adverbials in Northern Sotho, as shown by the examples in (11).

(11) Constituent questions for adverbials

- a. Mo-kgalabje o nyaka ngaka ka:e?
 CL1-old man SC1 look.for CL9.doctor where
 ‘Where is the old man looking for the doctor?’
- b. Kgarebe e bolela ma-aka ka:e?
 CL9.lady SC9 tell CL6-lie where
 ‘Where is the lady telling lies?’
- c. Mo-kgalabje o jwala mo-hlare ne:ng?
 CL1-old man SC1 plant CL3-tree when
 ‘When is the old man planting a tree?’

- d. O bona ngaka ne:ng?
 SC2.SG see CL9.doctor when
 ‘When do you see the doctor?’
- e. Kgarebe e bolela ma-aka bja:ng?
 CL9.lady SC9 tell CL6-lie how
 ‘How is the lady telling lies?’
- f. Mo-kgalabje o rekiša ma-swi bja:ng?
 CL1-old man SC1 sell CL6-milk how
 ‘How is the old man selling milk?’

The examples in (11) show that question words referring to adverbials are placed in the syntactic position where the adverbial occurs in declarative sentences. Examples (11a,b) shows this for local, (11c,d) for temporal, and (11e,f) for modal adverbials. The default position in which the adverbial appears is postverbally following the object. The examples in (12) are meant to exemplify that question words appearing in other positions than in situ render sentences ungrammatical.

(12) Illicit positions for non-subjects

- a. * Neng o bona ngaka?
 when SC2.SG see CL9.doctor
 Intend. ‘When do you see the doctor?’
- b. * Mang o bona?
 who SC2.SG see
 Intend. ‘Whom do you see?’
- c. * O bona ka mehla mang?
 SC2.SG see always who
 Intend. ‘Who do you always see?’
- d. * O fa eng mo-kgalabje?
 SC1 give what CL1-old man
 Intend. ‘What does he give the old man?’

The examples in (12a,b) show that question words cannot appear sentence-initially. The example in (12c) shows that appearing postverbally is not sufficient for the position of the object *wh*-word. An object *wh*-word appearing after a temporal adverbial is ungrammatical. The *wh*-word needs to appear verb-adjacently as any direct object does in Northern Sotho. More generally, (12c) shows that sentence-final is not a position for a question word. The example from (10c), repeated here as (12d) for convenience, shows that there is no obligatory verb-adjacent position for question words.

In sum, constituent questions are not distinguished from declaratives neither in syntax, nor in verbal morphology nor by the use of question particles. Neither is intonation reported to play a role. The only way constituent questions are differentiated from their declarative counterparts is by use of an interrogative word. Northern Sotho thus fits into the cross-linguistic pattern that constituent questions are less marked grammatically than polar questions are (at least for non-subjects), as the interrogative word unambiguously marks the relevant sentence as question (König & Siemund, to appear).

2.2.2 Subject constituent questions

Although Northern Sotho questions objects and adverbials in their canonical position, subjects cannot be questioned in their canonical preverbal position in the pragmatically unmarked case. As already mentioned above, there is a split in the distribution of the *wh*-word in constituent questions in Northern Sotho according to the grammatical function the questioned constituents fulfill. Subjects of transitive verbs are questioned by means of a cleft construction, as shown in the examples in (13).

(13) Cleft construction for subject questions

- a. Ké mang (yo) a nyaka-ng nga:ka?
 COP who RPRN.CL1 SC1 look.for-REL CL9.doctor
 ‘Who is looking for the doctor?’
- b. Ké mang (yo) a bolela-ng ma-a:ka?
 COP who RPRN.CL1 SC1 tell-REL CL6-lie
 ‘Who is telling lies?’
- c. Ké mang (yo) a rekiša-ng ma-swi ko toropo:-ng?
 COP who RPRN.CL1 SC1 sell-REL CL6-milk in town-LOC
 ‘Who is selling milk in town?’

Evidence for the cleft construction does not only come from the use of high-toned *ké* in sentence-initial position but also from the change in the verbal morphology. High-toned *ké* is the copula and also appears in other copula-constructions in Northern Sotho. With respect to verb morphology the subject marker *a* is used instead of *o* and furthermore *-ng* (or *-go* as an idiolectal variant) is suffixed to the final vowel of the verb stem.

Questioning the subject in its canonical preverbal position leads to ungrammaticality, as shown by the example in (14).

(14) Subject in preverbal position

- * Mang o nyaka ngaka?
 who SC1 look.for CL9.doctor
 ‘Who is looking for the doctor?’

The cleft construction, as used in subject questions in (13), is similar to subject relative clauses, the only difference being that the relative pronoun can be left out in subject questions (see also Kock 1997). In order to illustrate the relation between cleft sentence and subject relative clauses, the data in (15) show subject relative clauses.

(15) Relative clause (Ziervogel et al. 1969)

- a. Mo-šemane yo a šoma-go o
 CL1-young.man RPRN.CL1 SC1 work-REL SC1
 robala gae.
 sleep home
 ‘The boy who works sleeps at home.’
- b. Nna yo ke lema-go...
 PRN1SG RPRN.CL1 SC1.SG plough-REL
 ‘I that plough...’

In both examples (15a,b) the head noun is followed by a relative marker that agrees with the head noun in class membership. The following verb also agrees with the head noun in class. In case of a third person (see 15a), the subject agreement marker is *a* instead of *o* for noun class 1 in declarative clauses. Verbal morphology furthermore shows the relative clause suffix *-go* (or *-ng* as a dialectal variant).

As has been shown in this and the preceding section, there is a distinction according to grammatical functions when it comes to the formation of constituent questions and the placement of the question word. Section 2.2.4 will show that the asymmetry is explicable if one understands that the grammatical functions are closely related to default discourse functions in Northern Sotho. Before turning to this issue, however, the next section will shortly mention alternative structures of constituent questions in Northern Sotho.

2.2.3 Variation in the formation of questions

In section 2.2.1 it has been shown that non-subjects, i.e. objects and adverbials, are questioned in situ in Northern Sotho. However, non-subjects can also be questioned by means of cleft constructions, as shown in the examples in (16).

(16) Cleft constructions for non-subjects in Northern Sotho

- a. Ké mang o mo-kgalabje a mo nyaka:-ng?
 COP who RPRN.CL1 CL1-old.man SC1 OC1 look.for-REL
 ‘It is who that the old man is looking for?’
- b. Ké eng se mo-kgalabje a se jwala:-ng?
 COP what RPRN.CL9 CL1-old.man SC1 OC9 plant-REL
 ‘It is what that the old man is planting?’
- c. Ké neng mo mo-kgalabje a nyaka-ng nga:ka?
 COP when RPRN CL1-man SC1 look.for-REL CL9.doctor
 ‘It is when that the old man is looking for the doctor?’
- d. Ké kae mo mo-kgalabje a nyaka-ng nga:ka?
 COP where RPRN CL1-man SC1 look.for-REL CL9.doctor
 ‘It is where that the old man is looking for the doctor?’

The example in (16a) shows a cleft construction questioning a [+human] object, (16b) a [-human] object, (16c) a temporal adverb, and (16d) a local adverb. The cleft construction in the examples in (16) is clearly identifiable according to the characteristics established in section 2.2.2, namely the use of the high-toned copula *ké* and the change in verbal morphology both in subject agreement marker and in the additional suffix (-*go/-ng*). In object relative clauses, as in (16a,b), the relative complementizer cannot be left out, as is the case in subject questions. The relative complementizer in relative clauses appears initially and agrees with the head noun in class membership.

The use of the cleft construction for questioning non-subjects is, however, not the unmarked case. The cleft-construction is more marked pragmatically and hence more rarely used in Northern Sotho. In situ questions and cleft constructions cannot be used interchangeably for questioning non-subjects in Northern Sotho (for the semantics of object clefts see Zerbian 2006a).

For subjects, in situ questions are possible only in multiple questions and with echo-reading, as shown in the examples in (17).

(17) Subject question in situ

- a. Mang o dira eng?
 who SC1 do what
 ‘Who is doing what?’
- b. A: Mo-kgalabje o tlhokomela ngaka.
 CL1-old.man SC1 look.after CL9.doctor
 ‘The old man is looking after the doctor.’
 B: Mang o tlhokomela ngaka?
 who SC1 look.after CL9.doctor
 ‘Who is looking after the doctor?’

The example in (17a) shows that in multiple questions the question word questioning the subject can appear in situ (Machobane 2001). The mini-dialogue in (17b) shows that subject in situ constructions also occur when an echo-reading is intended.⁶

The examples in (18) show two other alternatives reported in the literature to question the subject in Northern Sotho, if subject is defined as the highest thematic role.

(18) Alternatives for subject questions

- a. Go tla mang?
SC17 come who
'Who is coming?'
- b. Pitša e thub-ile ke mang?
CL9.pot SC9 break.PASS-PST by who
'The pot was broken by who?'

The example in (18a) shows an impersonal construction. The verb does not show agreement with the logical subject of the sentence (i.e. with the constituent bearing the highest thematic role), but with class 17. Class 17 is an unproductive locative class. The subject appears following the verb. Impersonal structures as in (18a) can be used for questioning the logical subject only if the verb is intransitive. In their pragmatic connotation they are parallel to cleft sentences.⁷ The example in (18b) shows a passive structure. In the passive construction the agent is demoted and added as an optional *ke*-phrase. Also the passive construction can be used to ask for the logical subject as a pragmatically unmarked strategy.

The structures presented in (18) show in fact also show questions in which the question word appears in situ (see Demuth & Gruber 1995 for the subject-in-VP hypothesis for Sesotho relevant for example (18a)). An appearance as preverbal subject is, however, crucially excluded for the question words in (18). This observation leads to the next section that investigates the subject/object-asymmetry in Northern Sotho further by relating it to the discourse function that the preverbal subject is assigned.

As this overview of interrogatives in Northern Sotho focuses on the basic question strategy, the alternative strategies which show differing pragmatic

⁶ It is an open question why subject question words can appear in the position they occupy in declarative sentences in multiple questions. See Haida (in prep.) for an approach to multiple questions.

⁷ In transitive verbs, the impersonal construction is ungrammatical and the cleft structure is obligatory.

meaning (cf. 16, 17) will not be dealt with further. It has become clear from the examples provided that they involve a more specialised pragmatic meaning. Therefore the following section will come back to the split in the syntax of question formation on the basis of grammatical functions.

2.2.4 *The subject/object-asymmetry*

Section 2.2.1 and 2.2.2 illustrated that there is a basic difference in the formation of constituent questions in Northern Sotho which is related to the grammatical function the questioned constituent fulfils. Non-subjects are questioned in their canonical position, whereas subjects must not be questioned in their canonical preverbal position. Instead they are questioned by means of a cleft construction in the case of transitive verbs and postverbally in the case of intransitives. This section will discuss that this split is not arbitrary but that it can be explained by relating to the discourse function that preverbal subjects fulfil in Northern Sotho.

Such an asymmetry between subjects and non-subjects is quite common in the languages of the world. Some encode the difference in their morphology, like Greek, Latin, Sanskrit or Japanese. Others show a difference between subjects and objects in certain syntactic contexts. The contexts in which the syntax of subjects differs from that of objects is language-dependent. In Northern Sotho, it displays itself a.o. in questions.

In their typology Li & Thompson (1976: 461ff) distinguish four different kinds of languages according to the relation between topic and subject: (i) subject-prominent languages, (ii) topic-prominent languages, (iii) languages that are both subject-prominent and topic-prominent, and (iv) languages that are neither subject-prominent nor topic-prominent. Considering Northern Sotho with respect to the criteria they establish for subject versus topic, it turns out that the constituent in sentence-initial position in sentences like (1), here repeated as (19) for convenience, shows subject properties.

(19) Declarative sentences in Northern Sotho

- a. Mo-nna o ngwal-ela ba-sadi.
 CL1-man SC1 write-APPL CL2-woman
 ‘The man writes to the women.’
- b. Ba-sadi ba ngwal-ela mo-nna.
 CL2-woman SC2 write-APPL CL1-man
 ‘The women write to the man.’

Among the subject properties is that the subject is selected by the verb of the sentence, a property that is common to subjects but not to topics. Furthermore there is obligatory verb agreement with the sentence-initial constituent (in

structures as in (19)) which also is a common coding for subjects, according to Li & Thompson (1976).

However, Northern Sotho is not a purely subject-prominent language. It shows a further characteristic that, on the continuum from topic-prominent to subject-prominent languages, shifts it into the direction of topic-prominent languages (Li & Thompson 1976: 461ff): It is incompatible with focus. The examples in (20) illustrate the subject/object-asymmetry that has been described for constituent questions above by means of further focus-related contexts.

Example (20a) shows that in answers to questions the questioned element is focused in situ with non-subjects (here the object), whereas subjects are focused by means of a cleft construction, as shown in (20b). Negation also shows a subject/object-asymmetry as negated elements are said to be inherently focused and therefore creating a topic/focus conflict. Subjects are negated by means of a cleft (20c), non-subjects (here the object) are negated by verb morphology only, thus in situ, as shown in (20d). Focus sensitive particles like *only* create a conflict of topic and focus when they associate with the subject. Whereas with non-subjects (here the object) the particle follows the focused constituent in linear order (20e), the use of a cleft sentence is necessary with subjects, as in (20f).

(20) Asymmetry in focus-related contexts

- a. Mo-kgalabje o nyaka ngaka [toropo:-ng]_F.
 CL1-old man SC1 look.for CL9.doctor CL9.town-in
 ‘The old man is looking for the doctor in TOWN.’
- b. Ké mo-kgalabje a tlhokomela-ng nga:ka.
 COP CL1-old man SC1 look.for.REL CL9.doctor
 Lit. ‘It is the old man who is looking after the doctor.’
- c. Ga se mo-šemane o a rekišitše-ng borotho maabane,
 NEG CL1-boy RELPRN SC1 sell.PST-REL CL14.bread yesterday,
 ke mo-kgala:bje.
 COP CL1-old.man.
 ‘It wasn’t the BOY who sold bread yesterday, it was the OLD MAN.’
- d. Mo-šemane ga se a rekiše bo-rotho maabane, o
 CL1-boy NEG SC1 sell.PST CL14-bread yesterday SC1
 rekišitše bu:pi.
 sell.PST porridge
 ‘The boy didn’t sell BREAD yesterday, he sold PORRIDGE.’
- e. Mo-sadi o tliša ba-na fela se-kolo:-ng.
 CL1-woman SC1 bring CL2-children only CL7-school-LOC
 ‘The woman only brings the CHILDREN to school.’ (not the teenagers)
- f. Ke mo-sadi fela a tliša-ng ba-na se-kolo:-ng.

COP CL1-woman only SC1 bring-REL CL2-children CL7-school-LOC
 ‘Only the WOMAN brings the children to school.’ (not the man)

Additional evidence comes from a survey of discourse contexts which shows that the subject does not appear in sentence-initial position when it is discourse-new information. This is shown in (21a) for the beginning of stories or in (21b) for stage directions. The use of impersonal constructions at the beginning of stories underlines the observation that the preverbal subject is discourse-old. As in the beginning of stories all information is new, no grammatical subject surfaces.

(21) All-new contexts

- a. Go be go na le di-nonyana le di-phukubje
 SC17 was SC17 COP CL10-bird and CL10-jackal
 tš-eo di be-go di dula le-šoke-ng.
 CL10-REL CL10 was-REL CL10 live CL5-wilderness-LOC
 ‘There were birds and jackals that lived in the wilderness.’
 (Matabane 1998)
- b. Go tsena Matalebele. O a tšwa.
 SC17 enter PROP.NAME 3 A leave
 ‘Matalebele enters.’ ‘He is leaving.’ (Makwala 1958)

That the ungrammaticality of subject questions placing the *wh*-word in sentence-initial position is related to the syntactic position, is furthermore shown by the grammaticality of question structures in which the logical subject is questioned, but in which it does not appear in preverbal position. The examples in (18) are repeated here as (22) for convenience. Example (22a) shows the logical subject in postverbal position in the impersonal construction, example (22b) shows a passive structure where the agent is expressed by a prepositional phrase.

(22) Alternatives for subject questions

- a. Go tla mang?
 SC17 come who
 ‘Who is coming?’
- b. Pitša e thub-ile ke mang?
 CL9.pot SC9 break.PASS-PST by who
 ‘The pot was broken by who?’

For the Sotho languages, the generalization that subjects are topics has been brought forward by Louwrens (1979, 1981, 1991) and Demuth (1990, 1995).

The term topic is defined as either discourse-old constituent or aboutness-topic in the respective literature.

Further illustration for the strong restriction of subjects to topics comes from data on acquisition. Demuth (1990) observed that syntactic structures like passives are acquired earlier by children learning Sesotho than by English speaking children. She relates this observation to the fact that these constructions are a frequent means in Sesotho to focus or question the subject whereas in English intonation provides the possibility to keep the syntax simple.

To sum up, the subject position in Northern Sotho is not a pure subject position as, although the sentence-initial constituent definitely has subject properties, it bears the discourse function of topic. Therefore, the subject position is restricted to subjects in Northern Sotho, but subjects can only surface if they are topics. The generalization that the subject position is actually a subject-topic position in Northern Sotho can account for the subject/object-asymmetry observed in interrogatives. As the subject position is closely tied to a topic interpretation, it cannot be the position for a questioned element. Question words always ask for (discourse) new information and therefore do not fulfil the requirements to appear in sentence-initial position.

3 Conclusion

In this paper, data from Northern Sotho were presented that illustrate the formation of questions (in main clauses) in this language. It became clear that Northern Sotho follows cross-linguistic tendencies in marking interrogatives, such as raised intonation in polar questions and the use of question words for constituent questions.

However, also language-specific variation came to light, such as the presence of two question particles in connection with polar questions, variation in the syntactic placement of *na/naa* in polar questions, and the differentiation in the placement of question words according to grammatical functions.

This overview of Northern Sotho is not meant to be representative for Bantu languages in general. Although the subject/object-asymmetry is common to some Bantu languages, such as Kikuyu (Schwarz 2003) and Nguni (Sabel & Zeller 2006), others allow the subject to be questioned in preverbal position (e.g. Chichewa, Bresnan & Mchombo 1987). Furthermore, variation can also be found with respect to the position of the question words for non-subjects. Sentence-initial or -final position have also been reported for question words (e.g. Dzamba in Bokamba 1976, Kikuyu in Bergvall 1987). The observable microvariation makes Bantu languages an interesting field for further research.

4 References

- Bergvall, V. (1987). The Position and Properties of In Situ and Right-Moved Questions in Kikuyu. In: Odden, D. (ed.) *Current Approaches to African Linguistics*. Dordrecht: Foris Publications: 37-54.
- Bokamba, F. (1976). Question formation in some Bantu languages. PhD dissertation, Indiana University.
- Bresnan, J. & S. Mchombo, (1987). Topic, pronoun, and agreement in Chichewa. *Language* 63 (4): 741-782.
- Comrie, B. (1984). Russian. In: Chisholm, W. et al. (eds.) *Interrogativity: A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam: Benjamins.
- Cruttenden, A. (1997). *Intonation*. Cambridge: Cambridge University Press.
- Demuth, K. (1990). Subject, topic and Sesotho passive. In: *Journal of Child Language* 17: 67-84.
- Demuth, K. (1995). Questions, Relatives, and Minimal Projection. In: *Language Acquisition* 4 (1&2): 49-71.
- Demuth, K. & J. Gruber (1995). Constraining XP sequences. In: Manfredi, V. & K. Reynolds (eds.) *Niger-Congo Syntax and Semantics*. Boston: Boston University African Studies Center.
- Demuth, K. & M. Johnson (1989). Interaction between Discourse Functions and Agreement in Setswana. In: *Journal of African Languages and Linguistics* 11: 21-35.
- Doke, C. M. (1954). *The Southern Bantu languages*. London: Oxford University Press.
- Downing, L. (1996). *The tonal phonology of Jita*. Munich: Lincom Europe.
- Guthrie, M. (1967-71). *Comparative Bantu*. Farnborough: Gregg International.
- Hagman, R. (1977). *Nama Hottentot Grammar*. Bloomington, IN: Indiana University Press.
- Haida, A. (in prep.) Focus in Wh-questions. PhD thesis, University of Potsdam.
- Hyman, L. & Duranti, A. (1982). On the object relation in Bantu. In: Hopper, P. & S. Thompson (eds.) *Syntax and Semantics 15*. New York: Academic Press: 217-239.
- Jones, J., Louw, J. & J. Roux (2001a). Queclaratives in Xhosa: An acoustic analysis. In: *South African Journal of African Languages*, Suppl. 36: 3-18.
- Jones, J., Louw, J. & J. Roux (2001b). Perceptual experiments on queclaratives in Xhosa. In: *South African Journal of African Languages*, Suppl. 36: 19-31.
- Kock, J. (1997). The redundancy principle: relative descriptions in Northern Sotho. In: *South African Journal of African Languages* 17: 143-146.
- König, E. & P. Siemund (to appear). Speech act distinctions in grammar. In: Shopen, T. (ed.) *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press.

- Li, C. & S. Thompson (1976). Subject and Topic: A New Typology of Language. In: Li, C. (ed). *Subject and Topic*. New York: Academic Press.
- Lombard, D. (1976). Aspekte van Toon in Noord-Sotho. PhD dissertation, University of South Africa.
- Lombard, D. (1979). Duur en lengte in Noord-Sotho. *Studies in Bantoetale*: 39-69.
- Louwrens, L.J. (1979). Naamwoordfunksies in Noord-Sotho. PhD dissertation, University of Pretoria.
- Louwrens, L.J. (1981). The relevance of the notions given and new discourse information in the study of North Sotho syntax. In: *South African Journal of African Languages* 1: 21-49.
- Louwrens, L.J. (1987). Some pragmatic functions of interrogative particles in Northern Sotho. In: *South African Journal of African Languages* 7 (4): 121-124.
- Louwrens, L.J. (1991). *Aspects of Northern Sotho Grammar*. Pretoria: Via Afrika Limited.
- Louwrens, L.J., Kosch, I.M. & A.E. Kotzé (1995). *Northern Sotho*. Munich: Lincom Europe.
- Machobane, M. (2001). Questions in Sesotho. In: *Afrikanistische Arbeitspapiere* 66: 67-83.
- Makwala, S. (1958) *Kgashane*. Pretoria: Van Schaik.
- Matabane, A., Leshilo, M. et al (1998). *Ntwa ya dinonyana le diphoofolo*. Cape Town: Juta & Co.
- Poulos, G. & L.J. Louwrens (1994). *A linguistic analysis of Northern Sotho*. Pretoria: Via Afrika.
- Prinsloo, D.J. (1985). Semantiese analise van die vraagpartikels *na* en *afa* in Noord-Sotho. In: *South African Journal of African Languages* 5 (3): 91-95.
- Sabel, J. & J. Zeller (2006). Wh-question formation in Nguni. In: Mugane, J., Hutchinson, J. & D. Worman (eds). *Selected Proceedings of the 35th Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 271-283.
- Schwarz, F. (2003). Focus Marking in Kikuyu. In: Eckardt, R. (ed.) *ZAS Papers in Linguistics* 30: 41-118.
- Siemund, P. (2001). Interrogative constructions. In: Haspelmath, M., König, E. et al. (eds.). *Language Typology and Language Universals – An International Handbook*, Vol. 2. Berlin: Walter de Gruyter, 1010-1028.
- Statistics South Africa (2004). *Census 2001: Primary Tables South Africa: Census '96 and 2001 compared*. Pretoria: Statistics South Africa. Available on-line at: www.statssa.gov.za.
- Ultan, R. (1978). Some general characteristics of interrogative systems. In: Greenberg, J. (ed.). *Universal of human language*. Stanford, CA: Stanford University Press, Vol. IV, 211-248.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Zerbian, S. (2006a). Expression of Information Structure in the Bantu Language Northern Sotho. PhD thesis. Humboldt-University Berlin.

Zerbian, S. (2006b). High Tone Spread in the Sotho Verb. In: Mugane, J., Hutchinson, J. & D. Worman (eds.). *Selected Proceedings of the 35th Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 147-157.

Ziervogel, D., Lombard, D.P. & P.C. Mokgokong, (1969). *A Handbook of the Northern Sotho language*. Pretoria: Schaik.

zerbian@zas.gwz-berlin.de