

The Zulu locative prefix **ku-** revisited: a corpus-based approach

Gilles-Maurice de Schryver¹ and Rachéle Gauton^{2*}

¹ Research Assistant of the Fund for Scientific Research — Flanders (Belgium) and Department of African Languages, University of Pretoria, Pretoria 0002, South Africa

e-mail: gillesmaurice.deschryver@rug.ac.be

² Department of African Languages, University of Pretoria, Pretoria 0002, South Africa

* Corresponding author, e-mail: rgauton@postino.up.ac.za

Abstract: This article re-examines the distribution of the class 17 locative prefix **ku-** and its variants **kwi-** and **ko-** in the locativisation of nouns in Zulu. To this end an electronic corpus of 5 million running Zulu words — the University of Pretoria Zulu Corpus (PZC) — is queried. We indicate how PZC can be used to highlight previously under-emphasised and overlooked aspects of a seemingly well-documented language feature such as the class 17 locative prefix **ku-**. Analysing the output proffered by PZC — an organic corpus that has been organised chronologically, and that consists of a number of sub-corpora stratified according to genre — enables us to reach conclusions regarding not only the frequencies with which the variants **ku-**, **kwi-** and **ko-** of the class 17 locative prefix are used, but also regarding the possible changes in use that these prefixes have undergone with time. We also show how these prefixes relate to each other in the different sub-corpora.

Introduction

The nature, use and possible diachronic origin and development of the Zulu class 17 locative prefix **ku-** has been well-documented for over a century in various Zulu grammars such as those by Schreuder (1850), Grout (1893), Colenso (1903), Roberts (1909), Wanger (1917), Samuelson (1925), Stuart (1932), Mhlahdla and Mbata (1933), Doke (1955, 1973⁶), Engelbrecht (1956), Nyembezi (1956), Van Eeden (1956), Ziervogel *et al.* (1976), Taljaard and Bosch (1988), Poulos and Bosch (1997) and Poulos and Msimang (1998), as well as in studies undertaken by researchers such as *inter alia* Růžička (1959; 1960), Van der Poll (1970), Ziervogel (*s.d.* [1970?]; 1971), Grégoire (1975), Poulos (1985) and Gauton (1995).

In the light of such a large (and often detailed) body of existing analyses of this phenomenon, the value of revisiting the Zulu locative prefix **ku-** might well be questioned. In this article, however, this phenomenon will be examined from a different angle, and it will be shown how the utilisation of an electronic Zulu corpus can highlight previously under-emphasised and overlooked aspects of a seemingly well-documented language feature.

During the course of the research undertak-

en for the purposes of this article, it became clear that there are still other aspects of locativisation in Zulu that could benefit from being revisited by way of a corpus-based study. These aspects include *inter alia*:

- The interchange between the class 17 locative prefix **ku-** and the so-called 'discontinuous' locative morpheme **e-...-ini** in the derivation of locatives from [+human] nouns; and
- The use of the morphemes **o-...-ini** in the derivation of locatives from nouns in classes 11 and 14.

These particular issues will not, however, be discussed in any detail here, as they do not fall within the ambit of the current article and could possibly be the focus of a forthcoming endeavour. We will, nonetheless, discuss the use of the class 17 locative prefix **ku-** (and its variant form **kwi-**) in the derivation of locatives from nouns (other than those from classes 1/2, 1a/2a and [+human] nouns in class 6) in relation to the **e-/o-...-ini** locativisation strategy, i.e. in environments where the use of **e-/o-...-ini** would have been expected.

Before proceeding with a discussion on the utilisation of the class 17 prefix **ku-** in the derivation of Zulu locatives, we will first give an

indication of what an electronic corpus such as the University of Pretoria Zulu Corpus entails.

The University of Pretoria Zulu Corpus (PZC)

In a recent sequel of five articles De Schryver and Prinsloo laid the foundations of a discipline of corpus linguistics for the African languages (De Schryver & Prinsloo, 2000a; b; c; Prinsloo & De Schryver, 2001a; b).¹ In that sequel these scholars showed clearly that the compilation, querying and manifold applications of electronic corpora in various linguistic fields have become a reality for the African languages. They also indicated that most corpus work in South Africa revolves around the University of Pretoria (UP) (De Schryver & Prinsloo, 2000a: 96–97). This is still the case, with the sizes of corpora built at UP's Department of African Languages having doubled, tripled, and for some languages even risen tenfold since then.

As far as Zulu is concerned, the University of Pretoria Zulu Corpus (PZC) stood at 0.7 million running words (tokens) in May 2000. Uninterrupted and painstaking corpus-building by De Schryver and Dlomo resulted in a 5-million-token corpus just one year later — a sevenfold increase.² In order to give an idea of what '5 million Zulu words' mean, we can make the following comparison. The first edition of the by-now famous COBUILD series of dictionaries, the *Collins COBUILD English Language Dictionary* (COBUILD1) (Sinclair, 1987), was based on an English-language corpus of approximately 7.3 million words (Renouf, 1987: 7). Obviously, when comparing the sizes of *different-language* corpora, one cannot plainly use the number of words. Instead, we counted the number of words in several parallel texts in Zulu and English. Using only certified translations for those parallel texts revealed that for every word in Zulu, on average 1.32 words are needed in English. In other words, a corpus of 5 million Zulu words is equivalent to an English-language corpus of 6.6 million words. Stated bluntly: the *isiZulu National Lexicography Unit* (isiZulu NLU) now has the means — with PZC — to mimic the achievements of such a revolutionary dictionary as COBUILD1.

As the general framework of how to build African-language corpora has been dealt with in De Schryver and Prinsloo (2000a), we will limit the present discussion to a summation of the main characteristics of PZC.

- Firstly, PZC is an 'organic corpus', meaning that it '*must be allowed to grow and live if it is to reflect a growing, living language*' (Atkins, Clear & Ostler, 1992: 1). In simple terms this means that a method of successive approximations is employed in order to achieve some kind of balance and representativeness. Time and money constraints, however, did not allow for the building of an oral component so far. Compare in this context the design of the *Longman Lancaster English Language Corpus*, for which Summers argues that '*although the importance of the spoken language could hardly be overestimated, the logistical problems of capturing a sufficient large body of speech electronically were very real, unlike with written material*' (1993: 184). PZC is thus, by definition, a structured sampling of the written Zulu language.
- Secondly, PZC is a 'diachronic corpus', meaning that written material was excerpted from sources published from the 1930s on, up to the present. Studying the number of available Zulu works in an average bookstore or library, one notices a growth gaining momentum throughout the decades, reaching a peak in the 1980s. The 1990s though saw a recession, with publishing houses in South Africa struggling hard to get their publications on the market. Indeed, since the 1990s to date, many publications ended up being stored by the publishers, awaiting prescription. When the latter does not materialise, the stored material is pulped.³ At all times the growing organic PZC bore witness to the material available in an average bookstore or library, with the number of words per decade in the current PZC as shown in Table 1 and Figure 1.
- Thirdly, PZC is essentially 'genre-driven', meaning that written material was selected in

Table 1: Distribution of the number of words (tokens) per decade in the current PZC

Period	# words	%
1930s	142 896	2.86
1940s	156 492	3.13
1950s	348 150	6.96
1960s	795 610	15.91
1970s	1 242 635	24.85
1980s	1 785 876	35.71
1990s	529 797	10.59
Total	5 001 456	100.01

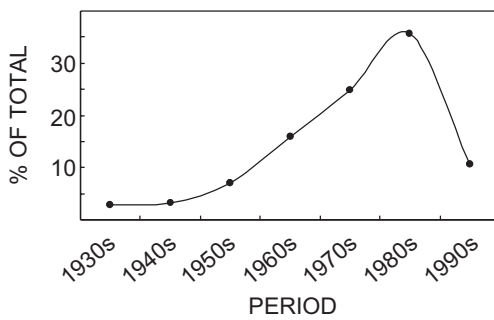


Figure 1: Allocation in % of the number of words (tokens) per decade in the current PZC

such a way as to reflect the spread of the genres in an average bookstore or library. A comparison with the *Longman Lancaster English Language Corpus* might again be useful. The design of that corpus was, in contrast, essentially *topic-driven* and revolved around ten broad subject areas, namely 40.0% fiction, 14.1% social science, 10.4% world affairs, 7.9% arts, 6.0% natural & pure science, 5.7% leisure, 4.7% belief & thought, 4.4% commerce & finance, 4.3% applied science, and 2.3% poetry, drama, humour (Summers, 1993: 192–193). The weight of the different sub-corpora in the current *genre-driven* PZC can be seen in Table 2 and Figure 2.

These three main features of PZC will prove to be of important significance throughout this article. Firstly, a collation of 5 million running words corresponds roughly to 300 books. Any query, and any claim based on that query, can therefore be considered as the average of those so-called 300 books. Secondly, the diachronic character of the corpus will enable us to see how the studied features change with time. And thirdly, we will be able to observe how those same features relate to one another in the different sub-corpora.

To proceed, however, a brief overview will first be given of the distribution of the Zulu locative prefix *ku-* as described in existing grammars and other sources.

The distribution of the class 17 locative prefix *ku-* in Zulu

Standard Zulu sources such as the grammars that were referred to earlier, agree that the prefix of locative class 17 *ku-* is used with the general locative meaning 'in, to, from, at, etc.' in order to derive locatives from (predominantly) [+human] nouns in classes 1/2 and 1a/2a, as well as from pronouns. Consider Example 1 in this regard.

Nyembezi (1956: 246) further points out that the prefix *ku-* can also be used in order to locativise [+human] nouns belonging to class 6

Example 1:

kumntwana	<	ku-(u)mntwana (cl.1)	'to, from, at, etc. the child'
kubafana	<	ku-(a)bafana (cl.2)	'to, from, at, etc. the boys'
kumama	<	ku-(u)mama (cl.1a)	'to, from, at, etc. mother'
komalume	<	k(u)-omalume (cl.2a)	'to, from, at, etc. the uncles'
kuwe	<	ku-we(na)	'to, from, at, etc. you'
kimi	<	ku-mi(na)	'to, from, at, etc. me' ⁶

Table 2: Weight of the different sub-corpora in the current *genre-driven* PZC

Sub-corpus	# words	%
Novels, Novelettes	2 098 836	41.96
Textbooks	889 175	17.78
Short stories, Essays, Readers ⁴	834 019	16.68
Dramas, One-act plays	357 073	7.14
Religious texts	343 842	6.87
Poetry	238 407	4.77
Oral literature, ⁵ Folklore, Legends	166 262	3.32
WWW files, Pamphlets	73 842	1.48
Total	5 001 456	100.00

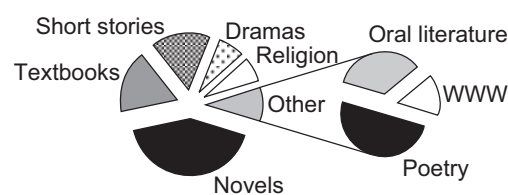


Figure 2: Pie diagram of the different sub-corpora in the current *genre-driven* PZC

(this includes the names of nations or clans), see Example 2.

Note that nouns such as these can also be locativised through the use of the locative morphemes **e-...-ini**; i.e. **emakhosikazini** in addition to **kumakhosikazi**, and **emaBhunwini** in addition to **kumaBhunu**. (This type of variation can also occur with certain nouns in classes 1/2.) As was stated in the introduction to this article, the occurrence of these variant forms in PZC, as well as the frequency of these forms in relation to each other, could constitute the topic of a separate study.

Wanger (1917: 208) indicates that this kind of variance is also found with the class 3 noun **umhlobo** 'friend, relative', which can be locativised either by way of the locative prefix **ku-** as in **kumhlobo** 'to, from, at, etc. a relative, friend' or by way of the morphemes **e-...-ini** as in **emhlotsheni** 'to, from, at, etc. a relative'. Wanger (1917: 209) states categorically that no nouns other than those belonging to classes 1/2, 1a/2a and the class 3 noun **umhlobo** 'friend, relative' can be locativised by utilising the prefix **ku-**. Further, Wanger (1917: 209) contends that the locativisation of [+human] nouns from class 6 by way of the prefix **ku-** (as has been discussed in Example 2), is the result of analogy with locativised forms from classes 1/2 or 1a/2a that precede these nouns in the particular conversation or context. Corpus queries proved this assumption to be totally unfounded.

It is furthermore generally accepted amongst grammarians writing on the Nguni languages that the class 17 locative prefix has a much more widespread distribution in languages such as Phuthi and Xhosa than is the case in Zulu. In Phuthi the prefix **ku-** can be used with nouns from *any* noun class, and not just with nouns from classes 1/2, 1a/2a, and

certain [+human] nouns in class 6, as has been indicated above with reference to Zulu. See Example 3 from Phuthi in this regard.

However, as early as 1909, Roberts indicates that in Zulu the class 17 locative prefix can be used with class 9 nouns such as **inkosi** 'king', resulting in the locativised form **kwinkosi** (< **ku-inkosi**) 'to, at, etc. the king' (Roberts, 1909: 46). Note that in cases such as these, consonantalisation takes place whereby the juxtaposition of the vowel **u** of **ku-** and the preprefix **i-** of the noun, results in the vowel **u** being replaced with the semi-vowel **w**.

During the late 1980s and early 1990s, this phenomenon started to receive renewed attention from grammarians writing on the Zulu language. The locativisation of nouns by means of the variant **kwi-** of the class 17 locative prefix, was described by Msimang and Shamase in 1988 as a so-called 'recent innovation', abounding in radio and television broadcasts (Msimang & Shamase, 1988: 111). Msimang and Shamase (1988: 113) indicate that the then Zulu Language Board rejected this usage of **kwi-** on the pretext that it constitutes Xhosa influence and that **kwi-** is therefore not a Zulu morpheme, but a Xhosa one. Yet Msimang and Shamase (1988: 113) are of the opinion that this is not necessarily the case, and argue that the change of **ku-** to **kwi-** before nouns commencing on an **i-** is completely in line with the morphophonological rules of the Zulu language. They therefore conclude that it is incorrect to discount **kwi-** as being a Xhosa locative prefix.

The prescriptive stance taken by the Zulu Language Board is to be expected when one considers that the ten African-language Language Boards of the time were the '*official sanctioning bodies in linguistic matters*' (Mtintsilana & Morris, 1988: 109). According to Trew (1994: 77) these bodies suffered not only

Example 2:

kumakhosikazi	<	ku-(a)makhosikazi (cl.6)	'to, from, at, etc. the women'
kumaBhunu	<	ku-(a)maBhunu (cl.6)	'to, from, at, etc. the Boers (Afrikaners)'

Example 3:

kumalume (cl.1a)	'to, at, etc. uncle'
kumuti (cl.3)	'to, at, etc. the village'
kwivu (cl.10)	'to, at, etc. the sheep'

(Van der Poll, 1970: 131–132)

from a lack of popular credibility, but also from 'obstinate linguistic puritism'. Trew's observations regarding the Language Boards' attitude towards the creation of terminology in the African languages, could equally apply to their stance on the use of the variant **kwi-** in Zulu:

The imperative of puritism is so strong in these institutions that any term, no matter how widely used, is liable to be excluded (and replaced with an ad hoc committee invention) if it is considered to be a 'loan word'. (Trew, 1994: 77)

Madiba argues that despite the contributions made by the Language Boards towards the development and standardisation of the African languages:

Unfortunately, the activities of these boards were undermined by their lack of legitimacy, as they were perceived as the agents of the apartheid government. (Madiba, 1998: 64)

According to Msimang and Shamase (1988: 113) the Zulu Language Board prescribed that instead of the variant **kwi-**, the prefix **ku-** be used also with nouns with the preprefix **i-**, even though:

- the change of **ku- + i- > kwi-** is completely in accordance with the morphophonological rules of the language; and
- the use of the locative prefix **ku-** could result in a violation of the agreement rules.

Consider in this regard Example 4.

Nevertheless, Msimang and Shamase (1988: 113–114) advocate that as the variant **kwi-** is a 'new locative prefix', it should be restricted to new Zulu terms, in particular brand names and other proper nouns where the use of the locative morphemes **e-...-ini** could change the phonetic structure of the word to such an extent that it becomes almost unrecognisable, see Example 5.

Canonici (1990: 55 *et seq.*) also refers to this phenomenon. He points out that with nouns in what he terms classes 3a and 9a, locatives are formed by use of either **ku-** or **e-...-ini**. Canonici's class 3a nouns are those [-human] nouns generally regarded as belonging to class 1a. Canonici (1990: 56) defines class 3a as representing traditional nouns signifying small animals and insects when not personalised, such as **unogwaja** 'hare', **umiyane** 'mosquito', etc., a few traditional nouns such as **ugwayi** 'tobacco', **umoba** 'sugarcane', **unyezi** 'moonlight', etc., and borrowed nouns of foodstuffs, mass nouns or relatively 'new' substances, such as **ushizi** 'cheese', **ubhontshisi** 'beans', **upende** 'paint', **uthayi** 'tie', **uKhisimusi** 'Christmas', etc. Compare the way in which nouns from this class are locativised, as shown in Example 6.

Canonici orders borrowed nouns, generally

Example 4:

Guqukela <i>ku-Omo i-Omo inamandla</i>	(Uncertainty as to the form of the noun – u-Omo (cl.1a) or i-Omo (cl.9) – and therefore a potential conflict with the agreement morpheme i- (cl.9) in inamandla)
Guqukela <i>ku-Omo u-Omo unamandla</i>	(No conflict between the noun and the agreement morpheme, but the noun i-Omo has been changed to u-Omo)
Guqukela <i>kwi-Omo i-Omo inamandla</i>	(No conflict between the noun and the agreement morpheme)

'Change to Omo [a brand of washing powder], Omo is strong'

(Msimang & Shamase, 1988: 113)

Example 5:

Instead of:	Msimang and Shamase suggest:	
e-Surfini	kwi-Surf	'in, etc. Surf [a brand of washing powder]'
e-Omeni	kwi-Omo	'in, etc. Omo [a brand of washing powder]'
eLexingtonini	kwiLexington	'to, at, etc. Lexington [a brand of cigarettes]'

Example 6:

umoba	'sugarcane'	>	kumoba	/	emobeni
ushukela	'sugar'	>	kushukela	/	eshukeleni
uthayi	'tie'	>	kuthayi	/	ethayini
ubhavu	'bath'	>	kubhavu	/	ebhavini

regarded as belonging to class 5, and which indicate new items, as belonging to class 9a. This class contains nouns such as **ihhovisi** 'office', **ibhodi** 'board', **ikhwaya** 'choir', **izu** 'zoo', **ibhayisikobho** 'cinema', etc., and these nouns are locativised by use of either **e-...-ini** or **ku-** (i.e. **ku-i-** > **kwi-**).⁷

Wilkes (1993: 173–174) agrees that the variant **kwi-** occurs especially with nouns signifying brand names and scientific and technical concepts, but points out that it can also be affixed to other nouns signifying plants and animals, e.g. **kwimpala** (< **ku-impala**) 'at (= near) the impala'. Wilkes ascribes speakers' reluctance to locativise nouns such as **i-Omo**, **i-Surf** and **i-Lexington** by way of the morphemes **e-...-ini** to basically two reasons which are probably based on practical considerations. Firstly, utilising **ku-** means that the resultant locativised form has less syllables than had **e-...-ini** been used. Locativising by way of **ku-** is therefore a more economical strategy. Secondly, and as is also argued by Msimang and Shamase (1988: 113–114), locativisation through the use of **e-...-ini** causes greater phonological changes to the noun in question than the use of **ku-**.

Lastly, Nyembezi (1956: 246) makes mention of another variant of the class 17 locative prefix, namely **ko-** that is affixed to nouns in class 2, see Example 7.

It is not immediately apparent why the locative prefix **ku-** would change to **ko-** when affixed to (class 2) nouns commencing on the preprefix **a-**. As was indicated earlier in Example 1, in the case of for example **kubafana** 'to, from, at, etc. the boys', class 2 nouns lose their preprefix **a-** when the locative prefix **ku-** is affixed to them, i.e. **ku-(a)bafana** > **kubafana**.⁸ Van Eeden (1956: 49) does, however, mention a little known (and often unacknowledged) morphophonological change found in Zulu whereby the combination **-ua-** or **-wa-** sometimes becomes **-o-**, as in: **umfelokazi** 'widow' a variant of **umfelwakazi**; **ulosi** 'smell' a variant of **ulwasi**; and **ulontane** 'delicious food' derived from **ulwantu** 'delicious or dainty food (a special treat)'. Wanger (1917: 71, 213) also men-

tions that in Zulu (as well as in Lala), the morpheme **-ku-** often appears as **-ko-**, and that the Zulu (locative) prefix **ku-** is often found as **ko-**.

To summarise, therefore, the literature indicates that when affixed to nouns, the class 17 locative prefix can take any of the forms **ku-**, **kwi-** or **ko-**, as determined by the nature of the noun following the prefix. In the next section, we will show how an electronic Zulu corpus of 5 million running words can be utilised to ascertain whether or not the various realisations of the class 17 locative prefix discussed so far can be verified. Concurrently, we will indicate how such a corpus can also shed more light on the frequency and distribution of the found morphemes in a written Zulu corpus that spans just over seven decades and a number of different genres.

Querying the University of Pretoria Zulu Corpus (PZC)

De Schryver and Lepota point out that:

Compared to intuition, informant elicitation and grammatical conjectures, the corpus has the big advantage in that it shows *real* language use. Any corpus query shows the 'attested and authentic usage *average*' of several hundreds of mother-tongue speakers. (De Schryver & Lepota, 2001: 16)

The results of such corpus queries can be presented in a multitude of ways by modern corpus query software packages, yet one of the most popular means to look at the proffered output is through the study of KWIC concordance lines. KWIC stands for 'keyword in context' and is basically the result of a smart way to automatically cull all instances of a particular word, phrase or morpheme — where each of those can take various wildcards — from an electronic corpus, and to present those culled strings with context on both sides. Together with the searched-for keyword (yet it is clear that it need not be a 'word') one can also specify context strings that should (or should not) appear within a specified horizon. Also, the number of KWIC lines and the type of sampling applied is coordinated by the person querying the corpus.

Example 7:

kobantu	<	ko-(a)bantu	'to, from, at, etc. the people'
kobelungu	<	ko-(a)belungu	'to, from, at, etc. the white people'
kobefundisi	<	ko-(a)befundisi	'to, from, at, etc. the (church) ministers'

PZC is currently analysed with WordSmith Tools, and a simple search for *ku*, *kwi*, and *ko* with the Concord tool (i.e. the tool that generates KWIC lines) immediately reveals that the three strategies discussed above are attested in written Zulu.⁹ In the sections that follow, we will therefore examine the three identified realisations of the class 17 locative prefix, i.e. **kwi-**, **ku-** and **ko-**, with reference to:

- their frequency and distribution;
- possible changes that might have occurred with time; and
- their distribution in the different sub-corpora, stratified according to genre.

The locative prefix **kwi-**

Querying PZC clearly indicates that there can be no doubt that the locative prefix **kwi-** is widely used in Zulu in the locativisation of loanwords and nouns signifying brand names and scientific and technical concepts. What is surprising, however, is the extent to which this prefix is used with Zulu nouns that are neither loanwords, nor express new scientific and technical concepts and brand names.

On querying PZC, the following distinctions were made between different types of nouns:

- Brand names; names of companies or institutions; titles of books; etc.
- Borrowed nouns in which the English spelling has been retained. Such words have not been transliterated, i.e. nativised in the sense that their phonology has been adapted to reflect the Zulu phonological system. In the case of words such as these, only a (noun class) prefix, usually **i(li)-** (or its plural form **ama-**) has been added.¹⁰
- Transliterations, i.e. new scientific and technical terms formed via a process of transliteration by adapting the phonological structure of the loanword to the sound system of Zulu.
- New scientific and technical terms formed through term-creation processes other than transliteration (such as coining, compounding, semantic transfer, etc.).
- None of the above, i.e. so-called 'normal' Zulu words that are neither loanwords nor do they signify previously unknown (new) scientific and technical concepts.

Compare in this regard Tables 3 through 7 which show concordance lines culled from the current PZC, each containing a few representative examples illustrating the noun types discussed above.

Table 3: A sample of concordance lines for the locative prefix **kwi-** affixed to nouns signifying brand names; names of companies / institutions; titles of books; etc.

e konakele. Wangena nezicabha **kwiMazda** yakhe ayeyethembile ngejubane ayelilungisel ze izincomo ngamazinga neziq **kwi-National Standards Body**; d. wenze izincomo ngenq phele; (Kuhunyushwe kukhishwe **kwi-Rand Daily Mail**, ka 14 Julayi 1983) 2. Abathi be thola ephoqekile ukuyosebenza **kwi-Sugar Laboratory yakwaHuletts**, iminyaka emibili. . . s (Kuhunyushwe kukhishwa **kwi-The Guns of Navarone** ebhalwe nguAlistair McLean,

Table 4: A sample of concordance lines for the locative prefix **kwi-** affixed to borrowed nouns in which the English spelling has been retained and only a noun class prefix added

elayo; ubacele ukuba bazifake **kwi-account** yakhe: Insipho yokuwasha noma iyiphi inh lela Kusho ukuthi sebesindile **kwi-Cholera**, abasoze bathelela abanye. Imininingwane efe? 6. Phawula ngabalingiswa **kwidrama** eniyenza nonyaka usho ukuthi ubani oyimbang zonke leyo mali ngizoyibamba **kwi-pay** wakho." "Nalapho ngisabonga Mnumzane. Angith ela ukuthi uyafuna yini ukuya **kwi-test** noma cha. Akukho muntu, ngisho nodokotela,

Table 5: A sample of concordance lines for the locative prefix **kwi-** affixed to transliterations

nti umoya omuhle ukhona njalo **kwi-atmosfera**. Amanzi ayimpilo ezitshalweni. a) Isit Uma kwenzeka kuxoxwa ngabantu **kwi-eseyi** ngabantu nje abangafani nabendaba emfushan ukuthi bethule imiqondo yabo **kwiKhabhinethi**. Umdwebo kaMnumzane u-Iaan Bekker iwo a sinyelela kancane siyongena **kwiipulatifomu** 4. Sathi siqeda ukuma, wabe esephansi ka zamthwala zamfaka phakathi **kwiambulense** yasho yathi ng . . . wi! abantu bavula

Table 6: A sample of concordance lines for the locative prefix **kwi-** affixed to new scientific and technical terms formed through term-creation processes other than transliteration

mlando. Chaza. 3. Abalingiswa **abakwinkondlo-mlando** behluke kanjani kwabendaba? Ivi a, ngazi nje ngokufunda ngayo **kwibalazwe**." "Awufuni sike sithi gqi ngakhona, kwaZw ezixazululo, bese nizithumela **kwigatsha** lenu. Izingumo noma izigxazululo zenu kofa gabavikela kanjani abantabami **kwingculazi**? Abazali abaningi bakhathazeka ngokuthi abantwana isiphuzo esibandayo **kwisigandisi**, utshela udadewenu aniphe amakhekhe nis

Table 7: A sample of concordance lines for the locative prefix **kwi-** affixed to 'normal' Zulu words that are not loanwords and do not signify previously unknown (new) scientific and technical concepts

o ngezimbambo. Ngehluka ngani **kwibululu** na? Nhloboni yobubi, nhloboni yobuhle. Eng uthola indlela yokwakha isihe **kwindaba** kamama. "Mhlawumbe uqinisile, ibhanoyi leth osi yaphendukisela amehlo ayo **kwihlabamkhosi**, amehlo ayo ayesezana nomkhonto ulol o. (iv) Izibongo zibalulekile **kwisizwe**, ngoba inkosi yabe ibongelwa ngokuhlabana k i lokhu. Ayibange-ke isadlula **nakwisambane** iyosazisa ngokuphumelela kwecebo. Kunal

In Table 8 the overall distribution of **kwi-** across the various identified noun types is illustrated.

As can be deduced from Table 8, well over a third of all the examples of nouns locativised by way of the prefix **kwi-** constitute 'normal' Zulu words. This truly astonishing finding illustrates the usefulness of an electronic corpus in linguistic study, as the occurrence of **kwi-** in this type of environment is an aspect of its distribution that has gone largely *unnoticed* up to now in the manifold descriptions of Zulu grammar.

The locative prefix **ku-**

As there seems to be general agreement amongst grammarians writing on the Zulu language that the locative prefix **ku-** is used to locativise nouns from classes 1/2, 1a/2a, [+human] nouns from class 6 and pronouns, and as this is also supported by the corpus, the use of **ku-** in such environments will not receive

any further attention here. For the purposes of this article, we therefore restricted ourselves to an examination of instances where the prefix **ku-** is used to locativise nouns from classes *other* than those just mentioned. As is clear from the descriptions of locativisation by means of the prefix **ku-** that were given earlier, the use of **ku-** in such other environments: (a) has been rejected outright by grammarians writing on Zulu, (b) has been described as exceptions to the rule, or (c) has in most cases not been mentioned at all. On querying PZC, however, many examples were found of the locative prefix **ku-** affixed to nouns other than those in classes 1/2, 1a/2a and [+human] nouns in class 6. It was also found that in these instances, **ku-** is used with largely similar types of nouns to what was observed in the case of **kwi-**. Compare Tables 9 through 13 in this regard, where representative examples are given of the use of **ku-** in the

Table 8: Distribution of the locative prefix **kwi-** across the noun types

Noun type	#	%
Brand names; names of companies / institutions; titles of books; etc.	43	18.22
Borrowed nouns in which the English spelling has been retained and only a noun class prefix added	14	5.93
Transliterations	31	13.14
New scientific and technical terms formed through term-creation processes other than transliteration	59	25.00
<i>Sub-total: Loanwords and new terms</i>	147	62.29
'Normal' Zulu words that are not loanwords and do not signify previously unknown (new) scientific and technical concepts	89	37.71
Total	236	100.00

Table 9: A sample of concordance lines for the locative prefix **ku-** affixed to nouns signifying brand names; names of companies / institutions; titles of books; etc.

yamaGatsha yama-Shopsteward, **kuma-BEC** kanye **nakuma-PEC**, bese zedluliselwa ehhovis enerating Bodies 22. Ubulungu **kuma-Standard Generating Bodies** 23. Ukuqokwa kwamalu hathi kahle ungafaka isikhalo **kuMonitoring and Complaints Department** ye-IBA ngokub , usebenza ukuqeqesha othisha **kuMultiple Perspectives Teacher Education**, eMichigan ngoNesi kufanele baphumelele **kuyiJunior-Certificate**. Noma zisekhona ezisavuma ama

Table 10: A sample of concordance lines for the locative prefix **ku-** affixed to borrowed nouns in which the English spelling has been retained and only a noun class prefix added

000, kanti unejubane elingeqi **kuma-120km** ngehora, kanye nejubane elikhululekile la ha indima ingabinamvumelwano. **Kumadrama** amaningi kaShakespeare sithola ubunkondlo anzwa ngobunyoningco kusukela **kuma-R39**. Kuqashelwa siqu. Inombolo yocingo 455443 k nye nesigaba esisentshenziswa **kuma-unit standard**; kanye o. ingxenye yokuzobhalwa o qeqeshwa ngemvelo emayunivesi **nakumatechnikon**. INingizimu Afrika idinga ososayensi

Table 11: A sample of concordance lines for the locative prefix **ku-** affixed to transliterations

khe. Wabuka uDlokwakhe edweba **ngakunombolo** 26. "Chaza bo phulofesa. Musa manje uku misakazweni **kumathelevishini**, **kumafilimi**, ezincwadini, **kumanyuziphepha** kanye namam kelwe **kumarekhodi** abasebenzi (**kumakhompyutha**), ukwenzela ukuthi lapho kubhalwa izi hephabhuku. Bathi nje bedlula **kumatholigethi** lapantsha isondo langemuva letekisi, gathi kukhona okwakungenelisi **kusitatimende** sikaZuma asenzayo ehovisi labaSeshi eS

Table 12: A sample of concordance lines for the locative prefix **ku-** affixed to new scientific and technical terms formed through term-creation processes other than transliteration

luhlobo lwephimbo lutholakala **kumabizo** ayisibanjalo. Kuvamisile ukuthi unkamisa wo olukhulu kangaka kithi sonke. **Kumaphephandaba** nasezindabeni esike sizizwe, buyatho we lolo hlaka seluyothunyelwa **kuMnyango** Wemfundo Nemfundiso ukuba luvunywe. Isilab lesine: Ukuzinikela kwentsha **kuzidakamizwa**. Iqembu lesihlanu: Izingozi emgwaqweni hola ubunkondlo esigqini sawo **nakumifanekiso-mqondo**. Kweminye imilolozelo lokho ok

Table 13: A sample of concordance lines for the locative prefix **ku-** affixed to 'normal' Zulu words that are not loanwords and do not signify previously unknown (new) scientific and technical concepts

phecelezi ngempophoma encane, **eyayikumfula** uMngeni - Albert Falls (Akuqondiwe uNog athi izoletha intokozo enkulu **kumhlaba** jikelele. Mhlawumbe ikhona intokozo esisedu ezifengqo. Isihlonipho. Lapha **kusihlonipho** kuthanjiswa inkulumo ngokusebenzisa iga nzisa ukubhala izincwadi kuya **kuzinkomo** nezimbuzi. Iholo lingaphindwa kabili kolih nathi ziyasithinta. Ushintsho **olukumazwe** amanye nathi luyasithinta. Kanti futhi, I

locativisation of nouns that are not generally regarded as being able to combine with this prefix. Note that in these tables, the same organisation of noun types has been followed as in the case of the locative prefix **kwi-** above.

In Table 14 the overall distribution of **ku-** across the various identified noun types is illustrated.

The findings regarding the distribution of

the locative prefix **ku-** as set out in this section again highlight the importance of a well-designed corpus in the undertaking of linguistic research. It is clear from these results that the locative prefix **ku-** is used much more widely with nouns other than those in classes 1/2, 1a/2a and [+human] nouns in class 6 than is currently indicated in Zulu grammars and other sources on the structure of the Zulu language.

Table 14: Distribution of the locative prefix **ku-** across the noun types

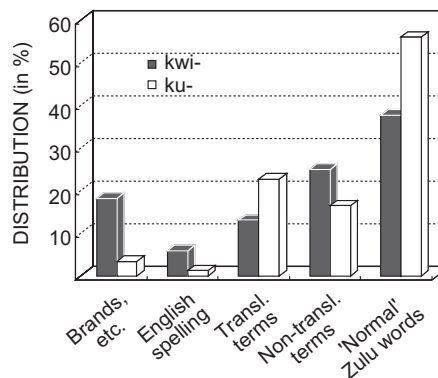
Noun type	#	%
Brand names; names of companies / institutions; titles of books; etc.	21	3.47
Borrowed nouns in which the English spelling has been retained and only a noun class prefix added	8	1.32
Transliterations	138	22.77
New scientific and technical terms formed through term-creation processes other than transliteration	100	16.50
<i>Sub-total: Loanwords and new terms</i>	267	44.06
'Normal' Zulu words that are not loanwords and do not signify previously unknown (new) scientific and technical concepts	339	55.94
Total	606	100.00

It is also obvious from Table 14 that whereas **ku-** in this environment is used to locativise brand names, loanwords and terms referring to new scientific and technical concepts in about 44.06% of the cases, more than half of the occurrences of **ku-** (55.94% to be precise), is with 'normal' Zulu words. This preponderance of affixing **ku-** (and **kwi-**) to 'normal' Zulu words is shown graphically in Figure 3.

The locative prefix ku- in environments where the variant kwi- would have been expected

Another interesting phenomenon that came to the fore during the course of our research into the distribution of the locative prefix **ku-**, is that this prefix is frequently used with nouns commencing on the preprefix **i-**, i.e. in environments where one would expect the variant **kwi-** to be found due to the morphophonological rule **ku- + i- > kwi-** which usually governs the juxtaposition of vowels in this environment in Zulu. Table 15 and Figure 4 give an overview of this phenomenon.

As can be seen from Table 15, just under half (48.02%) of the total number of occurrences found of **ku-** are in environments where the variant **kwi-** would have been expected according to the standard morphophonological rules of the language, i.e. in combination with nouns from classes 4 **imi-**, 5 **i(li)-**, 7 **isi-**, 8 **izi-**, 9 **in-/im-** and 10 **izin-/izim-**. For 'normal' Zulu words the use of **ku-** instead of **kwi-** even climbs to a staggering 69.62%. This might be indicative of Zulu writers' desire to try and stem the 'encroachment' of **kwi-** by using the locative prefix **ku-** instead of the variant **kwi-** – even in environments where **kwi-** would have been more usual. Consider the examples shown in Table 16 in this regard.

**Figure 3:** Distribution across the noun types of the locative prefixes **kwi-** and **ku-**

The sections that follow will now focus on the distribution of, and relation between, the locative prefixes **kwi-** and **ku-** across time, as well as with reference to the different (written) genres.

The distribution of the locative prefixes kwi- and ku- across time

As was indicated at the outset of this article, the diachronic character of PZC would enable us to examine how the frequency of the locative prefixes might have changed across time. In Table 17 and Figure 5, an indication is given of the distribution of the locative prefixes **kwi-** and **ku-** during the time-period spanned by the works contained in the corpus. In order to filter away the different sizes of the sub-corpora, or thus in order not to compare apples with oranges, the counts were 'normalised' per decade. As such, the fourth and sixth columns show the number of instances of **ku-** and **kwi-** 'per 100 000 words'.

Based on Table 17 and Figure 5, the following conclusions can be reached regarding the

Table 15: Distribution across the noun types of the locative prefix **ku-** in environments where the variant **kwi-** would have been expected

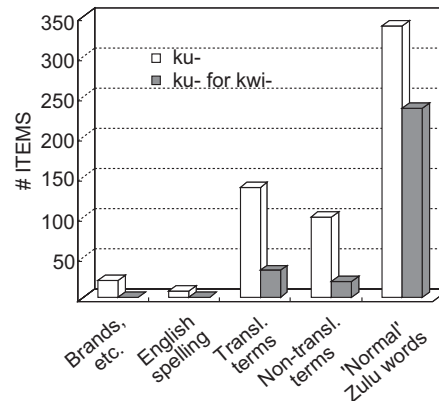
Noun type	# ku-	# ku- for kwi-	% ku- for kwi-
Brand names; names of companies / institutions; titles of books; etc.	21	1	4.76
Borrowed nouns in which the English spelling has been retained and only a noun class prefix added	8	0	0.00
Transliterations	138	34	24.64
New scientific and technical terms formed through term-creation processes other than transliteration	100	20	20.00
<i>Sub-total: Loanwords and new terms</i>	267	55	20.60
'Normal' Zulu words that are not loanwords and do not signify previously unknown (new) scientific and technical concepts	339	236	69.62
Total	606	291	48.02

distribution of **kwi-** and **ku-** across time:

- Except for a decrease in usage during the 1970s, the use of the locative prefix **kwi-** has risen steadily, with an accelerated rise during the 1980s and 1990s.
- Except for a relatively higher usage during the 1950s, the use of the locative prefix **ku-** has also risen steadily, yet with an exponential boom since the 1990s. This rise is particularly surprising as standard grammars of the Zulu language do not generally acknowledge occurrences of **ku-** in the environments as have been set out in this article. (As has been pointed out, these are environments that are largely similar to those that the variant **kwi-** occurs in.)
- Except for the 1980s, at all times the locative prefix **ku-** has a higher frequency of use (in the environments just referred to) than the locative prefix **kwi-**.

The locative prefix *ku-* in environments where the variant *kwi-* would have been expected

Another interesting phenomenon that needs to be examined, is the distribution across time of the usage of the locative prefix **ku-** in environments where **kwi-** would have been expected in terms of the normal morphophonological

**Figure 4:** Distribution across the noun types of the locative prefix **ku-** in environments where the variant **kwi-** would have been expected

rules of the language. Consider Table 18 and Figure 6 in this respect.

When reading Table 17 and Figure 5, together with Table 18 and Figure 6, the following can be observed regarding the relationship between the locative prefixes **kwi-** and **ku-** as found in similar environments:

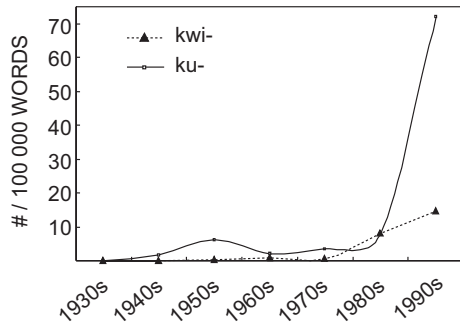
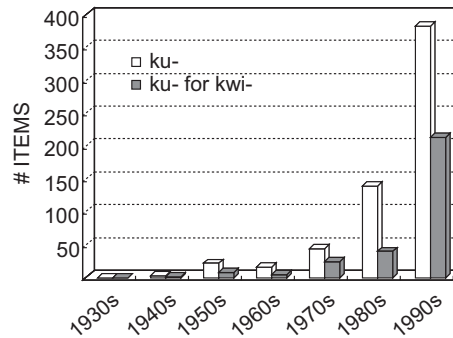
- It would seem that during the 1970s there was resistance to the use of **kwi-**, possibly because it was regarded as constituting Xhosa influence, with possibly a concomitant

Table 16: A sample of concordance lines for the locative prefix **ku-** in environments where the variant **kwi-** would have been expected

nze phezulu ukugcola okuphuma	kumisebenzi	yokumba emayini. Kumele sehlukaniselane
aqaphele kakhulu ngokutshala	kundawo	okutshalwe kuyo imithi. Lokhu kungaba yibang
uthi wenze isenzo esiyinhlaba	kusizwe	sonke samaZulu. Sigcwelegcwele: Ndabezitha.
ezi zingane zikwazi ukukhetha	kuzinto	ezilandelayo: ukungasheshi zingenele izindab
ncike kakhulu enkolweni kanye	nakunkambiso	yosikompilo lwesintu, lumumethe imfundi

Table 17: Distribution across the decades of the locative prefixes **kwi-** and **ku-**

Period	# words	# kwi-	# kwi- / 100 000 words	# ku-	# ku- / 100 000 words
1930s	142 896	0	0.00	0	0.00
1940s	156 492	0	0.00	3	1.92
1950s	348 150	1	0.29	22	6.32
1960s	795 610	7	0.88	16	2.01
1970s	1 242 635	7	0.56	44	3.54
1980s	1 785 876	144	8.06	139	7.78
1990s	529 797	77	14.53	382	72.10
Total	5 001 456	236	4.72	606	12.12

**Figure 5:** Distribution across the decades of the locative prefixes **kwi-** and **ku-****Figure 6:** Distribution across the decades of the locative prefix **ku-** in environments where the variant **kwi-** would have been expected**Table 18:** Distribution across the decades of the locative prefix **ku-** in environments where the variant **kwi-** would have been expected

Period	# words	# ku-	# ku- for kwi-	% ku- for kwi-
1930s	142 896	0	0	0.00
1940s	156 492	3	1	33.33
1950s	348 150	22	8	36.36
1960s	795 610	16	5	31.25
1970s	1 242 635	44	24	54.55
1980s	1 785 876	139	40	28.78
1990s	529 797	382	213	55.76
Total	5 001 456	606	291	48.02

increase in the use of **ku-**. The latter also seems to be evinced by the high incidence (54.55%) of **ku-** being used in the 1970s in environments where **kwi-** would have been expected according to the normal morphophonological rules of the language.

- The frequency of use of **kwi-** and **ku-** is very similar during the 1980s, which could explain why the degree of substitution of **ku-** in environments where **kwi-** would have been expected, is markedly lower (28.78%) than in

the previous decade. In this regard, it is also significant that the frequency of use of **kwi-** increased a massive fourteen-fold (as compared to the 1970s) during this time.

- During the 1990s, the frequency of use of **kwi-** and **ku-** again diverge with **ku-** being used five times more often than **kwi-**, coupled again with a significant increase (to 55.76%) of instances where **ku-** is used in environments where **kwi-** would have been expected.

The distribution of the locative prefixes **kwi-** and **ku-** across the sub-corpora

Another of the stated aims of this article is to ascertain how the occurrences of the locative prefixes are distributed across the different sub-corpora — these sub-corpora mainly being stratified according to genre. Consider Table 19 and Figure 7 in this regard.

Based on Table 19 and Figure 7, it can be concluded that:

- The locative prefixes **kwi-** and **ku-** (as used with largely similar types of nouns) are found least often in ‘conservative’ environments such as religious texts, and most often in modern environments such as pamphlets and Internet files. The use of **kwi-** and especially **ku-** in the latter genre is telling indeed.
- This first conclusion supports the finding (reached in the previous section) that these locative prefixes are increasingly being used — not only with brand names, loanwords and nouns signifying new scientific and techno-

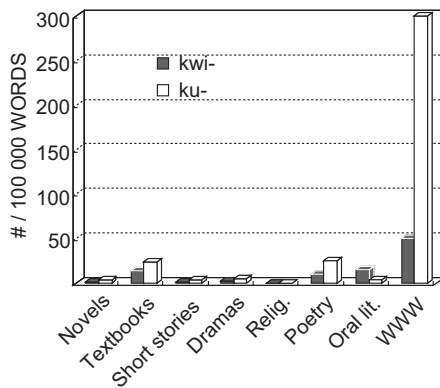


Figure 7: Distribution across the sub-corpora of the locative prefixes **kwi-** and **ku-**

logical concepts, but also with nouns expressing a wide range of other (non-technical) meanings.

- The locative prefix **ku-** is used much more frequently than **kwi-** for all genres. There is only one exception, i.e. for the genre oral literature, folklore and legends, yet here we must observe that PZC might be skewed towards the work of one author in this regard. Indeed, as 96% of the occurrences of **kwi-** in this genre come from the work of the same author, further study will be required.

The locative prefix **ku-** in environments where the variant **kwi-** would have been expected

It is again interesting to study those instances where **ku-** is used where one would have expected **kwi-**, this time as distributed across the sub-corpora. This is done in Table 20 and Figure 8.

Upon studying Table 20 and Figure 8, one arrives at one main conclusion:

- Throughout the genres, in roughly half of the cases **ku-** is used where one would have expected **kwi-**. For the genre poetry (and possibly oral literature, folklore and legends — yet cf. note on skewing above) **ku-** replaces **kwi-** even more often, while for the genre short stories, essays and readers this replacement is less frequent.

The occurrence of the locative prefixes **kwi-** and **ku-** as a replacement of **e-/o-...-ini**

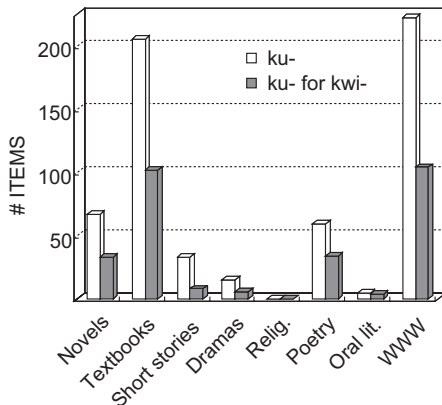
In Table 17 and Figure 5 we have seen that the locative prefixes **kwi-** and **ku-** are increasingly used over the years, with the use of **ku-** even showing an exponential growth. As was pointed out earlier, these locative prefixes are used

Table 19: Distribution across the sub-corpora of the locative prefixes **kwi-** and **ku-**

Sub-corpus	# words	# kwi-	# kwi- / 100 000 words	# ku-	# ku- / 100 000 words
Novels, Novelettes	2 098 836	11	0.52	67	3.19
Textbooks	889 175	124	13.95	205	23.06
Short stories, Essays, Readers	834 019	12	1.44	33	3.96
Dramas, One-act plays	357 073	6	1.68	15	4.20
Religious texts	343 842	0	0.00	0	0.00
Poetry	238 407	22	9.23	59	24.75
Oral literature, Folklore, Legends	166 262	24	14.44	5	3.01
WWW files, Pamphlets	73 842	37	50.11	222	300.64
Total	5 001 456	236	4.72	606	12.12

Table 20: Distribution across the sub-corpora of the locative prefix **ku-** in environments where the variant **kwi-** would have been expected

Sub-corpus	# words	# ku-	# ku- for kwi-	% ku- for kwi-
Novels, Novelettes	2 098 836	67	33	49.25
Textbooks	889 175	205	102	49.76
Short stories, Essays, Readers	834 019	33	8	24.24
Dramas, One-act plays	357 073	15	6	40.00
Religious texts	343 842	0	0	0.00
Poetry	238 407	59	34	57.63
Oral literature, Folklore, Legends	166 262	5	4	80.00
WWW files, Pamphlets	73 842	222	104	46.85
Total	5 001 456	606	291	48.02

**Figure 8:** Distribution across the sub-corpora of the locative prefix **ku-** in environments where the variant **kwi-** would have been expected

instead of the so-called ‘discontinuous’ locative morpheme **e-/o-...-ini** in the derivation of locatives from nouns from classes other than 1/2, 1a/2a and [+human] nouns in class 6. There are thus two valid additional research questions one could ask. Firstly, one might be interested to know whether or not the use of **e-/o-...-ini** in these environments also increases with time, and if so, how. Secondly, one might wish to know whether or not the use of **ku-** and **kwi-** as a replacement of the **e-/o-...-ini** locativisation strategy increases over the years, stays equal or perhaps even decreases.

Broad frequency indications in the form of *samples* from PZC will suffice to answer these additional questions. (As was indicated in the introduction to this article, space considerations do not allow for an *exhaustive* study of the **e-/o-...-ini** strategy.) Three different decades were sampled: the 1940s, the 1950s and the 1990s.

For those decades the ‘normalised’ number of occurrences of the **e-/o-...-ini** strategy is shown in the last column of Table 21, and graphically in Figure 9.

The sample decades shown in Table 21 and Figure 9 clearly indicate the trend, namely that also the **e-/o-...-ini** strategy is increasingly used over the years. From Figure 9 one sees that the trend is *linear*.

The key question now is whether or not, over the years, the use of **kwi-** and **ku-** versus **e-/o-...-ini** gains ground, stays equal or perhaps even decreases in those instances where **kwi-** and **ku-** replace **e-/o-...-ini**. Formulated differently, one wants to find out how the growth rates of the two strategies compare. This question is answered in Table 22 and Figure 10.

Table 22 and Figure 10 clearly indicate that, over the years, the use of the locative prefixes **kwi-** and **ku-** as a replacement of the **e-/o-...-ini** locativisation strategy increases *exponentially*. This contrasts sharply with the linear growth of the **e-/o-...-ini** strategy itself (cf. Figure 9). There can thus be no doubt that locativisation by means of the class 17 prefix **ku-** (and its variant **kwi-**) is a rising strategy in Zulu, being increasingly applied in the derivation of locatives from nouns from classes other than 1/2, 1a/2a and [+human] nouns in class 6, i.e. in environments where the use of the locative affixes **e-/o-...-ini** would have been expected.

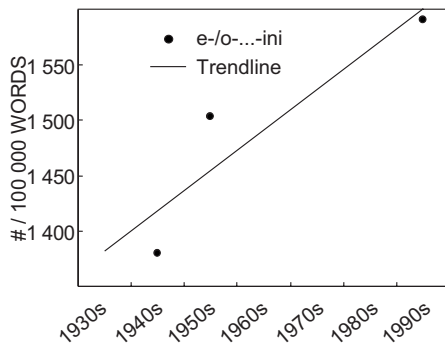
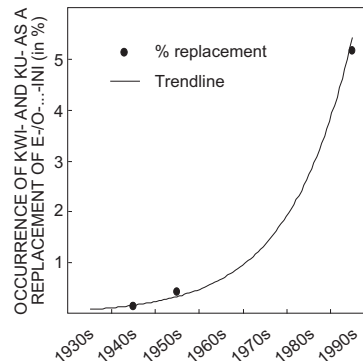
The locative prefix **ko-**

In PZC, 119 instances were found where the variant **ko-** of the class 17 prefix was used to derive locatives from nouns in class 2. A random sample of those is shown in Table 23.

We further found that **ko-** is used as an

Table 21: Distribution of the occurrence of **e-/o-...-ini** across three sample decades

Period	# words	# e-/o-...-ini	# e-/o-...-ini / 100 000 words
1940s	156 492	2 159	1 379.62
1950s	348 150	5 234	1 503.37
1990s	529 797	8 431	1 591.36

**Figure 9:** Distribution of the occurrence of **e-/o-...-ini** across three sample decades**Figure 10:** Distribution of **kwi-** and **ku-** as a replacement of the **e-/o-...-ini** strategy across three sample decades**Table 22:** Distribution of **kwi-** and **ku-** as a replacement of the **e-/o-...-ini** strategy across three sample decades

Period	# e-/o-...-ini + # kwi- + # ku-	# kwi- + # ku-	% replacement
1940s	2 162	3	0.14
1950s	5 257	23	0.44
1990s	8 890	459	5.16

Table 23: A sample of concordance lines for the locative prefix **ko-** affixed to nouns in class 2

uzobuza ngamathuba emisebenzi **kobahlobo** bakhe ababesebenza kwezinye izindawo. Waja aJohane kwavela ezulwini noma **kobantu** na? Ngiphenduleni." Bakhulumisana bodwa, bat elakubo nemfundiso ayeyithole **kobashokobezi** ababephenduka izimpimpi nxa sebebajwe o nasezindaweni ezinewozawoza **kobavakashi** basemadolobheni. Izitolo zazigcwele zich ya, kukhishwe izingane ekhaya **nakobakhelwane** ziyimemeze. Zazimpompoloza kudume ama

alternative to **ku-** to derive locatives from nouns in class 2, in the ratio of approximately 1:20. This is significant, as in the existing literature only Nyembezi (1956) explicitly refers to the usage of this locative morpheme in Zulu.¹¹ As not one of the modern grammarians writing on the Zulu language mentions this phenomenon, this gives rise to questions regarding the previous and current status of this variant in Zulu. By basing a linguistic study on a structured corpus such as PZC, previously under-emphasised and overlooked features of a specific phenom-

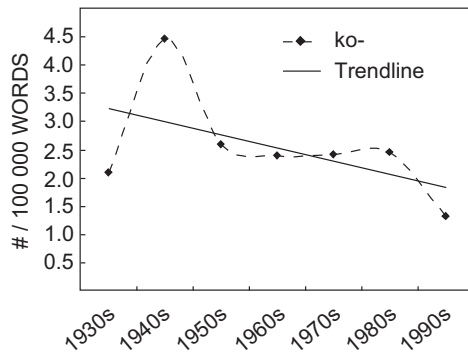
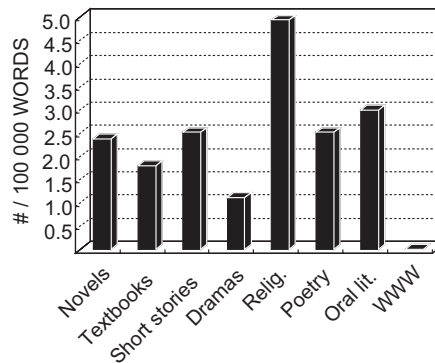
enon can be brought to light and questions such as these might be answered.

One can firstly study the distribution of the variant **ko-** across time. Table 24 and Figure 11 summarise the results of the corpus queries in this regard.

From Table 24 and Figure 11 one must conclude that the use of **ko-** shows a very stable pattern across the decades represented in PZC. This is especially true when compared to the use of **kwi-** and **ku-** across those same decades. Further, the 1940s experienced an

Table 24: Distribution across the decades of the locative prefix *ko-*

Period	# words	# <i>ko-</i>	# <i>ko-</i> / 100 000 words
1930s	142 896	3	2.10
1940s	156 492	7	4.47
1950s	348 150	9	2.59
1960s	795 610	19	2.39
1970s	1 242 635	30	2.41
1980s	1 785 876	44	2.46
1990s	529 797	7	1.32
Total	5 001 456	119	2.38

**Figure 11:** Distribution across the decades of the locative prefix *ko-***Figure 12:** Distribution across the sub-corpora of the locative prefix *ko-*

88% higher-than-average use of *ko-*, and the 1990s a decline, with a 45% lower-than-average use. Looking at the general trend, however, one sees that the variant *ko-* falls into disuse.

Secondly, when it comes to the distribution of *ko-* across the various sub-corpora, the data have been summarised in Table 25 and Figure 12.

From Table 25 and Figure 12, one must conclude that the use of *ko-* is relatively evenly-spread across the various genres. It is how-

ever significant to observe that the variant *ko-* occurs most often in a more 'conservative' genre such as religious texts, and is not found at all in the most modern sub-corpus containing pamphlets and Internet files.

Based on the results obtained from PZC, it can thus be concluded that the reasons why the variant *ko-* of the class 17 locative prefix is only mentioned in one grammar, and — significantly — one written by a mother-tongue speaker of Zulu at that, are:

Table 25: Distribution across the sub-corpora of the locative prefix *ko-*

Sub-corpus	# words	# <i>ko-</i>	# <i>ko-</i> / 100 000 words
Novels, Novelettes	2 098 836	50	2.38
Textbooks	889 175	16	1.80
Short stories, Essays, Readers	834 019	21	2.52
Dramas, One-act plays	357 073	4	1.12
Religious texts	343 842	17	4.94
Poetry	238 407	6	2.52
Oral literature, Folklore, Legends	166 262	5	3.01
WWW files, Pamphlets	73 842	0	0.00
Total	5 001 456	119	2.38

- This variant alternates with **ku-** in the formation of locatives in class 2 in the ratio of approximately 1:20. This could therefore be a contributing factor to it not being discussed in all but one of the Zulu grammars.
- The variant **ko-** has apparently fallen into disuse over time. This can be seen from the trendline in Figure 11, and may also be deduced from the fact that none of the pamphlets or Internet files (all of which stem from the late 1990s) contain even a single occurrence of this variant.

On the other hand, the variant **ko-** has always been there, i.e. throughout the decades, and has always been everywhere, i.e. throughout the genres (except for the very modern ones). From that point of view, omitting this locative strategy from a grammatical description is definitely an important oversight. Hence, the results presented in this section illustrate how a well-designed electronic corpus can highlight a previously overlooked phenomenon. In addition, querying such a corpus may give an interesting perspective on how that hitherto overlooked phenomenon changed with time and varies across the genres.

Conclusion

In this article we have shown how a structured electronic Zulu corpus can be used to highlight previously under-emphasised and overlooked aspects of a seemingly well-documented language feature such as the class 17 locative prefix **ku-**. The study was based on PZC — an organic 5-million-token Zulu corpus built at the

University of Pretoria — that has been organised chronologically, and that consists of a number of sub-corpora stratified according to genre. With PZC we were in a position to reach conclusions regarding not only the frequencies with which the variants **ku-**, **kwi-** and **ko-** of the class 17 locative prefix are used (e.g. respectively 606, 236 and 119 overall in 5 million words) but also regarding the possible changes in use that these prefixes have undergone with time.

The biggest changes arrived in the 1990s. On the one hand, the use of **ku-** (and to a lesser extent **kwi-**) grew exponentially, while the use of **ko-** fell into disuse. On the other hand, and concurrently, the use of **ku-** (and to a lesser extent **kwi-**) is many times more frequent in the sub-corpus containing the genres pamphlets and Internet files than in any other sub-corpus, while **ko-** does not occur in this sub-corpus at all. The variations across time and genre are summarised in Figures 13 and 14.

The other main findings of this study are:

- Not only is the locative prefix **kwi-** widely used in Zulu in the locativisation of loanwords and nouns signifying brand names and scientific and technical concepts, but well over a third of all the examples of nouns locativised by way of the prefix **kwi-** constitute 'normal' Zulu words that are neither loanwords, brand names, etc., nor do they signify previously unknown (new) scientific and technical concepts. The latter finding is an aspect of the distribution of the locative prefix **kwi-** that has gone largely *unnoticed* in the manifold existing descriptions of Zulu grammar.

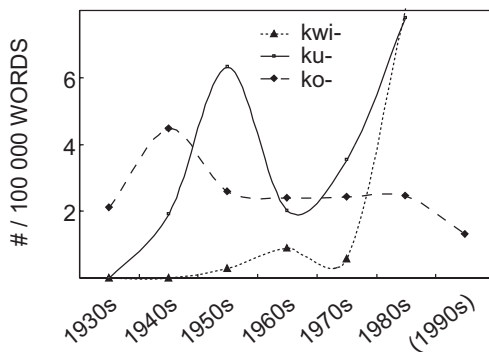


Figure 13: Summary of the distribution across time of **kwi-**, **ku-** and **ko-** (prior to the 1990s for **kwi-** and **ku-**)¹²

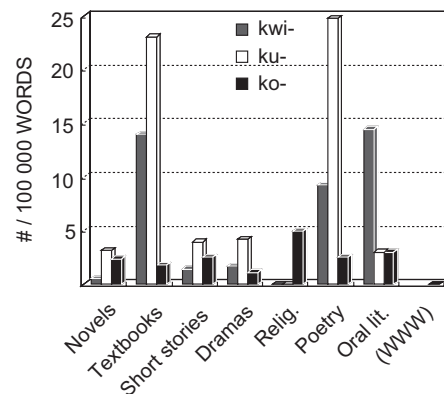


Figure 14: Summary of the distribution across the genres of **kwi-**, **ku-** and **ko-** (without taking pamphlets and Internet files into account for **kwi-** and **ku-**)¹³

- The locative prefix **ku-** is used much more widely with nouns *other* than those in classes 1/2, 1a/2a and [+human] nouns in class 6 than is usually indicated in Zulu grammars and other sources on the structure of the language.
- The locative prefix **ku-** is frequently used with nouns commencing on the preprefix **i-**, i.e. in environments where one would expect the variant **kwi-** to be found due to the morphophonological rule **ku- + i- > kwi-** which usually governs the juxtaposition of vowels in this environment in Zulu. Roughly half (48.02%) of the total number of occurrences found of **ku-** (with nouns other than those in classes 1/2, 1a/2a and [+human] nouns in class 6) are in environments where the variant **kwi-** would have been expected. For 'normal' Zulu words the use of **ku-** instead of **kwi-** even climbs to a staggering 69.62%. This might be indicative of Zulu writers' desire to try and stem the 'encroachment' of **kwi-**.
- Locativisation by means of the class 17 prefix **ku-** (and its variant **kwi-**) is an ascendant strategy in Zulu, being increasingly applied in the derivation of locatives from nouns from classes other than 1/2, 1a/2a and [+human] nouns in class 6, i.e. in environments where the use of the locative affixes **e-jo-...-ini** would have been expected.
- The locative prefix **ko-** alternates with **ku-** in the formation of locatives in class 2 in the ratio of approximately 1:20, and has furthermore apparently fallen into disuse with time. It is concluded that this probably explains why this prefix is not discussed at all in the majority of Zulu grammars. Yet, this does not alter the fact that omitting this locative strategy from a grammatical description is an important oversight.

Notes

- ¹ Since this article is being submitted for publication in a South African journal, necessary sensitivity with regard to the term 'Bantu' languages is exercised in our choice rather to use the term *African* languages. Keep in mind, however, that the latter includes more than just the 'Bantu Language Family'.
- ² The sizes of the corpora built at UP's Department of African Languages are in constant evolution. For the latest developments, we refer the reader to the Home Page of *ELC for ALL* (Electronic Corpora for African Language Linguistics): <http://www.up.ac.za/academic/libarts/afriLang/elcforall.htm>
- ³ Cf. Galloway (2001) for an in-depth discussion of the statistical trends in South African book publishing during the 1990s.
- ⁴ 'Readers' are school textbooks consisting of short sections of reading and literary material — often little stories. That is why we grouped them with 'short stories' and 'essays'. Further, 'short stories' and 'essays' usually appear together in one volume, i.e. half short stories, half essays.
- ⁵ 'Oral literature' refers to traditional Zulu literature such as praise poems, songs, sayings, etc. that have been documented, and is therefore not to be confused with a spoken (sub-)corpus.
- ⁶ Assimilation takes place in the case of this example, i.e. **ku-mi(na) > kiji** where the vowel **u** in the locative prefix **ku-** assimilates to the vowel **i** in the pronoun **mi(na)**. This also happens in the case of the forms **ku-thi(na) > kithi** 'to, from, at, etc. us' and **ku-ni(na) > kini** 'to, from, at, etc. you (plural)'.
- ⁷ Although Canonici states that (his) class 9a nouns may take **e-...-ini**, it must be noted that, like other loanwords, the nouns belonging to this class often take only **e-**, as in: **ebhayisikobho** 'in, at, to, etc. the cinema'; **ejele** 'in, to, etc. jail'; **ehhotela** 'in, at, to, etc. the hotel'.
- ⁸ Observe that a Zulu locativised noun such as **kobafana** 'to, from, at, etc. the boys', can only be derived from a form such as **ku-abafana**, and not (even if the morphophonological rules of the language should allow it) from e.g. a form such as **kwa-abafana** 'at the place, village, etc. of the boys', as the class 17 locative possessive concord **kwa-** 'at the place, village, etc. of' cannot be used with nouns in class 2. (This phenomenon is mentioned by Msimang and Shamase (1988: 110) and is also borne out by the corpus.)
- ⁹ For more information on *WordSmith Tools*, we would like to refer the reader to the Home Page of its creator, Mike Scott, at <http://www.lexically.net> (or its mirror site: <http://www.liv.ac.uk/~ms2928>).
- ¹⁰ It must be borne in mind that loanwords (in the singular form) are usually accommodated in class 5 in Zulu which has the prefix **i(li)-**. This means that affixing the locative prefix **ku-** to nouns such as these, will result in the sound change **ku- + i- > kwi-**.

¹¹ Wanger (1917: 213) mentions in a footnote that the (locative) prefix **ku-** often appears as **ko-** in Zulu, but he does not elaborate on this.

¹² Note concerning Figure 13: The **kwi-** and **ku-** values for the 1990s (14.53 and 72.10 respectively) cannot be successfully represented on the same graph. Hence, only the

ko- value (1.32) is shown for the 1990s.

¹³ Note concerning Figure 14: The **kwi-** and **ku-** values for the pamphlets and Internet files (50.11 and 300.64 respectively) cannot be successfully represented on the same graph. Hence, only the **ko-** value (0.0) is shown for the pamphlets and Internet files.

References

- Atkins BTS, Clear J & Ostler N.** 1992. Corpus design criteria. *Journal of Literary and Linguistic Computing* 7(1): 1–16.
- Canonici NN.** 1990. Subclasses of Zulu nouns. *South African Journal of African Languages* 10(2): 52–58.
- Colenso JW.** 1903. *First Steps in Zulu. Being an Elementary Grammar of the Zulu Language.* Pietermaritzburg: Vause, Slatter & Co.
- De Schryver G-M & Lepota B.** 2001. The lexicographic treatment of days in Sepedi, or when mother-tongue intuition fails. *Lexikos* 11: 1–37.
- De Schryver G-M & Prinsloo DJ.** 2000a. The compilation of electronic corpora, with special reference to the African languages. *Southern African Linguistics and Applied Language Studies* 18(1–4): 89–106.
- De Schryver G-M & Prinsloo DJ.** 2000b. Electronic corpora as a basis for the compilation of African-language dictionaries, Part 1: The *macrostructure*. *South African Journal of African Languages* 20(4): 291–309.
- De Schryver G-M & Prinsloo DJ.** 2000c. Electronic corpora as a basis for the compilation of African-language dictionaries, Part 2: The *microstructure*. *South African Journal of African Languages* 20(4): 310–330.
- Doke CM.** 1955. *Zulu Syntax and Idiom.* London: Longmans, Green & Co.
- Doke CM.** 1973⁶. *Textbook of Zulu Grammar.* Cape Town: Longman Southern Africa (Pty) Limited.
- Engelbrecht JA.** 1956. *Zoeloe-leerboek met Oefeninge, Leesstukke, Woordelys en 'n Aantal Raaisels en Spreekwoorde.* Johannesburg: Voortrekkerpers Beperk.
- Galloway F.** 2001. Statistical trends in South African book publishing during the 1990s. (Paper presented at a conference on Colonial and Postcolonial Cultures of the Book organised by the Institute for the Study of English in Africa (ISEA), Rhodes University, Grahamstown, 6–8 August 2001.)
- Gauton R.** 1995. *The Zulu Locative — A Diachronic Perspective.* D.Litt. thesis, University of Pretoria, Pretoria, South Africa.
- Grégoire C.** 1975. *Les locatifs en bantou* (Annales - Série in-8° - Sciences Humaines - n° 83). Tervuren: Royal Museum for Central Africa.
- Grout L.** 1893. *Grammar of the Zulu Language.* London: Kegan Paul, Trench, Trübner & Co. Ltd.
- Madiba MR.** 1998. Translating into African languages: the problem of terminology. In: Kruger A, Wallmach K & Boers M (eds) *Language Facilitation and Development in Southern Africa.* Pretoria: SATI. pp. 63–68.
- Mdhladhla GCS & Mbata AHS.** 1933. *A Zulu Manual for Native Primary Schools.* Pietermaritzburg: Shuter & Shooter.
- Msimang CT & Shamase S.** 1988. *IsiZulu Sezikhuthali. Ibanga 7. New Syllabus 1989.* Pretoria: De Jager-Haum.
- Mtintsilana PN & Morris R.** 1988. Terminography in African languages in South Africa. *South African Journal of African Languages* 8(4): 109–113.
- Nyembezi CLS.** 1956. *Uhlelo LwesiZulu.* Pietermaritzburg: Shuter & Shooter.
- Poulos G.** 1985. Typological trends in South-Eastern Bantu. *South African Journal of African Languages* 5(1): 17–23.
- Poulos G & Bosch SE.** 1997. *Languages of the World / Materials 50.* München: Lincom Europa.
- Poulos G & Msimang CT.** 1998. *A Linguistic Analysis of Zulu.* Cape Town: Via Afrika Limited.
- Prinsloo DJ & De Schryver G-M.** 2001a.

- Corpus applications for the African languages, with special reference to research, teaching, learning and software. *Southern African Linguistics and Applied Language Studies* 19(1–2): 111–131.
- Prinsloo DJ & De Schryver G-M.** 2001b. Monitoring the stability of a growing organic corpus, with special reference to Sepedi and Xitsonga. *Dictionaries: Journal of The Dictionary Society of North America* 22: 85–129.
- Renouf A.** 1987. Corpus development. In: Sinclair JM (ed) *Looking Up, An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London: Collins ELT. pp. 1–40.
- Roberts C.** 1909. *The Zulu-Kafir Language. Simplified for Beginners*. London: Kegan Paul, Trench, Trübner & Co. Ltd.
- Růžicka KF.** 1959. Locative formations in Bantu languages. *Archiv Orientální* 27(2): 208–250.
- Růžicka KF.** 1960. Locative formations in Bantu languages. *Archiv Orientální* 28(2): 181–219.
- Samuelson RCA.** 1925. *Zulu Grammar*. Durban: Knox Printing & Publishing Co.
- Schreuder HPS.** 1850. *Grammatik for Zulu Sproget*. Christiania.
- Sinclair JM** (ed). 1987. *Collins COBUILD English Language Dictionary*. London: Harper Collins Publishers.
- Stuart PA.** 1932. *A Zulu Grammar for Beginners*. Pietermaritzburg: Shuter & Shooter.
- Summers D.** 1993. Longman / Lancaster English Language Corpus — criteria and design. *International Journal of Lexicography* 6(3): 181–208.
- Taljaard PC & Bosch SE.** 1988. *Handbook of isiZulu*. Pretoria: JL van Schaik Ltd.
- Trew R.** 1994. The development of training models for African language translators and interpreters. In: Kruger A (ed) *New perspectives on teaching translators and interpreters in South Africa*. Pretoria: UNISA. pp. 73–102.
- Van der Poll JDP.** 1970. *Die Lokatief in Xhosa en die Ander Nguni-tale: 'n Deskriptiewe en Vergelykende Analise*. MA thesis, University of Stellenbosch, Stellenbosch, South Africa.
- Van Eeden BIC.** 1956. *Zoeloe-grammatika*. Stellenbosch: Die Universiteitsuitgewers en Boekhandelaars, (Edms.) Beperk.
- Wanger PW.** 1917. *Konversations-Grammatik der Zulu-Sprache*. Mariannhill: St. Thomas Aquins Druckerei.
- Wilkes A.** 1993. 'n Nuwe hipotese oor die herkoms van die lokatiefprefikse e-/o- in die Ngunitale. In: *Linguistica. Festschrift E.B. van Wyk. 'n Huldeblyk*. Pretoria: JL van Schaik Ltd. pp. 172–180.
- Ziervogel D.** s.d. [1970?] *Lokatiefvorming met Pre- en Suffikse in die Bantoetale*. D.Litt. thesis, University of Pretoria, Pretoria, South Africa.
- Ziervogel D.** 1971. The Bantu locative. *African Studies* 30(3–4): 371–384.
- Ziervogel D, Louw JA & Taljaard PC.** 1976. *A Handbook of the Zulu Language*. Pretoria: JL van Schaik Ltd.