Orature and Morpholexical Deconstruction as Lexicographic Archaeological Sites: Some Implications for Dictionaries of African Languages

Francis Matambirofa, Department of African Languages and Literature, University of Zimbabwe, Harare, Zimbabwe (francism@compcentre.uz.ac.zw)

Abstract: This article takes a multidisciplinary approach to African lexicographic practice. It has as its primary premise the assumption that without words there can be no dictionaries to compile and discuss. Owing to this fact, the article focuses on a specific strategy for collecting words which belong in a special category of their own, namely archaic or obsolete words. Such words are important because of the need to mark them as such in any general purpose dictionary. Most, if not all dictionaries of African languages seem not to have this category, giving the misleading impression that there are no such words in African languages. Apart from digging up archaic or what we have also referred to as artefact words, the article also argues that the words have a substantial and intrinsic etymological value. Thus they can be used in specialised etymological dictionaries of African languages or even in standard general dictionaries. The multidisciplinary aspect resides in the methodology proposed for the recovery of archaic words. It is considered necessary that disciplines such as oral literature, oral history, historical linguistics and to a limited extent theoretical linguistics and computational linguistics, and, symbolically, archaeology and lexicography itself, be brought to bear on the subject of inquiry. The article is also an attempt at working out a method which African lexicographers can employ as an instrument to dig up artefact words for etymological and other such purposes. Hopefully the method can be refined further.

Keywords: Orature, Morpholexical, Deconstruction, Reconstruction, Archaeological Sites, Lexicography, African Languages, Dictionaries, Diachronic, Synchronic, Historical Linguistics, Proverbs, Riddles, Idioms, Oral Traditions, Mnemonic, Etymology, Archaic, Shona, Deverbative Nouns, Verb Root, Semantic, Artefact Words

Opsomming: Mondelinge literatuur en morfoleksikale dekonstruksie as leksikografiese argeologiese terreine: 'n Aantal implikasies vir woordeboeke van Afrikatale. Hierdie artikel het 'n multidissiplinêre benadering tot die leksikografiese praktyk in Afrikatale. Dit het as primêre uitgangspunt die veronderstelling dat dit sonder woorde onmoontlik is om woordeboeke op te stel en te bespreek. Op grond van hierdie feit fokus die

Lexikos 11 (AFRILEX-reeks/series 11: 2001): 38-59

artikel op 'n spesifieke strategie vir die versameling van woorde wat tot 'n kategorie van hul eie behoort, naamlik argaïese of verouderde woorde. Sulke woorde is belangrik vanweë die noodsaaklikheid om hulle as sulks te merk in enige woordeboek vir algemene gebruik. Die meeste, indien nie alle woordeboeke van Afrikatale nie skyn nie hierdie kategorie te hê nie, wat die misleidende indruk wek dat daar nie sulke woorde in Afrikatale is nie. Behalwe om argaïese, of waarna ons ook as artefakwoorde verwys het, op te diep, word daar ook in hierdie artikel geredeneer dat die woorde 'n wesenlike en werklike etimologiese waarde het. Daarom kan hulle in gespesialiseerde etimologiese woordeboeke van Afrikatale of selfs in algemene standaardwoordeboeke gegee word. Die multidissiplinêre aspek lê in die voorgestelde metodologie vir die terugvind van argaïese woorde. Dit word as noodsaaklik beskou dat vakgebiede soos mondelinge literatuur, mondelinge geskiedenis, historiese linguistiek, in 'n beperkte mate teoretiese linguistiek en rekenaarlinguistiek, en, simbolies, argeologie en leksikografie self, aangewend word by die onderwerp van ondersoek. Hierdie artikel is ook 'n poging om 'n metode uit te werk wat leksikograwe van Afrikatale as 'n instrument kan gebruik om artefakwoorde op te diep vir etimologiese en ander doeleindes. Hopelik kan die metode verder verfyn word.

Sleutelwoorde: Mondelinge Literatuur, Morfoleksikaal, Dekonstruksie, Argeologiese Terreine, Leksikografie, Afrikatale, Woordeboeke, Diachronies, Sinchronies, Historiese Linguistiek, Spreekwoorde, Raaisels, Idiome, Mondelinge Tradisies, Mnemoniese Woorde, Etimologie, Argaïes, Sjona, Deverbatiewe Substantiewe, Werkwoordelike Stam, Semanties, Artefakwoorde

Introduction

This article deals with a sphere, the supporting evidence of which, chiefly coming from oral literature, is seldom the concern of linguistics. The broader sphere to which it belongs, is historical linguistics. The primary concern of this article is to develop an apparatus with which to extract and recover aspects of the diachronic state of Shona from its synchronic dimension for the benefit of African lexicography in general. Such a desire is motivated by the fact that Shona, in comparison with other Bantu languages, has a relatively short period of literacy from which to study its diachronic state. One could, broadly speaking, accord written Shona a little less than a century. While much effort to reduce the language into a written code was evinced by missionaries shortly before, but mainly after colonisation in 1890², this period is too short in terms of the scope of interest of this article.

Owing to this relatively short period of the written form of the language, we propose to use some disciplines of oral literature (henceforth orature) as lexicographic archaeological sites from which we will, in spite of the attendant limitations, glimpse into the past with the aim of extracting archaic words for lexicographic purposes. It is important to mark specified entries as archaic in Shona lexicography as well as in that of other African languages because this aspect of lexicography is erroneously being ignored. The only possible reason

for not paying attention to this category is to falsely assume that African languages are not changing and are therefore not casting off old words and usages which are worthy of the lexicographer's notice. But if they are, which is in fact the case, then African lexicographic practice must equally conform with and/or reflect standard international lexicographic practice where archaic words, just like other categories of entries, are indeed marked according to their status in different types of dictionaries.

Having noted the above, we hope to illustrate the point of the article from the major oral genres which are proverbs, riddles and idioms. The term archaeology, here used metaphorically, is borrowed from the branch of knowledge which deals with the cultural remains of a people. Using the term in its denotative sense, Finegan (1989: 255) points out:

Archaeological records enable researchers to make educated guesses about where our ancestors come from and where they migrated to, as well as how they lived and died.

The quest to glimpse into the past state of Shona, among other things, already suggests the lack of written records of the language, equally stretching back in the past as say English to, for instance, the times of Chaucer and Shakespeare. Benefiting from the long period of literacy of the English language, Finegan (1989: 252) is able to tell us:

About a thousand years ago, the English verb starve (Old English steorfan) meant simply "die" (by any cause); today starve refers principally to deprivation and death by hunger (or, by metaphorical extension, deprive of affection.)

Different writers (Fromkin and Rodman 1993, Finegan 1989, O'Grady et al. 1989, Aitchson 1995, etc.) all agree that all languages change imperceptibly over long periods of time. Referring directly to the importance of the written records of a language for historical research, Fromkin and Rodman (1993: 320), for instance, stress: "Many of the changes are revealed when languages have written records." Needless to say, Shona, alongside its sister languages, does not, as mentioned earlier, possess such written records from which such changes can be revealed.

On the Use of Orature or Oral Traditions

The study of the African past with *inter alia* the corroboration of oral records has become somewhat of an academic tradition in its own right. Many disciplines in the humanities take recourse to it in varying degrees. Included in these are history, anthropology, literature, cultural studies and to some extent archaeology. To these should be added linguistics and especially historical linguistics.

Before we look at how orature and oral traditions can benefit historical linguistic studies, it is necessary to define these terms. Orature can be described as a body of literary knowledge present in living traditions, the defining features of which comprise, among others, proverbs, tales, riddles, aphorisms, dictums, idioms etc. The term "oral tradition" is largely used in the domain of history. Vansina as quoted by Beach (1994: 251) defines it as "verbal messages which are reported statements from the past beyond the present generation ... spoken, sung or called out". Arguing from a purely historical point of view, Beach (1994: 244) informs us:

Oral traditions can be used to good effect to recover the pre-colonial past, providing that one does not expect too much.

A definite thread therefore links orature from the aspect of literary research and oral traditions to the study of history. The difference, if worth anything, is that oral traditions may only be viewed as specialised abstractions from the gamut of orature that is largely concerned with a people's history. In other words, oral traditions are a subgenre of orature and they therefore logically take on the broad defining aspects of orature itself.

Here we have established the characteristics of orature and it is therefore necessary to indicate its relevance to what we want to achieve. It should already be obvious that if indeed oral traditions can be put to "good effect to recover the pre-colonial past" then surely, apart from the recovery of historical and literary material, we can, *inter alia*, recover from it lexical materials for diachronic linguistic analysis. This is so because, after all, is it not the verbal and therefore linguistic base that conveys the literary material which is orature? In recovering the past from oral traditions or orature heed must however be taken of Beach's earlier cautionary comments where he says, "providing that one does not expect too much".

What to Recover and How to Recover it?

In this section, we want to give a brief outline of the following aspects on which this article will focus, viz.:

- (i) what we particularly aim to recover;
- (ii) the subgenres of orature targeted for the recovery procedure;
- (iii) the methodology of morpholexical deconstructive recovery;
- (iv) the methodology of reconstructive recovery; and
- (v) a cross-linguistic comparative analysis of lexical items.

From a linguistic point of view, we hope that lexical items especially can best be recovered *vis-à-vis* syntactic or phonological structures. We argue that the easiest to recover are unattached or freestanding lexemes as opposed to mor-

pholexically attached ones. For example, freestanding lexical items are **mwana** (child), **n'ombe** (cattle), **imba** (hut), **tsime** (well), etc. A host of lexical items are often connected in morpholexical word-formation processes such as derivation and inflection. In agentive nominal lexical items such as **musvuti** (smoker), **murimi** (father), **muroti** (dreamer), **muroyi** (witch) and **muurayi** (killer), we can recover related lexical items belonging to the grammatical category of verbs. A recovery demonstration will be given on two of these when we embark later in the discussion on its methodological aspect.

Fromkin and Rodman (1993: 322) point out that "no part of the grammar remains the same over the course of history". This indicates that an historical linguist can therefore look into the diachronic state of the language by means of a fairly long agenda dealing with phonology, syntax, morphology, semantics and pragmatics among others. The limiting factor can only be whether or not a particular area of interest is practically recoverable. Not all areas of linguistic interest can be studied with the same degree of success. Take for instance phonology, it is doubtful to what extent the "historical" tones on Shona historical names such as Munhumutapa, Changamire, Gatsi Rusere and many others can accurately be captured. There may indeed be no conjecture on the part of the present speakers that the tone they assign to different segments may after all be historically incorrect because there is no clue as to how the said names were pronounced in ancient times.

Moreover, it is not every aspect or subgenre of orature that passes for a productive, let alone plausible "linguistic archaeological site". With this in mind, our major aim is to unearth aspects of the language that specifically pertain to the lexical domain. It should therefore follow that we will identify subgenres of orature that have an inherent "lexical fossilisation" and "lexical preserving" tendency. Some such subgenres of orature are proverbs, riddles, idioms, incantations, songs, children's games and others. We shall, for our purposes in this article, however, largely explore proverbs. For illustrative purposes proverbs alone can, in our opinion, adequately stand in the stead of riddles and idioms because the latter two differ from the former mainly in terms of the proverb's extended length relative to that of riddles and idioms. It should, however, be henceforth assumed that where mention is made of proverbs, it shall implicitly be inclusive of riddles and idioms as well, unless the contrary is stated.

The respective "lexical or linguistic archaeological sites" that have been designated above, have the capacity to yield archaic words. We do not, however, work with the assumption of the *exclusive middle* that a word can only be either archaic or current. Rather, we take the pragmatic position that the lexicon can be put on a word-currency scale or continuum ranging from obsolete to current. Archaic words, on the one hand, would reside on the lower reaches of such a continuum, while words in current usage, on the other hand, would be found on its upper reaches. We shall explain, under methodology, the criteria determining the extraction of a particular word from say a proverb or a rid-

dle as a potential archaic word. We shall explicate methods through which we arrive at such an assertion. It is, at this juncture, necessary to answer the question why proverbs, riddles and idioms have been chosen as "lexical archaeological sites".

On the Choice of Proverbs, Riddles and Idioms

Although proverbs, riddles and idioms are categorically different in their form, they are, however, conceptually complementary in so far as they are all repositories and therefore carriers of ancient wisdom and lore. It is because of this that we shall, as indicated earlier, only concentrate on characterising the proverb, the characterisation of which, by extrapolation, shall be assumed to refer equally to riddles and idioms, at least in terms of their oral transmission and linguistic fixedness.

Bukenya (1994: 36-37) correctly observes that the term proverb has been defined "in various ways by different scholars". Some of those he lists are Kipury, Finnegan, Nandwa and Bukenya as well as Sunkuli and Miruka. Bukenya resists giving another definition but instead synchronises these various definitions to come up with what may be called the characteristic features of this genre. We shall abstract from these the characteristics of the proverb that best suit the argument of this article. The prominent features of proverbs show them as:

- short, terse or brief, and
- relatively invariable/fixed.

Bukenya (1994: 40) further points out that proverbs are "small poems". Poems, by their very nature, have an inherent mnemonic feature which supports our delineation of them as recorders of the archaic words of a language.

Pongweni (1989: 4) agrees with the view of the linguistic fixedness of the proverb when he makes the assertion regarding the "frozen character of the syntax of the proverb". He also quotes Hasan-Rokem who remarks that the proverb "bears many marks of a tight linguistic entity". Hasan-Rokem aptly captures the nature of the proverb as follows:

Even more than the idiom, the proverb is a fixed word-order not subject to inflectional changes according to linguistic context.

Lastly, the historicity of the proverb cannot be in doubt as most writers on the subject have emphasised the "time-immemorial" aspect of its existence. This comes out quite unmistakably through the characterisation of proverbs as "messages coded by tradition and transmitted from generation to generation" (Pongweni 1989: 5). Bukenya (1994: 12) describes the historicity of myths in a

manner that also applies to other genres of orature, proverbs included, when he writes:

Myth, to illustrate our argument, behaves like time: it has a tendency to be chronologically transcendental and can therefore defy the boundaries of historical time.

Proverbs, riddles and idioms as well as other subgenres of orature can be studied from many dimensions according to the research interests of the student or scholar. In respect of the above-mentioned characteristics and also from the viewpoint of linguistic research, the proverb or the riddle is most suited for what we hope to achieve because each of them relies for its transgenerative transmission on: (i) its relative linguistic fixity over long passages of time; (ii) its shortness which in turn renders it (iii) highly mnemonic due to its poetic form; (iv) its "chronologically transcendental tendency" owing much to its social regulatory function as well as (v) the respect accorded to griots, sages and all people of sound thinking. This by itself can act as a worthy incentive to carry it on in the people's collective memory over numerous generations.

The case we are making in respect of orature as a window into the past is, however, quite modest. We will indeed not claim, and neither can anyone else, to be able to use proverbs, idioms and other aspects of orature for reconstructing ancient Shona or any other Bantu language, dating say a thousand years back. It would be expecting too much from orature, especially if empirically better-placed disciplines such as archaeology, as Bourdillon (1987: 3) asserts, "can yield ... very limited [information] in the absence of literary deposits". Beach (1994: 21) provides some demarcation back in time regarding the reliability of Shona traditions as historical sources. In this regard, he (Beach 1994: 273) makes the following comments:

I can sum up by claiming that Shona oral traditions give us a reasonable basis for a history of the Zimbabwean plateau, but one going back only to about 1700 and often not as far.

There is therefore a limit to which one can successfully recover the past via orature. This should not, however, be misconstrued to be an underestimation of the relative wealth of information, especially in the absence of any written records, that one can still unearth, as will be demonstrated shortly, from this source.

On the Method of Recovery³

In an article of this kind, the ways along which we arrive at certain conclusions must be shown. This is necessary largely for three reasons. The first is that the elucidation of our methods enables independent judgment of the suitability of the outlined method. The second and corollary reason is that, should our proposed methods be judged correct, they can then be applied, not only to Shona, but also to other Bantu languages. The third and final reason is that if the proposed procedure is not sound, perhaps different and more effective ways can be found to achieve the same goal of extracting diachronic lexemes for the benefit of African lexicographic practice.

We adopt, as a broad methodological principle, what Katamba refers to as *discovery methods*. Katamba (1993: 23) makes some incisive remarks with regard to these methods when he notes:

The main principle used in the analysis of words is the principle of contrast. We contrast forms that differ (i) in phonological shape due to sounds used and (ii) in meaning broadly defined to cover both lexical meaning and grammatical function.

As indicated earlier, there are basically two targets from which we hope to obtain the desired lexical items. The first of these are free-standing lexemes as opposed to those we will disentangle from constructions of a higher order, such as verb phrase complexes and larger nominal phrases. Examples of such free-standing words have already been given from synchronic Shona. We will now proceed to explain each of these lexical targets in the following sections.

Our task is to justify the selection of certain words at the exclusion of others. Potential words are first selected from any relevant source, such as a book of Shona proverbs, riddles or idioms on the basis of being unfamiliar to the language archaeologist or lexicographer. Consider for instance the following proverbs from Hamutyinei and Plangger (1974: 394, 436 and 437 respectively):

- (a) Manga chena inoparira parere nhema
 (A white pumpkin betrays where the black one lies)
 - (b) **Sonzi** rakatuma nzou (An ant sends an elephant [on an errand])
 - (c) Muromo **chiponza** unopotsera (A mouth is like a knobkerrie which you can throw)

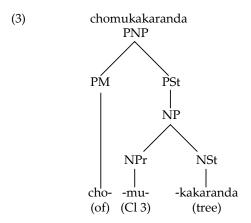
In (1) above, the words in bold type, manga, sonzi and chiponza, are unfamiliar to the writer of this article. After having identified such words, we propose, as the next step, that the existing dictionaries of Shona⁴ are checked. If the word is present, this might be an indication that the word is not necessarily archaic or this might be a suggestion that the word is becoming obsolete. It is also necessary to check the word with other speakers whose dialects may be different from that of the researcher's. The word can also be checked against the ALLEX Shona database and corpus.⁵ Failure to find the word in the corpus technically qualifies it for the designation "archaic". If, on the other hand, the word is in the corpus but showing a low frequency, i.e. after running a concordance of the word, then that word also qualifies as one becoming obsolete. If, however, the

particular word only occurs in the context of that proverb or riddle, then it can be considered archaic and can therefore be marked as such in the dictionary.

Apart from free-standing vocabulary items, we also find words that are part of larger constructions occurring in proverbs, riddles, idioms and other subgenres of orature. The following riddles contain such cases of potentially archaic words (see Pongweni and Chiwome 1995: 7, 46, 47 respectively):

- (2) (a) **Zifuridzo** mukombe une buri (The rumbling instrument (which is like) a ladle with a hole)
 - (b) Chinguri ndakazvarwa handisati ndamboona **bwabwa rechibvimo** (Since I was born, I have never seen the token of betrothal of my mother)
 - (c) Ndagumburwa nechitsiga **chomukakaranda** (I knocked my foot against a mukakaranda tree stump)

In (2)(a) **zifuridzo** is a noun belonging to class 21 as marked by the **zi-** prefix. From the main form of such a noun we can abstract **furidzo** as a potential archaic word. In (2)(b) and (c) potentially archaic words are encapsulated in possessive phrases. The possessive phrases **bwabwa**⁶ **rechibvumo** and **chomukakaranda** contain within them lexemes on which our interest is focused. In (3) below, we, by a binary deconstruction methodology, disentangle the noun of interest from the higher possessive phrase complex down to its respective morphemic components.



Through the above process, we have been able to isolate the target lexeme **mukakaranda**, a type of tree, from the PNP within which it was bound. This process is a function of the language type to which Shona belongs. From the five language types which Katamba lists⁷, it is evident that Shona is an agglutinating or agglutinative language. Katamba (1993: 57) describes an agglutinating language as a "kind of language [where] there tends to be a more or less

one-to-one matching of morphemes with morphs". He gives Turkish as a typical agglutinative language and in passing mentions Luganda as well. Being a Bantu language like Shona, Luganda belongs, according to Katamba's classification, to the agglutinative type, which, by inference, marks Shona as being agglutinative too. It is important to note that the kind of binary morpholexical deconstruction as demonstrated in (3) above is occasioned by the language type to which Shona belongs.

We also argue in this article that it is possible, in reverse order, to recover from the synchronic state of the language, aspects that cast some meaningful light on its diachronic state, i.e. in so far as the lexicon is concerned. Note that here, unlike what we posited earlier, we do not particularise orature as a "lexical archaeological site". In this particular argument, we posit that deverbatives can be a rich source of verbs that may no longer be in use although they may possibly have been, at one stage in the history of the language, commonly used.

It should be noted that our technique of morpholexical deconstruction corresponds with Dale's observation regarding the obvious derivational origins of deverbative nouns in Shona. Dale (1981: viii) clearly spells out how the Shona view the nature of deverbative derivational lexemes when he says in the introduction to his bilingual dictionary *Duramazwi: Shona-English Dictionary*:

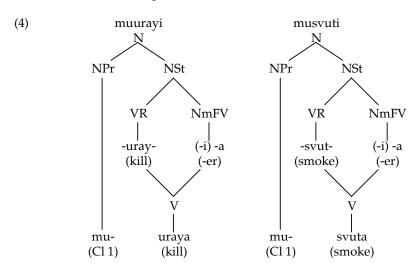
Deverbative nouns (e.g. *mubiki* from *-bika*) normally do not have the verb of origin suggested. This is because to the Shona speaker the relationship is very obvious.

We consider such lexical rederivations with the same belief as that of English speakers in the historical existence of words such as *sweveninges* (dreams), *lesynges* (lies), *cinn* (chin), *cirice* (church), *cuman* (come) as recorded from Chaucer's and Shakespeare's literary works (cf. O'Grady et al. 1989: 189 and 191) regardless of the paucity of evidence from the present state of English that such words in fact ever existed.

In the preceding we followed Dale's lead in respect of disentangling deverbative nouns to their minimal morphemic constituents. After having done so, we then "clinically" extracted and thereby recovered the seminal or nucleic verb root. Consider the verbal nouns **muurayi** (killer) and **musvuti** (smoker) in (4) below.

The branching diagram is a morpholexical deconstruction and disentanglement of a basic lexical form which is the underived (or seminal) root, being in this case **uraya** and **svuta** respectively and the agentive nominal forms derived thereof, which are **muurayi** and **musvuti**. The underived forms are what Katamba (1993: 92) refers to as "the kernel of the word". Note that in the derivation process the prefixal morpheme, **mu-** in either case, and the final vowel **-i**, are both viewed as the nominalising agent triggers (hence they are glossed as **-er**). There is no doubt, however, that **mu-** is relatively more significant in its nominalisation input to the verb, at least in so far as it assigns the class of the "new" word *vis-à-vis* the vowel, which is introduced phonologically

into the word-formation process.



It would not be necessary to draw elaborate branching diagrams for the range of deverbatives that fall within this category. What has been done in (4) methodologically suffices for the point we are making. In (5) below, we show a random list of lexical items situated within such a deverbative derivational morpholexical paradigm.

(5)	Deverbative Noun	Gloss	Verb Root	Original Verb	Gloss
	musika	stirring stick	-sik-	sika	stir
	chibhako	moulding stick	-bhak-	bhaka	mould sadza
	chigaro	chair	-gar-	gara	sit
	murimi	farmer	-rim-	rima	farm/plough
	musevenzi	worker	-sevenz-	sevenza	work
	mufemberi	prophet	-fember-	fembera	prophesy/guess
	chisvusvuro	breakfast	-svusvur-	svusvura	eat breakfast

The above nouns fit into the following verb-to-noun derivational schema:

(6) Cl Morpheme + VR + Nominal FV \Rightarrow Deverbative Noun

All the deverbative nouns in (5) are obvious synchronic examples that are part of the competence module of current speakers of Shona. The morpholexical deconstruction process that is worked out in (4) rederives, in reverse order, a verb from a noun. As has already been explained, the kernel for **muurayi** (killer) is the verb root **uray-** (kill). We are of the opinion that this process can be

applied to "deverbative" nouns that have lost their corresponding verb roots serving as a technique through which historical/archaic verbs can be defossilised. In a way, we basically reverse the process in (6) and this yields (7) whose input schema is provided below.

(7)
$$-$$
 Cl Morpheme $+$ VR $-$ Nominal FV \Rightarrow Diachronic Verb

It is evident from this input schema that our concern above is largely with the suspected diachronic or what we may call the diachronic "seminal" verb that has been fossilised⁸ within the synchronic deverbative noun. We may use a less controversial illustration such as the deverbative noun **mupakato** (handle) to illustrate this. A morpholexical deconstruction of this noun is done within the paradigm of (7), and it is represented as (8) below.

Note that this particular "rederived" verb is not necessarily diachronic as such. It is only employed for illustrative purposes. Speakers of Shona may agree with the writer of this article that the verb <code>pakata</code> (hold) is almost exclusively associated with a handle. Instead, in other contexts, speakers of Shona would use the verb <code>bata</code> (touch or hold) which is clearly close to <code>pakata</code>, both in terms of form and attested Bantu sound shifts. The verb <code>pakata</code> belongs almost solely within the context of the proverb <code>Kunzi pakata</code>, <code>hakuzi kunzi ridza</code> (To be asked to hold [a whip] is not the same as being asked to crack it). The occurrence of <code>pakata</code> in this proverb and seldom in other contexts, gives credence to our argument which singles out proverbs and other genres of orature as repositories of archaic words.

The efficacy of our methodology can be illustrated with lexical examples that are within the competence of synchronic native speakers of the language. Here we propose to analyse words intermediate in terms of archaic status, e.g. words that are within the bounds of "incontrovertible" etymological recoverability such as those given in (9) below:

(9) (a) mupakato (handle) (b) nhovo (shield) (c) museve (arrow) (d) pfumo (spear)

First it must be shown that all these lexical items are deverbative nouns since this is the *sine qua non* for their subjection to the morpholexical deconstruction procedure as outlined in (7). The next step, therefore, is to run the deverbative nouns through this procedure in order to arrive at their verb kernels. This is done in (10) below:

(10)		- Cl Morpheme	+ VR	- Nominal FV	\Rightarrow	Diachronic Verb
	(a)	– mu	+ pakat	- o	\Rightarrow	pakata
	(b)	- ∅	+ tav	- o	\Rightarrow	tava
	(c)	– mu	+ sev	– e	\Rightarrow	seva
	(d)	- ∅	+ pfum	- o	\Rightarrow	pfuma

The "recovered" verb **pakata** in (10)(a) has already been discussed. What is interesting to note, however, is that apart from **mupakato** (handle), the nouns **bakatwa** (knife), **mubatwa/nhapwa** (slave), **ubatwa** (slavery), **mubati** (hard worker) and **mbato** (tongs) are etymologically all derivatives of the verb **pakata**⁹.

All the rederived verbs in (10) which we give here in their infinitive mode as **kutava** (to shield/protect), **kuseva** (to dip into soup/relish), and **kupfuma** (to sew/pierce) are still current. The point of the argument is that, by a logical extension, this methodology can successfully be applied to deverbatives whose corresponding nucleic verbs are no longer in current use. However, with regard to the other rederived verbs in (10), we argue that **nhovo** (shield) is derived from the verb **tava** (to shield) and that the deverbative noun is phonologically different from the kernel verb owing to morphophonemic changes that are quite consistent and widespread with class 9 and other classes, such as 10 and 5¹⁰ which display similar morphophonological behaviour. We argue that nouns such as **teve** (protective charm), **nhovo** (shield), **tavira** (protect a home or person with charms against evil spirits) and **nhovo**¹¹ (protector) may have been derived semantically and etymologically from the kernel verb **tava**. To this end, we argue that the nouns **nhova** (fontanelle) and **dovo** (animal hide) are figuratively, though not literally, related to the verb **tava**.

Verbs such as **seva** and **pfuma** may not be so difficult to explain owing to the current sense of the verb **seva** which refers to dipping say a morsel of $sadza^{12}$ into soup or relish in the process of eating. Etymologically nouns such as **musuva** (morsel), **usavi** (relish), **museve** (arrow), **musevi** (?deeper — as in soup/relish), **museve** (sun's rays), **museve** (fast runner)¹³ as well as **seva** (to follow closely behind), it is argued here, are very closely related. There is a defining featural sense of \pm enter/or touch with (or be touched by) the tip or front end of some instrument. The notation \pm refers to the object as having a property of entering such as **musuva** and **museve**, or showing a property of being entered such as **usavi**, for instance.

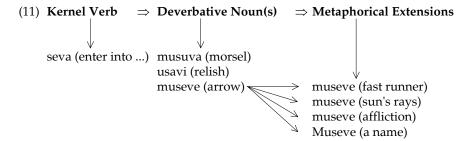
The other verb that has been rederived using the morpholexical deconstruction procedure as outlined in the schema in (7), is **pfuma**. We have already argued that **pfuma** as a verb means to pierce, i.e. in the infinitive mode. The noun **pfumo** from which we have derived **pfuma** could therefore be translated into English as a piercer, although the correct rendition of it is a spear. The sense of pierce is etymologically blended and extended both literally (i.e. denotatively) and abstractly (i.e. idiomatically) into the following words:

pfuma (sew), **pfumo** (spear), **pfumo** (army), **pfumo** (serious predicament) and **Mapfumo** (name). It is necessary to comment on the surfacing of different though related shades of meaning ascribed to a single diachronic kernel verb.

Semantically and phonologically the correlation between what we have referred to as kernel verbs and their derivatives do not have a mathematical one-to-one correspondence. In other words, the rederived verbs tava, pfuma and seva and their corresponding derivatives such as nhova, nhovo, pfumo, pfuma, museve and musuva do not have a readily discernible semantic and to some extent phonological correspondence. Traugott (quoted by Heine et al. 1991: 14) explains this phenomenon of semantic fading resulting from linguistic change when he makes the following observation:

Meanings [with time] tend to come to refer less to objective situations and more to subjective ones (including speaker point of view), less to the described situation and more to the discourse situation.

We argue that meaning changes from "objective situations" to "subjective ones" are evident in a continuum of semantic extensions which proceed from the verb to derived nouns and from the deverbative nouns to metaphors. In (11) below we illustrate this phenomenon using the rederived verb **seva**.



The diagram in (11) is largely self-explanatory as it is an illustration of semantic shifts from the verb **seva** to its derivative forms. **Seva** can be marked as the point of "objective situations" within the semantic shift continuum. The deverbative noun **museve** assumes the role of a semantic bridge across which metaphorical and therefore "subjective" meanings pass. In this process the metaphorical word retains its form. This does not only apply to **museve** in (11) but also to other nouns such as **nhovo** and **pfumo**.

It has already been said that language change is a multiple process. Among these changes are phonetic processes as well. Such processes are responsible for the phonological differences that sometimes occur between earlier and later forms of the words, apart from the differences that are cognates of the same phonological form. This is evident even in languages with a long literacy history such as, for instance, English. The following English

historical words and their corresponding synchronic forms already referred to are a case in point:

 $\begin{array}{ccc} (12) \ sweveninges & \Rightarrow & \text{dreams} \\ \ lesynges & \Rightarrow & \text{lies} \\ \ cinn & \Rightarrow & \text{chin} \\ \ cirice & \Rightarrow & \text{church} \end{array}$

Unquestionably it is difficult to decipher the phonological correspondences between the diachronic and synchronic forms of the words given in (12). It may in fact take somebody with a knowledge of English etymology rather than any ordinary speaker of the language to unravel the correspondences. This similarly applies to Shona. Any semblance of a relationship between tava and its derivatives such as **nhova**, teve, tavira, **nhovo**, and the names Matavire and Manhovo may not readily be seen. This may be analogous to the relationship that binds genetically related languages, for instance Bantu languages. To say that XiTsonga, Zulu, KiSwahili, Sotho and Shona are all Bantu languages, does not mean that speakers of these languages understand one another. Similarly, the relatedness of elements internal to a language may not be evident at first sight, due to what Heine et al. (1991: 15) describe as "adaptation, erosion, fusion, and loss" with respect to phonetic change.

The Cross-linguistic Comparative Aspect¹⁴

In a procedure such as the one we are advocating, one of the greatest frustrations is a lack of adequate feedback. Unearthing possible obsolete words and verbs in a language with no written records can easily be ridiculed as a worthless undertaking since there will be no records against which to check the authenticity of the findings. Trust in the methodological stringency corroborated (when one is fortunate) with cross-linguistic comparative evidence is sometimes the only proof on which to rely. Consider the examples in (13) below:

(1	3)	
---	---	----	--

DN	Gloss	-ClM	+VR	- Nm FV	⇒ DV	Gloss
chibako	tobacco pouch	– chi	+ bak	- o	⇒ baka/paka	put into
musika	stirring stick	– mu	+ sik	– a	⇒sika	spin/stir
muromo	mouth	– mu	+ rom	- o	⇒roma	bite
mugoti	cooking stick	– mu	+ got	– i	⇒gota	?cook
mugwaku	ladle	– mu	+ gwak	– u	⇒ gwaka	?to ladle
bakwa	stake of wood	$-\varnothing$	+ bak(w)	– a	⇒ baka/paka	stake/put into15
mbiko	hiding place	$-\varnothing$	+ bik	- o	⇒ bika	hide
muvi	bad person	– mu	+ v	– i	⇒vi	bad ¹⁶
muuya	good person	– mu	+ uy	– a	⇒uya	good ¹⁷
shomodzi	messenger	-Ø	+ shomodz	– i	⇒shomodza	announce ¹⁸

Most, if not all the verb artefacts, are decidedly unfamiliar to current speakers of Shona¹⁹. The recovered items can be a source of archaic words as well as a source of etymological derivatives of current words — information which is important in the dictionary-making process. The following pairs of verbs, for instance, make an interesting comparison:

In each case the "diachronic verb" is given first, followed by the current verb which is separated from the former by the notation ← which means "has changed to the following synchronic form".

Both the adjectives **-vi** and **uya** have more or less been lexicalised and fossilised within the grammatical noun category as used in (13). This further applies to **-vi** in **chivi** (sin) and **Chivi** (name of both a chieftainship and district of Masvingo province)²⁰, and to **uya** in **muuya** (good person). It is interesting to note that the same adjectives have also been fossilised in the idiom **mavi nemauya** (good and bad [words]). These words are almost obsolete outside the few usages that have been mentioned. It is on account of this that we technically qualify such words as archaic. Adjectives, as Fortune (1980) observes, are a long series of stems in that they are semantically compatible with a wide range of the nominal affixes of the noun classes²¹. The nominal lexicalisation and concomitant fossilisation of these adjectives are proof enough that they may be survivors from some past linguistic epoch.

As stated above, one way of getting feedback would be to check one's inventory of lexical artefacts with related languages. The need for such a procedure arises from the rationale that related languages might in some instances still be using "parent" or prototypical forms of words that may, however, have become synchronically obsolete in others, and vice versa. This is for instance the case with Shona and Nguni as well as Sotho-Tswana languages with regard to locatives. Meinhof gives Ur-Bantu locative prefixes respectively as pa-, kuand mu-. These locative prefixes or their cognates are by and large no longer productive in Nguni and the Sotho-Tswana languages. They have, at most, become lexicalised in locative nouns such as phezu (at/on the top of), phansi (down/underneath) and fatshe (down), hodimo (above), and morao (at the back)22. However, the same locative prefixes are highly productive in Shona as is evidenced by the existence of fully fledged noun classes with prefixes: 16 pa-, 17 ku- and 18 mu-. These locative prefixes, unlike in Nguni, are active and productive to a point where literally any other word, save for its intrinsic semantic constraints vis-à-vis a given locative morpheme, can be locativised through the usual preprefixing morpholexical process.

In the table under (13) there are some artefact verbs whose possible dia-

chronic existence can be proved from the suggested cross-linguistic perspective. The adjective **-vi**²³ (bad/evil) can be taken as an illustrative case. We have already said that **-vi** is a Shona adjective stem that has been rendered redundant and only survives lexicalised in fossil form in a few words as given above. On the contrary, this adjective stem, rendered as **-bi** in Nguni orthography, is highly productive in languages such as Ndebele and Zulu among others. The adjective phrases in (15) below are ample witness to this fact.

(15) Ndebele/Zulu Phrase²⁴ Gloss

umfazi omubi ugly/evil woman umoya omubi bad/evil heart indlu embi bad/ugly house umganu omubi ugly plate isitulo esibi bad/ugly stool ilizwe elibi bad country

It is clear from what has been shown above that cross-linguistic comparisons can sometimes be taken recourse to in order to validate findings in a related language. It is interesting to note that in the case of the above adjective **-bi**, Seto and Seto (1997: 476) argue that it is the same form as that which occurs in Papuan Austronesian languages of Southeast Asia whose forms of the adjective are given in (16):

(16)	Kewa	wae
	Malas	pi
	Pila	fiwi
	Moere	fia
	Mawak	fi:wi
	Hinihon	fi:wi
	Zulu	-bi

The argument of Seto and Seto (1997: 195) to justify the above comparison, which they in fact also apply to other different languages of the world, is as follows:

Chronologically significant lexical affinities [of different languages of the world] indicate the migratory derivation of the words for village life from a common ancestral homeland in tropical Mainland Southeast Asia.

The major obstacle to such a procedure, however, is the relative paucity of artefact words that are cross-linguistically attestable through a methodological paradigm such as the one performed in (15) and (16) above. If, however, Seto and Seto's claim is anything to go by, then one can perhaps look for cross-linguistic comparative evidence from much futher afield.

Conclusion

The foregoing discussion has been an attempt to harness methodologies from some different though related disciplines in the humanities for the advancement and enrichment of lexicography. Some of the disciplines which have been found useful for our purposes are historical linguistics, oral history and traditions as well as oral literature. All these have been identified as rich and specialised areas in which African lexicographers can excavate for archaic and etymologically significant words to be included in both general and specialised dictionaries.

We have in the foregoing discussion dwelt on archaic words and how to gather them since they are part of the raw material needed for the dictionarymaking process. We have delineated orature as a symbolically rich linguistic archaeological site. This is so because, as we have demonstrated, orature acts as a preservative of archaic word forms, due not only to its syntactic fixedness, but also due to its mnemonic proclivity and important cultural significance. We have, however, specifically singled out proverbs, riddles and idioms as some of the subgenres of orature that easily lend themselves to the procedure under discussion. Archaic words emanating from these sources can be marked in the same way as those in any dictionary. It is instructive to note that in the lexicographic traditions in contemporary African languages in Southern Africa, both in terms of the proliferating monolingual and the traditional bilingual dictionaries, lexicographers are quick to mark entries as loan-, taboo or offensive words, but very few dictionaries mark archaic or obsolete entries²⁵. Shall we, therefore, on the basis of this observation, arrive at the fallacious conclusion that African languages, unlike all other known languages, have no archaic or obsolete words?

Verbs specifically have been our main source of exemplification owing to their inherent capacity for prolific deverbative morpholexical noun formation processes. While the synchronic deverbative noun may, by historical fortuity, have lasted up to the present day, this study has, by implication, argued that the diachronic kernel or nucleic verb itself could over many generations have become obsolete to a point where it only remains preserved as a grammaticalised morpholexical fossil. A method of morpholexical deconstruction and disentanglement has been proposed. The efficacy of this method in extracting and recovering especially artefact verbs has been demonstrated. A similar method can be developed for other word forms depending on their idiosyncratic inflectional and derivational behaviour. In this regard, a method of dealing with possessive noun phrases has been suggested in (3). Here we have explored the methodology of verbal disambiguation by morpholexical deconstruction.

It has also been shown in the discussion that by employing the morpholexical deconstruction process, we can possibly recover an artefact verb or word to which other synchronic words can etymologically be traced back. This has been shown to be possible in (10) and (11). There are, however, cases in

which certain artefact words may be well beyond the scope of etymological recoverability. It has been suggested that a cross-linguistic comparative study may be undertaken as a way of ascertaining whether etymologically irrecoverable artefacts may be found in a related language or other languages. This was illustrated in (15) and (16). We have justified such an undertaking by observing that while certain diachronic word forms may have altogether vanished from the synchronic state of a target language, they may probably be found in the synchronic form of a sister language or, as Seto and Seto seem to suggest, any other language. Examples giving credence to such an exercise have been adduced by comparing Shona and Nguni as well as the Sotho-Tswana group of languages. Evidence for the cross-linguistic use of the adjective stem -bi has been seen to occur in some Papuan Austronesian languages. To this end, it has been demonstrated that, while the locative prefixes are no longer as active in the Nguni and Sotho-Tswana group of languages, they are nevertheless still very active and highly productive in the Shona language. The reverse was found to be true in respect of the adjective stem -bi, which is still highly active in Nguni while it has fallen dormant in Shona. Artefact word forms that are supported by the facts from related or other languages for that matter, may in all probability be authentic and should therefore not be dismissed as a figment of the historical linguist's or etymologist's imagination.

Notes

- I wish to acknowledge the valuable contributions and comments, some of which I have incorporated, that were made when I initially presented a version of this article as a paper in the Linguistics Department of the University of Zimbabwe in 1999. I single out the following for special mention: Mr Liveson Tatira, Dr Carolyn Harford, Prof. Norris Dembetembe, Mr Jairos Kangira, Mrs Hakani Moyo and Dr Ann Jefferies.
 - Apart from the contributions received from those named above, I want to mention the fact that a version of this article was also read as a paper during the Fourth International Conference on Lexicography (5-6 July 1999) that was organised by AFRILEX and hosted by the University of Pretoria in the Republic of South Africa. I note with pleasure the remarks of the keynote and opening address delivered by Prof. A.C. Nkabinde where he, among other issues, challenged Bantu lexicographers to explore the importance of historical linguistics in lexicographic work. I take the present effort as one of the many dimensions through which historical linguistics can fruitfully be brought to bear upon African lexicography. I acknowledge with thanks the useful questions, remarks and encouragement that I received from among the audience during my presentation. Apart from that, I extend my sincere thanks to two anonymous referees who made useful suggestions in terms of how to improve the general import of the article, some of which have been gladly incorporated into its final version.
- Rhodesia (the present-day Zimbabwe) was colonised by the British through Cecil John Rhodes's British South Africa Company.
- In this section, the following abbreviations will be used: PNP: Possessive Noun Phrase; NPr: Noun Prefix; NmFV: Nominalising Final Vowel; NSt: Noun Stem; VR: Verb Radical; N:

- Nominal; V: Verb; Cl: Class Affix followed by Gender Number; PM: Possessive Marker; NP: Noun Phrase; PSt: Possessive Stem; DN: Deverbative Noun; DW: Derived Word; Cl Morpheme: Delete Morpheme.
- 4. Existing Shona dictionaries include the following: Chimhundu et al. (1996), Dale (1981), Hannan (1981) and Barnes (1932).
- 5. ALLEX is the acronym for African Languages Lexical Project, which is a dictionary-making undertaking in partnership with collaborators from the universities of Oslo in Norway and Gothenburg in Sweden. The first monolingual Shona dictionary, *Duramazwi ReChiShona*, was compiled by the ALLEX Project. The ALLEX Shona corpus has at present over two million running words.
- 6. Note that **bwabwa**, though not our main focus for explanation by virtue of it being a standalone noun, is a word no longer in common use, if used at all.
- 7. Katamba gives the following as the five known language types: (i) analytic, (ii) agglutinating, (iii) inflecting, (iv) incorporating, and (v) infixing languages.
- 8. As we discuss the aspect of fossilisation, it is instructive to note that Heine et al. (1991: 15) discuss grammaticalisation as an aspect of language change. Here they state that "this process [of change] affects all levels of language structure". They distinguish between functional, morphosyntactic, and phonetic processes. In their discussion of the morphosyntactic which interests us, they outline the following as morphosyntactic processes of language change, i.e. permutation, compounding, cliticization, affixation and fossilisation (the emphasis is mine).
- 9. To this list we can also add mbati, for one caught red-handed, as in the idiom Mbavha imbati (A thief is caught in the act). This mbati, a deverbative as Shona speakers may agree, is exclusively used in the context of this idiom.
- 10. Fortune (1980) argues that diachronically class 5 had an overt syllabic prefix ri- but synchronically it is no longer overtly expressed except in a few cases. He further argues that the disappearance of the ri- morphemic prefix triggered off the wide-spread morphophonemic changes associated with nouns that fall in this class such as the systematic nominal stem onset consonant change g < k , b < p, f < bv and j < ch found between the singular and the plural forms respectively. The morphophonemic changes in classes 9 and 10 are associated with the influence of nasalisation which is quite common in these classes.</p>
- 11. Here nhovo (protector) is used as a metaphor, especially within the context of the Christian faith where God is often metaphorically referred to as such. Nhova (fontanelle) and dovo (animal hide) may possibly be figuratively related to the verb tava (shield). We would like to leave it to the reader to arrive at their own judgement on this matter.
- 12. Sadza is the staple food of much of Southern Africa. It is prepared from corn meal and made fairly thick so that it can be shaped it into a ball. Its name varies according to the particular language: it is called nsima in ChiChewa, isithwala in Ndebele and stywepap in Afrikaans.
- 13. Museve, as in sun's rays and fast runner, are taken from Chimhundu et al. (1996: 299).
- In this section, the following additional abbreviations will be used: DN: Deverbative Noun;
 Cl M: Class Marker/Prefix; DV: Diachronic Verb/Rederived Verb.
- 15. Note that **bakwa** is implicationally related to, if not the same as, **chibako**.
- 16. -vi is decidedly not a verb root, but an adjective stem.
- 17. Uya is the opposite of -vi.

- 18. Shomodza, we infer, can mean to announce or to deliver a message.
- Most speakers of Shona do not agree with this claim. Some, however, begin to see some resemblances only after being given certain explanations.
- Masvingo, one of the major provinces of Zimbabwe, lies within the south and south western zone of the country.
- 21. Officially Shona has 21 noun classes but technically they are actually 20. Class 20 was originally assigned to some Kalanga nouns. This provision became redundant after the recognition of Kalanga as a separate language from Shona, therefore leaving Shona with 20 and not 21 noun classes.
- The locatives phezu and phansi are taken from Taylor (1994) and the locatives fatshe, hodimo and morao from Moeketsi (1994).
- 23. Note that the difference between **-vi** and **-bi** is only orthographic while they are phonologically and semantically identical. Shona orthographises the phoneme through the spelling **-vi** while the other languages which were examined, use the form **-bi**.
- 24. I would like to thank Ms. Karren Dube (of the Department of African Languages and Literature, University of Zimbabwe) who made some valuable spelling corrections to the words in the table
- 25. I have verified with Mr. Samukele Hadebe (of the Department of African Languages and Literature, University of Zimbabwe) that both Zulu and Ndebele dictionaries do not mark archaic words. In an informal discussion at the Fourth International AFRILEX Conference, Prof. Nkabinde also confirmed that Zulu dictionaries, in spite of their long tradition, have up to the present not marked entries as being archaic. The same practice prevails in Shona lexicography. I have, therefore, arrived at the conclusion that the picture may in all probability be the same for other African languages in Southern Africa.

References

Aitchison, J. 1995. Linguistics: An Introduction. London: Hodder and Stoughton.

Barnes, B.H. 1932. Vocabulary of the Dialects of Mashonaland. London: The Sheldon Press.

Beach, D.N. 1980. The Shona and Zimbabwe 900-1850: An Outline of Shona History. Gweru: Mambo Press

Beach, D.N. 1994. A Zimbabwean Past: Shona Dynastic Histories and Oral Traditions. Gweru: Mambo

Bourdillon, M.F.C. 1987. The Shona Peoples. Gweru: Mambo Press.

Bukenya, A. (Ed.). 1994. Understanding Oral Literature. Nairobi: Nairobi University Press.

Chimhundu, H., N.E. Mberi, R.B. Batidzirai, M.B. Chitauro, E. Mangoya, A. Masasire and F. Matambirofa. 1996. *Duramazwi ReChiShona*. Harare: College Press.

Dale, D. 1981. Duramazwi: A Basic Shona-English Dictionary. Gweru: Mambo Press.

Finegan, E. 1989. Language: Its Structure and Use. Orlando: Harcourt Brace and Co.

Finnegan, R. 1970. Oral Literature in Africa. Oxford: Oxford University Press.

Fortune, G. 1980. Shona Grammatical Constructions, Part 1. Harare: Mercury Press.

Fromkin, V. and R. Rodman. 1993. *An Introduction to Language*. Fort Worth: Harcourt Brace Jovanovich College Publishers.

- Hamutyinei, M.A. and A.B. Plangger. 1974. Tsumo-Shumo: Shona Proverbial Lore and Wisdom. Gweru: Mambo Press.
- Hannan, M. 1981. Standard Shona Dictionary. Harare: College Press.
- **Heine, B., U. Claudi and F. Hunnemeyer.** 1991. *Grammaticalization: A Conceptual Framework.* Chicago: University of Chicago Press.
- Katamba, F. 1993. Morphology. Houndmills: The Macmillan Press Ltd.
- **Louwrens, L.J.** 1992. The Conceptualisation of Spatial Relationships as Expressed by Locative Structures in Northern Sotho. *South African Journal of African Languages* 12(3): 107-111.
- **Moeketsi, R.** 1994. *The Versatile Nature of the Southern Sotho Demonstrative*. Unpublished paper presented at a Workshop on Conceptions of Space in Language, Cognition and Social Interaction in Southern Africa held at the University of Botswana, January 1994.
- O'Grady, W., M. Dobrovolsky and M. Aronoff. 1989. Contemporary Linguistics: An Introduction. New York: St. Martin Press.
- Pongweni, A.J.C. 1989. Figurative Language in Shona Discourse. Gweru: Mambo Press.
- Pongweni, A. and E. Chiwome. 1995. Zvirahwe Zvakare Nezvitsva: Traditional and Modern Shona Riddles. Eiffel Flats: Juta Zimbabwe.
- Seto, G. and F. Seto. 1997. Ancient Tropical Mainland Southeast Asia: Linguistic and Cultural History, Vol. 2. Mandaluyong: Africa-Asia Educational and Cultural Publications.
- Spencer, A. 1991. Morphological Theory: An Introduction to Word Structure in Generative Grammar. Oxford: Basil Blackwell Ltd.
- **Taylor, J.R.** 1994. *The Syntax and Semantics of the Zulu Locative*. Unpublished paper presented at a Workshop on Conceptions of Space in Language, Cognition and Social Interaction in Southern Africa held at the University of Botswana, January 1994.