UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

Modal expressions in Xhosa, part II : Possibility

Crane, Thera Marie; Savi, Stefan; Slater, Onelisa; Bernander, Rasmus

2025-12-19
Musée royal de l'Afrique Centrale
http://hdl.handle.net/10138/625357
Crane, T M, Savi, S, Slater, O & Bernander, R 2025, 'Modal expressions in Xhosa, part II : Possibility', Africana Linguistica, vol. 31, pp. 31–67. https://doi.org/10.2143/AL.31.0.3295058

Downloaded from Helda, University of Helsinki institutional repository. https://helda.helsinki.fi
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.

Modal expressions in Xhosa, part II: Possibility

Thera Marie Crane, Stefan Savić, Onelisa Slater & Rasmus Bernander

This study reports on modal expressions of possibility in Xhosa, a Bantu language of South Africa. Combining analysis of occurrences in a small corpus and the native-speaker intuitions of one of the authors, we describe patterns and restrictions in the distribution of four markers of possibility: auxiliary(-like) **kwazi** and **nakho**, and prefixes **noku-** and **nga-**. Results confirm the generalisation that further grammaticalisation (e.g. auxiliary > affix) corresponds to more subjective meanings. Our corpus data further suggest the significance of the category of existential possibility in Xhosa and its position in pathways of semantic change. Source lexical and constructional meanings are also significant in the usage patterns of Xhosa modal expressions. Finally, our results suggest that semantic expansion from participant-internal possibility (i.e. ability) functions to participant-external (circumstantial or goal-based) may occur via purpose and result clauses.

Keywords: corpus linguistics, modality, Nguni Bantu, possibility, grammaticalisation and subjectification

Introduction

This is the second article of a two-part study on modal expressions in Xhosa, (Guthrie classification code S41; endonym **isiXhosa**), a Bantu language spoken in South Africa. In the first article (Crane *et al*. 2024), we gave background to our study and discussed our findings for necessity expressions. In this article, we describe Xhosa possibility expressions in detail.

We examine the six most frequent modal expressions found in our Xhosa corpus: the auxiliary(-like) **kwazi** (1) and **nak(h)o** (2), and the prefixes **noku-** (3) and **nga-** (4).[1]

(1) **U-ya-kwazi uku-fund-a**.
SP$_1$-DJ-POSSIB INF-read-FV
'[S/]he can read.' (Tshabe & Shoba 2006: 38)

(2) **… u-nako uku-bhalis-el-a uku-vot-a**.
SP$_{2SG}$-POSSIB INF-register-APPL-FV INF-vote-FV
'… you can register to vote.' (Roux *et al.* 2001, corpus example)[2]

(3) **Ndi-noku-y-a e-khaya nokuba a-ndi-na-mali**.
SP$_{1SG}$-POSSIB-go-FV LOC-5.home even_if NEG-SP$_{1SG}$-COM-9.money
'I am allowed to go home even if I do not have any money.'
(Mini *et al*. 2003: 416)

(4) **U-nga-lim-a imifino phaya**.
SP$_{2SG}$-POSSIB-cultivate-FV 4.vegetable there
'You can grow vegetables there.' [because there is sunlight] (constructed)

The article is organised as follows: In Section 1, we give a brief background to the study and define and exemplify the modal (possibility) categories we use in our analysis. Section 2 then delves into the properties and corpus distributions of the possibility markers **kwazi** (Section 2.2), **nak(h)o** (Section 2.3), **noku-** (Section 2.4), and **nga-** (Section 2.5). The article concludes with a discussion of our key findings, as well as some open questions.

1. Background

This section gives a very brief background on our study. For extensive theoretical background, as well as details on our corpus and methodology, see the first part of this study (Crane *et al*. 2024).

1. Throughout the paper, we gloss all possibility markers as POSSIB (and necessity markers as NEC) so as not to pre-impose an analysis on their functions.

2. For the remainder of this article, corpus examples are attributed according to the original corpora from which they are drawn (see Crane *et al*. 2024 for corpus details). Other examples are marked with their source or as (our own) constructed examples.

We define modality as the expression of necessity and possibility. Our modal taxonomy is an extended version of that found in van der Auwera & Plungian (1998; see also e.g. Nauze 2008). The modal categories we distinguished are briefly defined and exemplified below.

'Participant-internal (PI) possibility' refers, basically speaking, to ability (van der Auwera & Plungian 1998: 80). Abilities can be physical or mental, inherent or acquired. We used the diagnosing question in (5) to code participant-internal possibility. A constructed illustrative example is given in (6).

(5) *Does it reflect the immediate subject's physical / mental / inherent / learnt capabilities or needs?*

(6) **uLucky** **u-ya-kwazi** **uku-balek-a** **ngokukhawuleza**.
1A.Lucky SP_1-DJ-POSSIB INF-run-FV quickly
'Lucky can run fast.' (constructed)

'Participant-external (PE) possibility' describes situational factors that make something possible for discourse participant(s) (van der Auwera & Plungian 1998: 80), including teleological possibility, describing possible means to a goal; and more general circumstantial or situational possibility. In order to best understand the markers' distribution, we attempted to distinguish teleological from situational possibility, where the former refers to a means for accomplishing a goal, while the latter does not do so explicitly (and may refer to the goal itself). In practice, of course, this distinction is not always straightforward to draw.

Example (7) gives our diagnosing question for teleological possibility, exemplified for Xhosa in (8).

(7) *Does it reflect a possible means to fulfil a participant's (or participants') goals?*

(8) **Uku-suk-a** **e-Rhawuti-ni** **u-si-y-a** **e-Kapa**
INF-depart-FV LOC-J.-LOC SP_{2SG}.SIT-EPEN-go-FV[3] LOC-CT
u-nga-thath-a **uN1**.
SP_{2SG}-POSSIB-take-FV 1A.N1
'To drive from Johannesburg to Cape Town, you can take the N1.' (constructed)

The question we used to diagnose more purely situational possibility is given in (9) and illustrated with a Xhosa example in (10). Note that (10) refers to the possibility

3. Like other Southern Bantu languages, Xhosa has several distinctive verbal paradigms, traditionally called 'moods' in studies of Southern Bantu. What we gloss as 'situative' is often referred to as the 'participial mood' and marks non-main-clause verbs that are interpreted as co-temporal with the main-clause predicate (see Creissels 2024 for discussion of these forms and the related terminology, with examples from Tswana; Creissels refers to the relevant forms as 'circumstantial'). These forms can be tonally and morphologically distinct from main-clause indicative forms. For the sake of consistency, we indicate situative forms in the glosses of subject markers, even when these are not morphologically distinct. We indicate some other subordinate and subjunctive forms that are not segmentally distinct in our glosses at the end of the form.

of achieving a goal itself (growing vegetables) rather than to a means to a goal (like taking the N1 in order to get to Cape Town in example 8) – even though growing vegetables is likely a means to an additional, unspoken goal (like saving money or having access to fresh and healthy produce).

(9) *Does it reflect what is possible given external situational constraints?*

(10) **U-nga-lim-a** **imifino** **phaya**.
SP$_{2SG}$-POSSIB-cultivate-FV 4.vegetable there
'You can grow vegetables there.' [because there is sunlight] (constructed)

'Deontic possibility' refers to what is allowed according to rules, norms, or the speaker's subjective judgment (van der Auwera & Plungian 1998: 81). We used the diagnosing question in (11); example (12) gives an illustrative Xhosa sentence.

(11) *Does it reflect the speaker's or societal [etc.] rules or expectations?*

(12) **Ngoku** **umnumzana** **uMutenda** **u-ya-kwazi** **uku-qhub-a**
now 1.mister 1A.Mutenda SP$_1$-DJ-POSSIB INF-drive-FV
ngokusemthethweni.
legally
'Mr Mutenda is now able (allowed) to drive legally' (constructed)

'Epistemic possibility' involves a knowledge- or belief-based assessment of what is possibly true (van der Auwera & Plungian 1998: 81). We diagnosed epistemic possibility using the question in (13). An example of epistemic possibility is given in (14).

(13) *Does it reflect the speaker's [or matrix clause subject's] knowledge, predictions, beliefs?*

(14) **uSipho** **u-noku-b-a** **u-s-e-restyu**.
1A.Sipho SP$_1$-POSSIB-be-FV SP$_1$.SIT-EPEN-LOC-9.restaurant
'Sipho might be at the restaurant.' [based on what we know] (constructed)

'Existential possibility' is also sometimes referred to as 'quantificational possibility'. This category, if described at all, is given various theoretical analyses (see e.g. Narrog 2012: 10; Portner 2009: 213-220; Tomozawa 2009). Following Tomozawa (2009), we consider existential possibility to indicate that a state of affairs 'sometimes' occurs, that is, that the potentiality is part of an entity's or category's characteristics. In our understanding, existential modality differs from participant-internal and from participant-external modality, as well as from epistemic modality, in that existential possibility is based on generalisations about an entity or class and its characteristics (i.e. things that 'sometimes' happen) rather than referencing a participant's abilities or specific situational possibilities.[4] For

4. Narrog's (2012: 10) describes existential modality as "the inanimate counterpart to the typically animate participant-internal modality". *Pace* Narrog, we view inanimacy as a common but not universal property of existential modality.

example, *my toddler can be exhausting* expresses existential possibility, whereas *my toddler can walk* expresses participant-internal possibility. Existential possibility is not directly addressed by van der Auwera & Plungian (1998), but we found it useful in categorizing some of the corpus data.

We used the diagnosing question for existential possibility given in (15), with a constructed illustrative example shown in (16).

(15) *Can the modal verb be substituted with 'sometimes'?*

(16) **Iklasi** **ye-fiziksi** **i-ya-kwazi** **uku-b-a** **nzima.**
9.class PP_{9}.CONN-9.physics SP_{9}-DJ-POSSIB INF-be-FV difficult
'Physics class can be [i.e. is sometimes] difficult.'(constructed)

Table 1 summarises the modal taxonomy for possibility used in this study.

Participant-Internal (PI)	Participant-external (PE)		Deontic	Epistemic	Existential
	Teleological	Situational			
Abiliy	Possible means to goal; some conceptual overlap with situational	Possibility given situational factors; some conceptual overlap with teleological	Permission; often treated as a subcategory of PE	Proposition judged as possibly holding	Proposition sometimes holds; semantic relationship to other categories unclear [to us]

Table 1. Modal taxonomy used in this study

Instantiations of modal expressions may be associated with one or more modal flavours.

A brief word on terminology is also in order. We refer to the modal forms, **kwazi**, and **nakho** using the (atheoretical) term 'auxiliary-like', because they may not fit traditional syntax-based definitions of auxiliaries (see Crane *et al*. 2024: 46-48 for discussion), but they nevertheless have important structural differences from the prefixes **noku-** and **nga-**. It may be noted that despite acting predicatively, the **na** in the **nakho** and **noku-** construction(s) is not a verb originally but a comitative preposition. The same might hold true also for **nga-**, which may have developed from a non-verbal element (see Section 2.5). Non-verbal predication is a robust phenomenon across Bantu (Downing & Marten 2019) and is prominent in Nguni languages.

Due to our verbal focus (including items that apparently originated in non-verbal predication), we do not consider possibility adverbs like **mhlawumbi** 'maybe'. Furthermore, although the line is tricky to draw, we attempted to restrict our corpus searches to markers that we consider primarily modal in nature. As a result, we exclude grammatical forms and constructions we know to be polysemous, with modal meaning being one of their functions, but, in our judgement, not a primary one. For example, verbal morphology such as subjunctive marking and the 'neuter' derivational suffix **-ek** can give rise to modal interpretations in certain contexts.

The modal functions of these grams are important topics for future study (see e.g. Veselovsky *et al*. 2023 for a corpus study of the subjunctive in Ruruuli-Lunyala, JE103). However, given their minimal morphological structure and heavy polysemy, they were not feasible candidates for the current corpus study.

We also exclude lexical verbs with single-flavour modal meanings, like **vunyelwa** 'be allowed to', used for deontic possibility (permission).

Our corpus and methodology are described in Crane *et al*. (2024). We coded each modal (and postmodal) use of each corpus example for its (post)modal function(s) and other factors; see Crane *et al*. (2024: 50-52) for details. As noted in that article, we make a distinction "between what we judged to be the 'primary' sense or senses, which seemed to us to be the main semantic function(s) of the marker in that context, and 'secondary' senses, which seemed to us to be plausible additional interpretations, but not the main intended meaning." (2024: 51).

2. Xhosa expressions of possibility

2.1. Introduction and overview

This section provides a general overview of the distributional patterns of the markers and their modal uses, based on intuition and corpus findings. Subsequent sections describe the history, syntactic features, and distribution and usage patterns of each marker in turn.

The third author's intuitions of the markers' felicitous uses in positive-polarity contexts are illustrated in Table 2. Brackets around the checkmarks indicate interpretive complexities, as discussed below.

	PI	PE		DEONTIC	EXISTENTIAL	EPISTEMIC
		Teleo	SIT			
kwazi	✓	(✓)	✓	✓	✓	
nak(h)o	(✓)	✓		✓	✓	✓
noku-	(✓)	✓		(✓)	✓	✓
nga-	(✓)	✓		✓	✓	✓

Table 2. Overall functional distribution of flavours in positive-polarity contexts, by possibility marker, in Xhosa, based on the third author's intuitions

According to the third author's intuitions, each of the markers is at least marginally compatible with all tested modal flavours, with a single exception: in positive-polarity contexts, **kwazi** cannot have straightforward epistemic meaning (17a).[5] However, the expression of epistemic impossibility with negated **kwazi** is felicitous (17b).

5. The third author notes that (17a) would sound better, although still not entirely natural, if **ngelixesha** 'at this time' were added to specify; this conceivable but unlikely usage may be (a pathway for) incipient semantic change. Such examples seem to be rare at best and are not found in our corpus.

(17) a. **#uSipho u-ya-kwazi uku-b-a-s-e-restyu.**
1A.Sipho SP_1-DJ-POSSIB INF-be-FV-EPEN-COP-9.restaurant
Intended: 'Sipho might be at the restaurant.' [based on what we know] (constructed)

b. **uSipho a-ka-kwazi uku-b-a-s-e-restyu.**
1A.Sipho NEG-SP_1-POSSIB INF-be-FV-EPEN-COP-9.restaurant
'Sipho can't be at the restaurant.' [based on what we know] (constructed)

The third author also finds **kwazi** to be of questionable felicity in the teleological participant-external contexts we tested, with polarity effects possibly playing a (lesser) role here, as well; this finding is supported and clarified by corpus data (Section 2.2).

The third author further notes that markers **nakho**, **noku-**, and **nga-** do not express pure ability in participant-internal contexts, but seem to invoke other situational factors, as well. For example, a translation of 'Lucky can run fast' using any of these markers would imply not only that Lucky has the ability to run fast, but also that additional situational factors are involved, such as good shoes or even Lucky's desire to run quickly. This intuition is also reflected in the corpus data (Sections 2.3-2.5): these markers are not used in canonical ability contexts, although they sometimes refer to ability more broadly construed.

Finally, **noku-**, although it can be used with deontic possibility as one of several senses, does not easily express straightforward permission in the third author's judgment. Instead, **noku-** invokes additional non-deontic meaning in permission contexts, as illustrated in (18). According to the third author's intuitions, (18) would imply that Mr. Mutenda also has the option to drive illegally.

(18) **#Ngoku umnumzana uMutenda u-noku-qhub-a**
now 1.mister 1A.Mutenda SP_1-POSSIB-drive-FV
ngokusemthethweni.
legally
Intended: 'Now Mr. Mutenda can drive legally.' (constructed)

Again, the corpus aligns with this intuition, with only one straightforwardly deontic positive-polarity occurrence of **noku-**[6] and a single negative-polarity deontic use.

Although these important meaning patterns already emerge, Table 2 might suggest a functional distribution that is only slightly differentiated by modal expression. However, as we show in Sections 2.2-2.5, the corpus shows usage patterns that diverge rather significantly, both in the distribution of form-flavour pairs and in the syntactic and semantic environments in which these pairings appear.

6. Additionally, this deontic occurrence employs the more straightforwardly decomposable form **ube nokuvota**, made up of a subjunctive verb followed by a comitative-marked infinitive verb and meaning, literally '[so that] you [may] be with voting'. This may even be a slightly different construction than subject-marked prefixal **noku-**. See Section 2.4.1, example (36) and surrounding discussion.

We now turn to discussion of the possibility markers as they occurred in the corpus. Figure 1 shows the number of occurrences (with modal function) of each marker in our corpus.

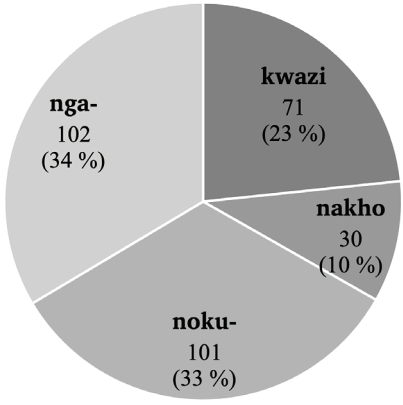


Figure 1. Total number of occurrences of each possibility marker in corpus (total = 304)

Figure 2 shows the overall distribution of modal flavours in the corpus distribution of flavours (both those coded as primary and as secondary senses) in corpus examples related to possibility, situated in order of frequency. The numbers shown indicate the total number of occurrences of each sense, with some occurrences associated with multiple senses (and therefore adding up to more than the 304 occurrences noted in Figure 1). The most frequent flavour found was participant-external (PE), with 176 corpus examples having at least a plausible participant-external interpretation, whether we judged this as the primary interpretation. The next most frequent possibility flavour was participant-internal (PI, ability), with 106 examples conveying at least some sense of (in)ability. Epistemic possibility uses were also common (found in 70 examples), with other flavours used rather less frequently.

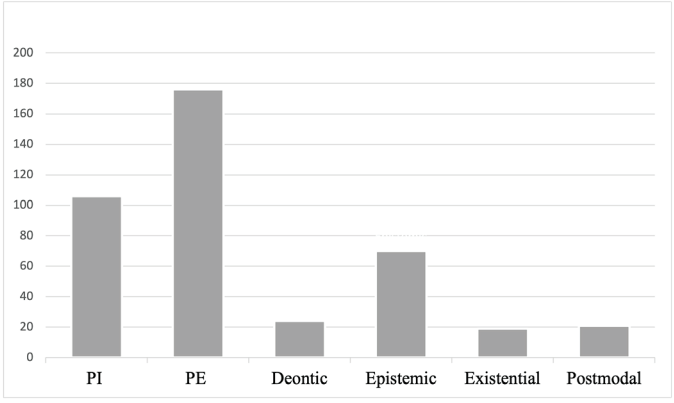


Figure 2. Overall distribution of possibility flavours, both primary and secondary senses

If the modal flavours were evenly distributed across the modal markers, we would, based on Figure 2, expect participant-external readings to be associated with the largest share of occurrences of each marker, and so forth. Figure 3 shows, however, that this is not the case. Figure 3 depicts the (normalised) distribution of flavours associated with occurrences of each possibility marker, whether we judged them to be primary or secondary. For example, participant-internal senses made up nearly 60% of senses associated with **kwazi**. (Recall again that the numbers include all modal flavours each corpus example plausibly invokes, so they add up to more than the total occurrences of each marker.)[7] Although these numbers are inherently approximate, Figure 3 clearly shows that as markers become more grammatical and less lexical, the relative role played by the participant-external and participant-external meanings (i.e. "root" modality) decreases, with more subjective (epistemic, existential) and even postmodal uses playing a bigger role.[8]

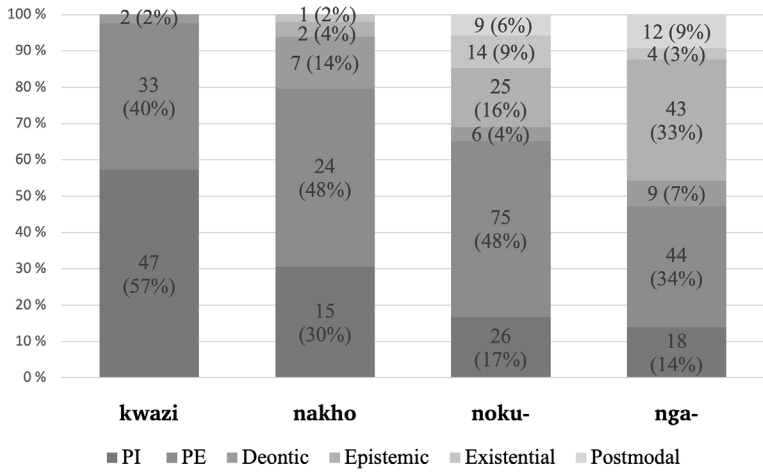


Figure 3. Total modal senses (both primary and secondary) associated with each possibility marker, as a proportion of total senses. The numbers in the graph showthe number of occurrences of each sense with the respective marker, along with the percentage (rounded) of total senses this represents.

7. It is regrettable that we cannot more straightforwardly show senses as a proportion of total occurrences (rather than, as here, senses as a proportion of total senses associated with the markers across all occurrences), but it would be less informative to attempt to graph only the primary senses here, because so many occurrences were ambiguous between two, three, or more senses, without any being obviously primary. We deal with the reading(s) judged as primary in the sections dedicated to specific modal markers.

8. Deontic modality is sometimes subjective, for example, when it refers to a speaker's personal assessment of what is right and proper. We have not found obvious associations between more subjective (or speaker-oriented) deontic modality (in contrast to externally imposed rules) and more advanced grammaticalisation, but the total number of deontic possibility senses in the corpus is too small to investigate the issue thoroughly.

2.2. Kwazi

2.2.1. Form and source

The first possibility marker, **kwazi**, is transparently derived from the verb **azi** 'know' (Proto-Bantu ***jíjìb**; see Bastin *et al*. 2002). The form **kwazi** arises via an object marking prefix **ku-**, agreeing with the (class 15) infinitive main verb, as illustrated in (19).

(19) **U-ya-kw-azi** **uku-fund-a**.
SP$_1$-DJ-INF-know INF-read-FV
'[S/]he can read.' (Tshabe & Shoba 2006: 38)

Bybee *et al*. (1994: 190) note that a verb meaning 'know' is "the most commonly documented lexical source for ability" in their cross-linguistic sample. Nevertheless, 'know' seems to be a rare possibility source in Eastern Bantu languages outside Zone S. For the eastern part of Eastern Bantu, Bernander *et al*. (2022) only list a handful of attestations of 'know'-derived possibility auxiliaries, while also mentioning that these are typically marginal markers in the modal systems of which they are a part. Moreover, it seems that none of these are cognates with **kwazi**.[9] In Nguni languages, in contrast, as well as in some other zone S languages, such as Xitsonga (S53), 'know' is widely used as a modal verb. In some of these languages, 'know' verbs are restricted largely to abilities that are learnt (this appears to be the case in, for example, Xitsonga, siSwati S42, and Sindebele S408, possibly due to influence from Sotho-Tswana languages), while in other languages, like Xhosa, no distinction between learnt and inherent abilities is drawn with 'know' as a modal verb (Crane *et al*. 2025).

There is evidence that **kwazi** is undergoing reanalysis as a more dedicated, monomorphemic modal marker. First, like the third author, a Xhosa speaker consulted on the issue does not allow straightforward modal meanings for 'bare' **azi** without **ku-** (Kristina Riedel, p.c.), suggesting lexical differentiation between **azi** and **kwazi**.[10] The second piece of evidence comes from the conjoint/disjoint distinction (see e.g. Savić 2020: 102-113). Simply put, 'disjoint' forms, marked with the **ya-** prefix in the present tense, indicate that material following the verb does not share syntactic constituency with it; 'conjoint' forms are used when the following material shares constituency with the verb. Agreeing object marking on the verb

9. However, there are possibility adverbs, originally derived from this verb stem, attested in Pangwa (G64), Nande (JD42) and Ndali (M301); See Bernander (2025).

10. The third author allows ability readings with bare **azi** (as expected with the lexical meaning 'know') but considers them pragmatically marked. For example, she judges that **uLucky wazi ukudlala ipiyano** 'Lucky can play the piano', with the bare **azi** in a conjoint verb form, can be used with the connotation that Lucky spends an inordinate amount of time playing the piano, to the detriment of everything else; this form would be used in a hyperbolic sense, as in something that might be said by a parent to chastise a child. The exclusive sense (Lucky plays/can play the piano and doesn't play or do anything else) is likely due to the shared constituency indicated with the conjoint form, as in (20), but the hyperbolic, chastising sense seems, at least in this case, to relate to the use of the bare **azi** form. Follow-up work on this topic, with a broader speaker sample, is obviously needed.

can only be used with disjoint morphology if the object it agrees with follows the verb, and conjoint forms cannot take object markers that are coreferential with overt objects (see e.g. Savić 2020: 104-106 and references therein). **Kwazi**, however, can be used in conjoint form followed by an infinitive object, which would not be expected if **ku-** represented an agreeing object marker (20).

(20) **uLucky** **u-kwazi** **uku-dlal-a** **ipiyano**.
1A.Lucky SP_1-know INF-play-FV 9.piano
'Lucky can play the piano [and not something else].' (constructed)

2.1.2. Corpus statistics and analysis

We judged 71 corpus occurrences of **kwazi** as modal. We excluded occurrences that did not have verbal complements (i.e., knowing something). We included borderline contexts that still seemed to have a strong sense of the lexical meaning 'know', but which were followed by verbs and had a plausible ability reading. We judged six occurrences (four positive polarity, two negative) as being borderline lexical in this way. Because the line between lexical and modal is inherently impressionistic, difficult to draw, and influenced by translation, we included all six in our analysis, acknowledging that others might draw different conclusions. Some examples of these borderline cases are given in (21) and (22).

(21) [**Ngaba la makhosikazi aza kubafundisa njani abantwana babo ukuziphatha kakuhle babengoomama namakhosikazi andilisekileyo ukuba wona**]
'[How are these women going to teach their children to behave well and be future mothers and dignified wives if]
a-wa-kwazi **uku-zi-phath-a** **kakuhle** **ku-qal-a?**
NEG-SP_6-know/POSSIB INF-REFL-carry-FV well INF-start-FV
they can't (/don't know how to) behave themselves in the first place?'
(Roux *et al.* 2001)

(22) [**Ndingathi nje abadlali baseYurophu banxanelwe impumelelo ngaphezu kwethu**. **Ngokomzekelo ndahamba kweli ndingumdlali waphambili ohlaselayo**,]
'[I can say that European players are so much hungrier for success than us. For example, I left this place [SA] as an offensive player,]
ndi-nga-kwazi **uku-buy-a**
SP_{1SG}.SIT-NEG-know/POSSIB INF-return-FV
not knowing (how) to drop back
[**ndizokuncedisa abadlali basemva yakuphuncuk' ibhola**.]
[to help defend when we lost the ball.]' (Snyman *et al.* 2012)

All 71 occurrences of **kwazi** had animate subjects, a pattern not seen in any of the other modal markers we examined, suggesting persistence of the meaning 'know', which tends to be associated with sentient beings. Of the 71 occurrences,

45 appeared in positive-polarity contexts, while 26 were marked for negation. We separate positive- and negative-polarity occurrences of **kwazi**, since previous work has shown that – like with English **can** – the distribution of compatible modal flavours differs across polarity for **kwazi** in Xhosa and other Nguni languages (Crane *et al*. 2025). Most significantly, as discussed in Section 2.1, negative epistemic uses of **kwazi** are felicitous (17b), while their positive-polarity counterparts are not straightforwardly so (17a). This contrast did not surface in our corpus investigation, however. We attribute this situation to two factors. First, **kwazi** may not be the first choice for expressing epistemic impossibility; second, the corpus contained very few epistemic impossibility contexts to begin with.

Figure 4 shows our corpus coding regarding the flavour distributions of **kwazi** in positive- and negative-polarity contexts. Note that unlike in the summary charts comparing markers, we report here only on the primary senses of each occurrence, that is, the main modal flavour(s) we judge to be associated with the marker's use in context, rather than plausible additional interpretations that we do not judge to be the main meaning.

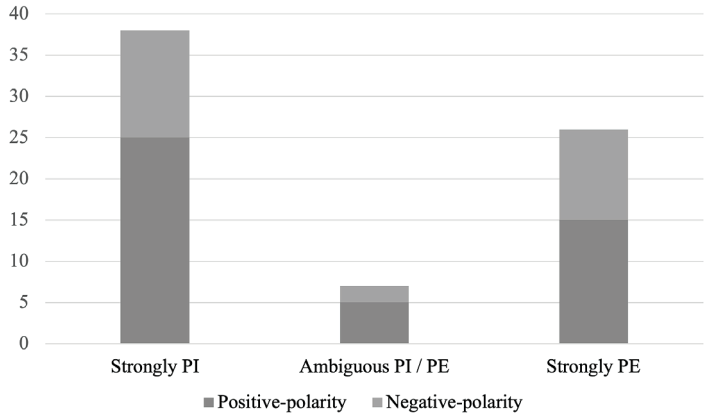


Figure 4. Distribution of primary senses of **kwazi** (y-axis indicates number of occurrences as a primary meaning)

The participant-internal-participant-external distinction is not always easy to determine, and indeed, in most cases, ability to do something is a function of both characteristics of the participant and characteristics of the situation. An illustrative example of this challenge is shown in (23): the ability to participate in the economy is explicitly linked to personal but situationally acquired "tools". Many examples we ranked as primarily participant-internal, with an additional participant-external sense, referred to the ability to cope in particular situations; note that **kwazi** is also used with similar function elsewhere in (23) (**bakwazi ukumelana noboni babo** 'be able to cope with their sins').

(23) [**iChrysalis Academy inikezela ngenkqubo yokuthintelwa kolwaphulo-mthetho ekuhlaleni, ijolise "kulutsha olusemngciphekweni". Injongo yale nkqubo ke kukuxhobisa abantu abaselula ukuba ba<u>kwazi</u> ukumelana noboni babo kwanokubanika**]
'[Chrysalis Academy offers a community-based crime prevention program, targeting "at-risk youth". The purpose of this program is to empower young people to be able to cope with their sins and to give them]
izixhobo **zo-ku-<u>kwazi</u>** **uku-thabath-a** **inxaxheba**
8.tool PP$_8$.CONN-INF-POSSIB INF-take-FV 9.rightful_share
ku-qoqosho
17-11.economy
the tools to be able to participate in the economy
[**ze babe ngabona bantu babonisa ithemba lempucuko kuluntu olukwiingingqi oluhlala kuzo**.]
[and become the people who show the hope of civilisation in the communities in the areas where they live.]'
(Snyman *et al.* 2012)

Example (24), in contrast, shows an example in which there simply was not enough additional context to resolve the question of participant-internal vs. participant-external modality. Even when the example was read within its broader context,[11] it was not clear whether the speaker was referring primarily to opportunities to act in plays, or to an early ability to do so; it is possible that both senses are intended.

(24) **<u>Nda-kwazi</u>** **ndi-se-lula** **uku-dlal-a**
SP$_1$.REMPST-POSSIB SP$_1$.SIT-PERS-young INF-play-FV
e-midlalw-eni **ye-qonga**.
LOC-4.play-LOC PP$_4$.CONN-5.stage
'I was able to play in plays at a young age.'
(Roux *et al.* 2001)

Figure 4 suggests that the distribution of **kwazi** skews only mildly towards participant-internal contexts (approximately 33% of positive-polarity and 42% of negative-polarity occurrences of **kwazi** were judged as primarily participant-external). However, closer investigation unmasks an important imbalance in **kwazi**'s distribution. In positive-polarity contexts, **kwazi** occurs as a participant-external possibility marker mostly in purpose/result clauses (25, 26),[12]

11. When he was working at Radio Mmabatho, he was chosen as the Best Presenter at the Marang Awards in 1991 and 1993. He also won the Best TV Actor Award for playing the role of the talkative Papankie in Bop TV's Sepelong. "<u>I was able to play in plays at a young age</u>. As a boy I played in many plays at school and on weekends. People who watched the games I played in loved me. When I passed my Matric, I did speaking and acting lessons at the Mmabana Cultural Center in Mmabatho and I got a certificate," he said.

12. Following Diessel (2001: 443): "RESULT clauses describe a consequence or conclusion derived from the main clause; and PURPOSE clauses indicate the intention or goal of the activity denoted in the main clause".

and almost exclusively in subordinate clauses. Of the 20 positive-polarity examples potentially associated with participant-external meaning (15 judged as primarily PE and five as ambiguous PI/PE), 12 occur in purpose/result clauses. Six others occur in other types of subordinate clauses, such as 'if'/'whether' clauses, 'because' clauses,[13] and relative clauses.

(25) **Abafana basemlungwini abakholelwa ncam ekugcineni imali yabo ebhankini. Bakholelwa bona kukusoloko beyibona eyabo imali**]
'[The young men of the organisation don't really believe in keeping their money in the bank. They believe in [keeping] their money [where they can see it]]
ukuze **ba-kwazi** **uku-theng-a** **iinkomo**
so_that SP$_2$-POSSIB.SBJV INF-buy-FV 10.cattle
so that they can buy cattle
[**bakubuyela emakhaya. lziganeko zasekapa azibanga nakumfundisa nto umathayi.**]
[when they return home. The events of Cape Town did not even teach him anything.]'
(Roux *et al.* 2001)

(26) [**UmPhathi woNyulo uya kulungisa isaziso esinika iinkcukacha kwaye enze iikopi ezaneleyo**]
'[The Electoral Officer will prepare a notice giving the details and make enough copies]
ukuze **inqununu** **ye-sikolo** **i-kwazi** **uku-nik-a**
so_that 9.principal PP$_9$.CONN-7.school SP$_9$-POSSIB.SBJV INF-give-FV
so the school principal can give
[**umfundi ngamnye esikolweni jonga ifomu P2.**]
[each student in the school a P2 form.]'
(Eiselen & Puttkammer 2014)

With negative polarity, 13 contexts clearly or potentially dealt with participant-external possibility. These examples occurred in a variety of clause types, but there wasn't any obvious tendency towards purpose/result clauses as with the positive-polarity occurrences. In addition to circumstantial participant-external readings, at least two occurrences seemed to have plausible deontic readings. In (27), for example, the woman's circumstantial inability to claim her late fiancé's property is due in part to legal considerations.

(27) ["**Umntakwethu wasweleka sele elobole yaye eza kutshata.**]
'["My brother passed away when he already paid lobola and he was about to get married.]
Umfazi **wa-khe_** **a-zange** **a-kwazi** **ukw-enz'**
1.wife PP$_1$.CONN-POSS$_1$ NEG.SP$_1$-PST.NEG SP$_1$-POSSIB.SBJV INF-do[FV]

13. One of these does not involve 'bare' **kwazi**, but rather **kwazi** inflected with the possibility prefix **nga-** (**ungakwazi**).

ibango
5.claim
His wife couldn't make a claim
[**kwizinto zakhe awayezishiyile. Kwakuzakuba bhetele ukuba wayefumene irisithi," utsho uMpho**.]
[for the things he had left behind, it would have been better if she had got a receipt", said Mpho.]'
(Snyman *et al.* 2012)

Therefore, **kwazi** occurs unrestrictedly only in participant-internal contexts in the corpus; exceptions either have negative polarity or occur, with few exceptions, in subordinate clauses, most of which are purpose/result clauses. Of the two positive polarity occurrences not in relative clauses, at least one is ambiguous between participant-internal and participant-external readings (24).[14] This distribution aligns with participant-internal possibility as the primary, and diachronically prior, meaning of **kwazi**, as predicted by its lexical meaning.

A functional overlap between the participant-external domain and purpose (and some result) clauses have been noted for Bantu elsewhere. In her study on modality in Shangaci (P312), Devos (2008) analyses verb forms in purpose clause constructions as expressing PE possibility, based on the notion that the actualisation of the event described in the purpose clause is dependent on the event depicted in the main clause, i.e., on external circumstances. In Xitsonga, purpose/result clauses also play an important role in allowing **tiva** 'know' to be used in non-participant-internal contexts (Crane *et al*. 2025).

2.3. **Nako/nakho**

2.3.1. Forms and source

Xhosa has two auxiliary-like forms, **nako** and **nakho**, which seem to be variants of the same expression. The *Greater Dictionary of isiXhosa* (Mini *et al*. 2003: 70) defines **-khó** as "the isiTshiwo ["standard"] form of **-ko** used after **na-** to indicate ability to do something". Mini *et al*. (2003: 131) further suggest that **khó** is the more conservative form. McLaren (1975: 91) defines **nakho** as 'be able to, can'.

Oosthuysen (2016: 111), analyses this phonological variation as an instance of morphological variation, with **ko** representing infinitive noun class 15 and **kho** locative class 17. This line of thinking ultimately relies on the crucial fact that **nako** ~ **nakho**, albeit showcasing auxiliary-like functions, does not originate in a lexical verb. Ultimately, it consists of the comitative preposition **na** (common throughout Bantu and reconstructed to Proto-Bantu; see Bastin *et al*. 2002) further indexed with a pronoun (from the 'pronominal prefix' series of Bantu concord markers; see van de Velde 2019), where **ko** and **kho** conform to class 15 and 17, respectively (cf. similar examples from other noun classes, like **na-ye** of class 1, **na-bo** of class 2

14. The other speaks of an opportunity: **Ndiza ku<u>kwazi</u> ukusebenzisana nabantu abahlelelekileyo ndithabathe inxaxheba kokwenziwa luluntu** 'I will be <u>able</u> to work with disadvantaged people and participate in community development'.

or **na-zo** of class 10, etc).[15] As suggested by the dictionary discussion cited above, the **nako** ~ **nakho** distinction, even if it originated in different noun class pronouns, may now be dialectal rather than grammatical.

For the sake of simplicity of analysis, and since the meaning conveyed by these different formal realisations is identical, we will from this point on refer to both discussed variants jointly as **nakho**.

As shown by Andrason (2019) in his thorough analysis of the manifold functions associated with the comitative **na** in Xhosa, it may, when used together with subject prefixes, convey a predicate-like ("copulative") possessive function. In this role, it may also operate on an infinitive verb as its possessed complement.
Example (28) illustrates these points, while also containing the pronominal suffix attached to the comitative.

(28) **Ndi-na-ko** **konke** **uku-thobek-a** **kwentliziyo**.
SP_1-COM-$PRON_{15}$ 15.all 15/INF-be_humble-FV PP_{15}.CONN.9.heart
'I have all humility of heart.' (Oosthuysen 2016: 111)

We are inclined to believe that constructions like (28), that is, where **na** functions predicatively, coding an agent subject possessing a nominalised activity, constitute the source construction of the modal function of **nakho**. From a typological perspective, it is important to point out that the reading of the source construction is not 'I have to X' (a type of auxiliary construction commonly associated with obligation; see Bybee *et al*. 1985, 1994: 183-186; Bybee & Pagliuca 1985 - also in an Eastern Bantu context; see Bernander *et al*. 2022), but rather 'I have (it) X', i.e., 'I control X'. As Bybee *et al*. (1994) note, the important issue here is not the meaning of the semantically quite empty 'be with/have'-copula(tive), but rather the actual meaning of its (de-)verbal complement. In English, the infinitive comes with a "projected sense" – i.e., the future-oriented 'have to X' reading where the subject is projected to engage in the event depicted – which easily lends itself to a reading of obligation. However, as we see in (28) above where the subject is coded as currently being active and controlling the event depicted, that is not (always) the case with the infinitive class 15 in Xhosa (nor in Bantu in general).

Most likely, the form of the modal prefix **noku**- (discussed in Section 2.3), originates from a similar source construction where a predicatively used **na** operates on an infinitive (/deverbal noun). The difference lies in the absence of the

15. Xhosa also has the forms **kho(na)** 'exist, be present, be there, etc.', an existential non-verbal predicate (see Bloom Ström 2020 for extensive discussion; see also Andrason 2019; Mini *et al.* 2003: 70); and **khona**, an 'absolute emphatic' pronoun (see Andrason 2019; Guthrie 1970: 19); cf. similar examples from other noun classes like **yena** (class 1), **bona** (class 2), **zona** (class 10) and consequently **nayena, nabona**, **nazona** when occurring with comitative **na**. Some other Nguni languages (e.g. Southern Ndebele S407 and Swazi S43) have a form **k(g)hona** 'be able to', which appears to be a borrowing from Sotho-Tswana languages (Crane *et al*. 2025). **Nak(h)o** in Xhosa does not appear to be related to the modal verb **k(g)hona,** nor have we found any corpus occurrences of **nakhona**, with or without modal meaning, so we do not consider any of these forms to be directly relevant to our analysis.

pronominal 'accretions' suffixed to **na**, which are replaced by the direct combination of the comitative and the infinitive.

Nakho can be inflected with (at least) subject marking (29), negation (30), and the persistive prefix (31).

(29) **Zam-a kangangoko u-nako, u-za ku-phumelel-a.**
try-FV to_the_utmost_extent SP_{2SG}-POSSIB SP_{2SG}-FUT INF-succeed-FV
'Try as hard as you can, you will succeed.'
(Snyman *et al.* 2012)

(30) **Umqeshi a-ka-nakho ukuba:**
1.employer NEG-SP_1-POSSIB COMP
'An employer may not:
[**anganika umntu utywala njengendlela yokumthundezela ukuba amsebenzele**; **…**]
[give someone alcohol as a way to induce him to work for him; …]'
(Eiselen & Puttkammer 2014)

(31) **a-ka-se-nakho uku-zi-sebenz-el-a, u-jing-a kum**
NEG-SP_1-PERS-POSSIB INF-REFL-work-APPL-FV SP_1-depend_on-FV 17.$PRON_{1SG}$
'as he is no longer able to work he depends on me (for support)'
(Tshabe & Shoba 2006: 817)

It also appears on its own (at least in the most common orthographical variant found in the corpus) after forms of the copula **ba** 'be' (32).

(32) [**Ngoko ke kuya kufuneka ube kanti umelene nezi meko zimbini ngaphambi kokuba**]
'[Therefore, you will need to have resistance to these two conditions before]
u-b-e nakho uku-fuman-a unyango
SP_{2SG}-BE-SBJV POSSIB INF-get-FV 11.treatment
you can receive treatment
[**ngamachiza adodobalisa ulwamvila lwale ntsholongwane kaGawulayo**.]
[with drugs that weaken the immune system of the AIDS virus.]'
(Roux *et al.* 2001)

In addition, it can be prefixed with **nga**- (33).

(33) **"U-nga-nakho uku-wu-qabelis-a loo mcimbi,**
SP_{2SG}-POSSIB-POSSIB INF-OP_3-solve(problem)-FV DEM_3.MED 3.matter
'Would you be able to discuss the matter
[**njengoko ndithemba ukuba selekumfikelele uMagusha?"**
[as I hope it has reached Magusha?]'
(Roux *et al.* 2001)

2.3.2. Corpus statistics and analysis

Nakho appeared 30 times in our corpus data (14 times as **nakho** and 16 as **nako**). It appeared in negative polarity contexts 18 of these times, and with positive polarity only 12 times (three times as **nakho** and nine as **nako**). It therefore skews remarkably to negative contexts compared to the other modal expressions in our sample. This tendency is also noted by Oosthuysen (2016: 318).

Unlike **kwazi**, we found **nakho** with a few (two or three)[16] inanimate subjects. Still, **nakho** clearly prefers animate subjects in our corpus sample (27 or 28 out of 30 occurrences). If our analysis of the source construction of **nakho** (Section 2.3.1) is correct, the tendency towards animate subjects may be another example of persistence, since in the non-modal source construction the subject is a possessor, and possessors tend to be animate.

Figure 5 shows the distribution of corpus uses of **nakho**. Note again that unlike in the summary charts in Section 2.1, we focus here only on what we judge to be the primary readings. Corpus uses seem to fall mostly into participant-external (circumstantial, teleological) territory, with deontic functions also important. Some participant-internal uses were also found, but most of them occurred in contexts that are ambiguous between participant-internal and participant-external. Although the corpus distributions of **nako** and **nakho** differ with respect to polarity and flavour, the overall low numbers suggest that this is coincidental. Similarly, the apparent differences between positive- and negative-polarity contexts in terms of flavour (where PI and deontic flavours appear more likely to be negated than are PE or ambiguous PI/PE contexts) are likely to be coincidental, especially given that four of the negative-polarity deontic examples are from the same text excerpt.

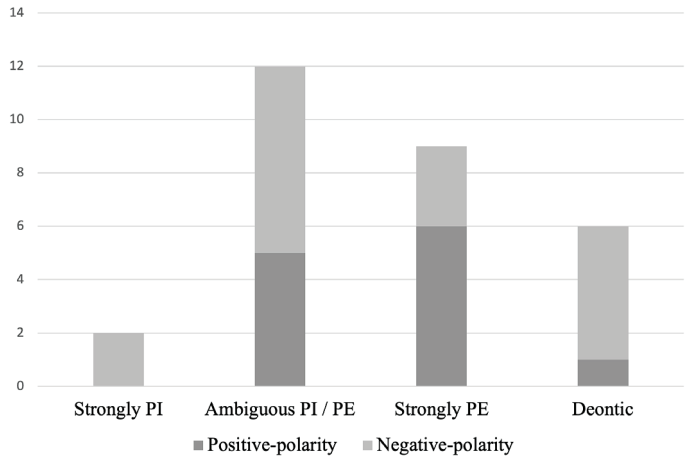


Figure 5. Distribution of primary senses of **nakho** (y-axis indicates number of occurrences as a primary meaning)

16. One of the examples referred to a participant's heart and its inability to calm down in certain circumstances.

Only two strongly participant-internal uses were found (34, 35), the latter of which conveys a very subjective meaning, not typically associated with straightforward abilities.

(34) **Ukuba** **a-ka-nako** **uku-fund-a** **ku-funeka**
if NEG-SP$_1$-POSSIB INF-read-FV SP$_{17}$-NEC
u-m-fund-el-e **ingxelo**.
SP$_1$-OP$_1$-read-APPL-SBJV 9.report
'If he can't read you have to read the report to him.'
(Eiselen & Puttkammer 2014)

(35) [**Wayenolwamahashe unina kanxele**,]
['Nxele's mother could argue about (?) / fight like (?) horses,]
e-nge-nakho **mpela** **uku-m-bukel-a** **umntwana**
SP$_1$.SIT-NEG-POSSIB completely INF-OP$_1$-watch-FV 1.child
wa-khe **e-yob-a**.
PP$_1$.CONN-POSS$_1$ SP$_1$.SIT-faint-FV
she could not watch her child die.' (Roux *et al.* 2001)[17]

Deontic examples are given in (36) and in (30) above. **Nakho** has a significantly higher proportion of deontic examples than the other markers, but, since four of the six come from the same text, it seems unlikely to us that **nakho** is more specialised for deontic than the other markers. Instead, deontic possibility meanings appear to be generally rare in the corpus.

(36) [**Ukuba uneminyaka eli-17 kodwa uza kuba uneminyaka eli-18 ngaphambi komhla woku-1 kuNovemba**, **ekuyimini yonyulo**,]
'[If you are 17, but you will turn 18 before 1 November, the day of the election,]
u-nako **uku-bhalis-el-a** **uku-vot-a**.
SP$_{2SG}$-POSSIB INF-register-APPL-FV INF-vote-FV
you can register to vote.' (Roux *et al.* 2001)

Based on this distribution, we analyse **nakho** as a general circumstantial marker, used for enabling circumstances related to participants (participant-internal), situations (participant-external circumstantial/teleological), or authority (deontic). It is not typically used in straightforward ability contexts, although it may be more common in negated PI (inability) contexts (see also examples 31 – ambiguous between PI and PE – and 34 above). The third author judges the use of **nakho** (as well as all other markers besides **kwazi**) in contexts intended as participant-internal to imply that the referenced situation does not depend solely on the participant's abilities, but also on other circumstantial factors, as described in Section 2.1.

17. Example (35) comes from a literary source that employs a great deal of metaphorical language, and our translations of **wayenolwamahashe** and **eyoba** are tentative. The context refers to mother animals' reactions to the suffering of their offspring and seems to juxtapose that imagery with the image of Nxele, who is being beaten because he lost his father's horses, and of his mother, who is begging for the beatings to stop. In any case, the translation uncertainties do not affect the interpretation of the possibility expression **engenakho**.

We found some potential bridging contexts to epistemic uses. Example (37) illustrates such a context, also demonstrating the challenges of categorising modal expressions when used with abstract subjects. The hypothetical solution described as being possible is so because of its internal properties as well as because of circumstantial factors, but the description is also at least somewhat epistemic, as it is based on the speaker's knowledge, predictions and beliefs.

(37) **[Eyona nto ingundoqo kule nto ziinzame ezisendleleni zokuqinisekisa ukuba abo baphambili eSomalia bayancokola ukuze bafumane isisombululo esiquka umntu wonke**]
'[Central to this are the ongoing efforts to ensure that those at the forefront of Somalia are talking to find an inclusive solution]
n-e-si-<u>nako</u> **ukw-enzek-a**
COM-REL-SP$_7$-POSSIB INF-happen-FV
that can also be realised / that is also viable
[**ngokubhekiselele kwimfuneko yokufezekisa uxolo lukazwelonke**.]
[in light of the need to achieve national peace.]'
(Eiselen & Puttkammer 2014)

Although the corpus did not produce clearly existential or epistemic uses of **nakho**, the third author finds them to be very natural, as in the existential context in (38) and the epistemic context in (39). Thus, the absence of these functions from our corpus seems to be due to the small number of occurrences of **nakho** overall, rather than to its intrinsic properties.

(38) **Amafu** **a-<u>nakho</u>** **uku-b-a** **ma-khulu** **okanye** **(a-<u>nakho</u>**
6.cloud SP$_6$-POSSIB INF-be-FV COP$_6$-big or SP$_6$-POSSIB
uku-b-a) **ma-ncinci**.
INF-be-FV COP$_6$-small
'Clouds can be big or small.' (constructed)

(39) **uSipho** **u-<u>nakho</u>** **uku-b-a-s-e-restyu**.
1A.Sipho SP$_1$-POSSIB INF-be-FV-EPEN-LOC-9.restaurant
'Sipho might be at the restaurant.' (constructed)

2.4. **Noku-**

2.4.1. Forms and source

Noku- is a prefix that appears before the verb stem, as in (40). It also appears before any agreeing object markers. **Noku**- seems to derive from the construction SP-**na**-INF, where SP is the subject prefix and **na** is a comitative preposition used predicatively to indicate possession. **Uku**- is the infinitive prefix, with which the vowel in **na** coalesces to form **noku-**. This coalescence is seen clearly in (41), where instead of **noku**- being directly prefixed with a subject marker the subjunctive-marked auxiliary **ube** '[so that] you [may] be' is followed by a comitative-marked infinitive

verb **na-uku-vota** > **nokuvota**, literally 'so that you may be with voting'. In this case, comitative **na** does not seem to be used predicatively.[18]

(40) **Ndi-noku-y-a** **e-khaya** **nokuba** **a-ndi-na-mali**.
SP$_{1SG}$-POSSIB-go-FV LOC-5.home even_if NEG-SP$_{1SG}$-COM-9.money
'I am allowed to go home even if I do not have any money.'
(Mini *et al*. 2003: 416)

(41) **Umthetho** **u-thi** **ku-funeka** **u-b-e** **ne-minyaka**
3.law SP$_3$-say SP$_{17}$-NEC SP$_{2SG}$-be-SBJV COM-4.year
e-li-18 **ukuze** **u-b-e** **noku-vot-a**.
REL$_4$-COP$_5$-5.eighteen so_that SP$_{2SG}$-be-SBJV POSSIB-vote-FV
'The law says that you should be 18 in order to [so that you can] vote.'
(Roux *et al.* 2001)

Its negated form is **naku**- (42), with the vowel change at least originally reflecting the absence of the augment vowel **u**- on nouns (including verbal infinitives) in negative-polarity contexts. However, the third author also uses **noku**- in negated forms, as in (43), a usage we did not observe in our corpus results. Although use of the augment under negation in Xhosa is variable and may relate to contrastive focus, the augment is obligatorily dropped after predicative **na** under negation (Bloom Ström & Miestamo 2024). The use of **noku**- under negation therefore suggests a type of micro-variation indicating increasing non-compositionality and thus a further step in morphologisation.

(42) **umntana** **a-ka-naku-wu-thwal-a** **umthwalo** **o-nzima**
1.child NEG-SP$_1$-POSSIB-OP$_3$-carry-FV 3.load REL$_3$-heavy
kangaka
such
'a child cannot carry such a heavy load' (Mini *et al*. 2003: 971)

(43) **A-wu-noku-thath-a** **iLongStreet** **uku-y-a** **kuloSipho**
NEG-SP$_{2SG}$-POSSIB-take-FV 9.Long_St INF-go-FV 17.Sipho's_place
namhlanje.
today
'You can't take Long St to Sipho's house today.' (constructed)

We are inclined to believe that prefixal **noku-** is interrelated with the auxiliary-like modal **nak(h)o**, but that these forms (and their variants) constitute two slightly different source constructions with **nakho** involving an additional (class 15 or 17) pronominal morpheme, whereas **noku**- is attached directly to the (class 15) infinitive verb. These structural differences could explain the functional differences between these two markers as well. That is, to some extent we are dealing with a case of

18. As alluded to in Section 2.1, the form in (42) may represent a different construction than prefixal **noku**-. The possibility reading here may also be contributed by the subjunctive marking on the auxiliary.

'polygrammaticalisation' (Craig 1991), where (partly) similar source material yields different grammatical outcomes.

2.4.2. Corpus statistics and analysis

The modal prefix **noku-**, with 101 occurrences, was more common in the corpus than either **kwazi** or **nakho**. Its meanings were also more diffuse. **Kwazi** and **nakho** cluster around participant-internal and participant-external readings, respectively. In contrast, **noku-**, while primarily used in participant-external possibility contexts, also appeared in existential and epistemic contexts, as well as with several postmodal functions. The few (arguably) participant-internal readings of **noku-** did not refer to prototypical abilities. The distribution of what we classified as the primary interpretations of each occurrence is given in Figure 6. Although the figures seem to skew strongly towards participant-external readings, this distribution should be interpreted with attention to the overall distribution of modal interpretations in the corpus, where primarily participant-internal are largely associated with **kwazi** and participant-external meanings make up the bulk of the rest of the senses (see Figure 2). Only **nga-** (Section 2.5) has a larger share of epistemic readings than **noku-**.

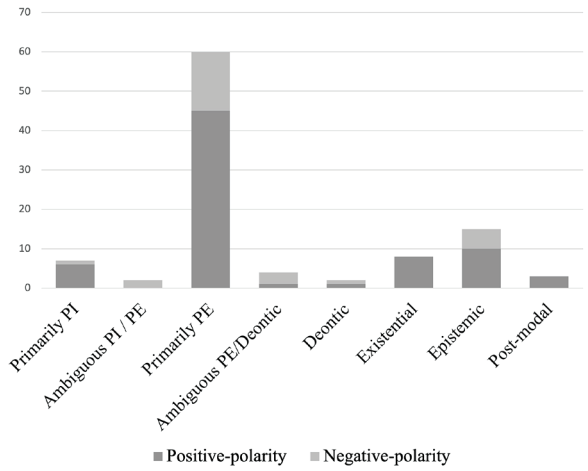


Figure 6. Distribution of primary senses of **noku-** (y-axis indicates number of occurrences as a primary meaning)

Of **noku-**'s 101 occurrences, 74 were positive in polarity, and 28 were negative. At least 27 had inanimate or abstract subjects, making up a significantly higher percentage of occurrences than with either **kwazi** or **nakho**. This pattern can be interpreted as another instance of semantic bleaching correlating with formal grammaticalisation, especially if, as we posit, **noku-** and **nakho** arose from a similar source construction involving possession, but the former, as a verbal prefix, is more grammatically integrated. The large number of occurrences of **noku-** with inanimate subjects made for more challenging coding and analysis, because

flavour distinctions often collapse with inanimate subjects, making multiple interpretations feasible. For example, the sentence in (44), where the modal marker on the relative-clause verb 'keep' refers to the subject noun 'things' (class 10), can be argued to have interpretations that are participant-internal (properties of the beneficial foci), existential (things that sometimes happen), teleological (enabling the goal of staying out of trouble), and possibly even epistemic (belief or prediction about the results of appropriate foci).

(44) [**Umyalezo endiwunika ulutsha lwelali yam nolwaseMzantsi-Afrika ngokubanzi ngowokuba maluziyeke izenzo ezingalunganga luzimisele ezincwadini zalo nakwezinye izinto**]
'[The message I give to the young people of my village and South Africa as a whole is that they must stop doing wrong things and start focusing on books and all the things]

e-zi-noku-lu-gcin-a **lu-xakek-ile**.
REL-SP$_{10}$-POSSIB-OP$_{11}$-keep-FV SP$_{11}$-be_busy-PFV.DJ
that can keep them busy [out of trouble].' (Roux *et al.* 2001)

However, even disregarding examples with inanimate or abstract subjects, **noku-** occurrences were harder than either **kwazi** or **nakho** to neatly slot into modal flavours, especially in negative-polarity contexts. We judged fully a third of negative-polarity occurrences to be associated with at least three possible modal flavours.[19]

As we alluded to above, even the occurrences that we coded as primarily participant-internal do not refer to prototypical abilities; instead, they mostly seem to refer to the nature of participants, from the perspective of another participant (here, the addressee) and their goals, desires, or expectations as in (45) and (46).[20]

(45) **Ugqirha** **no-mongikazi** **ba-noku-tsho ukuba** **w-enzakal-e**
1A.doctor COM-1A.nurse SP$_2$-POSSIB-say COMP SP$_{2SG}$-hurt-PFV.CJ
njani.
how
'The doctor and nurse can tell how you got hurt.' (Roux *et al.* 2001)

19. **Noku**- was challenging to code in general: at the first pass, we disagreed significantly (at least two completely different answers) about 47 of the 101 cases, and it was only through much discussion and reconsideration of modal categories that we arrived at the tentative coding of "primary" modal interpretations illustrated in Figure 6.

20. A reviewer notes that they get primarily participant-external readings with these examples, as they describe what is possible given facts about the world. We still judge them as primarily participant-internal, as they refer to the abilities of doctors and nurses, although, as we note, the speaker's goal is not to describe these abilities but to mention them as means of achieving the participant's goals.

(46) [**Xa engafuni yena ukuba niye kucetyiswa, yiya nokuba uwedwa Kwiiofisi zeFAMSA okanye zeFamily Life niya kufumana abantu abayiqondayo imeko yenu**]
'[If he does not want you to be advised, go alone to the FAMSA or Family Life offices, you will find people who understand your situation (and)]
a-ba-noku-ni-nced-a.
REL-SP_2-POSSIB-OP_{2PL}-help-FV
who can help you.' (Roux *et al.* 2001)

Noku- was also used in some contexts that were – at least arguably – postmodal. Postmodal uses included conditional (/counterfactual) and concessive. In most cases, more than one modal flavour is associated with the marker, and postmodal functions are supported by other linguistic signals such as tense-aspect-mood marking, but our data at least suggest ongoing grammaticalisation to postmodal functions (see also Andrason 2019 on concessive uses in general of the comitative **na**, the partial etymon of **noku-**). An example is given in (47), where the concessive meaning is formed by **noku-** together with persistive **se-** and with the verb **enzeka** 'happen', which frequently occurs in expletive form in epistemic contexts.

(47) **Ku-se-nokw-enzek-a**
SP_{17}-PERS-POSSIB-happen-FV
'It may be
[**ukuba ungomnye wabona bantu badala ehlabathini, kodwa eli xhegokazi lisomelele, kwaye lithi liseza kuphila iminyaka emininzi ezayo**.]
[that she is one of the oldest people in the world, but this old lady is still strong and she says she is still going to live many more years.]'
(Snyman *et al.* 2012)

Example (48) shows an apprehensive-like use with epistemic **noku-**, again employing the persistive.

(48) [**Khuthaza iqabane/amaqabane akho ukuba aye kufumana unyango eklinikhi njengoko**]
'[Encourage your partner/partners to go to get medical attention at the clinic]
u-se-noku-phind-a **w-o-sulelek-e** **futhi**
SP_{2SG}-PERS-POSSIB-do_again-FV SP_{2SG}-FUT-be_infected-SBJV again
ku-funeka **u-phind-e** **w-a-nyang-w-a**.
SP_{17}-NEC SP_{2SG}-do_again-SBJV SP_{2SG}-CONS-treat-PASS-FV
as you can [lest you] get infected again and need to be re-treated.' (Eiselen & Puttkammer 2014)

Based on the corpus distribution, we classify **noku-** as a general possibility marker, used for all flavours except for straightforward ability. This analysis is confirmed by the third author's intuitions, which furthermore support the idea that **noku-** is underspecified for flavour and therefore invites multiple (non-ability) readings, as

suggested by our difficulty in coding **noku-** occurrences for primary flavour (see the discussion in Section 2.1).

2.5. **Nga-**

2.5.1. Forms and source

The "potential" morpheme **nga**-, like **noku**-, is prefixed to the macrostem (that is, before the verb root and object prefix, if present). It is pronounced with high tone (see e.g. Oosthuysen 2016: 316), but tone is not marked in Xhosa orthography. Xhosa has several **nga** morphemes with various functions – the string **nga** appeared 1982 times in our corpus[21] – so the third author, a native speaker of Xhosa, extracted **nga**- occurrences based on her intuitions of its function. A constructed example of potential **nga**- with (situational) participant-external meaning is shown in (49). Although **nga**- is frequently glossed as 'potential' (POT), in this article we gloss it POSSIB like the other markers in order to remain agnostic about its meanings and functions.

(49) **U-nga-lim-a** **imifino** **phaya**.
SP$_{2SG}$-POSSIB-cultivate-FV 4.vegetable there
'You can grow vegetables there.' [because there is sunlight] (constructed)

Reflexes of **nga**- associated with the marking of possibility as well as closely related conditional and concessive functions, are common throughout Southern Bantu and Eastern Bantu at large (cf. Nurse & Philippson 2006; Nurse 2008: 28, 54-55, 251; Nurse & Devos 2019), and can likely be linked to a reconstructable element ***ngà** '(be) like' (Bastin *et al*. 2002). Although there is disagreement on whether ***nga** was originally a verb or not, the discussion in Creissels (2017) seems to suggest that it was predicative verb-like uses of **nga** that led to the development of verb-internal possibility prefixes, at least in Southern Bantu languages like Xhosa. Furthermore, Creissels (2017) posits a pathway of semantic extension progressing from the similative ('(be) like') source meaning to possibility, via the reading 'suitable' (see also Crane *et al.* 2024 on the necessity verb **fanele**, with a similar etymology).

According to several grammars of Xhosa, potential **nga**- is negated as **nge**- (Oosthuysen 2016: 317; Einhorn & Siyengo 1993: 45) or in auxiliary form as **ngekhe** (Oosthuysen 2016: 317). These forms are found in other Nguni languages (see e.g. Crane *et al*. 2025), but in our corpus, we did not find any occurrences of **ngekhe** (we did find one occurrence of **ngeke**, which co-occurred with **kwazi**), and, to the best of our knowledge, only one of modal **nge**- (50), with somewhat vague circumstantial flavour, as in *I can't say because my knowledge does not allow it, it's not proper to guess,* and so forth.[22]

21. Morphemes with this shape are used to indicate, among other functions, as an instrumental or to mark negation.

22. We do not consider this a straightforward ability reading; the speaker has the physical and mental capacity to say something but lacks the situational ability.

(50) **A-ndi-<u>nge-</u>tsho** **Georgina**.
NEG-SP$_{1SG}$-POSSIB.NEG-say 1AGeorgina
'I can't say, Georgina.
[**Mhlawumbi neemali zokwakha iblorho yodidi ephakamileyo zazingekho ngaloo maxesha**.]
[Perhaps the money to build a high-class bridge was not available in those days.]' (Roux *et al.* 2001)

The third author accepts negation of **nga-** with **nge-** but prefers to use negated forms of **noku-** (often negated as **naku-**) rather than **nge-**, a pattern seemingly supported in the corpus data. We therefore only consider positive-polarity occurrences of **nga-**, having found just one candidate for **nge-** as an impossibility marker (50).

2.5.2. Corpus statistics and analysis

Somewhat surprisingly, considering its purported status as a general-purpose modal (see Nurse & Devos 2019: 227-228), we only found **nga-** with clear modal function 102 times in our corpus, that is, with almost exactly the same frequency as **noku-**. Given that we mostly dealt with positive-polarity **nga-** occurrences, **nga-** does seem to be more frequent than the other modal forms, but not extravagantly so. In terms of animacy, 67 **nga-** forms had animate subjects, 14 were inanimate, and a further 22 were the **(i)ngaba** auxiliary (discussed with example 53), which we did not classify for animacy. Other than those 22, which were always epistemic, there was no significant correlation between animacy and modal flavour.

As with **noku-**, it was difficult to categorise occurrences of **nga-**, and we only judged 64 of its occurrences as expressing a single modal flavour and no postmodal functions. Furthermore, fully 37 of these single-flavour occurrences are epistemic, with 21 judged as exclusively participant-external, three as participant-internal, and three as deontic. In addition to the modal uses, we judged nine occurrences to be primarily postmodal, with functions like optative/desiderative and concessive. Figure 7 shows the distribution of what we judged as 'primary' flavours, that is, our impression of the flavour(s) most strongly associated with each utterance.

Although it is found with every modal function in our taxonomy, uses of **nga-** cluster around participant-external and epistemic senses. Many of the epistemic uses employ a special construction, discussed below. As with **noku-**, even the 'strongly' participant-internal readings are not prototypical descriptions of ability. Instead, they tend to also convey a sense of epistemicity (51) or describe abilities or properties conducive to reaching someone's goal (52).

(51) [**Ndisebenze noMambush iminyaka esi-8 ngoku ukuba bekukho nje abahlanu abafana naye kweli qela lam**]
'[I've worked with Mambush for eight years now, if there were just five [people] that were like him on my team]
a-ku-kho **klabhu** **ibi-<u>nga-</u>s-oyis-a**
NEG-SP$_{17}$-EXIS 9.team SC$_9$.PST.IPFV-POSSIB-OP$_{1PL}$-surpass-FV
e-hlabath-ini.
LOC-5.world-LOC
there would be no team in the world that could beat us.' (Roux *et al.* 2001)

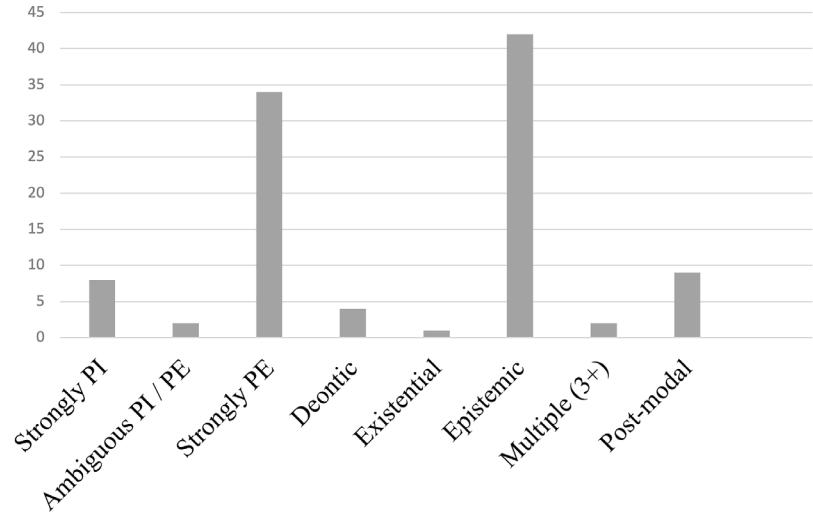


Figure 7. Distribution of primary senses of **nga-** (y-axis indicates number of occurrences as a primary meaning)

(52) **Kodwa** **iimpuku** **zona** **zi-lula** **ngoko**
but 10.mouse PRON$_{10}$ SP$_{10}$-light as
zi-nga-zi-joj-a **iziqhushumbisi** **zi-nga-qhushumb-i**.
SP$_{10}$-POSSIB-OP$_{10}$-sniff-FV 10.explosive SP$_{10}$-NEG-explode-NEG.SBJV
'But mice are light[weight] so they **<u>can</u>** sniff out the explosives/mines without them [the mines] exploding.' (Snyman *et al.* 2012)

Nga- also frequently co-occurs with the auxiliary **ba** 'be', either with the class 9 prefix **i**- (**ingaba**) or without a prefix (**ngaba**), always with epistemic meaning. These auxiliary forms serve as face-saving 'softeners', or hedges, preceding questions, so that questions that may be less comfortable for the addressees (such as 53) come across as less direct and more polite. A similar construction **ingathi** (not exemplified here) uses **thi** 'say, do'.

(53) **I-nga-b-a** **u-khe** **u-cing-e** **ngoku-thath-a**
SP$_{9}$-POSSIB-be-FV SP$_{2SG}$-ever SP$_{2SG}$-think-SBJV INSTR.INF-take-FV
isithembu?
7.polygamy
'**<u>Could</u>** it be that you ever think about polygamy?' (Snyman *et al*. 2012)

Most epistemic uses of **nga-** in the corpus employ these or similar constructions, but a few used lexical verbs, as in the negated example (54).

(54) [**Kodwa iAustralia itsho abalandeli banexhala lokuba**]
'[But fans of Australia are worried that]
u-nga-nga-zalisek-i **umnqweno** **wa-bo**.
SP_3-POSSIB-NEG-fullfil-NEG.SBJV 3.wish PP_3.CONN-$POSS_2$
their dreams might not come true.' (Roux *et al.* 2001)

Nga also occurs as an auxiliary, meaning 'appear to, pretend to…' which, when followed by another verb marked with potential **nga-**, indicates 'wish' (Mini *et al*. 2003: 501), as in (55). We coded these constructions based on our interpretations of the **nga-** marked verb's role (in 55, for example, participant-external possibility) noting also the optative or desiderative sense (see Section 2.2); **nga-** in this role can refer to either speaker or participant wishes – or to both, as in (55), where the speaker is the first-person participant.

(55) **Ndaku-gqib-a** **uku-fund-a** **ndi-nga**
$TEMP_{1SG}$-finish-FV INF-study-FV SP_{1SG}-AUX/POSSIB
ndi-nga-fund-el-a **ubutitshala** **okanye** **ubongikazi**.
SP_{1SG}-POSSIB-study-APPL-FV 14.teachEr or 14.nurse
'When I finish school I wish I could study towards a teaching or nursing degree.' (Snyman *et al.* 2012)

2.5.3. Double modals

A few corpus occurrences employed two different possibility modals, one a prefix and one auxiliary-like, which we believe to be common in Xhosa. For instance, negation of **kwazi** is often expressed in a complex construction involving another modal marker. Examples from the corpus with **noku-** (negated as **naku-**) and **kwazi** are given in (56, 57). In (56), the main modal meaning seems to be circumstantial and teleological, no straightforward participant-internal sense.

(56) **A-ndi-naku-kwazi** **uku-hlawul-a** **imali**
NEG-SP_{1SG}-POSSIB-POSSIB INF-pay-FV 9.money
yo-ku-yi-lungis-a **le** **ndlu**.
PP_9.CONN-INF-OP_9-fix-FV DEM_9.PROX 9.house
'I cannot pay the money to fix this house.' (Snyman *et al.* 2012)

In (57), two modal flavours seem to be activated: negated **noku-** in the auxiliary **akanakuze** refers to the (epistemic and circumstantial) impossibility that something will happen in the future, while **akwazi** refers to having the (participant-internal) ability to work.[23]

23. In fact, the scope of negation seems to be ambiguous here, since the expression can be translated as "can/will never" or as "might not be".

(57) [**Ubunzima okanye ubunzulu boko kukhubazeka kuko okuxela ukuba umsebenzi lo**]
'[The severity or the extent of this disability is what determines whether the worker]
a-ka-naku-ze **a-kwazi** **ku-phind-a**
NEG-SP$_1$-POSSIB-then_do SP$_{1SG}$-POSSIB.SBJV INF-do_again-FV
a-sebenz-e
SP$_1$-work-SBJV
will/might never be able to work again
[**okanye kuya kuba nzima kusini na ukuphinda afumane omnye umsebenzi**.]
[or whether it will be difficult to find another job].'
(Eiselen & Puttkammer 2014)

The co-occurrences we have observed so far in the corpus and in elicitation/introspection include **nga**-+**kwazi** (58), **noku**-+**kwazi** (56, 57, 59), and **nga**-+**nakho** (60).[24] Generally, both modal markers appear to make a semantic contribution. For example, in (58), **nga**- contributes a participant-external flavour to the description of the 'ability' to pay. In (59), **noku**- adds a situational flavour to the physical possibility indicated by **kwazi**. In (60), repeated from (33), **nga**- contributes an epistemic element ('is it possible') to the situational ability contributed by **nakho**. Because each of these elements can also express several modal flavours on its own, the specific contributions vary according to context.

(58) [**Ukungahlawuli imirhumo yakho akuthethi kuthi ipolisi ivaliwe**]
'[Not paying your premiums doesn't mean the policy is closed]
kuba **u-nga-kwazi** **uku-phind-a** **u-yi-hlawul-e**
because SP$_{2SG}$-POSSIB-POSSIB INF-do_again-FV SP$_{2SG}$-OP$_4$-pay-SBJV
because you can pay them again
[**xa ivuma imigwaqo yepolisi yakho**.]
[if the terms of your policy allow it .'] (Snyman *et al.* 2012)

(59) **A-wu-nokwazi** **uku-thath-a** **iLong** **St**
NEG-SP$_{2SG}$-POSSIB.POSSIB INF-take-FV 9.Long _St
uku-y-a **kuloSipho** **namhlanje.**
INF-go-FV 17.Sipho's_place today
'You can't take Long St to Sipho's house today.' (constructed)

(60) **"U-nga-nakho** **uku-wu-qabelis-a** **loo** **mcimbi,**
SP$_{2SG}$-POSSIB-POSSIB INF-OP$_3$-solve(problem)-FV DEM$_3$.MED 3.matter
'Would you be able to discuss the matter,
[**njengoko ndithemba ukuba selekumfikelele uMagusha?"**]
[as I hope it has reached Magusha?]' (Roux *et al.* 2001)

24. According to the third author, **unokuba nakho** is also possible with modal meaning, although **noku**- cannot be directly prefixed to **nakho**.

The non-participant-internal flavour contributed by **nga-** in double-modal constructions can be seen in the contrast between (61) and (62). In (61), with an explicit ability meaning (contrasted with deontic possibility), **nga-** cannot be prefixed to **kwazi**. In (62), where the deontic reading is specifically invoked, both **ngakwazi** and bare **kwazi** are felicitous, but **ngakwazi** is preferred by the third author.

(61) **Umnumzana uMutenda u-ya-kwazi / *a-nga-kwazi**
1.mister 1A.Mutenda SP_1-DJ-POSSIB / SP_1-POSSIB-POSSIB
uku-yi-qhub-a imoto kodwa a-ka-na-yo ilayisenisi.
INF-OP_9-drive-FV 9.car but NEG-SP_1-COM-$PRON_9$ 9.licence
'Mr Mutenda can drive a car, but he doesn't have a licence.' (constructed)

(62) **Ngoku umnumzana uMutenda a-nga-kwazi /**
now 1.mister 1A.Mutenda SP_1-POSSIB-POSSIB /
u-ya-kwazi uku-qhub-a ngokusemthethweni.
SP_1-DJ-POSSIB INF-drive-FV legally
'Mr Mutenda is now able to drive legally' (constructed; **angakwazi** preferred to **uyakwazi**)

A systematic exploration of all combinatorial possibilities for possibility markers and associated senses is beyond the scope of this article, but these examples suggest that further study will be fruitful.

Conclusions

This two-part study of expressions of possibility and necessity in Xhosa sheds light on several significant features of Xhosa modality. The corpus data and the third author's intuitions largely confirm the semantic paths laid out in van der Auwera & Plungian (1998) shown in (63), which is repeated from Crane *et al*. (2024: 45).

(63) **[participant-internal] > [participant-external (> deontic)] > [epistemic]**

More structurally grammaticalised modal markers have also travelled further down the semantic pathways, illustrating the connections between grammaticalisation and subjectification. For example, inflectional possibility forms **noku-** and **nga-** are used more with subjective modal functions like epistemicity, and with postmodal functions, whereas the more lexical **kwazi** is used mostly with more straightforward participant-internal meanings.

However, the semantic and diachronic relationship between deontic and participant-external possibility in Xhosa is less clear than predicted in van der Auwera & Plungian (1998), where deontic possibility is considered a subdomain of participant-external possibility, and, as we understand the model, participant-internal modals are expected to develop general participant-external uses before developing more specific deontic functions. Deontic possibility senses of **kwazi** appear

to be more felicitous than (teleological) participant-external senses, with most participant-external uses of **kwazi** in the corpus having non-main-clause structure, most frequently associated with purpose ('so that') clauses in positive-polarity contexts. That is, the use of **kwazi** in teleological participant-external contexts shows signs of being at least partially more restricted than its use in participant-internal contexts. Purpose clauses seem to be a bridging context into the "participant-external", circumstantial domain, independent of deontic uses. Similar patterns are noted in Crane *et al*. (2025) for other South African languages, as well. More data are needed to explore these patterns further, and to understand the role of subjectification with **kwazi** in the deontic and participant-external domains. For example, are non-subjective deontic uses of **kwazi** that relate to externally imposed rules more felicitous than those related to the speaker's subjective assessment of what is allowable? Or does the strong ability reading associated with **kwazi** make it ambiguous and therefore dispreferred in main-clause contexts? Negation also makes a difference in **kwazi**'s flavour compatibility (similarly to English *can*) and its interactions with negation should be studied more thoroughly.

The study also confirms the conceptual separation of participant-internal modality from participant-external 'circumstantial' modality. Participant-internal modality is a meaningful category in Xhosa, as shown by the prevalence of **kwazi** in straightforward ability contexts in the corpus, and – as described in the first part of this study (Crane *et al*. 2024) – by the third author's preference for **funeka** participant-internal necessity contexts. The other possibility markers occurred in some corpus examples we judged as primarily participant-internal, but these did not describe straightforward abilities. We did not find any straightforwardly participant-internal necessity examples in the corpus, but the third author finds **fanele** infelicitous in such contexts. The participant-internal/participant-external distinction is not made in some theoretical literature on modality (e.g. Kratzer 1991), but Xhosa and other Bantu languages provide evidence for its importance and distinctiveness: see, for example, Gluckman & Bowler (2020) on Logoori (JE41, Kenya and Tanzania), Bernander *et al*. (2022) on other Eastern Bantu languages, and Crane *et al*. (2025) on a selection of Southern Bantu languages, including Xhosa.

Within the category of participant-internal modality, we coded for physical vs. mental and learned vs. inherent abilities, but we did not find any relevant distinctions in Xhosa, although these sub-categories do seem to play a role in some related Nguni languages (Crane *et al*. 2025). Instead, an important distinguishing feature seems to be between prototypical abilities, mostly expressed with **kwazi**, and references to possibilities that may be technically associated with a participant's abilities (such as a doctor's ability to diagnose), but which serve a discourse function of expressing non-participant-internal possibility, for example situational or teleological ability (for example, a doctor's diagnosis as a means of getting proper treatment). These latter kinds of "participant-internal" meanings tend to be associated with the other markers, if those markers are used to express PI possibility at all.

Existential possibility may also be a meaningful category in Xhosa, as evidenced by existential senses in the corpus associated with **nakho**, **nga-**, and (especially) **noku-**. As noted in Section 1, different modal taxonomies treat this category in quite different ways; we used replacement with 'sometimes' as a diagnosing test.

Narrog (2012) notes a relationship between participant-internal and existential modality, both dealing with properties of entities (or classes), but connections with epistemic modality (predictions based on what is known) may also play a role. Our study was not able to determine with certainty the place of existential modality in processes of semantic change in Xhosa, since its use is felicitous with all possibility markers we examined. Therefore, the question remains: are existential meanings more like participant-internal possibility, but extended to generic subjects (or characteristics of subjects), or are they more like epistemic readings, but focused on knowledge of general trends rather than making a prediction about a specific situation? The corpus patterns, in which existential possibility meanings appear more frequently with more grammaticalised markers (and not at all with **kwazi**, in our coding), suggest that existential possibility uses correlate other more subjective modal functions and therefore might not be an early pathway of semantic change (i.e. participant-internal > existential). However, the third author's intuitions allow existential uses of **kwazi** and the corpus patterns may relate to the preference for animate subjects with **kwazi** rather than to its incompatibility with existential meaning. The third author's intuitions further suggest that epistemic and existential modality also need to be kept conceptually separate, as **kwazi** can express existential but not epistemic possibility.

As noted in Crane *et al*. (2024), we did not find a clear role for existential necessity in the modal expressions we examined, although **fanele**, which can express a greater range of modal meanings than **funeka**, did have a few occurrences with something like existential necessity as one of several senses.

The corpus data reflect other differences in flavour distribution between possibility and necessity, as well. Not surprisingly, participant-internal possibility is a far more salient category than participant-internal necessity. Our data also suggested a greater role for deontic necessity (requirements) than deontic possibility (permission).

The persistence of lexical meaning (and the meaning of constructions such as comitative plus infinitive) is evident throughout our corpus results, with the possible exception of **nga-**. In addition to the prevalence of senses associated with original lexical meanings, patterns related to subject animacy were also observed in the realm of possibility, with more lexical markers of possibility (**kwazi** and **nakho**) being used more frequently with animate subjects, suggesting persistence of original lexical 'know' as indicating mental ability and of the comitative as coding possession and hence subject control.

Our study highlights the usefulness of multi-pronged studies of semantic topics. The corpus data, even though our corpus was relatively small, shed significant light on the usage patterns of the markers we examined. The speaker intuitions of the third author allowed us to supplement this data with information about the functional boundaries of each expression, even for less common uses that did not appear in the corpus data. Published general grammatical descriptions and other documentary material such as the *Greater Dictionary of (isi)Xhosa* (Pahl *et al*. 1989; Mini *et al*. 2003; Tshabe & Shoba 2006) helped us to determine which markers to explore in the first place.

In addition to the issues highlighted above, further targeted study should investigate the modal contributions of non-dedicated verbal inflectional and derivational patterns (e.g. subjunctive forms and "neuter" suffixes), ideally based on both elicitation and the use of a fully parsed corpus. Similarly, the semantics and functional limits of more-specific modal expressions such as **vunyelwa** 'be allowed to' should be explored. Functionally, the categories of existential possibility and necessity, and their place in Xhosa and other Southern Bantu modal systems, require a closer look. Most significantly, more extensive corpus studies are called for. Ideally, these would involve a larger dataset for more statistical precision, including greater attention to information structural issues and the inclusion of spoken data in addition to written sources.

Abbreviations and glosses

1A	class 1a
2SG	2nd person singular (etc.)
3	class 3 (etc.)
APPL	applicative
COM	comitative
COMP	complementiser
CONN	connective
COP	copula
DEM	demonstrative
DJ	disjoint
EPEN	epenthetic form
EXIS	existential (non-verbal copula)
FV	final vowel
NF	infinitive
IPFV	imperfective
LOC	locative
MED	medial (distance)
NEC	modal necessity marker
NEG	negative
OP_9	object prefix (class 9, etc.)
PE	participant-external
PERS	persistive
PFV	perfective
PI	participant-internal
PL	plural
POSS	possessive
POSSIB	modal possibility marker
PP	possessive prefix
PRON	pronoun
PST	(near) past
REFL	reflexive
REL	relative
REMPST	remote past

SBJV	subjunctive
SG	singular
SIT	situative
SP_1	subject prefix (class 1, etc.)
SUBORD	subordinate form
SUPERL	superlative
TEMP	temporal marker

Acknowledgements

We would like to thank the editors and staff of *Africana Linguistica* for their wonderful help throughout the publication process. We also extend heartfelt thanks to two anonymous reviewers for their detailed and thoughtful comments, which greatly improved this paper. Rasmus Bernander's contribution was funded by a Riksbankens Jubileumsfond grant (grant number P23-0101), here gratefully acknowledged. Thera Crane's contribution was funded by the Finnish Research Council.

References

Andrason, A. 2019. 'The conjunctive coordinator NA in Xhosa. Its categorial status and map of polyfunctionality'. *Italian Journal of Linguistics* 31(1): 3-60.

Bastin, Y., Coupez, A., Mumba, É. & Schadeberg, T. (eds). 2002. *Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3*. Tervuren: Royal Museum for Central Africa, online database. http://www.africamuseum.be/collections/browsecollections/humansciences/blr

Bernander, R. Devos, M. & Gibson, H. 2022. 'Modal auxiliary verb constructions in East African Bantu languages'. *Journal of African Languages and Literatures* 3: 22-85.

Bernander, R. 2025. 'Epistemic possibility adverbs in East African Bantu - a typological-comparative overview'. *Journal of African Languages & Linguistics* 46(1): 71-127.

Bloom Ström, E.-M. 2020. 'The existential copula in Xhosa in relation to indefiniteness'. *Studies in African Linguistics* 49(2): 213-240.

Bloom Ström, E.-M. & Miestamo, M. 2024. 'The use of the augment in Nguni languages - a marker of referentiality?'. In E.-M. Bloom Ström, H. Gibson, R. Guérois & L. Marten (eds), *Approaches to morphosyntactic variation in Bantu*. Oxford: Oxford University Press, pp. 215-242.

Bybee, J. & Pagliuca, W. 1985. 'Cross-linguistic comparison and the development of grammatical meaning'. In J. Fisiak (ed.), *Historical semantics, historical word formation*. Berlin: Mouton, pp. 59-83.

Bybee, J., Perkins, R. & Pagliuca, W. 1994. *Evolution of grammar: Tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Craig, C.G. 1991. 'Ways to go in Rama: A case study in polygrammaticalization'. In E.C. Traugott & B. Heine (eds), *Approaches to grammaticalization, vol. 2: Types of grammatical markers*. Amsterdam: John Benjamins, pp. 455-492.

Crane, T.M., Mabena, M.P., Lubambo, R., Nkwinika, C., Sibisi, M. & Slater. O. 2025. 'What can be said? Variation among expressions of modal possibility in a South African language cluster'. *Studies in Language* 49(1): 44-92. DOI: https://doi.org/10.1075/sl.22074.cra

Crane, T.M., Savić, S., Slater, O. & Bernander, R. 2024. 'Modal expressions in Xhosa, part I: Necessity'. *Africana Linguistica* 30: 39-70.

Creissels, D. 2017. 'Similarity, suitability, and non-epistemic modalities (volitionality, ability, and obligation)'. In Y. Treis & M. Vanhove (eds), *Similative and equative constructions: A cross-linguistic perspective*. John Benjamins Publishing Compan, pp. 79-89.

Creissels, D. 2024. 'The circumstantial forms of Tswana verbs and the notion of participle'. *Studies in African Linguistics* 52(1): 25-34.

Devos, M. 2008. 'The expression of modality in Shangaci'. *Africana Linguistica* 14: 3-37.

Diessel, H. 2001. 'The ordering distribution of main and adverbial clauses: A typological study'. *Language* 77(3): 433-455.

Downing, L & Marten, L. 2019. 'Clausal morphosyntax and information structure.' In M. Van de Velde, K. Bostoen, D. Nurse & G. Philippson (eds), *The Bantu Languages [2nd ed]*. London: Routledge, pp. 270-307.

Einhorn, E. & Siyengo, L. 1993. *Xhosa: A Concise Manual*. Cape Town: College of Careers.

Eiselen, E.R. & Puttkammer, M.J. 2014. 'Developing text resources for ten South African languages'. *Proceedings of the 9th International Conference on Language Resources and Evaluation*, Reykjavik, Iceland: 3698-3703.

Gluckman, J. & Bowler, M. 2020. 'The expression of modality in Logoori'. *Journal of African Languages and Linguistics* 41 (2): 195-238. DOI: https://doi.org/10.1515/jall-2020-2010

Guthrie, M. 1970. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. *Vol. 4.* Farnborough: Gregg Press.

Kratzer, A. 1991. 'Modality'. In A. von Stechow & D. Wunderlich (eds), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*. Berlin: De Gruyter, pp. 639-650. DOI: https://doi.org/10.1515/9783110126969.7.639.

McLaren, J. 1975. *A new concise Xhosa-English dictionary.* 4th ed. Cape Town: Longman.

Mini, B.M., Tshabe, S.L., Shoba, F.M. & van der Westhuizen, P.N. 2003. *The Greater Dictionary of isiXhosa. Vol 2: K to P*. Alice: University of Fort Hare.

Narrog, H. 2012. *Modality, Subjectivity, and Semantic Change: A Cross-Linguistic Perspective.* Oxford: Oxford University Press.

Nauze, F. 2008. *Modality in Typological Perspective*. Institute for Logic, Language and Computation, Amsterdam: Universiteit van Amsterdam.

Nurse, D. 2008. *Tense and Aspect in Bantu*. Oxford: Oxford University Press.

Nurse, D. & Devos, M. 2019. 'Aspect, tense and mood'. In M. Van de Velde, K. Bostoen, D. Nurse & G. Philippson (eds), *The Bantu Languages [2nd ed]*. London: Routledge, pp. 204-236.

Nurse, D. & Philippson, G. 2006. 'Common tense-aspect markers in Bantu'. *Journal of African Languages and Linguistics* 27: 153-194.
Oosthuysen, J. 2016. *The grammar of isiXhosa.* Stellenbosch: SUN Press.
Pahl, H.W., Pienaar, A.M. & Ndungane, T.A. 1989. *The greater dictionary of Xhosa*. Vol 3: Q to Z. Alice: University of Fort Hare.
Portner, P. 2009. *Modality.* Oxford University Press.
Roux, J.C., Louw, P.H. & Botha, E.C. 2001. 'African Speech Technology telephone speech databases'. *Proceedings of the 12th Annual Symposium of the Pattern Recognition Association of South Africa* (PRASA), University of Pretoria, pp. 115-119.
Savić, S. 2020. 'Tense and aspect in Xhosa'. PhD dissertation, Rhodes University, Makhanda.
Snyman, D., van Huyssteen, G.B. & Daelemans, W. 2012. 'Cross-lingual genre classification for closely related languages'. *Proceedings of the 22nd Annual Symposium of the Pattern Recognition Association of South Africa*, 132-137.
Tshabe, S.L. & Shoba, F.M. 2006. *The greater dictionary of isiXhosa*. Vol 1: A to J. Alice: University of Fort Hare.
Tomozawa, H. 2009. 'The instantiative function of can and existential modality'. *English Linguistics* 26(1): 150-159.
van de Velde, M. 2019. 'Nominal morphology and syntax'. In M. Van de Velde, K. Bostoen, D. Nurse & G. Philippson (eds), *The Bantu Languages*, 2nd edition. London: Routledge, pp. 237-269.
van der Auwera, J. & Plungian, V.A. 1998. 'On modality's semantic map'. *Linguistic Typology* 2: 79-124.
Veselovsky, A., Rubinstein, A., Witzlack-Makarevich, A. & Namyalo, S. 2023. 'The subjunctive in Ruruuli-Lunyala (Bantu): A corpus-based study'. Presentation at Workshop on Modality in Underdescribed Languages: Methods & Insights. Humboldt-Universität zu Berlin, 20-22 April 2023.

Authors'addresses

Thera Marie Crane
Department of Languages
University of Helsinki, Finland
Department of African Languages
University of Pretoria, South Africa
thera.crane@helsinki.fi

Stefan Savić
Deutsches Seminar
University of Zurich, Switzerland
stefan.savic@ds.uzh.ch

Onelisa Slater
Department of Linguistics and Applied Language Studies
Rhodes University, South Africa
onelisaslater@outlook.com

Rasmus Bernander
Department of Languages and Literatures
University of Gothenburg, Sweden
rasmus.bernander@sprak.gu.se

Résumé

Cette étude porte sur les expressions modales de possibilité en xhosa, une langue bantu d'Afrique du Sud. En combinant l'analyse des occurrences dans un petit corpus et les intuitions d'un des auteurs de l'article, qui est locuteur natif, nous décrivons les schémas et les restrictions dans la distribution de quatre marqueurs de possibilité : les (sortes d') auxiliaires **kwazi** et **nakho**, et les préfixes **noku-** et **nga-**. Les résultats confirment la généralisation selon laquelle une grammaticalisation plus poussée (par exemple, auxiliaire > affixe) correspond à des significations plus subjectives. Les données de notre corpus suggèrent en outre l'importance de la catégorie de possibilité existentielle en xhosa et sa place dans les chaines du changement sémantique. Les lexiques sources et les sens constructionnels sont également importants dans les schémas d'utilisation des expressions modales en xhosa. Enfin, nos résultats suggèrent que l'expansion sémantique des fonctions de possibilité internes aux participants (capacité) vers des fonctions externes aux participants (circonstancielles ou basées sur un objectif) peut se produire via des propositions de but et de conséquence.