

# On the use of idioms for testing focus\*

Jenneke van der Wal

Leiden University Centre for Linguistics

**Abstract:** Parts of idioms can be used to test whether focus is an inherent part of meaning of a construction, because focus triggers alternatives and parts of idioms are not able to generate alternatives in their idiomatic meaning. This is illustrated with examples from Bantu languages, specifically for two types of the conjoint/disjoint alternation (focus-based and constituency-based), focus on augmentless nouns, and the use of idioms in clefts. Two aspects should be taken into account when testing a language's focus strategies with idioms: first, whole idioms may still be focused but parts of idioms cannot, and second, parts of 'useless task' idioms are transparent and scalar, thus allowing alternatives and therefore these cannot be used reliably in testing focus constructions.

**Keywords:** focus, idioms, transparency, scalarity, information structure

## 1 Introduction

While information structure increasingly forms part of linguistic descriptions, and the field of information structure is rapidly growing, it remains difficult to establish the precise semantic-pragmatic interpretations of information structural strategies such as focus. Apart from the necessary spontaneous speech and longer stretches of text or discourse, elicitation is useful in this area to identify whether a given linguistic strategy inherently encodes focus, or carries a mere implication of focus. To help researchers in this area, the Questionnaire on Information Structure (QUIS, Skopeteas et al. 2006) and the QUISsem for semantics (Renans et al. 2011) were designed for systematic elicitation on topic, focus, and contrast, and I discuss various tests for focus in a paper 'Diagnosing focus' (2016). One of the tests discussed there involves idioms, and in this paper I want to expand on how the test works, illustrating from the Bantu languages, and share a new insight on the type of idioms that can and cannot be used in this test.

The logic of the idioms test is the following. A generally accepted definition of focus is that it triggers a set of alternatives (Rooth 1985, 1992, 1996; Krifka 2008), and asserts that the predicate is true for the referent mentioned. For example, focusing the object in 'I drank TEA' triggers a set of alternatives for 'tea': {coffee, lemonade, gin, ...} and asserts that out of those alternatives, my drinking is true for tea. Therefore, it should be impossible to focus referents that cannot trigger alternatives. Crucially, parts of idioms are a class of such unfocusable items (see e.g. Cruse 2000: 73). Parts of idioms cannot be focused because no alternatives are available for parts of idioms in their idiomatic reading. Instead, when alternatives are generated, these will only refer to the literal meaning and not the idiomatic one. This formalises Fraser's (1970) reasoning for the ungrammaticality of idioms in English clefts:

---

\* I wish to thank my collaborators Allen Asimwe, Patrick Kanampiu, Amani Lusekelo, and Aurelio Simango, the speakers of Rukiga, Kĩtharaka, Kinyakusa, and Changana who helped us gain insights into their languages, the participants of the idiom workshop in Potsdam November 2015, and the comments of two anonymous reviewers as well as Lisa Matthewson. Any errors remain my responsibility. This research is supported by NWO Vidi Grant 276-78-001 through the BaSIS 'Bantu Syntax and Information Structure' project.

“Implicit [...] is the assumption that no part of the idiom has retained any literal interpretation, if it ever had any in this construction. That is, we maintain that no part of the idiom actually contributes to the semantic interpretation of the expression, once the idiom has been formed. It follows from this assumption that to [cleft] one part of the idiom, in this instance a noun phrase, is to impute to the noun phrase some semantic integrity which it does not have”. (Fraser 1970: 33)

To illustrate, consider the English idiom ‘to beat around the bush’, meaning ‘to avoid talking about something unpleasant’. Focusing one part in a cleft results in the loss of the idiomatic meaning:

- (1) a. John beat around the bush = John avoided talking about something unpleasant  
 b. It was the bush that John beat around ≠ It was something unpleasant that John avoided  
 (Fraser 1970: 32)

Focusing *the bush*, like in the cleft in (1b), if it makes sense at all, triggers alternative items he may have beaten around instead. This prevents the object bush from being interpreted as part of the idiom together with the verb.

Considering these properties, we can use idioms as tests for focus constructions. If a linguistic strategy that is suspected to encode focus can felicitously be used with an idiom (i.e. retaining its idiomatic interpretation), it shows that the strategy is not a dedicated marker of focus on the affected phrase. Instead, it may be that such a marked construction is not a dedicated focus strategy at all, or that focus is on a larger constituent with alternatives for the whole idiom.

Fanselow and Lenertová (2011) use this test to investigate whether so-called ‘subpart of focus fronting’ in German and Czech is indeed a focus construction. It is said that the object can be fronted when it is itself in narrow focus (2a), but also when it is part of the focus, in VP focus and sentence focus (2b).<sup>1</sup>

- (2) a. What did you see there?  
 [Eine LaWIne] haben wir gesehen!  
 a avalanche have.1PL 1PL.PRO see.PST.PTCP  
 ‘An avalanche have we seen!’  
 b. What happened?/What’s new?  
 [Einen HAsen] habe ich gefangen.  
 a.ACC rabbit.ACC have.1SG 1SG.PRO catch.PST.PTCP  
 ‘I caught a rabbit.’ German (Fanselow and Lenertová 2011: 172, 174)

The question is whether such fronting is inherently linked to focus. If it were, we would expect the idiomatic reading to disappear when fronting part of an idiom. In contrast, if the idiomatic reading is still present, we can conclude that fronting is not inherently linked to focus. Considering that the idiomatic interpretation is accepted in (3), Fanselow and Lenertová conclude that ‘subpart of focus

<sup>1</sup> Numbers refer to noun classes, or to persons when followed by SG or PL. High tones are indicated by an acute accent; low tones are unmarked. All-caps indicates stress. I follow the Leipzig Glossing Rules and use following additional abbreviations: AUG ‘augment’, CJ ‘conjoint verb form’, DJ ‘disjoint verb form’, EXH ‘exhaustive’, FV ‘final vowel’, IP ‘intensifying particle’, lit. ‘literal meaning’, N.PST ‘near past’, OM ‘object marker’, PERS ‘persistent’, PRO ‘(strong) pronoun’, SM ‘subject marker’, T ‘tense’, YPST ‘yesterday past’.

fronting’ is not a focus construction and that movement to the initial position in German cannot be motivated by focus features.

- (3) [Den GARaus] hat er ihm gemacht.  
the.ACC *garaus* have.3SG 3SG.M.PRO 3SG.M.DAT make.PST.PTCP  
‘He killed him.’ German (Fanselow and Lenertová 2011: 176, adapted)

In this paper, I want to further illustrate how idioms can be used as a focus diagnostic. In sections 2 and 3, I discuss two focus marking strategies found in the Bantu languages, showing how idioms can be used to diagnose whether focus is indeed an inherent aspect of the strategy. In section 4, it is shown that unacceptability judgements of idioms in clefts may not tell us anything about the syntactic properties of the construction (movement, reconstruction) but instead be due to the incompatibility of parts of idioms with focus. Finally, section 5 shows a special subclass of idioms referring to ‘useless tasks’ which do allow focusing – I propose that such idioms are transparent and have a scalar aspect of meaning which allows the generation of alternatives and hence focusing.

Unless otherwise indicated, the data in this paper come from databases in the BaSIS project. The data have been gathered in collaborative fieldwork with Allen Asiimwe for Rukiga, with Patrick Kanampiu for Kîtharaka, with Aurélio Simango for Changaná, with Ernest Nshemezimana for Kirundi, and with Amani Lusekelo and Simon Msovela for Kinyakyusa, whose co-authorship of the data is indicated for the relevant examples. The databases are stored as Online Language Databases accessed via Dative (<https://www.dative.ca/>) and will be publicly accessible later. We followed the methodology of the BaSIS project (available via <https://bantusyntaxinformationstructure.com/methodology/>), including both spontaneous and elicited data, though the idiom judgements used here were all elicited. Finding idioms in a given language is generally quite hard; asking for euphemisms and taboo areas (death, defecation, certain diseases) can form an incentive, as well as feelings and emotions which are often expressed as associated with body parts, in languages in general (Wierzbicka 1999) as well as in Bantu languages (e.g. in Makhuwa ‘the heart is white’ for happiness, ‘the heart walks’ for being nauseous, or ‘the head is hard’ for stupidity).

## 2 Conjoint/disjoint typology

In this section, I show how idioms have been used to help distinguish two subtypes of one phenomenon associated with focus: the conjoint/disjoint alternation.

A specific trait of some eastern and southern Bantu languages is the pairing of conjugational categories called ‘conjoint’ (CJ) and ‘disjoint’ (DJ). These verb forms encode the same tense/aspect semantics, but differ in their relation with what follows the verb. The Makhuwa example in (4) shows three characteristics of the alternation (see Van der Wal 2017 for further general discussion on the alternation, and Van der Wal 2009, 2011 for in-depth discussion of Makhuwa):

1. the morphology of the inflected verb differs for the CJ and DJ form;
2. the CJ verb form cannot appear sentence-finally;
3. and there is a difference in interpretation of the element following the verb.<sup>2</sup>

<sup>2</sup> Often the conjoint form does not have a specific morphological marker. Rather than assuming and glossing a null morpheme as ‘conjoint’, I have indicated the form of the verb in front of the example, as CJ or DJ.

- (4) a. CJ Nthíyáná o-c-aalé \*(nramá).  
 1.woman 1SM-eat-PFV.CJ 3.rice  
 ‘The woman ate *rice*.’
- b. DJ Nthíyáná o-hoó-cá (nráma).  
 1.woman 1SM-PFV.DJ-eat 3.rice  
 ‘The woman ate (rice).’

Makhuwa (P31, Van der Wal 2011: 1735)<sup>3</sup>

For Makhuwa, I have shown that the choice between the CJ or DJ verb form is dependent on focus, specifically that the CJ verb form expresses exclusive focus on the element directly following the verb (Van der Wal 2011) – in (4a) the object ‘rice’ is focused, to the exclusion of other possible alternatives that may have been eaten.<sup>4</sup> However, for other languages it has been observed that the CJ and DJ verb form “may in certain contexts differ in meaning but for the most part are used in different syntactical positions” (Ziervogel and Mabuza 1976:174 on Swati). Van der Spuy (1993) and others after him analyse the CJ/DJ alternation in Zulu as determined purely by *constituency*: when the verb is final in the vP constituent, it takes the DJ form; when it is not final, i.e. when some element follows within the vP, the verb takes a CJ form. Buell (2005) shows this in a range of circumstances, one of which is illustrated in (5): despite the fact that the adverb *kahle* ‘well’ in the second clause is not in focus (as can be deduced from the preceding clause), the verb *ngicula* ‘I sing’ takes the CJ form. Following Van der Spuy, Buell argues that this is because *kahle* is phrased in the same constituent as the verb, and the verb form in Zulu is determined purely by whether the verb is constituent-final or not.

- (5) CJ A-ngi-dans-i kahle, kodwa ngi-cul-a kahle.  
 NEG-1SG.SM-dance-FV well but 1SG.SM-sing-FV well  
 ‘I don’t dance well, but I sing well.’ Zulu (S42, Buell 2005: 64, 66)

This suggests that there is crosslinguistic variation, with the alternation being determined either by focus or by constituency (Van der Wal 2017). Establishing which type of alternation a given language has can be tricky, but idioms can be used as a diagnostic to distinguish a focus-based alternation from a constituency-based one.

If the CJ form encodes focus on the following element, idioms involving an object should not retain their idiomatic reading. In contrast, if only constituency plays a role, then the CJ form is expected to be licensed irrespective of the interpretation of what follows the verb. Therefore, in a language with a constituency-based alternation, the idiom should be acceptable with a CJ form. Based on this diagnostic, we can distinguish Matengo and Kirundi on the one hand, from Zulu and Changanana on the other hand. Matengo and Kirundi both require the DJ form for an idiomatic reading and do not allow the CJ form, as shown in (6) and (7) respectively (see also Bostoen and Nshemezimana (2017: 410)). This shows that the alternation is focused-based in these languages.

<sup>3</sup> Bantu languages have been classified by Guthrie (1948) into geographical areas; I provide the Guthrie number for each language according to the update by Maho (2009).

<sup>4</sup> Theoretically, one can distinguish between exclusive focus, in which some alternatives must be excluded, from exhaustive focus, in which all alternatives must be excluded. A broader array of focus tests can distinguish the two. Idioms merely test for focus tout court.

- (6) a. CJ Ju-a-som-aje      mwikindamba.  
           1SM-PST-read-CJ 18.7.hut  
           × ‘He didn’t have a formal education.’  
           ✓ ‘He studied in a HUT/STUDIED IN A HUT.’
- b. DJ Ju-a-som-iti      mwikindamba.  
           1SM-PST-read-PF 18.7.hut  
           ‘He didn’t have formal education.’  
           lit. ‘He studied in a hut.’<sup>5</sup>

Matengo (N13, Yoneda 2017: 437)

- (7) a. CJ Dufashe            imbwá      amabóko.  
           tu-fát-iyé            i-m-bwá    a-ma-bóko  
           1PL.SM-hold-PFV AUG-9-dog AUG-6-arm  
           × ‘We really landed ourselves in a nasty situation.’  
           ✓ ‘We hold the dog by the arms.’
- b. DJ Turafáshe            imbwá      amabóko.  
           tu-ra-fát-iyé            i-m-bwá    a-ma-bóko.  
           1PL.SM-DJ-hold-PFV AUG-9-dog AUG-6-arm  
           ‘We really landed ourselves in a nasty situation.’  
           lit. ‘We hold the dog by the arms.’

Kirundi (JD62, data Nshemezimana and Van der Wal)

On the other hand, Zulu and Changana happily use a CJ form with an idiomatic reading, as illustrated in (8) and (9), respectively. This shows that the conjoint verb form in these languages does not encode focus on the postverbal element, and hence we can diagnose the alternations in these languages as constituency-based.

- (8) CJ Ilanga li-kipha      umkhovu e-tshe-ni.  
           5.sun 5SM-extract.CJ 1.zombie LOC-9.stone-LOC  
           ‘It’s really hot.’  
           lit. ‘The sun brings the zombie out of the stone’

Zulu (S42, Claire Halpert p.c.)

- (9) a. CJ Ko:ndlo rí-phukw-é      ncê:le.  
           5.rat 5SM-fail-PFV.CJ 9.hole  
           ‘The thief was caught.’  
           lit. ‘The rat failed (to reach) the hole.’
- b. DJ Ko:ndlo rí-phukw-ílé      nce:le.  
           5.rat 5SM-fail-PFV.DJ 9.hole  
           ‘The thief was caught.’  
           lit. ‘The rat failed (to reach) the hole.’

Changana (S53, data Simango and Van der Wal, unclear what the interpretational difference is)

<sup>5</sup> When indicating the literal meaning, I am assuming this is also a possible reading.

For the conjoint/disjoint alternation, the unfocusability of parts of idioms has thus helped to identify the defining factor for the choice of verb form: either postverbal focus or the constituent-final position of the verb. Together with the other diagnostics used in Van der Wal (2017), this leads to new insights for the individual Bantu languages with respect to how deeply information structure is anchored in the grammar – focus features can be claimed to play a direct role in the focus-based alternation, but the possible information-structural effects in the constituency-based alternation can only be secondary.

### 3 Augmentless nouns

Idioms can also help identify the circumstances under which the so-called ‘augment’ (to be explained just below) can be omitted in Luganda. Furthermore, when tested with the CV augment in Kinyakyusa, it is explicitly shown that the test only works with *parts* of idioms (not idioms as a whole).

In Luganda, nouns can appear with or without an augment – an initial vowel preceding the noun class prefix:

(10)	a-ba-ana	__-ba-ana	o-mu-sajja	__-mu-sajja
	AUG-2-children	2-children	AUG-1-man	1-man

Luganda (JE15, Saudah Namyalo p.c.)

Luganda is one of the languages that is ‘default augmented’, meaning that an accurate account of the presence/absence of the augment is best captured as listing the environments in which the augment is absent. Hyman and Katamba (1993) show that there are two factors licensing the absence of the augment: negation and focus. Following a negative verb, the augment is always absent, but after an affirmative verb, there is a choice to be made. Hyman and Katamba (1993) and Van der Wal and Namyalo (2016) show that nouns that lack the augment are in exclusive focus in an affirmative clause.

A second focus strategy in Luganda is the immediate after verb (IAV) focus position. As shown in (11), inherently focused question words such as *ki* ‘what’ can only appear postverbally in the verb-adjacent position (see among others Watters 1979, Buell 2009, Van der Wal 2009, on the IAV focus position; Hyman and Katamba 1993 and Van der Wal and Namyalo 2016 for IAV in Luganda).<sup>6</sup>

- (11) a. O-mu-sómésa y-a-w-á                      á-b-áana      e-m-mére.  
           AUG-1-teacher 1SM-PST-give-FV AUG-2-child AUG-9-food  
           ‘The teacher gave the children food.’
- b. O-mu-sómésa y-a-w-á                      kí      á-b-áana?  
           AUG-1-teacher 1SM-PST-give-FV what AUG-2-child  
           ‘The teacher gave the children food.’
- c. \*O-mu-sómésa y-a-w-á                      a-b-áana      kí?  
           AUG-1-teacher 1SM-PST-give-FV AUG-2-child what  
           int. ‘What did the teacher give the children?’

Luganda (JE15, Van der Wal and Namyalo 2016: 357)

<sup>6</sup> Luganda does not have the conjoint/disjoint alternation.

The idiom diagnostic now shows that the two focus strategies differ: part of idioms may freely appear in the IAV position, but it is impossible to omit the augment on the object-part of an idiom, as in (12b).

- (12) a. E-my-aaka gy-aa-li gi-mu-wubidde a-ka-taambaala.  
 AUG-4-years 4SM-PST-be 4SM-1OM-wave.PFV AUG-12-handkerchief  
 ‘He is very old.’  
 lit. ‘The years waved a handkerchief at him.’
- b. #E-my-aaka gy-aa-li gi-mu-wubidde –-kataambaala.  
 AUG-4-years 4SM-PST-be 4SM-1OM-wave.PFV AUG-12-handkerchief  
 × ‘He is very old.’  
 ✓ ‘The years waved a handkerchief at him.’ (non-sensical)  
 Luganda (JE15, data Namyalo and Van der Wal)

We deduce that omission of the augment is indeed a dedicated focus strategy, but the postverbal position, if it encodes focus at all, must show a rather underspecified or broad type of focus – indeed SVO order can also express VP focus.

VP focus in fact brings out another aspect to take into account when using idioms for testing focus. While it is true that parts of idioms cannot trigger alternatives in their idiomatic meaning, there is no restriction on the whole idiom being in focus. This became explicit when we tested idioms with the so-called ‘CV augment’ in Kinyakyusa. In Kinyakyusa, nouns can be preceded by a V augment, as in (13), or what has been called a CV augment, as in (14). Lusekelo and Van der Wal (2021) argue that the CV augment is in fact not an augment, but a marker of exhaustive focus, as indicated in the glossing and translations of (13) and (14) (see Lusekelo and Van der Wal 2021 for a range of diagnostics showing exhaustivity).

- (13) a. **u**-mu-ndu  
 AUG-1-person  
 ‘a/the person’
- b. **a**-ma-bifu  
 AUG-6-banana  
 ‘(the) bananas’
- (14) a. **ju**-mu-ndu  
 EXH-1-person  
 ‘only a person’
- b. **ga**-ma-bifu  
 EXH-6-banana  
 ‘only bananas’
- Kinyakyusa (M31, data Lusekelo, Msovela, Van der Wal)

Given the exhaustive focus interpretation, we would predict idioms to not be able to take the CV marker. The acceptance of the idiom in (15) is therefore unexpected at first sight, but if we look at the context, we see that the example is only possible if interpreted with the contrast/exclusion on the level of the VP. That is, the set of alternatives is generated for the meaning of the whole idiom, in this case other activities that he could do instead of drinking in the bar, or even other VP idioms.

- (15) a. ku-koma a-ma-isi  
 15-hit AUG-6-water  
 ‘to get drunk’  
 lit. ‘to hit water’
- b. As soon as he gets up in the morning, he drinks, and straight from work he goes to the bar.  
 I-ku-kom-a ga-ma-isi.  
 1SM-PRS-hit-FV EXH-6-water  
 ‘He is only getting drunk.’  
 lit. ‘He is only hitting water.’

Kinyakyusa (M31, data Lusekelo, Msovela, Van der Wal)

This stresses an important methodological point: the diagnostic power is limited to *parts of* idioms in focus, and as soon as the whole idiom is in the scope of focus, alternatives *are* available. When applying the idiom test for focus we should therefore always take into account the context (as is generally the case, especially for information-structural data).

#### 4 Clefts

As an additional observation, I want to point out how the effect of focus on idioms should be taken into account when using idioms in syntactic tests. Idioms are a typical test used to establish whether relativised and clefted constituents have been moved from a lower position in which they were generated (e.g. as the complement of the verb), or instead been generated in the high position (no movement); see a summary of the debate about relative clauses in Webelhuth et al. (2018). The retention of an idiomatic reading has been used as an argument in favour of a movement analysis of cleft constructions and relative clauses. The reasoning is as follows: if idioms are stored as phrases,<sup>7</sup> the idiom pieces should be lexically inserted together to generate the idiomatic reading.

However, the infelicity of an idiom in a cleft does not necessarily mean that it is base-generated (no movement): if focus is inherent to a cleft, the incompatibility with the unfocusable part of an idiom will also lead to rejection of the sentence.

For Lubukusu, Wasike (2007) shows that it is not possible to focus part of an idiom in a cleft:

- (16) a. Nanjekho a-a-ara chi-njekho.  
 1.Nanjekho 1SM-PST-break 10-laughter  
 ‘Nanjekho laughed loudly.’  
 lit. ‘Nanjekho broke laughter.’
- b. #Chi-li chi-njekho ni-cho Nanjekho a-a-ara.  
 10-be 10-laughter REL-10 1.Nanjekho 1SM-PST-break  
 × ‘Nanjekho laughed loudly.’  
 ✓ ‘It is laughter that Nanjekho broke.’ (non-sensical)

<sup>7</sup> This is a point of debate, and it is generally accepted that not all idiomatic expressions are stored as wholes, but a larger or smaller part of the expression forming an ‘idiomatically combining expression’ (Wasow et al. 1983, Nunberg et al. 1994) with separately stored lexemes.





## 5 How useless tasks can be focused

The use of idioms as a diagnostic for focus depends on the inability of parts of idioms to create a set of alternatives in their idiomatic reading. There is, however, a class of idioms that apparently does allow focus, and that should therefore not be used as a diagnostic. I first illustrate this class of idioms and then reason through why they do allow focus.

New data from Kĩtharaka show a split into two types of idioms under focus. The first type loses its idiomatic meaning in a cleft, in line with the data we saw above, whereas the second type retains the idiomatic interpretation, with the cleft adding intensity (doing something *completely* useless). The former is illustrated in examples (19)-(21), and the latter in (22)-(24).<sup>89</sup>

- (19) a. N' á-á-tw-eere mû-tí.  
           FOC 1SM-PST-climb-PFV 3-tree  
           ‘She became pregnant.’  
           lit. ‘She climbed a tree.’
- b. Í mû-tí mw-aarí á-tw-eete. *cleft*  
           FOC 3-tree 1-girl 1SM-climb-STAT.PFV  
           × ‘The girl became pregnant.’  
           ✓ ‘It’s a tree that the girl climbed.’
- (20) a. Kĩ-thaká gĩ-kũ-nunk-a m-bwé.  
           7-bush 7SM-PRS-stink-FV 9-fox  
           ‘There is imminent trouble.’  
           lit. ‘The bush is stinking (of) fox.’
- b. I kĩ-thaka gĩ-kũ-nunk-a m-bwe. *cleft*  
           FOC 7-bush 7SM-PRS-stink-FV 9-fox  
           × ‘It is trouble that is imminent.’  
           ✓ ‘It is the bush that is stinking (of) fox.’
- c. I m-bwe kĩ-thaka gĩ-kũ-nunk-a. *cleft*  
           FOC 9-fox 7-bush 7SM-PRS-stink-FV  
           × ‘It is trouble that is imminent.’  
           ✓ ‘It is (of) the fox that the bush is stinking.’
- (21) a. Tĩri î-gũ-úk-a.  
           9.dust 9SM-PRS-rise-FV  
           ‘Things are finished.’  
           lit. ‘Dust has risen.’

<sup>8</sup> In Kĩtharaka, the particle *ni* glossed as FOC has various functions (see Muriungi 2005, and Abels and Muriungi 2008). Combined with the tense marker *-kũ-* it expresses the present progressive, as in (21a) and (22a), in a cleft it marks the focus, and its origin is still visible in its use as a copula. The particle appears as *i-* before a consonant, and *n-* before a vowel.

<sup>9</sup> The verb in (24) *kũina* refers to both singing and dancing (as traditionally the two go together).

- b. I tîri y-á-ûk-á. *cleft*  
 FOC 9.dust 9SM-PST-rise-FV  
 × ‘It is things that are finished.’  
 ✓ ‘It is dust that has risen.’
- (22) a. N' ú-gû-kím-a rû-úyî na n-tîrí.  
 FOC 2SG.SM-PRS-pound-FV 11-water with 9-mortar  
 ‘You are engaged in a useless task.’  
 lit. ‘You are pounding water with a mortar.’
- b. Í rû-úyí ú-gû-kím-a ná ñtîrí. *cleft*  
 FOC 11-water 2SG.SM-PRS-pound-FV with 9-mortar  
 ‘You are engaged in a (really) useless task.’  
 lit. ‘It is water that you are pounding with a mortar.’
- (23) a. kû-thaamb-i-a n-gúkú ma-gûrû  
 15-wash-IC-FV 9-chicken 6-legs  
 ‘to engage in a useless task’  
 lit. ‘to wash a/the chicken the legs’
- b. Ii n-gúkú û-kû-thaamb-i-a ma-gûrû. *cleft*  
 FOC 9-chicken 2SG.SM-PRS-wash-IC-FV 6-legs  
 ‘You're doing something (really) useless.’  
 lit. ‘It's a/the chicken you're washing the legs.’
- (24) a. N' ú-kw-ín-îr-a n-já.  
 FOC 2SG.SM-PRS-sing-APPL-FV 9-outside  
 ‘You're doing something useless.’  
 lit. ‘You are singing outside.’
- b. I n-já û-kw-ín-îr-a. *cleft*  
 FOC 9-outside 2SG.SM-PRS-sing-APPL-FV  
 ‘You're doing something (really) useless.’  
 lit. ‘It is outside that you are singing.’

Kîtharaka (E54), data Kanampiu and Van der Wal

As is immediately clear, the idioms of type 2 all mean ‘do something useless’. The fact that they retain the idiomatic reading can be understood if we look at the alternatives triggered by focus: the object ‘water’ in (22) can be placed on a scale of ‘usefulness of pounding’, with alternatives such as millet or maize ranking higher, and water being close to the useless end of the scale. Importantly, there are alternatives that can be considered for interpretation of the idiom. When these alternatives are present, the low value of usefulness of the given referent is highlighted, leading to the intensified interpretation. The arguments in the type 1 idioms, such as *mûti* ‘tree’ in (19), do not have viable alternatives linked to the idiomatic meaning, and therefore only retain the literal meaning when alternatives are required in a cleft.

There are two observations I want to discuss before the semantic analysis. The first observation is that pseudoclefts and reverse pseudoclefts resist the idiomatic interpretation across the board, as shown in (25), (26) and (27) below.

- (25) Kî-thaka i-ky-ó gí-kú-nunk-a m-bwe. *reverse pseudo*  
 7-bush FOC-7-PRO 7SM-PRS-stink-FV 9-fox  
 × ‘There is imminent trouble.’  
 ✓ ‘The bush is what is stinking (of) fox.’

- (26) N-gúkú n-îy-ó û-kû-thaamb-i-a má-gûrû. *reverse pseudo*  
 9-chicken FOC-9-PRO 2SG.SM-PRS-wash-IC-FV 6-legs  
 × ‘Something useless is what you did.’  
 ✓ ‘The chicken is what you washed the legs.’

- (27) Kû-rá û-ku-in-îr-a i njá. *pseudocleft*  
 17-DEM.DIST 2SG.SM-PRS-SING-APPL-FV COP 9-outside  
 × ‘Something useless is what you did.’  
 ✓ ‘Where you are singing is outside.’ Kîîtharaka (E54) data Kanampiu and Van der Wal

The possibilities for each idiom are summarised in Table 1. The unacceptability of (reverse) pseudoclefts can be understood if we consider their syntactic structure: (reverse) pseudoclefts are copular constructions, one part of which is a free relative. This means that the parts of the cleft are not generated as one chunk, and hence cannot receive their idiomatic interpretation. The reference of the free relative is calculated independently, creating a presupposition with only the literal meaning: for (27), there exists a place where you are singing (and this is identified as outside).

**Table 1:** Idioms in Kîîtharaka

idiom	meaning	cleft	pseudocleft
to climb a tree	‘to become pregnant’	×	×
the bush stinks of fox	‘trouble is imminent’	×	×
dust has risen	‘things are finished’	×	×
the handle became hot	‘things got out of control’	×	×
to pound water	‘useless task’	✓	×
to bathe the chicken the legs	‘useless task’	✓	×
to sing outside	‘useless task’	✓	×
to sieve water from melons	‘useless task’	✓	×
to plant hulled sorghum	‘useless task’	✓	×

A second interesting fact is that in Makhuwa, 1670 kilometers further south so not neighbouring Kîîtharaka, we find a similar split in idioms. Here, too, the idiom for ‘doing something useless’ retains the idiomatic reading under focus, whereas the idiom ‘to stretch the legs’ for ‘to die’ does not. As shown in section 2, the conjoint form in Makhuwa expresses focus on the element directly following the verb. In line with the expectations, the object in a regular idiom as in (28) loses its

idiomatic reading with the conjoint form, but the ‘useless task’ idiom in (29) does retain its idiomatic interpretation: since more useful alternatives to write on than water are easily available, the conjoint form is still accepted.

(28) a. DJ O-hoó-kóla mwétto.  
 1<sub>SM-PFV</sub>.DJ-stretch-FV 3.leg  
 ‘S/he died.’  
 lit. ‘S/he stretched the leg.’

b. CJ O-o-kol-alé mwettó.  
 1<sub>SM-PFV</sub>.DJ-stretch-PFV.CJ 3.leg  
 × ‘S/he died.’  
 ✓ ‘S/he stretched the leg.’

(29) A friend is asking you to lend them money, you say you can’t, they keep asking.

a. DJ W-oo-lép-á m-maátsí=ni vó!  
 2<sub>SG.SM-PFV</sub>.DJ-write-FV 18-6.water=LOC IP  
 ‘You are wasting your time!’  
 lit. ‘You have written on water!’

b. CJ O-lep-alé m-maátsí=ni vó!  
 2<sub>SG.SM</sub>-write-PFV.CJ 18-6.water=LOC IP  
 ‘You are wasting your time!’ (emphatically, annoyed)  
 lit. ‘You have written on water!’

Makhuwa (P31)

The fact that it is precisely the ‘useless task’ idiom that allows focusing suggests that the division between the two types of idioms is rooted in an inherent semantic property of these idioms, not a language-specific quirk.

So what might that property be? Nunberg (1978) and Nunberg et al. (1994) argue that idiomatic expressions differ in their semantic decomposability, which is by now an accepted standpoint (see also the overviews in Fellbaum 2011, 2014, for example). Three semantic properties that have been used to distinguish idiomatic expressions (Nunberg et al. 1994: 498):

- conventionality: how predictable is the idiomatic meaning based on the conventional meaning of the words in isolation in a particular language environment?
- opacity (or transparency): to what extent is the original motivation for use recoverable?
- compositionality: to what extent can the phrasal meaning, once known, be derived from the idiom’s parts?

We can clearly see how idioms like ‘climb a tree’ (for being pregnant) or ‘stretch the legs’ (for dying) are more conventional and less transparent than ‘pound water’ or ‘write on water’ (for doing something useless). Nevertheless, the latter are still non-compositional, as the idiomatic meaning does not follow from the meaning of the parts. Instead, the useless tasks fall into a category of metaphors (Horn 2003) or ‘quasi-metaphorical’ idioms (Glucksberg 2001). These “literally refer to situations, actions, or events that epitomize a class of situations, actions, or events” (Glucksberg

2001: 72). For example, washing the chicken’s legs is a prototypically useless action (as it will run around in the dirt immediately afterwards), and can thus metaphorically be used to refer to other useless actions: it is *as if* you are washing the chicken’s legs.

We may thus hypothesize that only non-compositional or non-transparent idioms reject focusing, and transparent or (quasi)metaphorical idioms allow it (see also experiments by Wierzba 2016, and Frey 2005:160-161, who notes that “a contrastive interpretation of a constituent in an idiom is only possible when in its idiomatic use it contributes an identifiable, compositionally processable part of the overall meaning of the idiom”). Nevertheless, idioms such as ‘to study in a hut’ (for not having a formal education) and ‘to break laughter’ (for laughing loudly) are quite transparent, and can even be said to be metaphorical (it is as if he’s studied in a hut) – yet they resist focus. I therefore propose that the ‘useless task’ idioms, in addition to being transparent, need to have a scalar aspect of meaning, in order to facilitate the generation of alternatives in the idiomatic interpretation. All useless task idioms combine an action with a referent that is at the low extreme of usefulness for the given action, making it easy to generate alternatives that are higher on the scale of usefulness. The presence of these alternatives makes it all the more clear that the asserted referent is at the lowest point of usefulness, hence deriving the intensive interpretation under focus. I conclude that ‘useless task’ idioms are transparent and scalar, that this enables the generation of alternatives, and hence that they are acceptable in focus constructions.

## 6 Summary and further research

In this short paper, I have shown how parts of idioms can be used as a test for focus, illustrating two types of the conjoint/disjoint alternation (focus-based and constituency-based), focus on augmentless nouns, and the use in clefts. Two aspects were pointed out that should be taken into account when testing a language’s focus strategies with idioms: first, whole idioms may still be focused but parts of idioms cannot, and second, parts of ‘useless task’ idioms can still be focused and hence these cannot be used reliably in testing focus constructions.

The analysis proposed for ‘useless task’ idioms predicts that other idioms involving scalarity may also allow focusing, and in fact Kĩtharaka shows one such idiom: to kill an elephant with a baobab fruit, meaning to do something very difficult in an easy way. The elephant being very big, it ranks low on the scale of likelihood of being killed by a relatively little baobab fruit (see Figure 1).<sup>10</sup> There are plenty of easier or more likely alternatives that can be generated, and indeed this expression retains its idiomatic meaning in a cleft, as shown in (30).<sup>11</sup>



Figure 1: Baobab fruit

- (30) a. N’ á-rá-urag-íré n-jógú na kî-ráám̄ba.  
 FOC 1SM-YPST-kill-PFV 9-elephant with 7-baobab.fruit  
 ‘S/he’s done something very difficult in an easy way.’  
 lit. ‘S/he killed an elephant with a baobab fruit.’

<sup>10</sup> Picture by Leonora Enking, via Flickr.

<sup>11</sup> For those who are wondering how one can kill an elephant in this way: you can block its trunk with the baobab fruit.

- b. I n-jógú a-rá-úrág-íre na kî-ráám-bá. *cleft*  
 FOC 9-elephant 1SM-YPST-kill-PFV with 7-baobab.fruit  
 ‘It’s an elephant s/he killed with a baobab fruit.’  
 ‘It’s something difficult s/he’s done in an easy way.’

Kîtharaka (E54), data Kanampiu and Van der Wal

More research is needed to establish the precise dependencies between the various semantic properties (conventionality, opacity, compositionality, and scalarity) of idioms and their behavior under focus. A test that would be particularly interesting is internal modification of the idiom, which is only possible for decomposable idioms (thanks to a reviewer for pointing this out; see Nunberg et al. 1994 and Ernst 1981). Unfortunately, the data to test this (‘climb a high tree’ or ‘washing the chicken’s dirty legs’, for example), are not available at this point.

It will be interesting to explore the focus behavior of the ‘useless task’ idioms in other languages, as well as extending the diagnostic to potential other idioms that are more semantically transparent and involve a scalar aspect of meaning.

## References

- Abeillé, Anne 1995. The flexibility of French idioms. A representation with lexicalized tree adjoining grammar, 15–42. In M. Everaert et al. (eds.), *Idioms. Structural and Psychological Perspectives*. Hillsdale, NJ: Erlbaum.
- Bargmann, Sascha & Manfred Sailer. 2018. The syntactic flexibility of semantically nondecomposable idioms. In S. Markantonatou & M. Sailer (eds.), *Multiword expressions: Insights from a multi-lingual perspective*, 1–29. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.1182587>.
- Bostoën, Koen and Ernest Nshemezimana. 2017. The conjoint/disjoint alternation in Kirundi (JD62): a case for its abolition. In J. van der Wal & L. M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*, 390–425. Berlin: De Gruyter.
- Buell, Leston C. 2005. *Issues in Zulu morphosyntax*. PhD dissertation, University of California Los Angeles.
- Buell, Leston C. 2009. Evaluating the immediate postverbal position as a focus position in Zulu. In M. Matondo, F. Mc Laughlin & E. Potsdam (eds.), *Selected Proceedings of the 38th Annual Conference on African Linguistics: Linguistic Theory and African Language Documentation*, 166–172. Somerville, MA: Cascadilla Proceedings Project.
- Cruse, D. Alan 2000. *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.
- Ernst, Thomas. 1981. Grist for the linguistic mill: Idioms and ‘extra’ adjectives. *Journal of Linguistic Research* 1. 51–68.
- Fanselow, Gisbert & Denisa Lenertová. 2011. Left Peripheral Focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory* 29. 169–209.
- Fellbaum, Christiane. 2011. Idioms and Collocations. In: Maienborn, C. et al. (eds.) *Handbook of Semantics*. Berlin: de Gruyter.
- Fellbaum, Christiane. 2014. The syntax and grammar of idioms. In: T. Kiss and A. Alexiadou (eds.), *Handbook of Syntax*, 776–902. Berlin: de Gruyter.
- Fellbaum, Christiane. 2019. How flexible are idioms? A corpus-based study. *Linguistics* 57(4). 735–767.
- Fraser, Bruce 1970. Idioms within a transformational grammar. *Foundations of Language* 6. 22–42.
- Frey, Werner. 2005. Zur Syntax der linken Peripherie im Deutschen. In Franz Josef d’Avis (ed.), *Deutsche Syntax: Empirie und Theorie (Göteborg Germanistische Forschungen 46)*, 147–171. Göteborg: Acta Universitatis Gothoburgensis.
- Glucksberg, Sam. 2001. *Understanding figurative language. From metaphor to idioms*. Oxford: Oxford University Press.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages*. London: Oxford University Press.
- Horn, Georg M. 2003. Idioms, metaphors and syntactic mobility. *Journal of Linguistics* 39. 245–273.
- Hyman, Larry M. & Francis X. Katamba. 1993. The augment in Luganda: syntax or pragmatics? In Mchombo, Sam (ed.), *Theoretical aspects of Bantu grammar*, 209–256. Stanford: CSLI.
- Lusekelo, Amani and Jenneke van der Wal. 2021. The augment and exhaustivity in Kinyakyusa. Online talk at WOCAL10, June 2021, available via <https://osf.io/jc6vm/>.
- Maho, Jouni F. 2009. NUGL Online: The online version of the new updated Guthrie list, a ref-



- erential classification of the Bantu languages (Available online at <http://goto.glocalnet.net/mahopapers/nuglonline.pdf>)
- Nunberg, Geoffrey, 1978. *The Pragmatics of Reference*. Indiana University Linguistics, Bloomington.
- Nunberg, Geoffrey, Ivan Sag, and Thomas Wasow. 1994. Idioms. *Language* 70(3). 491-538.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27. 53-94.
- Renans, Agata, Malte Zimmermann & Markus Greif. 2011. *Questionnaire on Focus Semantics*. Working Papers of the SFB 632. Potsdam: Universitätsverlag Potsdam.
- Rooth, Mats. 1985. *Association with focus*. Amherst, MA: University of Massachusetts. (PhD dissertation).
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75-116.
- Rooth, Mats. 1996. Focus. In Leppin, S. (ed.), *The handbook of contemporary semantic theory*, 271-297. Oxford: Blackwell.
- Skopeteas, Stavros, Ines Fiedler, Sam Hellmuth, Anne Schwarz, Ruben Stoel, Gisbert Fanselow, Caroline Féry & Manfred Krifka. 2006. *Questionnaire on Information Structure (QUIS)*. Working Papers of the SFB632 (ISIS) 4.
- Van der Spuy, Andrew. 1993. Dislocated noun phrases in Nguni. *Lingua* 90. 335-355.
- Van der Wal, Jenneke. 2009. *Word order and information structure in Makhuwa-Enahara*. Utrecht: LOT.
- Van der Wal, Jenneke. 2011. Focus excluding alternatives: conjoint/disjoint marking in Makhuwa. *Lingua* 212 (11). 1734-1750.
- Van der Wal, Jenneke and Saudah Namyalo. 2016. The interaction of two focus marking strategies in Luganda. In D. L. Payne, S. Pacchiarotti & M. Bosire (eds.), *Diversity in African languages*, 355-377. Berlin: Language Science Press.
- Van der Wal, Jenneke. 2017. What is the conjoint/disjoint alternation? In J. van der Wal & L. M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*, 14-60. Berlin: Mouton de Gruyter.
- Wasike, Aggrey. 2007. *The left periphery, wh-in-situ and A-bar movement in Lubukusu and other Bantu languages*. PhD, Cornell University.
- Watters, John. 1979. Focus in Aghem. In L. M. Hyman (ed.), *Aghem grammatical structure*. Los Angeles: University of Southern California.
- Webelhuth, Gert, Sascha Bargmann & Christopher Götz. 2018. Idioms as evidence for the proper analysis of relative clauses. In M. Krifka & M. Schenner, *Reconstruction Effects in Relative Clauses*, 225-262. Berlin, Boston: De Gruyter.
- Wierzba, Marta. 2016. An experimental view on the syntactic flexibility of German idioms. Presentation at Budapest-Potsdam-Lund Colloquium, Budapest. Slides via [http://www.martawierzba.de/wierzba\\_2016\\_slides\\_budapest.pdf](http://www.martawierzba.de/wierzba_2016_slides_budapest.pdf)
- Wierzbicka, Anna. 1999. *Emotions Across Languages and Cultures: Diversity and Universals*. Cambridge: Cambridge University Press.
- Yoneda, Nobuko. 2017. Conjoint/disjoint distinction and focus in Matengo (N13). In J. van der Wal & L. M. Hyman (eds.), *The conjoint/disjoint alternation in Bantu*, 426-452. Berlin: De Gruyter.
- Ziervogel, Dirk & Enos J. Mabuza. 1976. *A grammar of the Swati language*. Pretoria: J.L. van Schaik.