Stellenbosch Papers in Linguistics Plus, Vol. 62(2), 2022, 81-111
doi: 10.5842/62-2-894

The role of vowel length and pitch in Xhosa sentence type intonation

Eva-Marie Bloom Ström

Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, Sweden
E-mail: eva-marie.strom@gu.se

Abstract
This paper is a first study of intonation across different sentence types in Xhosa[1]. A recent increase in intonation research of African languages, including Bantu, has shown that intonation involves the manipulation of several distinct prosodic features such as vowel length and voice, next to pitch (Downing and Rialland (2017a)). Xhosa is, as many other African languages, a tonal language, meaning that the use of pitch for intonational purposes interacts with the use of lexical and grammatical tone. Moreover, other means are employed than pitch rise to distinguish different intonational phrases. For example, previous research shows that polar questions in Xhosa are indicated by reducing the lengthening of the penultimate vowel of a phrase, which is long in declaratives (Jones 2001). This paper expands on such previous studies and includes the intonation of different kinds of questions, in order to sketch a more complete picture of Xhosa intonation. Instead of being experimental, it makes use of examples recorded for non-intonational purposes, representing more natural language use. The study shows that the manipulation of penultimate lengthening plays an important role in distinguishing different phrases, in combination with declination. There are also indications of final lowering of pitch. Furthermore, devoicing the last vowel to a whisper indicates end of the utterance. Understanding the interaction of tone, intonation and phonological phrasing is important for understanding the grammatical structure and discourse pragmatics of Bantu languages.

Keywords: intonation, prosody, Bantu, Xhosa, vowel lengthening, phonological phrasing

1. Introduction

The study of the intonation of African languages, including Bantu languages, has brought important insights to the typology of intonation. The idea that question prosody involving high pitch is a near universal has for example been refuted based on intonation studies of African languages (Rialland 2007). Further analysis of more languages from different African language

[1] Xhosa (or isiXhosa) is a language of the Nguni subgroup of the Bantu language family. Varieties in this subgroup are mostly mutually intelligible and the subgroup is often treated as a dialect continuum in the linguistics literature. The examples in this paper are from different parts of the Eastern Cape and cannot always be considered standard Xhosa. Although Xhosa is referred to as a unit in this paper, the aim is not to describe intonation in the standard but rather to include variation. The regional source of the examples is evident from the information about the speaker in the footnote for each example.

families is needed in order to establish how widespread these results are and what other ways the languages employ to express different kinds of questions. This will help to improve linguistic theories of intonation in tonal languages. Intonation in Bantu languages involves the manipulation of a diversity of prosodic features such as pitch, vowel length and voice, and has received increased attention in recent years (Downing and Rialland 2017a). This paper would not have been written were it not for the inspiration of Laura Downing, during a seminar series on prosody at the University of Gothenburg, and also through the volume that she co-edited with Annie Rialland.

As far as the southern Bantu language Xhosa is concerned, polar questions involve a reduction of penultimate lengthening, as shown in Jones (2001) and Jones and Roux (2003). This important work, which was limited to polar questions consisting of one word, will be reviewed in section 4.2.1 and forms the starting point for this paper. Apart from this, information on intonation in the language is scarce. Therefore, this paper aims to expand on existing work and to sketch a more complete picture of Xhosa intonation, including declaratives and different types of questions. The analysis is based on full sentences, in contrast with the one-word utterances examined in previous work on Xhosa intonation mentioned above. The study is not experimental and the data used was only partly collected for prosodic purposes. The analysis is therefore preliminary. At the same time, this means that the analysis is based on a rich set of semi-natural data, see section 4. This paper contributes to our knowledge of intonation in Bantu languages and specifically of languages belonging to the Nguni subgroup of southern Africa. Hopefully it can serve as a basis on which further studies of intonation in Xhosa and other Nguni languages take place.[2]

Although the term intonation most commonly refers to systematic changes in pitch at the phrase or clause level, there are other features that contribute to the expression of higher-level linguistic functions, such as the distinction between questions and statements. These features include duration and voice quality (Zerbian 2010). In this paper, the analysis bases itself on a broad definition of intonation, including any non-pitch features that have functions similar to pitch changes in other languages (Zerbian 2010).

The use of pitch and other correlates of intonation in a systematic way is very important in human communication and carries a wealth of functions. It can for example distinguish between sentence types, indicate the demarcation of syntactic borders and signal information structure such as new and given in discourse (Zerbian 2010). In this paper, I did not consider intonation as used to express emotion and attitudes, referred to as paralinguistic features in Ladd (2008). The main interest here is on intonation as used to convey distinctions between sentence types such as questions and statements.

[2] I am grateful to Ntombifikile Futu, Thembani Onceya, Nkosekaya Hlitane, Andiswa Bukula, Sikhuthele Ncamani, Nqobile Jojo and Lumka Mdledle (3 female, 4 male) for contributing data for this analysis. The speakers are in their (late) 20s, apart from Lumka Mdledle who is in her 60s. I thank Laura Downing, Yasuko Nagano-Madsen, Nancy Kula and especially Seunghun Lee for guidance, help with Praat and useful comments on previous drafts, and two anonymous reviewers for extensive and helpful comments. The study took place as part of the project 'The role of the verb phrase and word order in the expression of definiteness in Bantu languages' funded by the Swedish Research Council (Dnr. 2017-01811). I thank the Swedish Research Council as well as the following organisations for travel funding: Åke Wiberg, Birgit and Gad Rausing, Kungliga Vetenskapsakademien, Gunvor and Josef Anér, Knut and Alice Wallenberg, Wilhelm and Martina Lundgren, and Magnus Bergvall.

The study shows that relevant categories in Xhosa intonation are pitch, vowel length and devoicing of final vowels. Xhosa statements are characterized by penultimate lengthening and declination. Pitch decline was not systematically measured in this study, but based on the visual inspection of the recordings included, declination is found to be accompanied by final lowering in declaratives. Polar interrogatives (yes-no) differ from statements in that the penultimate vowel of the utterance is not lengthened to the same degree, as shown by measurements. Moreover, there is an overall raised pitch in polar questions. In some cases, this combine with a pitch raise on the penultimate vowel, but not always. Content interrogatives (*wh*-questions) are characterized by an initial pitch register increase, which results in the declination being sharper than with declaratives. Penultimate lengthening is also slightly less than in declaratives. A rising intonation towards the end of the question is optional but found to be rare.

This paper is organised as follows: in section 2.1, a background to the tonal system of Xhosa is given, based on existing sources. The same physical property of fundamental frequency underlies the pitch differences used to convey tonal as well as intonational information (Gordon 2016), and it is therefore important to understand the tonal system of a language. In section 2.2, the methods used in this study are explained. Section 3 gives a brief background to intonation in (southern) Bantu. We then move on to the important aspects of penultimate lengthening and prosodic structure, as well as devoicing and downtrends in section 3. In section 4 the sentence types which are covered in this study are presented: declaratives, polar questions and content questions. Section 5 summarises the results, including a table with the sentence types and their intonational characteristics. The paper is concluded in section 6.

2. Background

2.1 The tonal system of Xhosa

This section gives a background to the tonal system in Xhosa, based on previous analyses in existing literature, specifically Claughton (1983), Downing (1990), Jokweni (1995), Cassimjee (1998), and Cassimjee and Kisseberth (1998). This background is necessary in order to understand changes in pitch that are due to intonational factors. Due to space limitations, it covers only the most important tonal processes. Readers are advised to consult the above sources for more complete accounts.

Xhosa and many other Bantu languages have a system of two level tones, high (H) or low (L), a system that has also been reconstructed for Proto-Bantu (Downing 2011). This system is characterized by the presence or absence of underlying high tones, i.e. a word is either underlyingly toneless, or it has one or more underlying high tones (Jokweni 1995). Syllables that are not assigned a high tone on the surface, are low (Jokweni 1995). In this kind of tone system, also characterized by phonological processes such as high tone shift/spread and deletion, the high tones are referred to as tonally active (Zerbian and Barnard 2008) and are the only ones that are marked in the examples by acute accents. Low tones are predictable from context and default (Downing 1990, Downing 2011). The vowel contributing a High tone (H) is underlined throughout the chapter.

Nouns in many Bantu languages including Xhosa can have a tonal contrast on every syllable. The augment contributes a H which shifts to the right in longer stems. Therefore, nominal tone can be elicited with the frame *akukho…* ‘there are no’, as the augment is dropped in this context

(Claughton 1983). An example is the noun *a̲mádoda* 'men' where the H is contributed by the augment; the stem is low: *madoda* (Claughton 1983: 5).

On a verb stem, however, there is generally only one underlying H possible, on the initial syllable of the root (Cassimjee and Kisseberth 1998, Downing 2011). On the surface, high tones in Xhosa verbs can be shown to originate from underlyingly high toned verb stems, and from certain prefixes that contribute a H. Due to a number of tone rules, they surface elsewhere mainly by rules of tone shifting, but in some regions also tone spreading (Cassimjee 1998). Moreover, there is a rule that deletes a final H in a word (Jokweni 1995). In short, Xhosa can be said to have a tone system characterized by mobility of the High tone (Cassimjee and Kisseberth 1998).

The general outcome of the tone rules is predictable, however. See below for alternations in the phonetic realization. Table 1 gives examples with low and high-toned verb stems. Firstly, the examples are presented with 1st and 2nd person subject markers, since these are tone-less. The present disjoint morpheme *ya-* does not contribute a H tone either. In the column 'Present disjoint', we see that the low toned verbs remain low such as in example (1a), and that the high toned verbs keep the high that the verb root contributes, such as in (2a) (Cassimjee and Kisseberth 1998).

The 3rd person singular subject prefixes, as well as all object prefixes, contribute a H tone. As we can see in the column 'Present disjoint with H on subject or object prefix' in table 1, this H surfaces on the penult or antepenult, in the case of a low toned verb (Downing 2003: 5). A high from the subject prefix surfaces on the antepenult *ya-*, if available, such as in (1b) *bayába:la*. If there is an object marker, the H contributed by this prefix surfaces on the penult such as in (1c) *niyawabá:la*. When verb stems are trisyllabic or longer, the H surfaces on the antepenult (Cassimjee and Kisseberth 1998), such as in (1f) *bayaxóle:la*.

In the case of a H toned verb, the result is somewhat different. The H surfaces on the antepenult when the stems are four syllables or longer (2h), and on the penult in bi- and trisyllabic stems such as (2f) *báyaboní:sa*. We also see that a H subject marker remains H. A high tone object marker and a H verb stem results in one H, such as in (2c) *niyabamé:ma*. In the analysis of Cassimjee and Kisseberth (1998), this is because the high toned object prefixes and the high toned verb stems are in one High tone Domain (HD), and the subject marker is in a separate HD. Only one H surfaces per HD, as standard Xhosa exhibits tone shift (Cassimjee and Kisseberth 1998):[3]

[3] The following abbreviations are used in this paper: numbers 1, 2, etc. refer to noun class numbers. SM1 thus refers to the subject marker of noun class 1, which equals the 3rd person singular (as do the other singular noun classes). AS= associative; CJ=conjoint; COM=comitative, COP=copulative; DJ= disjoint; FV= final vowel; NEG= negative; OM= object marker; PASS= passive; POT= potential; PRO=pronoun; PRT= participial; Q= question particle; RC= relative concord; REC= recent past; REL= relative marker; SM=subject marker. Vowels which are underlyingly high are underlined, and high tone is indicated by acute accent. Parentheses indicate that the vowel is not heard (devoiced). Square brackets are used to indicate phonological phrasing. Half-long and long vowels are followed by a dot or a colon, respectively, see section 2.2 for definition. Conjoint and disjoint refer to two morphologically distinct forms of the same tense-aspect-mood (TAM). The conjoint form is used when there is a close phonological bond between the verb and what follows, and the disjoint form is used elsewhere, including at the end of phrase (van der Wal and Hyman 2017). Every example is followed by a code, indicating the initials of the speaker, the date of the recording, and the source of the example, see section 2.2 Method.

Table 1. High and low toned verbs, with and without object prefix. Examples from Cassimjee and Kisseberth (1998)

	Present disjoint			**Present disjoint with H on subject or object prefix**	
Low tone verb stem					
(1) a)	*ndi-ya-ba:la* SM1SG-DJ-count	'I count'	1b)	*ba-yá-ba:la* SM2-DJ-count	'they count'
			1c)	*ni-ya-wa-bá:la* SM2PL-DJ-OM6-count	'you (pl.) count them'
			1d)	*bá-ya-wa-bá:la* SM2-DJ-OM6-count	'they count them'
1e)	*ndi-ya-xole:la* SM1SG-DJ-forgive	'I forgive'	1f)	*ba-ya-xóle:la* SM2-DJ-forgive	'they forgive'
1g)	*ni-ya-shukumi:sa* SM2PL-DJ-shake	'you (pl.) shake'	1h)	*ni-ya-wa-shukúmi:sa* SM2PL-DJ-OM6-shake	'you (pl.) shake them'
High tone verb stem					
(2) a)	*ni-ya-mé:ma* SM2PL-DJ-invite	'you (pl.) invite'	2b)	*bá-ya-mé:ma* SM2-DJ-invite	'they invite'
			2c)	*ni-ya-ba-mé:ma* SM2PL-DJ-OM2-invite	'you (pl.) invite them'
			2d)	*bá-ya-ba-mé:ma* SM2-DJ-OM2-invite	'they invite them'
2e)	*ndi-ya-boní:sa* SM1SG-DJ-show	'I show'	2f)	*bá-ya-boní:sa* SM2-DJ-show	'they show'
2g)	*ni-ya-bonísi:sa* SM2PL-DJ-show.clearly	'you (pl.) show clearly'	2h)	*bá-ya-bonísi:sa* SM2-DJ-show.clearly	'they show clearly'

Although the tone rules are regular, it should be kept in mind that variation can be found in the phonetic output. When a H spreads to the right, all vowels between the underlying H and the target H can become H (Downing 1990). Only the right-most H (the target) is obligatorily H, however, according to Claughton (1983). If any of the preceding vowels becomes L, all vowels to the left of this must also be low. This variation between spread and shift is reported to depend on the region concerned (Cassimjee and Kisseberth 1998).

There are also nominal and verbal stems that intrinsically end in an H, despite the rule mentioned above which deletes a final H in the terminology of Jokweni (1995), or violate a constraint 'Nonfinality' (Cassimjee and Kisseberth 1998: 53). Also, the role of depressor consonants needs to be considered when discussing Xhosa tonology, as certain classes of consonants interfere with the tonal patterns in that they, in onset position of a syllable, lower

the tone of the following vowel (Cassimjee and Kisseberth 1992). Depressor consonants in Xhosa are for example /z/ and /v/ (Jokweni 1995), for a complete list see Claughton (1983: 84).

The tonal rules presented here form the basis of the tonal system, although other details have been omitted. Xhosa also exhibits grammatical tone connected to certain tenses and particles (Cassimjee 1998).

2.2 Method

The analysis is based on a selection of ca. 60 phrases, from a database of 1222 phrases. The recordings were made in different areas in the Eastern Cape in South Africa and are based on translational and transformational elicitation[4], as well as on stimuli (story boards). The elicitation sessions were based on different questionnaires with the aim to investigate information structure and phonological phrasing, but also for example relative clauses. They were not primarily aimed at investigating intonation but contain declarative clauses as well as different interrogatives (260 of the phrases in the database are interrogatives) and were found suitable for a preliminary analysis of intonation in Xhosa. These have been followed up by a few interviews for prosodic purposes only.

The use of language examples from a wider database of examples rather than from an experimental prosodic study can be advantageous. When speakers are aware that a study targets prosodic differences, they might over-emphasise such contrasts. This kind of interference is avoided by using semi-naturalistic data.

On the other hand, certain acoustic correlates of intonation are hard to compare between utterances when the segmental content is not the same or similar, as the segments and length of an utterance can influence pitch. Also, speakers are of different ages and sexes. Therefore, a quantification of the declination of pitch across utterances, of the register expansion and of any final lowering, could not be undertaken. This will have to await future, more controlled, measurements.

Quantification of vowel length measurements was considered more viable. Vowel length has been measured in Praat, following segmentation criteria as presented in Turk et al. (2012). A vowel following a stop therefore starts at oral release and includes any voiceless aspiration in the case of voiceless stops, in order to make the segmentation comparable with that of fricatives and voiced stops (Turk et al. 2012). The consonant constriction release is measured from the onset of the first release burst, a [b] in the following illustrating figure. Any preceding nasal in a nasal+consonant combination (such as in *ntoni* 'what?') is not included. The onset of the next consonant constriction, a nasal in the example in the figure, is measured from the end of F2 energy of the vowel. We therefore get the following measurement of the vowel [a] in *bobani* 'who':

[4] Transformational elicitation is when, for example, the interviewer asks the speaker to give the negative counterpart of an affirmative utterance, or the past tense of a verb in the present, without translation from another language.

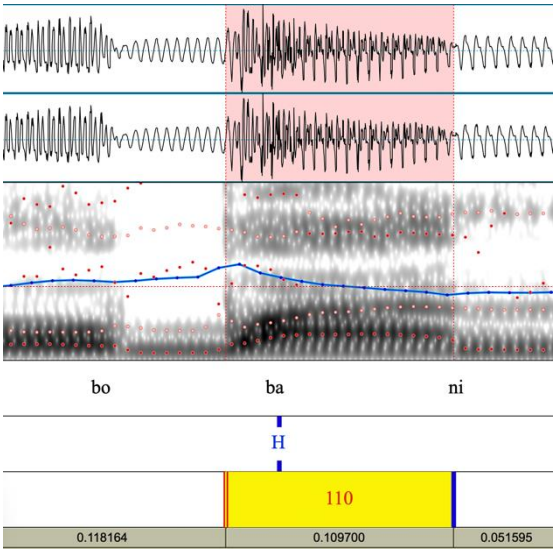

Figure 1. Vowel measurement criteria used in the study

The mean values of vowel length as well as standard deviation were calculated for all examples of all speakers. These measurements will be presented in the relevant section for declaratives, polar questions and content questions respectively, and summarized in 5.

Based on these measurements and compared to publications that mention vowel length in Xhosa and other Nguni languages (Jones et al. 1998b, Zeller et al. 2017), short vowels in Xhosa are under 100 milliseconds (ms). The penultimate vowels of declaratives are 200ms and above, which are considered long and are indicated with a colon (V:). Vowels that are longer than short vowels, in the 100-200 range, are referred to as half-long, and are marked with a single dot (V˙). This is in line with what Doke (1954) has pointed out for the whole Southern Bantu group (apart from Shona). These measurements need to be tested further and are merely indications. Similar observations were made by Bennett and Perkins (2016). They found preliminary evidence for a three-way distinction in vowel length, between non-penult, word-penult and utterance-penult vowels. Systematic further phonetic studies of vowel length in the (Southern) Bantu languages are however largely lacking.

When an example from a certain speaker is presented in the text for the first time, a few details of the informant are given in a footnote: sex, age and the place where this person lives (which happens to be the same as place of growing up in the case of all these speakers). Initials and date of recording are given following each example. This is followed by a letter E for elicitation and S for stimuli. The data collection took place as part of a larger project which included the collection of corpus data. Such examples have a letter M for monologue or D for dialogue, but have not been included in this paper.

The Xhosa used in the Eastern Cape today shows little geographical variation, especially among younger speakers (Bloom Ström 2018). It could be that such variation can still be found in tonal and intonational aspects of the language. This has been kept in mind during analysis but has not been found to play a role in the distinction between declaratives and interrogatives.

The recordings were made using a Zoom H4n Pro. Tone and intonation were analysed by acoustic impression combined with studying the F_0 contour in Praat. Tonal analysis was compared with expected underlying tones based on sources such as Claughton (1983), Jokweni (1995), Cassimjee and Kisseberth (1998) and the three volumes of the Greater Dictionary of isiXhosa (Pahl et al. 1989-2006), in which tone is marked. It should be noted that the sources sometimes give different tonal patterns for the same word, possibly due to regional or other variations.

3. Acoustic correlates of intonation in (southern) Bantu and Xhosa

This section introduces acoustic characteristics of utterances that play a role in characterizing different intonational phrases in Xhosa. We will return to these when examining the data in section 4. This section also forms the background to the knowledge of intonational phenomena in southern Bantu languages derived from existing literature.

In recent years there has been an increase in studies of intonation in African languages including Bantu languages. Mention of intonation in publications on Bantu languages have until recently been mainly impressionistic in nature, as pointed out by Jones et al. (1998b). For example, Zulu is reported to have higher pitch towards the end of the interrogative sentence in combination with the reduction in penultimate lengthening (Taljaard and Bosch 1988: 52). By comparing available data, Rialland (2007) showed in a comparative study that many African languages do not have high pitched polar question prosody, formerly considered a (quasi) universal (Bolinger 1978). Other intonational means such as vowel length or the suppression thereof, a final Low tone and breathy termination either complement high pitch question markers or occur as the only means of distinguishing a declarative statement from a question (Rialland 2007). As pointed out by Rialland (2007), the availability of reliable data for a comparative study of intonation in African languages was limited, especially from the Bantu group, considering its size. Subsequent to this, a number of studies of intonation in Bantu languages were carried out. This surge in interest has resulted in a recent volume on intonation in African languages including some Bantu languages (Downing and Rialland 2017a).

Studies of intonation in zone S[5] languages include Northern Sotho (Zerbian 2006b), Tswana (Zerbian 2017) and Shekghalahari (Hyman and Monaka 2011) as spoken in Namibia and Botswana. See also Zerbian and Barnard (2008) for an overview of studies in the phonetics of intonation in South African Bantu languages up to that point. Other Eastern-Southern Bantu languages include Chichewa, in which intonational phrases are indicated by boundary tones, lengthening and manipulation of downdrift (Downing et al. 2004). Boundary tones are single tones associated with the end (or sometimes beginning) of an intonational phrase (Ladd 2008: 88) and boundary low tones (L%) are relevant in several Bantu languages, meaning that the end of the phrase is marked by final pitch lowering (e.g. Kula and Hamann 2017, Patin 2017).

In the following two sub-sections, penultimate lengthening as well as devoicing are introduced. Changes in pitch in the different sentence types will be noted in the relevant section for each sentence type in (4).

[5] According to the referential system as proposed by Guthrie (1967/71).

3.1 Penultimate lengthening and phonological phrasing

Xhosa and other southern Bantu languages do not have lexically contrasting vowel length (Jokweni 1995, Zerbian and Barnard 2008). The – sometimes considerable – lengthening of the penultimate syllable in a phrase has however been noted by authors writing on Xhosa and other Southern Bantu languages (see Doke 1954 *et seq*). The lengthening of the penultimate vowel of an utterance was sometimes considered to be automatic (Cassimjee and Kisseberth 1998). Not all utterances have the same degree of penultimate lengthening however, as will be shown below for polar questions. Instead, lengthening is an intonational feature and different kinds of phrases are distinguished by penultimate lengthening and the suspension thereof. In this respect Xhosa follows the same pattern as several other Bantu languages, especially in the Eastern and Southern Bantu zones (Hyman and Monaka 2011, Hyman 2013, Zerbian 2017).

In the case of Xhosa, parts of the intonational patterns of the language have been relatively well studied; namely the intonation of Xhosa polar questions. These are referred to as queclaratives, when not segmentally marked in any way, in a doctoral dissertation and related publications (Jones et al. 1998a, 1998b, Jones 2001, Jones and Roux 2003). In these studies, including perceptual experiments, penultimate lengthening was shown to play an important role. We will return to this in section 4.2.1. In addition, other Bantu languages of zone S suspend penultimate lengthening in interrogatives (Gowlett 2003). Declaratives in Shekghalahari exhibit penultimate lengthening combined with a pitch fall when the last two syllables have identical tones. Shekghalahari restricts penultimate lengthening more than several other Bantu languages according to Hyman and Monaka (2011) and suspends the lengthening not only in polar and content questions but also in imperatives and hortatives, for example. According to Zerbian (2007), southern Bantu languages such as Northern Sotho, Zulu and Xhosa use a raised overall pitch and reduced penultimate lengthening to indicate polar questions as compared to declaratives. This is confirmed in the present study.

As indicated in the introduction, intonation can serve to distinguish sentence types, which is the main focus of this paper. Intonation also has a grammatical function in that it can indicate syntactic borders. Lengthening as an intonational feature has this function in many Bantu languages. In combination with different word orders and grammatical agreement, it correlates to notions of information structure such as what is given and what is new in discourse. Such notions of information structure (for example focus) are not directly expressed through intonation by means of pitch manipulation compared to many European languages. In English for example, contrastive focus is indicated by high pitch and stress on the first syllable in *Rónald made the hamburgers* (Chafe 1976: 33). However, pitch changes have been shown to be less important (if at all) in the expression of focus for several Bantu languages such as Chichewa (Kanerva 1990). Rather, phonological phrase boundaries play a role as evidenced by tone and vowel length. No prosodic cues at all were found for in situ focus in Northern Sotho (Zerbian 2006a, 2007). For Xhosa, the phonological phrase has been claimed to map with the syntactic organization of an utterance and to be relevant for the expression of information structure (Jokweni 1995). The application of tone rules in Xhosa such as tone shift has been shown to be bounded by such prosodic domains (Jokweni 1995) and as pointed out previously by Beach (1924) who refers to the domains as breath groups.

According to Jokweni (1995), then, penultimate lengthening is an indication of the end of a phonological phrase. For example, there is penultimate lengthening on a complement to a verb, when the verb and the complement phrase together (-3). The verb needs to take the conjoint form, in such cases (this form is unmarked in the present). On the other hand, there is penultimate lengthening on the verb when this verb does not phrase together with what follows (-4). In such cases, the disjoint form of the verb, indicated by the morpheme *ya-*, is needed. The verb shows agreement with the following object, which is considered dislocated (Jokweni 1995: 31)[6]:

(3) *(ba-vúl'* *ín-cwa:dí)*
SM2-open 9-book
'They open the book.'

(4) *(bá-ya-yi-vú:l')* *(ín-cwa:dí)*
SM2-DJ-OM9-open 9-book
'They open it, the book.'

However, there are no published studies giving phonetic details of lengthening as indication of the phonological phrase. In a conference presentation, reporting from a limited study[7], Bennett and Perkins (2016) show that the penultimate vowels of words are significantly longer than other vowels in those words. They also show that penultimate vowels of utterances are significantly longer than penultimate vowels of words. However, the penultimate vowels of presumed phonological phrases, such as in *uNolu uyalib**o**na] ixolo]* 'Nolu saw it, the treebark', with a dislocated object, could not be shown to be significantly longer than penultimate vowels of a word that is not followed by a phonological phrase break, such as in *uNolu ub**o**na ixolo]]* 'Nolu sees the treebark'. They therefore find three levels of length; short word internal vowels, semi-long penultimate vowels of words, and long penultimate vowels of utterances. Following this, there is so far no phonetic evidence to show that examples similar in structure to examples (-3) and (-4) are differentiated through vowel length. There is also no evidence for an accumulation of length that depends on how final a penult position is.

A larger study of lengthening in syllables which are penultimate in a supposed phonological phrase in Zulu, does present evidence of such phrasing, but also shows that this lengthening does not always take place. Zeller et al. (2017) set out to determine whether penultimate lengthening as an indication of the phonological phrase is relevant also in the future and remote past tenses that do not have a distinction between a conjoint and a disjoint form of the same TAM (see footnote 3). Their data set consisted of a total of 126 utterances, from three speakers. Sentences are of similar structure to those in Jokweni (1995) and Bennett and Perkins (2016), with subject, verb (object-marked or not) and object noun. Their conclusion is that penultimate lengthening does on average correlate with the phrase-finality of a verb form. The difference between assumed phrase-final penultimate vowels, and those that are not, is greater in the forms which are morphologically marked for conjoint and disjoint (present), than in forms which do not mark this distinction (remote past and future). The phrase-final penultimate syllable is not always lengthened, however. Penultimate lengthening is not an absolute cue to the end of a phonological phrase.

[6] Jokweni indicates the phonological phrasing with parentheses. An apostrophe indicates vowel elision.
[7] 2 speakers, 16 sentences per speaker.

That the correlation between penultimate lengthening and phonological phrasing is not straightforward has also been noted in other publications (Bloom Ström 2017), and is a topic for future research. In Zulu also, Cheng and Downing (2009) show that different word orders with left- and right- dislocated elements do not always correspond neatly with penultimate lengthening as an indicator of the phonological phrase. See also van der Wal (2017), section 4.3.

Since the examples used for the present study do not come from controlled data of similar structure, I will not make claims about the phrasing of the different utterances. The results from the studies mentioned above make it clear that we need to be cautious about interpreting length: a lengthened penultimate vowel is not immediate evidence that we are dealing with the end of a phonological phrase. The length of any vowel in the examples, which is not short, will be indicated in the annotation.

3.2 Devoicing

Devoicing has been reported as indicating the end of utterance, often in combination with a boundary low tone (L%), in several Bantu languages (Patin 2017, Rialland and Aborobongui 2017). In Tswana, devoicing of the vowel, or even the whole syllable, takes place when the utterance-final syllable is low toned, while it is less prominent in final high toned syllables and rarely spans the whole syllable (Zerbian 2017). As Downing and Rialland (2017b) pointed out, the use of devoicing to signal boundaries is a phenomenon which needs to be further researched.

Xhosa is a very clear example of a Bantu language with devoicing, and the phenomenon can be established in various examples used in this paper; declaratives, polar and content questions. The last vowel of the utterance is always of reduced amplitude, although this can be to slightly different degrees.

Devoicing is illustrated here with the last word *ku-la-makati* 'with the cats' cut from the sentence *kwenzeka ntoni kulamakati?* 'What are you doing with the cats?' (NJ150517E). In the last syllable, we identify the release of the alveolar consonant [t] (the sharp vertical line in the spectrogram where the last syllable starts), with following aspiration. As for the final vowel, amplitude is low, there is no vocal fold vibration and the vowel is reduced to a whisper:

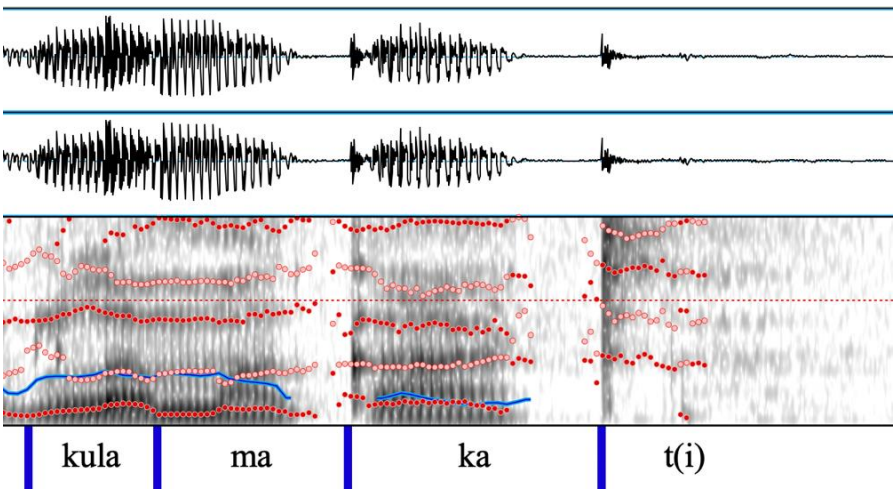

Figure 2. Devoicing

It will be noted in the data when devoicing does not take place to the same degree as exemplified here.

4. Intonational phrases in Xhosa - declaratives and questions

In this section, the different sentence types included in this study are examined, including examples and spectrograms, and their intonational analysis. The results are summarized in 5.

4.1 Declarative statements; penultimate length, devoicing and declination

The following is an example of a declarative, with a subject noun phrase; a subject-marked verb in the present conjoint form, and an object noun phrase:

(5)	*abántwa·na*	*ba-lé·qa*	*iká:t(i)*
	2.children	SM2-chase-FV	5.cat
	'(The) children are chasing a/the cat.'		[TO160516E[8]]

In all figures in this paper, I give the transcription and the surface tones in the annotation tiers, as well as the measurement of the penultimate vowel of the utterance. Also, the length of non-short vowels (above 100ms) is indicated.

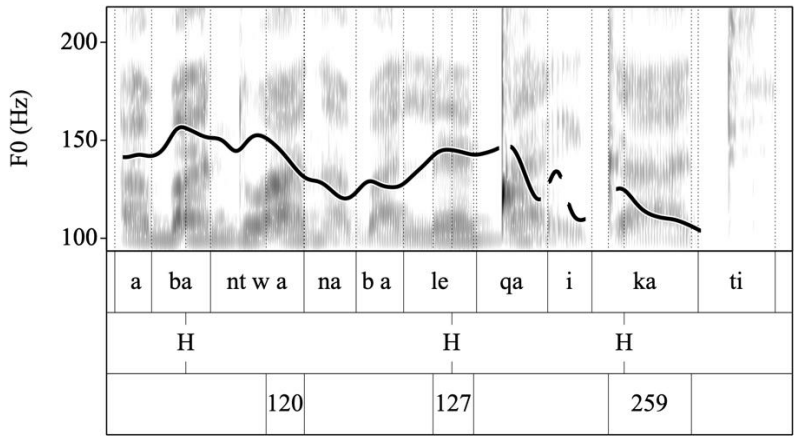

Figure 3. abantwana baleqa ikati

In example (-5) the underlying high of the augment shifts to the antepenult, yielding *abántwana*. This high tone appears (acoustically and auditorily) to carry over to the penultimate syllable of the subject, which would mean that this is a high tone spread rather than shift. The subject marker of noun class 2 *ba-* contributes a H, that shifts to the low toned verb *-leqa* 'chase'. A high tone contributed by a high toned prefix shift to penult or antepenult when the verb stem is L-toned in Xhosa, see section 2.1, and in the case of disyllabic stems, this is the penult. Underlying tones on *í-káti* is HHL.

The speech in this example is rather slow, and penultimate vowels in each word are slightly long (approximately 120ms in the subject), but the penultimate vowel of the object noun is

[8] Male, 26, Makhanda/Grahamstown.

much longer, indicating the end of the utterance. The end of the utterance is also indicated by the devoicing of the last vowel of the utterance: in *ikati* 'cat', the last vowel is barely a whisper, as indicated by the lack of vertical striations that correspond to vibrations of the vocal cords. See also section 3.2.

As mentioned in section 3.1, the penultimate vowel or syllable of an utterance has been reported as long in Xhosa. This is confirmed in the present study through measurements, with the following results:

Table 2. Length of penultimate vowel in declaratives

	Mean penultimate length, n= 26	Standard deviation
Declarative	217 ms	42 ms

The end of unmarked declarative statements is most often associated with a falling intonation in the languages of the world, which is explained by the natural decline in pitch over the duration of an utterance (Gordon 2016). Such falling intonation is seen also in Xhosa and exemplified in (-5). I propose that the language exhibits *declination* in declaratives; a gradual downsloping of the fundamental frequency (Downing and Rialland 2017b) which is assumed to be a phonetic effect (Connell and Ladd 1990, Connell 2001). Therefore, the first high tone in 5 has a higher frequency than the second high, which has a higher frequency than the last high. The last high, in *ikati* 'cat', is even further lowered, which is an indication that the decline is complemented by *final lowering*; an additional lowering of the pitch register at the end of an utterance (Downing and Rialland 2017b). According to Connell and Ladd (1990), this lowering is more abrupt than in the case of declination and is confined to phrase and utterance ends. This indication would need to be further corroborated in future research, with measurements of final pitch lowering.

Declination is also evident in an utterance that consists of a verb phrase with all low tones such as with *-bala* in example (-6). The subject prefix *ni-* 'you pl.´ is underlying low, as is the disjoint morpheme *ya-*:

(6) *ni-ya-ba:l-a*
SM2PL-DJ-count-FV
'You (pl) are counting' [TO160516E]

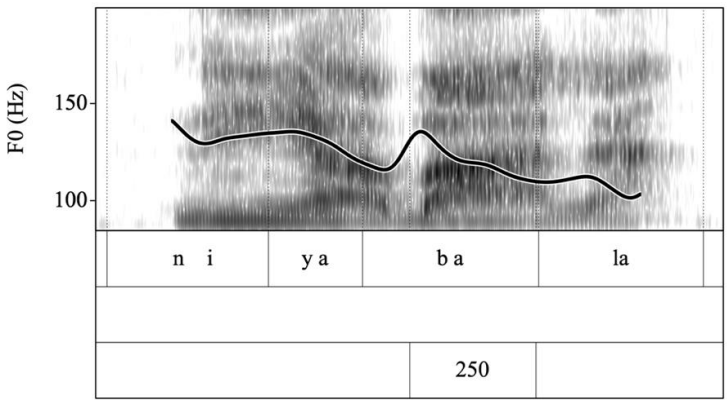

Figure 4. niyabala

In Figure 4 the fundamental frequency can be seen to gradually decrease. The final lowering appears to be less evident in an example containing only low tones. In this example, all syllables are carefully pronounced and therefore somewhat long. The last vowel has a lower amplitude than the preceding vowels but is not completely devoiced.

The final fall in pitch is also illustrated in the following example, with a sequence of high and low tones. The augment of the noun and the subject marker of noun class 4 contribute high tones. The examples from this speaker exhibit high tone spread rather than shift, see also example -14. She originates from and lives in the region on the coast, to the south of Port St. Johns.

(7) *ímí·thi í-ya-shúkú·m-(a)*
4.trees SM4-DJ-shake-FV
'The trees are shaking.' [NF151210E[9]]

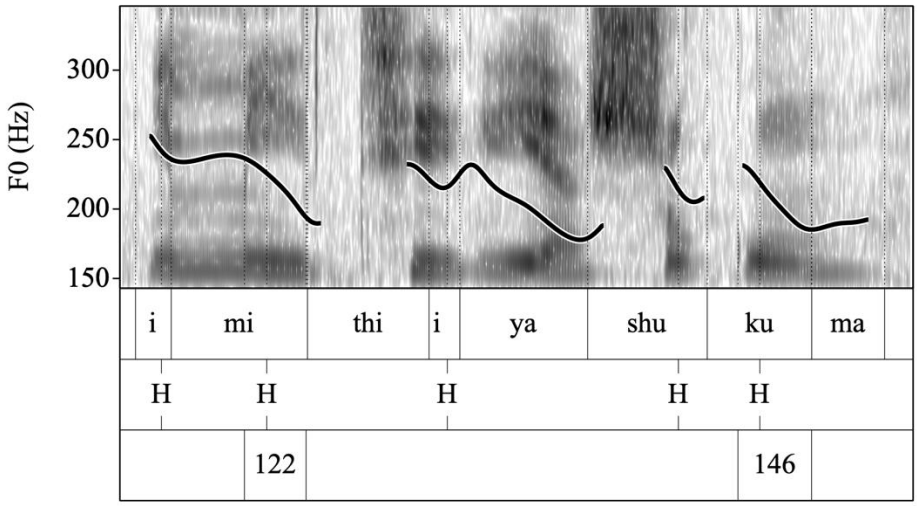

Figure 5. imithi iyashukuma

In a longer utterance, in which several intonation phrases follow each other, the last penultimate vowel is not necessarily longer than the preceding ones. Importantly, however, only the very last vowel is devoiced, showing that this is a clear indication of end of utterance:

(8) *inkwé·nkwé i-khwél-é emthi:ní] ya-wá phá:ntsi] ya-bónwa*
9.boy SM9-climb-REC.CJ 3.tree.LOC SM.PST9-fall down SM.PST9-see.PASS
ba-ba:nt(u)]]
2COP-2.people
'A boy has climbed the tree, he fell down, he was seen by people.' [SN150927S][10]

[9] Female, 29, Nqileni.
[10] Male, 30, Baleni.

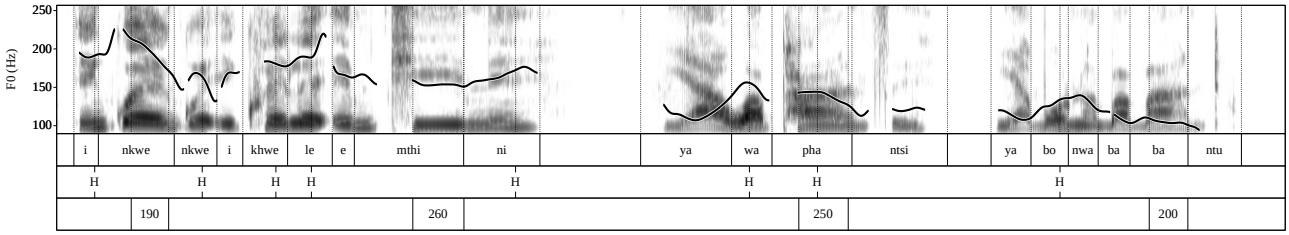

Figure 6. inkwenkwe ikhwele emthini, yawa phantsi, yabonwa babantu

The decline over the whole utterance is very clear, with the high vowels becoming increasingly lower. Final lowering appears to only occur at the end of the whole utterance. The last vowel of each intonation phrase remains somewhat high to indicate that more is following. There is lengthening on the penultimate vowel of the subject *inkwenkwe*, indicating that it could be considered to be in its own phonological phrase.

Based on these and other recordings of elicited declarative sentences, the Xhosa language indicates declaratives by means of a combination of decline, penultimate lengthening, final vowel devoicing and (tentatively) final lowering.

4.2 Interrogatives

4.2.1 Polar questions

Polar questions query the truth value of a corresponding declarative statement and are also referred to as yes/no questions. The intonation of polar questions in the languages of Africa has been shown to be diverse (Rialland 2007, Downing and Rialland 2017b). Rising intonation or high pitch in polar questions is not universal, even when no syntactic reorganization of the sentence is needed to indicate the question (Rialland 2007). Pitch does have importance in the expression of polar questions in many Bantu languages, however. For example, the polar question as a whole is pronounced with a higher overall pitch in Northern Sotho (Zerbian 2006b) and the super-high pitch on a penultimate or ultimate syllable is the only distinction between a declarative and a polar question in Shingazidja (e.g. Patin 2017). The cancellation of penultimate lengthening as a marker of questions is found mainly in languages spoken in the south of Africa and is often combined with pitch expansion (Rialland 2007). For Tswana, Zerbian (2017) reports that polar questions are characterized by the suppression of penultimate lengthening, in combination with an expanded pitch range. A short penultimate syllable in questions is reported for Southern Sotho (Doke and Mofokeng 1967) as cited in Jones (2001: 27) as well as for Northern Sotho (Ziervogel 1976).

Xhosa patterns with the other southern Bantu languages. Polar questions do not necessarily have a rising intonation. A rising intonation occurs in some of the examples in the corpus, see (-11), and this possibility could be the reason that polar questions in Xhosa have previously been reported to have rising intonation (Jokweni 1995: 83). They do have a higher overall pitch than declaratives (see below), but suppression of penultimate lengthening is the most distinct feature of polar questions or ‘queclaratives’, as they are referred to in Jones (2001). In a sequence of studies of Xhosa (Jones et al. 1998a, 1998b, Jones 2001, Jones and Roux 2003), one-word utterances were compared when used as statements vs. questions. These utterances, 858 in total from 11 mother tongue speakers, were copulative predicates such as *ngumntu* ‘it is

a person' (*ng* + *umntu*) (Jones et al. 1998b, Jones 2001: 15). The difference in penultimate length is the most important indication of such questions, although every syllable in the question was in fact shorter (Jones et al. 1998b). The duration of the whole word was on average 34,2% longer in statements than in questions (Jones et al. 1998b). In a perception experiment, questions were increasingly perceived as declaratives when the length of the penultimate syllable was manipulated and lengthened. Interestingly, though, declarative phrases were not increasingly perceived as questions when the penultimate syllable was manipulated and shortened (Jones et al. 1998a). Pitch was also shown to be important in that, apart from a raised overall register, the penultimate or final vowel has a raised pitch for a majority of the words in their sample (Jones et al. 1998b). The correspondence is not absolute, and it is pointed out that some statements have a rising contour and some polar questions a falling contour (Jones 2001).

In the present study, penultimate vowels of longer utterances (mostly 2-3 word utterances) than in the above-mentioned studies were measured. Moreover, the data was not experimental in set up, as mentioned in section (2.2). Nevertheless, the results of these previous studies were corroborated. The penultimate vowel in declaratives is on average 34% longer than that of a penultimate vowel in a polar question:

Table 3. Length of penultimate vowel in polar questions

	Mean penultimate length, n= 15	Standard deviation
Polar question	162 ms	33 ms

The following exemplifies a polar question. The recent past disjoint form *-gqíbile* 'finish' is underlyingly high and there is also a H contributed by the object marker *wu-*. However, the voiced click *gq* [ᵍ!]is a depressor consonant, the effects of which gives a falling contour on *-ile*:

(9) *u-wu-gqib-í·le* *um-sébé·nzi?*
SM2SG-OM3-finish-REC.DJ 3-work
'Have you finished the work?' [NH150428E[11]]

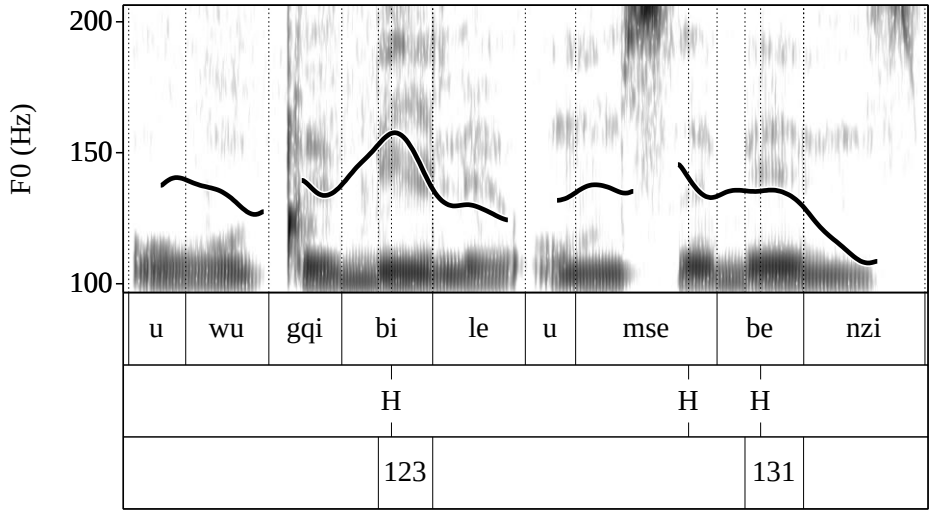

Figure 7. uwugqibile umsebenzi?

[11] Male, 21, East London.

There is declination in the polar question, and the devoicing we saw with declaratives is evident here also. This example also shows evidence of final lowering. However, final lowering is not found throughout for the polar questions and needs to be further investigated before any conclusions can be drawn.

The polar questions can therefore be said to be characterized by suppressed penultimate lengthening, declination and devoicing. If we compare the question in (-9) with the answer in (10), which is a full sentence answer rather than simply a 'yes', the pitch tracks are strikingly similar. However, there is a clearly raised overall pitch register in the question. Surprisingly, penultimate lengthening is shorter in the answer than in the question, in this case. The answer, as expected for declaratives, shows evidence of final lowering:

(10) *ndi-wu-gqib-í·le* *úm-sébé·nzi*
SM1SG-OM3-finish-REC.DJ 3-work
'I have finished the work.' [NH150428E]

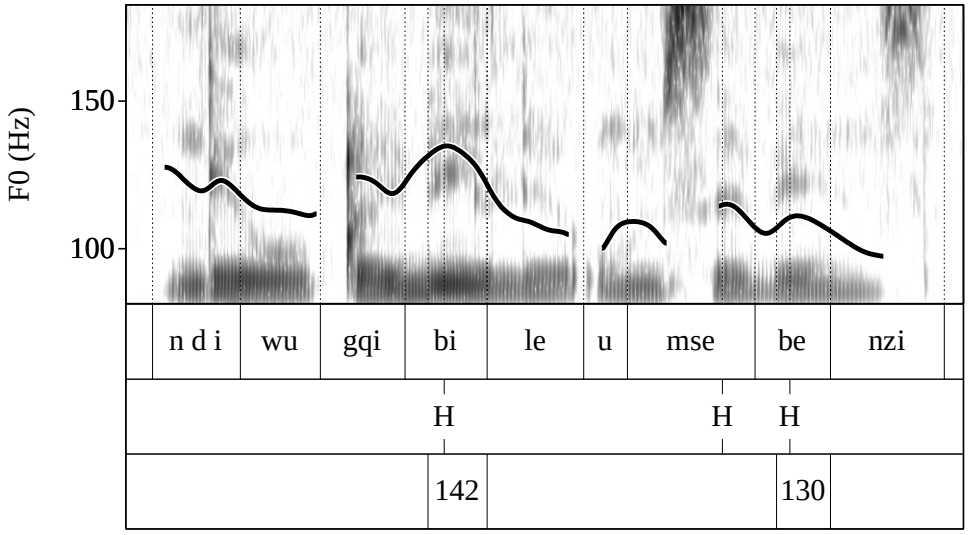

Figure 8. ndiwugqibile umsebenzi

Other polar questions do not exhibit a fall, but rather a slight rise:

(11) *u-zé·* *na-yo* *imá·li?*
SM2SG-come-REC.DJ COM-9PRO 9.money
'Did you come with the money?' [NH150512E]

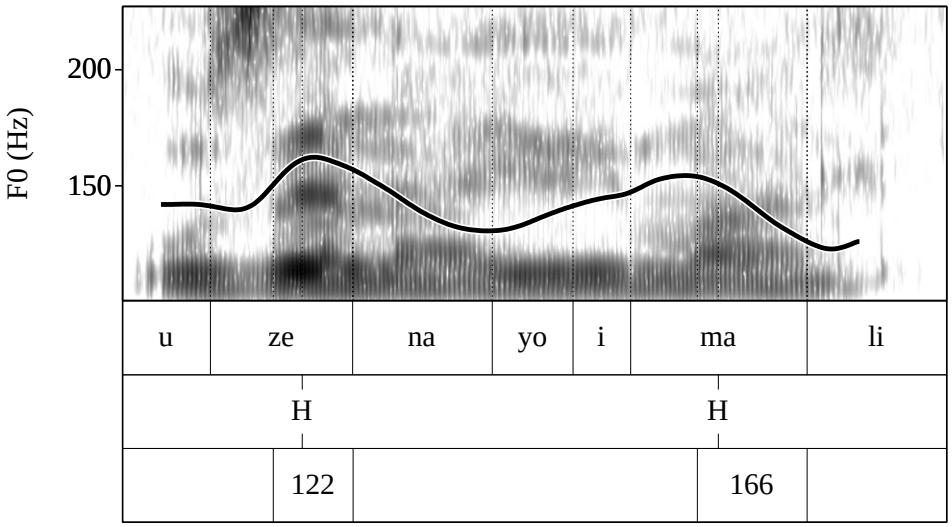

Figure 9. uze nayo imali?

This is expected based on previous findings, as the penultimate vowel of polar questions often has raised pitch, although this was statistically found not to be the most important correlate of polar question intonation (Jones et al. 1998b). The pattern is confirmed in other recordings from the database. Devoicing takes place also here.

The answer to the question in (-11) is included here (-12) to further illustrate the expanded pitch register in the question, well above 150 Hz, in comparison with the answer, which lies below 150 Hz:

(12) *ndi-zé·* *na·-yo*
SM2SG-come-REC.DJ COM-9PRO
'I have come with it.' [NH150512E]

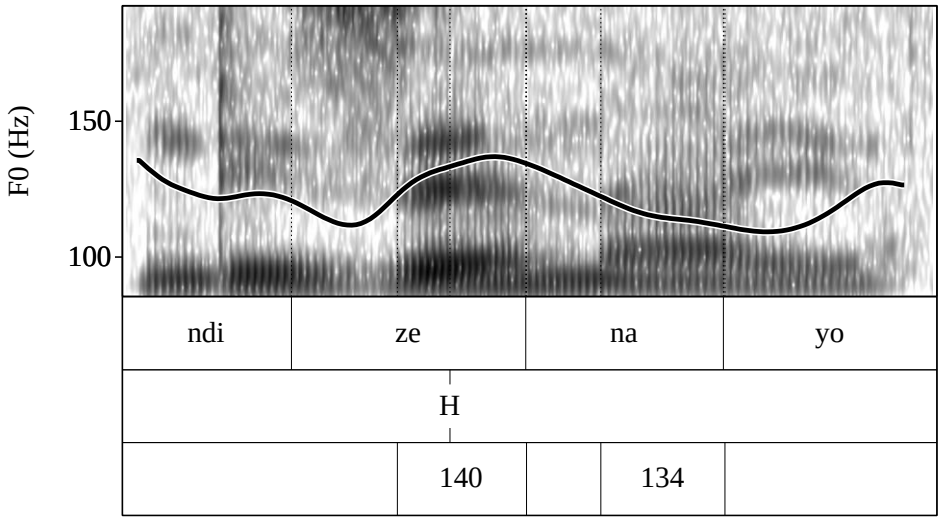

Figure 10. ndize nayo

As seen in this example, a long final vowel characterizes the recent past conjoint, in combination with a high, or falling, tone, see for example Cassimjee (1998: 201-207).

Interestingly, a high rise (or perhaps suspension of final lowering) occurs in all examples with the potential marker *nga-*, which indicates a question. The *nga-* prefix has been shown to have an underlying high tone (Cassimjee 1998: 144-150):

(13) *ndi-ng<u>á</u>-nge:n-a?*
SM1SG-POT-enter-FV
'May I enter?' [AB190227E[12]]

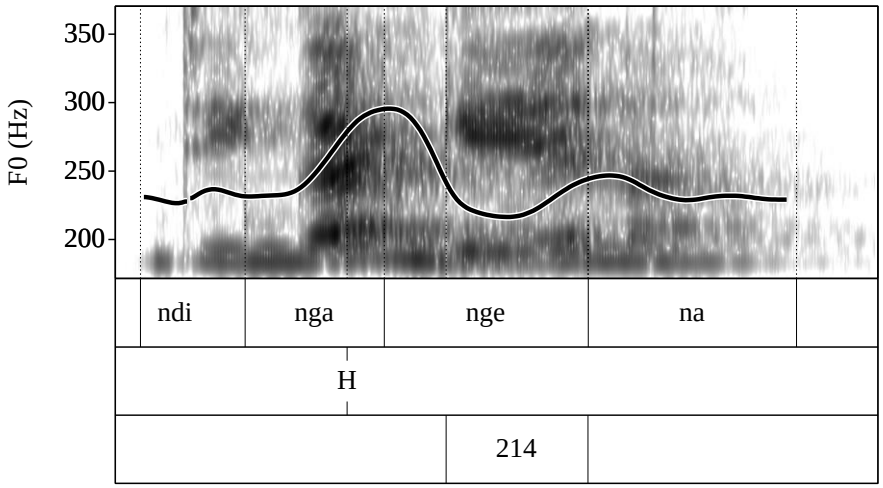

Figure 11. ndingangena?

There is less final devoicing in this example, although spectral energy is lower in the last vowel than in other vowels. Also in the following example with *nga-*, there is a clear question intonation with an overall raise in pitch, and a high pitch towards the end of the question:

(14) *ndi-nga-s<u>í</u>-f<u>ú</u>má˙n-a* *<u>í</u>sís<u>é</u>˙l(o)?*
SM1SG-POT-OM7-get-FV 7.drink
'Can I have the cold drink?' [NF161212E]

[12] Female, 26, Port Elizabeth.

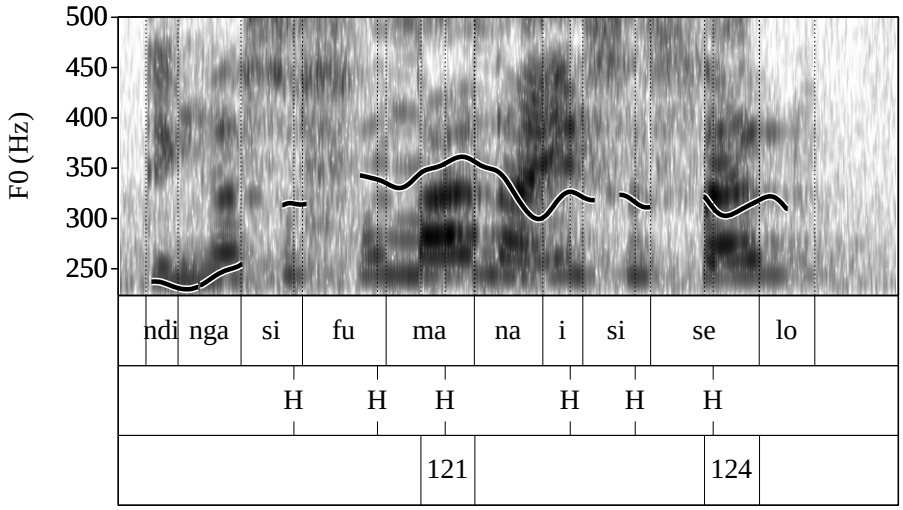

Figure 12. ndingasifumana isiselo?

In conclusion, polar questions in Xhosa show suspension of penultimate lengthening in combination with a raised overall pitch register. There is declination and perhaps even final lowering in the polar questions which are not segmentally marked as such (see example -9). Devoicing of the final vowel marks the end of the utterance.

4.2.2 Content questions

In a compilation of studies of intonation in African languages, content questions were often found to have similar intonation profile to assertions, which is cross-linguistically common (Downing and Rialland 2017b). In Northern Sotho, for example, penultimate lengthening is not supressed in content questions and the overall pitch register is not raised. Rather, the content question word, which occurs in situ, is the only distinguishing factor between a declarative and a content question when an object or an adverb is questioned. Subjects are questioned through cleft constructions (Zerbian 2006b).

Content questions differ from declaratives in Xhosa, however. Most importantly, there is an initial pitch register expansion which makes the question start high. There is no clear evidence for final lowering, but declination is more marked due to the high start. The penultimate lengthening is similar to that in declaratives, although there appears to be more variation in length. As with other sentence types, a content question ends in a devoiced last vowel.

Table 4. Length of penultimate vowel in content questions

	Mean penultimate length, n= 20	Standard deviation
Content question	200 ms	50 ms

Content questions can be formed in Xhosa with in-situ question words, such as in the following example. The verb needs to take the conjoint form, and the object argument is cross-referenced on the verb. The object marker contributes a H, but in the following example, it does not shift or spread, possibly due to the depressor consonant *gq*. The underlying high or falling tone of the recent past perfective conjoint form *-gqibê* of the verb ‘to finish’ (Pahl et al. 1989-2006) is

lowered although it is still slightly raised and marked as high here. Declination affects the pitch contour, with a gradual lowering of the pitch, and there is devoicing of the last vowel:

(15) *u-wú-gqib-é* *niːni* *umsébeːnz(i)?*
SM2SG-OM3-finish-REC.CJ when 3.work
'When did you finish the work?' [NH150428E]

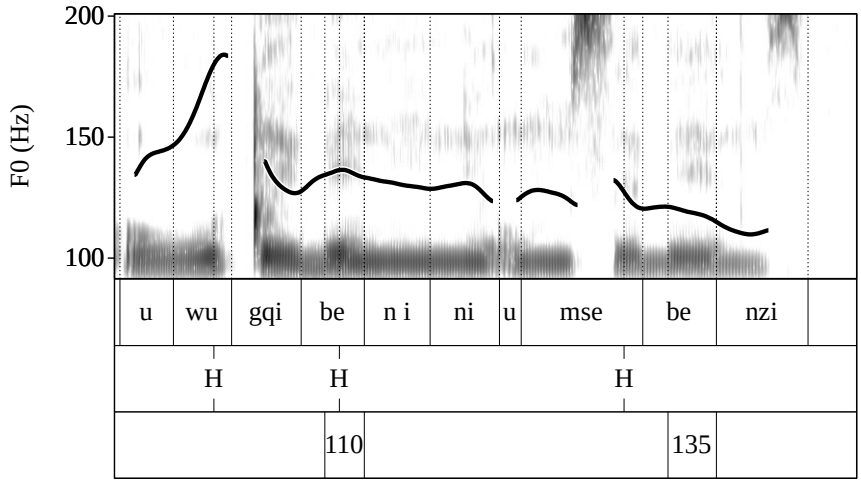

Figure 13. Uwugqibe nini umsebenzi?

A similar pattern is seen in the following example. There is penultimate lengthening of the question word *njani* 'how', indicating that it phrases with the preceding verb and that the following object noun *umbona* is in its own phonological phrase. As in many other Bantu languages, the object in Xhosa can be argued to be right-dislocated in examples such as (-15) and (-16), as evidenced also by object marking on the verb:

(16) *u-wú-phéːk-eː* *njaːni* *úmboːn(a)?*
SM1-OM3-cook-REC.CJ how 3.maize
'How did she cook the maize?' [NH150505E]

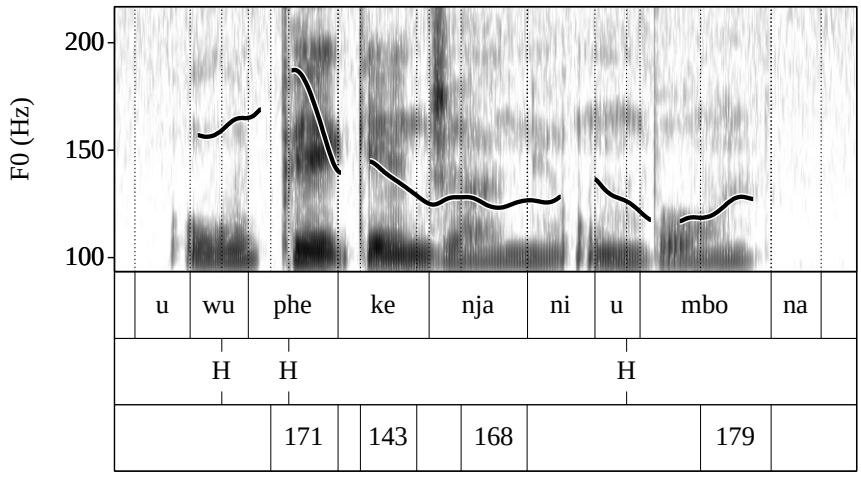

Figure 14. uwupheke njani umbona

The object marker of noun class 3 contributes a high tone. The initial pitch register expansion combines with declination, so that the high of *umbona* ‘maize’ is much lower than the high on the verb. There is no evidence of an overall expansion of the pitch register, however. The last vowel of the utterance is devoiced. The initial register expansion becomes clear when compared to the answer:

(17) *u-wu-phék-e·* *ngomlí·l(o)*
SM1-OM3-cook-REC.CJ on.the.fire
‘She cooked it on the fire.’ [NH150505E]

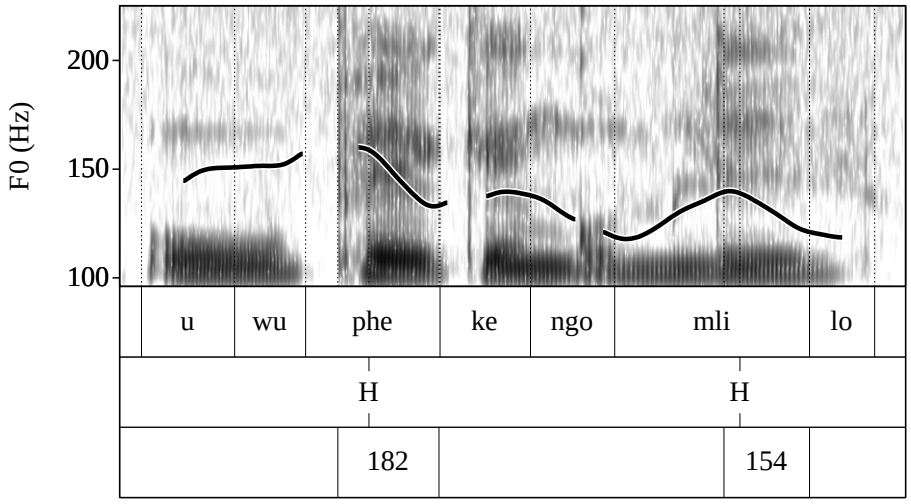

Figure 15. uwupheke ngomlilo

Subjects are often questioned in an inversion construction with expletive agreement (using the subject marker of the locative noun class 17), and with the question word immediately following the conjoint form of the verb (-18). Focussed subjects also appear in this position in Xhosa (-19). A literal translation of the question would approximate ‘there is running who?’:

(18) *kú-gijim-á·* *ba:n(i)?*
SM17-run-FV who
’Who is running?’ [SN150925E[13]]

[13] Male, 30, Baleni.

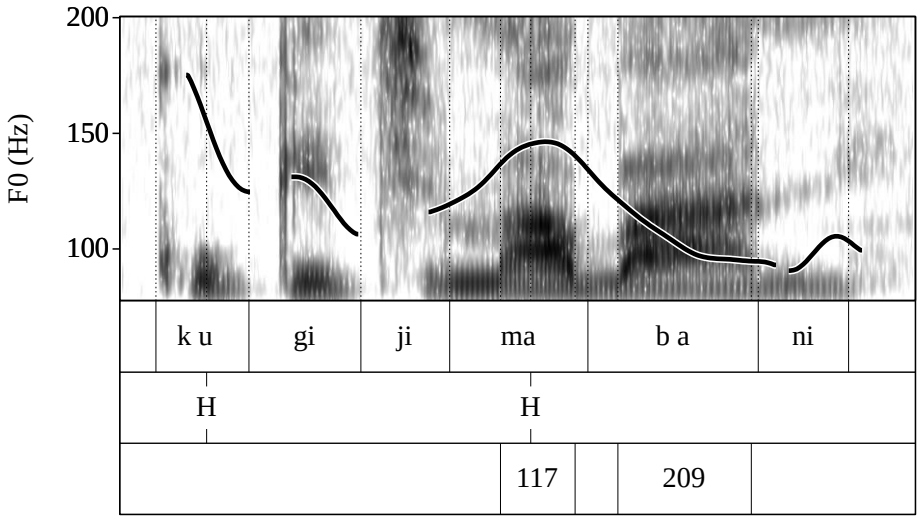

Figure 16. kugijima bani?

If we compare the pitch registers of the question (-18) and the answer (-19), we see that they are quite similar. However, the question starts higher which gives a more marked declination. We also find that the last vowel of the underlyingly low toned verb *-gijima* 'run' is high and somewhat lengthened in the question. In the answer, the last vowel of the verb is also high but not lengthened:

(19) *ku-gijim-á* *uXóla·n(i)*
SM17-run-FV 1a.Xolani
'Xolani is running.' [SN150925E]

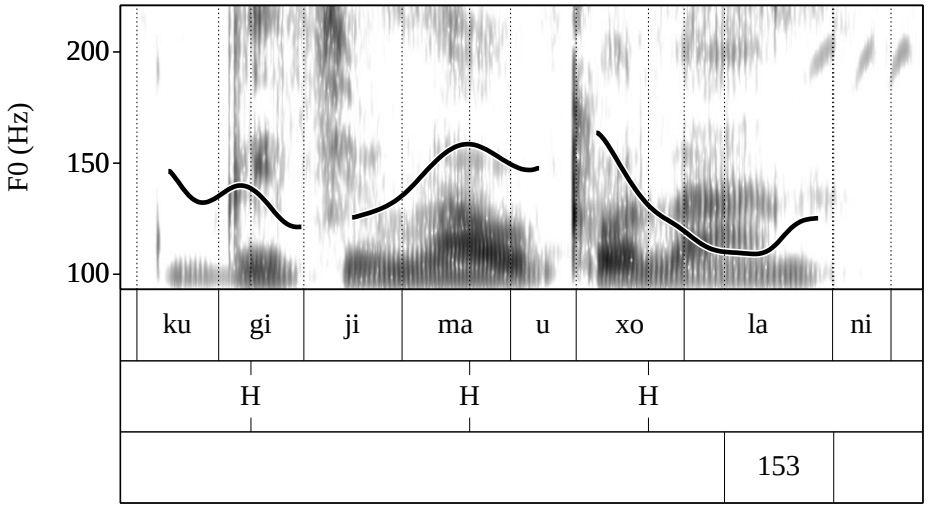

Figure 17. kugijima uXolani

This H on the last vowel of the verb is possibly due to grammatical tone of the inversion construction with expletive agreement, in combination with a certain degree of lengthening. This is noted in many examples in my data but is a question that I will not pursue further here (see also example -20).

In certain examples we also note a rising end of the intonational phrase, with a high on the last vowel. Such examples are rare in the data, however. The following exemplifies this:

(20) *kwa-khálá· ntó·n(í)?*
SM.PST17-call what
'What is calling/crying?' [LM161207][14]

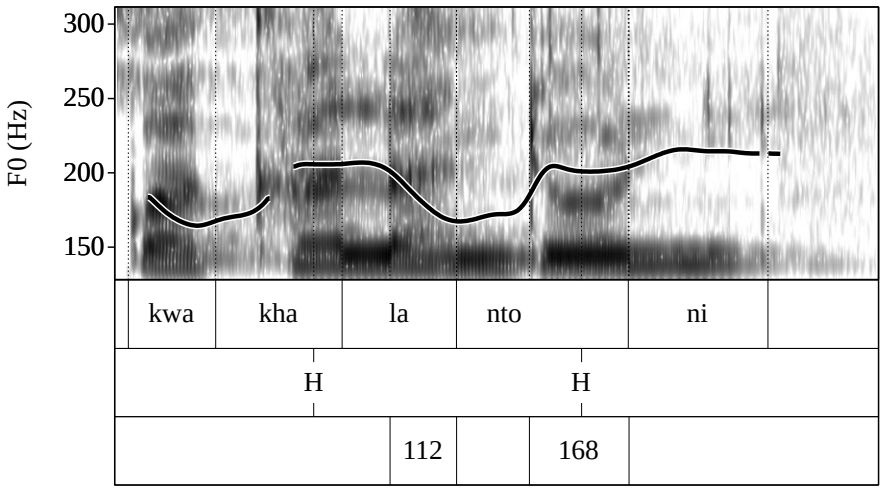

Figure 18. kwakhala ntoni

Content questions are also often formed with the clitic *phi* 'where' (-21) and *ni* 'what' (-22). That it is a clitic is evidenced by the last vowel of the verb it is clitisized to becoming lengthened, since it is the penultimate vowel in the phrase:

(21) *ba-fúnd-a: = ph(í)?*
SM2-learn-FV = where
'Where do they study?' [NF161212E]

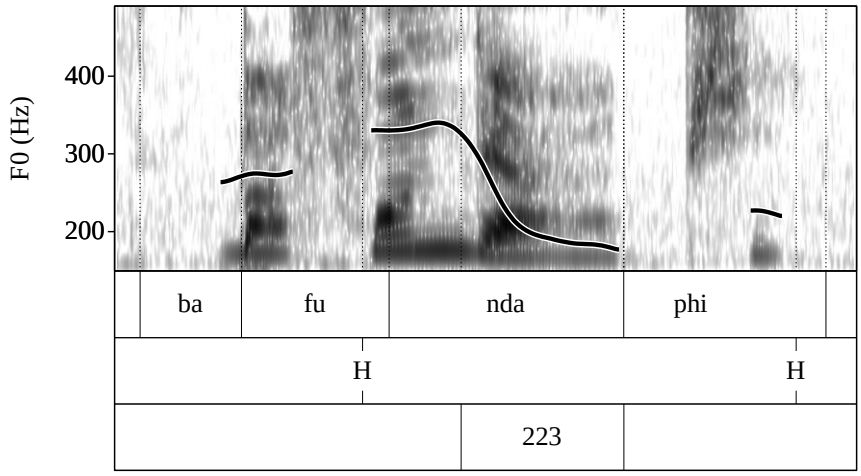

Figure 19. bafunda phi?

[14] Female, 63, Mnyameni

The initial pitch register expansion is present but less evident in the examples with a clitic, such as also when the clitic ni 'what' is attached to the verb:

(22) *b-énza:=ni* *abantwa·n(a)?*
SM2-do=Q 2.children
'What are the children doing?' [NJ150517E]

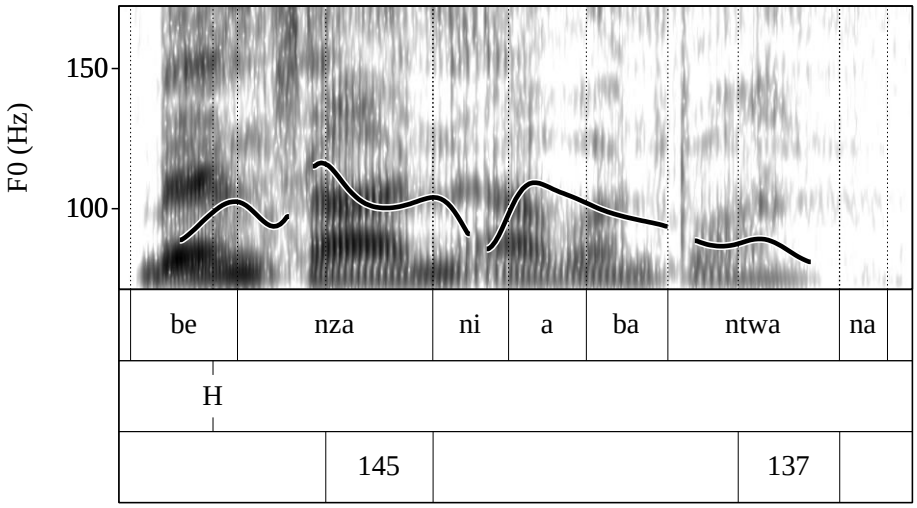

Figure 20. benzani abantwana?

Summarizing this section then, content questions have initial pitch register expansion and declination can be quite sharp due to this. There is penultimate lengthening, although not as marked as in many declaratives. There is also devoicing. In some instances there is a rise in pitch towards the end of the question.

5. The interaction between vowel length, downtrends and raised pitch

The following table summarises the penultimate length of different sentence types in Xhosa:

Table 5. Penultimate vowel length across sentence types

	Mean penultimate length	Standard deviation
Declarative	217 ms (n= 26)	42 ms
Polar question	162 ms (n= 15)	33 ms
Content question	200 ms (n= 20)	50 ms

The findings from the overall analysis are summarized in the following table. We can see that a rise in pitch towards the end is optional in all questions. It occurs more frequently with polar questions in my data than with content questions. Many polar and content questions instead show declination. In some cases there appears to be final lowering, but since the existence of final lowering cannot be concluded from this limited study, polar and content questions have a question mark in the table. Declaratives are tentatively concluded to have final lowering. Declination is very clearly present in declaratives.

Table 6. Sentence types with their intonational characteristics

	Long penult of intonation phrase	**Register expansion**	**Decline**	**Final lowering**	**High pitch at end of intonation phrase**	**Devoicing of last vowel**
Declarative	+	-	+	+	-	+
Interrogative						
Polar questions	-	+	+/-	?	Optional	+
Content questions	+/-	+ (initial)	+	?	Optional	+

The Xhosa intonation fits nicely into the bigger picture of intonation in southern Bantu languages. Suspension of penultimate lengthening is an important indication of non-declarative sentences. In Xhosa, penultimate lengthening is clearly suspended in polar questions, and perhaps also slightly in content questions. As in several other Bantu languages such as Shekgalagari, most of the intonation marking is not tonal (cf. Hyman and Monaka 2011).

An important indication of end-of-utterance is the devoicing of the last vowel, to a whisper or making it completely inaudible. This finding has implications for the phonotactics of Xhosa, which is, together with many other Bantu languages, considered not to have closed syllables. As the last vowel of many of the utterances in the present study can be considered to be completely inaudible, I propose that Xhosa indeed has closed syllables.

6. Conclusions and directions for further research

The intonation patterns of Xhosa show an interesting interplay between pitch and vowel length. Declaratives as well as most content questions are characterized by different downtrends and penultimate lengthening. In addition, polar questions are not necessarily accompanied by a raise in pitch. Rather, the penultimate lengthening is suppressed in polar questions and the pitch register is increased overall. Content questions are characterized by an initial increase in the pitch register, resulting in a sharper decline than in other phrases. An important finding of this paper is the complete devoicing of the last vowel of utterances, resulting in closed syllables.

As mentioned in the section on method, this analysis is based on auditory and manual acoustic analysis of a limited set of examples of each phrase type, with the aim of closing the knowledge gap on intonation in Xhosa and other Nguni languages. In the first place, this analysis would need to be complemented by analysing a larger number of examples for each phrase type and for different speakers. A quantification of the declination of pitch is desirable, as well as of the register expansion and final lowering. Furthermore, there are many more phrase types that need analysis, such as imperatives. I hope that this paper has contributed to a foundation for continuous work on intonation in Xhosa and Nguni.

References

Beach, D.M. 1924. The science of tonetics and its application to Bantu languages. *Bantu studies and general South African anthropology* 2: 75-106. https://doi.org/10.1080/02561751.1923.9676175

Bennett, W.G. and J. Perkins. 2016. Degrees of penultimate lengthening in isiXhosa. Paper presented at the SAMWOP-5, University of Free State.

Bloom Ström, EM. 2017. What the giant tells us about agreeing post-verbal subjects in Xhosa. *Stellenbosch papers in linguistics Plus* 52: 73-100. https://doi.org/10.5842/52-0-710

Bloom Ström, EM. 2018. Linguistic and sociolinguistic aspects of variation in the Eastern Cape: complexities of Xhosa language use. *Studia Orientalia Electronica* 6: 90-120. https://doi.org/10.23993/store.75138

Bolinger, D. 1978. Intonation across languages. In J.H. Greenberg (ed.) *Universals of human language. Volume 2 Phonology*. Stanford: Stanford University Press. pp. 471-524.

Cassimjee, F. 1998. *Isixhosa tonology: an optimal domains theory analysis*. München: Lincom Europa.

Cassimjee, F. and C.W. Kisseberth. 1992. The tonology of depressor consonants: evidence from Mijikenda and Nguni. *Annual Meeting of the Berkeley Linguistics Society* 18: 26. https://doi.org/10.3765/bls.v18i2.1543

Cassimjee, F. and C.W. Kisseberth. 1998. Optimal domains theory and Bantu tonology: a case study from Isixhosa and Shingazidja. In L.M. Hyman and C.W. Kisseberth (eds.) *Theoretical aspects of Bantu tone*. Stanford: Center for the Study of Language and Information (CSLI), Stanford University. pp. 33-132. https://doi.org/10.1017/s002222679925774x

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects and topics. In C.N. Li (ed.) *Subject and topic*. New York: Academic Press. pp. 27-55.

Cheng, L.L. and L.J. Downing. 2009. Where's the topic in Zulu? In H. de Hoop and G. van Bergen (eds.) *Special issue on Topics cross-linguistically, The Linguistic Review* 26: 207-238. https://doi.org/10.1515/tlir.2009.006

Claughton, J.S. 1983. The tones of Xhosa inflections (ed.) Department of African Languages. Grahamstown: Rhodes University.

Connell, B. 2001. Downdrift, Downstep, and Declination. *Typology of African Prosodic Systems Proceedings*.

Connell, B. and R.D. Ladd. 1990. Aspects of Pitch Realisation in Yoruba. *Phonology* 7: 1-29. https://doi.org/10.1017/s095267570000110x

Doke, C.M. 1954. *The Southern Bantu languages*. London: Oxford University Press for the International African Institute (IAI).

Doke, C.M. and S.M. Mofokeng. 1967. *Textbook of Southern Sotho*: Maskew Miller Longman.

Downing, L.J. 1990. Local and metrical tone shift in Nguni. *Studies in African Linguistics* 21: 261-317. https://doi.org/10.32473/sal.v21i3.107431

Downing, L.J. 2003. Stress, tone and focus in Chichewa and Xhosa. In R-J. Anyanwu (ed.) *Stress and Tone – The African Experience*, Frankfurter afrikanistische Blätter 15: 59-81.

Downing, L.J. 2011. Bantu tone. In M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds.) *The Blackwell companion to phonology*, Chapter 114. Malden, MA: Wiley-Blackwell. https://doi.org/10.1017/s0952675713000079

Downing, L., J., A. Mtenje and B. Pompino-Marschall. 2004. Prosody and Information Structure in Chichewa. In S. Fuchs and S. Hamann (eds.) *Papers in Phonetics and Phonology*. Berlin: ZAS. pp. 167-187. https://doi.org/10.21248/zaspil.37.2004.248

Downing, L.J. and A. Rialland. 2017a. *Intonation in African Tone Languages*. Berlin and Boston: De Gruyter Mouton.

Downing, L.J. and A. Rialland. 2017b. Introduction. In L.J. Downing and A. Rialland (eds.): *Intonation in African Tone Languages*. Berlin and Boston: De Gruyter Mouton. pp. 1-16. https://doi.org/10.1515/9783110503524-001

Gordon, M.K. 2016. *Phonological Typology*. Oxford: Oxford University Press.

Gowlett, D. 2003. Zone S. In D. Nurse and G. Philippson (eds.) *The Bantu languages*. London & New York: Routledge. pp. 609-638

Guthrie, M. 1967/71. *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages*. London: Gregg International.

Hyman, L.M. 2013. Penultimate lengthening in Bantu. Analysis and spread. In B. Bickel, L.A. Grenoble and D.A. Peterson (eds.) *Language Typology and Historical Contingency: In honor of Johanna Nichols*. John Benjamins Publishing Company. pp. 309-330. https://doi.org/10.1075/tsl.104.14hym

Hyman, L.M. and K.C. Monaka. 2011. Tonal and Non-tonal Intonation in Shekgalagari. In S. Frota, G. Elordieta and P. Prieto (eds.): *Prosodic categories: production, perception and comprehension*. Dordrecht: Springer Verlag. pp. 267-290. https://doi.org/10.1007/978-94-007-0137-3_12

Jokweni, M.W. 1995. Aspects of IsiXhosa phrasal phonology. Doctoral dissertation. Urbana-Champaign: University of Illinois.

Jones, J. 2001. Queclaratives in Xhosa: An acoustic and perceptual analysis. Doctoral dissertation. Stellenbosch: University of Stellenbosch.

Jones, J, J.A. Louw and J.C. Roux. 1998a. Perceptual experiments on Queclaratives in Xhosa. *South African journal of African languages 16, supplement* 36: 19-32. https://doi.org/10.1080/10118063.1998.9724400

Jones, J., J.A. Louw and J.C. Roux. 1998b. Queclaratives in Xhosa: an acoustic analysis. *South African journal of African languages supplement* 36: 3-18. https://doi.org/10.1080/02572117.2003.10587220

Jones, J. and J.C. Roux. 2003. Acoustic and perceptual qualities of queclaratives in Xhosa. *South African journal of African languages* 23: 223-236. https://doi.org/10.1080/02572117.2003.10587220

Kanerva, J.M. 1990. *Focus and phrasing in Chichewa phonology*. New York: Garland Publishing.

Kula, N.C. and S. Hamann. 2017. Intonation in Bemba. In L.J. Downing and A. Rialland (eds.) *Intonation in African Tone Languages*. Berlin and Boston: De Gruyter Mouton. pp. 321-363. https://doi.org/10.1515/9783110503524-010

Ladd, R.D. 2008. *Intonational phonology*. Cambridge: Cambridge University Press. 2nd edition.

Pahl, H.W., A.M. Pienaar, T.A. Ndungane, B.M. Mini, S.L. Tshabe, F.M. Shoba and P.N. van der Westhuzien. 1989-2006. *The greater dictionary of isiXhosa, vls. 1-3*.

Patin, C. 2017. Tone and Intonation in Shingazidja. In L.J. Downing and A. Rialland (eds.) *Intonation in African tone languages*. Berlin and Boston: De Gruyter Mouton. pp. 285-320. https://doi.org/10.1515/9783110503524-009

Rialland, A. 2007. Question prosody: an African perspective. In C. Gussenhoven and T. Riad (eds.) *Tones and Tunes: Studies in Word and Sentence Prosody*. Berlin and Boston: De Gruyter Mouton. pp. 35-62. https://doi.org/10.1515/9783110207569.35

Rialland, A. and M.E. Aborobongui. 2017. How intonations interact with tones in Embosi (Bantu C25), a two-tone language without downdrift. In L.J. Downing and A. Rialland (eds.) *Intonation in African Tone Languages*. Berlin and Boston: De Gruyter Mouton. pp. 195-222. https://doi.org/10.1515/9783110503524-007

Taljaard, P.C. and S. Bosch. 1988. *Handbook of isiZulu*. Pretoria: J.L. van Schaik.

Turk, A., S. Nakai and M. Sugahara. 2012. Methods in Empirical Prosody Research. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter and J. Schließer (eds.) *Acoustic Segment Durations in Prosodic Research: A Practical Guide*. Berlin and Boston: De Gruyter Mouton. pp. 1-28. https://doi.org/10.1515/9783110914641

van der Wal, J. 2017. What is the conjoint/disjoint alternation? Parameters of crosslinguistic variation. In Jenneke Van Der Wal and Larry M. Hyman (eds): *The conjoint/disjoint alternation in Bantu*. Berlin and Boston: De Gruyter Mouton. pp. 14-60. https://doi.org/10.1515/9783110490831-002

Van der Wal, J. and L.M. Hyman. 2017. *The conjoint/disjoint alternation in Bantu*. Berlin and Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110490831

Zeller, J., S. Zerbian and T. Cook. 2017. Prosodic evidence for syntactic phrasing in Zulu. In J. van der Wal and L.M. Hyman (eds.) *The conjoint/disjoint alternation in Bantu*. Berlin and Boston: De Gruyter Mouton. pp. 295-328. https://doi.org/10.1515/9783110490831-011

Zerbian, S. 2006a. Expression of information structure in Northern Sotho. Doctoral dissertation: Humboldt University.

Zerbian, S. 2006b. Questions in Northern Sotho. *Linguistische Berichte* 208: 385-405.

Zerbian, S. 2007. Investigating prosodic focus marking in Northern Sotho. In K. Hartmann, E. Aboh and M. Zimmermann (eds.) *Focus Strategies in African Languages: The Interaction of Focus and Grammar in Niger-Congo and Afro-Asiatic*. Berlin and Boston: De Gruyter Mouton. pp. 55-79. https://doi.org/10.1515/9783110199093

Zerbian, S. 2017. Sentence intonation in Tswana (Sotho-Tswana group). In LJ. Downing and A. Rialland (eds.) *Intonation in African Tone Languages*. Berlin and Boston: De Gruyter Mouton. pp. 393-433. https://doi.org/10.1515/9783110503524-012

Zerbian, S. 2010. Developments in the Study of Intonational Typology. *Language and Linguistics Compass* 4: 874-889. https://doi.org/10.1111/j.1749-818x.2010.00233.x

Zerbian, S. and E. Barnard. 2008. Phonetics of intonation in South African Bantu languages. *Southern African Linguistics and Applied Language Studies* 26: 235-254. https://doi.org/10.2989/salals.2008.26.2.5.569

Ziervogel, D. 1976. *Handboek van Noord-Sotho*. Pretoria: Van Schaik.

Appendix:

List of sentences used for measurements of penultimate length of the utterance. All sentences have been recorded twice (except one):

1. Declaratives
abantwana baleqa ikati
niyabala
imithi iyashukuma
kude kukhale inkhukhu
kwakhala inkhukhu
ndiyabala

niyabala
uyabala
baleqa amakati
kulima uThandi
kulime uThandi
balima amathanga
uSipho uthanda umphokoqo

2. Polar questions
uwuqibile umsebenzi?
ndingangena?
ndingasifumana isiselo?
ingaba ndingafumana isiselo?
ndingaboleka isitixo?
uze nayo imali?
bakhona abantu?
bathathu abantu enqanaweni?

3. Content questions
kugijima bani?
kwakhala ntoni?
kwenzeka ntoni kulamakati?
uthanda bani?
kulima bani?
uwupheke njani umbona?
uwugqibe nini umsebenzi?
uyele ntoni edolophini?
ulibele ntoni?
kulime bani?