Studies of the Department of African
Languages and Cultures, No 44, 2010
ISSN 0860-4649

Iwona Kraska-Szlenk
University of Warsaw

Morphophonological reduction in Swahili: the pressure of frequency and lexical diffusion

Abstract

The article discusses a number of morphophonological alternations in Swahili which vary as to their scope and degree of regularity. It is argued that the allomorphy between syllabic and non-syllabic variants of some morphemes is due to phonological reduction which affects high frequency lexical items first and gradually extends on others by way of lexical diffusion. The analysis is supported by comparing the data of Standard Swahili to less conservative colloquial varieties and non-standardized local dialects, in which reductions are more advanced. The analysis contributes to a better comprehension of synchronic allomorphy and it also sheds light on the mechanism of diachronic change.

1. Introduction

This article[1] focuses on phonological reduction in Swahili which is observed in various morphophonological environments. The process brings about a lexically specified allomorphy of grammatical morphemes which alternate between syllabic and non-syllabic variants. I will demonstrate that this kind of reduction is principally conditioned by high frequency of particular words. The reduction starts in few lexical items of high frequency initiating an allomorphy rule. When there is a potential domain of the rule's application in other contexts, it may extend on less frequent words by way of lexi-

[1] Parts of this paper were read at the 6 th World Congress of African Linguistics in Cologne, August 17-21, 2009. I thank members of that audience for their comments.

cal diffusion. Ultimately, the allomorphy may reach a stage of regularity when it applies to the entire lexicon. The dialectal data provide extensive evidence that reductions limited to a few lexical items in the Standard language have a much wider scope in non-standardized dialects. Within Standard Swahili, there is also some variation with more reductions occurring in less formal registers being an indication of a diffusion process *in statu nascendi*. Before discussing the Swahili data, I will briefly present a general perspective to the reduction processes of this kind as discussed in the literature.

2. Frequency factor in sound change and lexical diffusion

Reductions of the phonological form tend to occur when these forms are used more often than others. The correlation between reduction and high frequency, observed by many linguists at various times, was initially looked upon as a completely sporadic and unpredictable phenomenon. At a later time, it started to be regarded as a relatively ordinary state of affair either from the perspective of a possible way of diachronic change or as a factor in synchronically observable variation. Witold Mańczak in the early 1960s first formulated a coherent hypothesis of the "irregular phonetic development caused by frequency" which, according to him, represents a third (in addition to "regular" and "analogical") path of change in sound structure and which takes place on a large scale in all languages. Mańczak has been supporting his thesis by abundant cross-linguistic evidence in numerous publications up to this day, drawing arguments from the variety of data and using different methodologies (see, among others, Mańczak 1965, 1969, 1977, 1978, 2008). A few English examples below, coming from a much wider corpus of similar facts, provide a brief illustration of his data and arguments.

The comparison of the parallel, i.e. regular and irregular developments of the same words/affixes reveals that the latter characterizes only items of high frequency, cf. English regular *one* and irregular *a(n)*. Among English monosyllabic words ending with *-ave*, *-ay*, *-een*, *-f*, *-ine*, respectively, only the most frequent words show reductions, namely, *have*, *say*, *been*, *if*, *mine (>my)*. The initial Germanic **h-* remained in over one thousand words and was lost only in the

pronoun *it* which is the most frequent of all *h*-initial words. Similarly, **spr-* remained in all words except in the most frequent *speak*. What is also characteristic of reductions caused by high frequency is that they often take place in words with similar meanings in different languages unlike in the case of all other irregularities being very much language-specific. For example, the word for 'speak' has undergone irregular reduction not only in English, but also in French, Italian, Polish, Russian or Swahili (see later in this paper) and likely in other languages, because words with this meaning typically belong to the most frequent lexicon. Apart from W. Mańczak's work, diachronic reduction related to frequency has been observed in numerous studies on grammaticalization, because a form more general in meaning (more grammaticalized) is used more often (see, among others, Bybee 1985, Bybee *et al*. 1994, Heine 1993, Traugott and Heine 1991), cf. English examples as *gonna < going to, 'll < will, n't > not* etc.

Once a reduction initiated in the context of higher frequency begins, it may propagate onto other words by way of an analogical process referred to as 'lexical diffusion'. This possible path of development has been empirically observed for various kinds of phonological changes in progress not necessarily limited to reduction, and has been argued to present a parallel, alternative mechanism to the regular sound change postulated by the Neogrammarians (cf. Bermúdez-Otero 2007, Bybee 2000, 2001, Hock 1986, Hooper [Bybee] 1976, Labov 1972, 1994, Wang 1977, among others). Each stage of the gradual process: the emergence of a change and its extension on other lexical items is accompanied by a period of variation between an old form and a new one until the complete elimination of the former.

The studies of synchronic variation have also brought about plenty of evidence for the positive correlation between high frequency and reduction of form (e.g. Bybee 2000, 2001, 2007, Hay 2003, 2007, Pierrehumbert 2001). For example, the positional reduction of schwa in English is most advanced in words of extremely high frequency, as e.g. *ev**e**ry* or *ev**e**ning* (complete vowel elision), relatively advanced in words of mid frequencies as e.g. *mem**o**ry, sal**a**ry* or

*summ**a**ry* (syllabic *r*) and it does not occur in low frequency words, as e.g. *mamm**a**ry* or *artill**e**ry*. Similarly, the reduction of the prefix *un-* shows a gradient continuum correlated with the frequency of prefixed words, being more significant in high frequency words as e.g. *unfortunately* or *uncertain* than in low frequency words as e.g. *unboring* or *unbiased*.

Taking into account the fact that frequency of use constitutes one of the major factors in lexical reduction, a question arises, why it should be so? In the literature, there have been two basic lines of explanation proposed, which can be called a mechanistic view and a functional view. Under the first approach, reductive changes associated with high frequency are directly related to higher predictability of more frequent words and morphemes which causes an increase in overlapping of articulatory gestures (cf. Bybee 2007:11). The alternative latter approach, advocated in W. Mańczak's work, assumes a more general motivation of reduction which goes back to Zipf's (1935) statistical laws and the principle of the inversed correlation between the frequency and size of a linguistic unit. This functional view can be supported by an additional argument based on the observation that frequency-related reductions may have an abrupt character, as in the case of acronyms or truncation, which are not due to the mechanical gesture overlap.

Frequency-related reductions discussed in this paper include mostly vowel weakening occurring in grammatical affixes. I will interpret the resulting difference in syllabicity as a binary distinction with two allomorphic forms schematized as *CV~C(G)*, i.e. a full form with the vowel and a reduced non-syllabic form with the vowel either completely deleted or preserved as a glide. Perhaps detailed instrumental studies could show phonological significance of some phonetic intermediate stages gradually interpreted. I will leave it as a question for the future research.

Throughout this paper I will use the Standard Swahili orthography for the notation of the Standard language as well as dialectal forms, for which the rule of the thumb is (roughly): consonants as in English, vowels as in Italian. Special symbols include: *j* for the voiced palatal stop, *ny* for the palatal nasal, *ng'* for the velar stop.

Underlying indicates a dental stop (t̲) and a colon is used for long vowels (a:). Even though I use 'Swahili' as the name of the Swahili language conventionalized in English, I will refer to the non-standard dialects using their original names, i.e. with the inclusion of the class prefix *ki~chi*.

3. Standard Swahili reduction and allomorphy

All languages provide examples of idiosyncratic reductions which occur in frequently used expressions, such as, for example, greetings or forms of address, cf. English *goodbye* reduced from the phrase *God be with you* or Spanish *usted* 'sir' from *vuestra merced*. Reductions of this kind are spectacular in that large portions of original structures are deleted or fused, but since they are completely unsystematic, they have no consequences for the language system as such. Occasional blendings often show abrupt idiosyncratic reductions, too, cf. *brunch* from *breakfast & lunch* or *broccoflower* from *broccoli & cauliflower*. More systematic abrupt reductions occur in truncation as well as in various kinds of acronyms and abbreviations, cf. English: *Pam* from *Pamela*, *GB* from *Great Britain*, etc. Such formations are structured according to particular prosodic patterns, which are quite general and may be mapped on new lexical items, but the deletion processes involved in them do not constitute sound change. Swahili examples of similar kinds are shown in (1) and consist of: a highly reduced case of a greeting in (1a), a blending in (1b) and two kinds of acronyms in (1c) and (1d).

(1) Abrupt reductions in acronyms, abbreviations, truncation etc.
a. *shikamoo* 'kind of greeting'
< *ninashika miguu (moo) yako* 'I am holding your feet'
b. *chajio* 'supper'
< *chakula cha jioni* 'evening food'
c. *BAKITA*
< *Baraza la Kiswahili la Taifa* 'National Swahili Council"
d. *CCM*
< *Chama cha Mapinduzi* 'Revolutionary Party'

While idiosynctratic reductions as in (1) cannot be source of lexical diffusion, certain grammaticalization processes tend to occur in chains. In Swahili, as in many other Bantu languages, grammaticalization of verbal constructions to tense/aspect markers often follows the same path, illustrated in (2), which represents a change in progress (the structure being reduced is marked in bold). Some other verbs in Swahili have undergone the grammaticalization process completely and reached the final stage of (2d) when a monosyllabic tense/aspect marker represents a remnant of the lexical verbal root, as in the case of the future marker *ta* diachronically deriving from the verb *taka* 'want' or the perfective marker *me* from the verb **mala* 'finish'. Still other Swahili verbs are at the beginnig stage of (2a), when the verb occurs as an auxiliary, as e.g. the verbs *pata* 'receive' or *wahi* 'do on time', both used in a more grammaticalized meaning of 'manage, have an occasion (to do something)'. Even if these and other similar verbal constructions undergo the same development as *kwisha* 'finish' in (2), it would be hard to consider these cases as "diffusion" comparable to lexical diffusion in phonological processes.

(2) Grammaticalization of Swahili verbs to tense/aspect marker

a. Nimekwisha kusoma. 'I have finished reading.'
b. Nimekwishasoma. 'I have finished reading/have already read.'
c. Nimeshasoma. 'I have already read.'
d. Nisha**soma**. (*dial., coll.*) 'I have already read.'

I will proceed now to the issue of phonologically conditioned reductions which are frequency-sensitive and occur in grammatical morphemes creating the above-mentioned allomorphy pattern: CV~C(G). The reduction takes place in a specified phonological environment, namely, before a vowel and not before a following consonant. But it is limited to only some vowel-initial and lexically specified environments. In all cases, the words in which reduction occurs have high token frequency, higher than other words with the same affixes. Whenever possible, the quantitative data of the Stan-

dard language have been drawn from the electronic Helsinki Corpus of Swahili, specifically, the part *Books*, consisting of literary texts, and the contents of the newspaper *Nipashe* (jointly). I will limit the presentation of the Standard Swahili data to a few examples only; more discussion of this issue together with a formal analysis in the Optimality Theoretic framework can be found in Kraska-Szlenk (2007, 2009). It should be noted that in addition to the lexically specific reductions treated in this section, the same alternation pattern, i.e. a vowel before a consonant and a glide or zero before a vowel, characterizes a large portion of Swahili morphophonology, where it appears as a completely regular process. I will return to this problem later in this section.

The first set of examples illustrates lexical reductions in Swahili personal and noun class markers positioned directly before a verbal root. There are three morphological environments in which these markers and the verbal stem are adjacent: subjunctive (hortative) forms with subject markers, infinitive (gerund) forms with class 15 marker *ku* and verbs containing object prefixes. In all these three environments, personal and noun class markers regularly occur in their full syllabic forms *CV* or *V*[2], whether the subsequent verbal stem starts with a consonant or a vowel. Exceptional reductions occur only in one case of a subject marker and in several cases of class 15 marker, which I discuss in turn below. No reduction of object prefixes is observed in Standard Swahili, cf. the following examples (object prefix underlined): *wali<u>ni</u>ambia, wali<u>ku</u>ambia, wali<u>tu</u>ambia, wali<u>wa</u>ambia* 'they told me/ you/ us/ them'.

The subjunctive (hortative) forms are illustrated with the first person singular subject marker *ni* and the first person plural marker *tu* positioned before vowel-initial stems. As seen in the following examples, the hiatus remains at the prefix-stem juncture, cf. *niimbe* 'let me sing', *nione* 'let me see', *niuze* 'let me sell', *niendelee* 'let me continue', *tuimbe* 'let's sing', *tuone* 'let's see', *tuuze* 'let's sell', *tuendelee* 'let's continue'. There is only one verb of this structure in which the subject marker is reduced to the non-syllabic form, name-

[2] Only third person singular object marker has the form *m~mw*.

ly, *twende* ‘let’s go’. It would be hard to determine on the basis of the Helsinki Corpus created from the written texts that *twende* has high token frequency, because this particular form occurs mostly in the spoken language. However, everyday experience clearly shows that Swahili *twende* and its equivalents with the meaning ‘let’s go’ in other languages belong to extremely frequent expressions.

The class 15 prefix *ku* is also realized in the full form before vowel-initial stems, cf. *kuimba* ‘to sing ’, *kuona* ‘to see’, *kuuza* ‘to sell’, *kuendelea* ‘to continue’. The reduction of the prefix to the non-syllabic *kw* takes place in several cases. The lexically specified allomorphy *ku~kw* is illustrated in (3) where words with the reduced variant of the prefix are listed together with their frequency figures based on (the part of) the Helsinki Corpus and some other infinitives with the full variant of the prefix are included for comparison.

The first example, in (3a), is the infinitive of the mentioned earlier verb of motion, cf. *kwenda* ‘to go’, although a few cases of *kuenda* are found in the corpus, too. Another verb *kuisha* ‘to finish, end’ and its reduced variant *kwisha* occur in variation, cf. (3b). This verb has developed into a more grammaticalized meaning, as mentioned earlier in (2) above; it has also extended into the adverb ‘afterwards, then’. In this latter case, it has an even more reduced form *kisha* which occurs in variation with *kwisha*, but not *kuisha*. The corpus figures[3] for all three variants demonstrate that the shorter the form, the more frequent it is. While all cases of *kuisha* carry the meaning of infinitive/gerund and all cases of *kisha* are adverbs, it is sometimes hard to distinguish between verbal and adverbial usages of *kwisha*, because of the existence of intermediate stages where both interpretations are possible, cf. the corpus examples: an infinitive usage, e.g. *pombe sasa inaelekea kwisha* ‘the beer now is coming to an end’ (lit. ‘to end’) and *baada arusi kwisha* ‘after the wedding ended’, an adverbial usage, e.g. *kwisha maamkio* ‘after greetings’, undetermined, e.g. *kwisha kusema maneno haya* ‘having said/after

[3] The Helsinki Corpus figures are surprisingly low for this frequent verb. This is due to the spelling convention and incorporating *kwisha* into the verb complex, cf. the earlier examples in (2).

saying these words'. A similar process of grammaticalization affected also the frequent verb meaning 'to begin' which occurs in its full form in the infinitive as *kuanza* 'to begin', but is reduced to *kwanza* in the meaning of 'first, firstly', cf. (3c). The most frequent lexical item in this group is the complementizer *kwamba* in (3d) which is etymologically related to the gerund (class 15) of the obsolete verb **amba* 'to say', still used in Swahili in its derived forms, e. g. *ambia* 'tell (someone)', *ambiana* 'say to each other'. The remaining examples in (3) demonstrate that other verbs occur in the full form of the infinitive prefix *ku*. In some cases, the reduction is not possible on the phonological grounds, as indicated in (3e-f) examples, since a sequence of **kwo* is generally prohibited in Swahili; if it were to take place, the glide would have to delete, too, as it happens in other contexts (cf. *kote* 'all-class 15/17, **kwote*), but such forms do not exist, either. On the other hand, it could be argued that the reduction in the previous cases of (3a-d) has a phonological motivation and is encouraged by the factor of stress, because the glided vowel occurs before the syllable carring stress (which in Swahili regularly falls on the penult). The following examples in (3g-e) with the full realization of the prefix before disyllabic stems starting with *e-* or *i-* demonstrate that the position of stress can not be the only cause here. The example of *kuendelea* 'to continue' in (3j), which contains the same root as *kwenda* 'to go' shows that the reduction is not a property of particular lexical roots. The final example in (3k) contains a highly grammaticalized, but not very frequent infinitive *kuelekea* 'to go towards; towards' with no reduction observed; this shows that the grammaticalization process is not a condition of the prefix's reduction, if not on par with high frequency.

(3) Frequencies of *ku~kw* allomorphs in Helsinki Corpus (joint *Books* & *Nipashe*)

a. *kwenda* 2260	*kuenda* 10	'to go'
b. *kwisha* 44	*kuisha* 14 *kisha* 719	'to end/then'
c. *kuanza* 945	*kwanza* 2493	'to begin'/'first'
d. *kwamba* 10496	*kuamba* 0	'that'
e. *kuona* 973	*kwona* 0, *kona* 0	'to see'
f. *kuomba* 458	*kwomba* 0, *komba* 0	'to ask'

g. *kuimba* 91	*kwimba* 2[4]	'to sing'
h. *kuishi* 590	*kwishi* 0	'to live'
i. *kuepa* 6	*kwepa* 0	'to avoid'
j. *kuendelea* 886	*kwendelea* 0	'to continue'
k. *kuelekea* 384	*kwelekea* 0	'to go towards; towards'

The next example of an irregular reduction involves the negation prefix *ha* which is positioned before a personal or noun class marker in certain types of negated verbs. It occurs in the full form before all noun class markers, whether consonant or vowel-initial, e.g. class 7: *(kiti) hakikuanguka* '(the chair) did not fall', class 8: *(viti) havikuanguka* '(the chairs) did not fall', class 3: *(mti) haukuanguka* '(the tree) did not fall', class 4: *(miti) haikuanguka* '(the trees) did not fall', class 11: *(ukuta) haukuanguka* '(the wall) did not fall'. However, the negation prefix reduces to *h* before vowel-initial second and third person singular markers, cf. 2nd sg /ha+u/ > *hu: (wewe) hukuanguka* '(you (sg) did not fall', 3rd sg /ha+a/ > *ha*: *(yeye) hakuanguka* '(s)he did not fall'. Notice that the reduction in the second person marker creates an exact minimal pair with respect to the lack of reduction in the classes 3 and 11, since each of these markers consists of the vowel /u/. An irregular reduction (fusion) also takes place in the first person singular form, cf. 1st sg /ha+ni/ > *si*: *(mimi) sikuanguka* 'I did not fall'. It is also interesting to observe that there is no comparable idiosyncratic shortening in plural personal subject markers, cf. 1st pl /ha+tu/ > *hatu*: *(sisi) hatukuanguka* 'we did not fall', 2nd pl /ha+m/ > *ham*: *(nyinyi) hamkuanguka* 'you (pl) did not fall', 3rd pl /ha+wa/ > *hawa*: *hawakuanguka* 'they did not fall'. To summarize, all irregular reductions involving the negation marker *ha* occur in the singular personal verbs only. This fact is clearly related to the frequency of use, since verbs with personal subjects have much higher text occurrence than verbs with non-personal subjects. Also, the singular number is used more often than the plural as it has been confirmed by various kinds of cross-linguistic data (e.g. Green-

[4] Apart from the place name Kwimba.

berg 1966). The case of the negation markers, however, differs from the data discussed earlier in this section, because the reduction of the negation markers is regularized to the whole category of the verbs with the singular personal markers. On the basis of the cases discussed previously as well as cross-linguistic evidence, we can hypothesize that the negation marker originally weakened in particular lexical items of the highest frequency first and was then generalized to the whole category.

In addition to the above data, irregular allomorphy motivated by frequent criteria characterizes some other morphemes. The future marker alternates as *ta~taka*, with the longer allomorph preserved only in relative clauses which are a context of low frequency, cf. *atafanya* '(s)he will do' and *atakayefanya* '(the one) who will do'. In class 4 *mi~my~m* variation, the shorter allomorph tends to appear with some adjectives of high frequency, but not with nouns and low frequency adjectives, e.g. *mingi* 'many', *miiba* 'thorns', *miovu* 'bad' (cf. Kraska-Szlenk 2007, 2009 for details).

It is well known that apart from idiosyncratic irregular reductions discussed thus far in this paper, Swahili has a number of regularized morphophonemic processes which apply across-the-board in particular contexts. All of such alternations are based on the mentioned earlier pattern: a full variant (*CV* or *V*) before a consonant and a reduced variant (*C* or *CG*) before a vowel. The following examples (of reduced variants only marked in bold) of noun class markers illustrate: before the tense marker *a*, e.g. *gari **l**aibwa* 'the car is stolen', before the associative particle *a*, e.g. *miti **y**a machungwa* 'orange trees', with possessive pronouns, e.g. *nyumba **y**ao* 'their house', with the relative pronoun *o*, e.g. *kitabu nili**ch**okinunua* 'the book which I bought', with the pronoun *ote*, e.g. *habari **z**ote* 'all news'. What is interesting about all such regularized cases of allomorphy is that it takes place in highly grammaticalized, and therefore frequent contexts. It can be hypothesized that such general patterns diachronically arose from idiosyncratic reductions similar to the ones previously discussed in this section and by lexical diffusion and analogy extended onto all lexical items co-occurring with these morphemes.

4. Advanced morphophonological reductions in non-standard dialects.

The non-Standard dialects of Swahili exhibit much more reductions in various morphophonological environments. I will illustrate some of them starting with the contexts mentioned before in section 2. The data are cited after the general sources as Bertoncini (1999) and Nurse and Hinnebusch (1993), and more detailed sources as Maganga (1991) for Kimakunduchi, Kipemba and Kitumbatu, and Kisseberth and Abasheikh (1976) for Chimwiini. Since the sources of the dialectal data usually do not include comments on frequency and regularity of the reductions, no attempt will be made to determine how advanced the process of the lexical diffusion in particular dialects is and in most cases I will limit myself to showing the examples.

Many Swahili dialects demonstrate considerable reductions of personal subject markers before vowel-initial stems. For some dialects, non-syllabic variants appear in the sources as the only possible pronunciation (e.g. Kimakunduchi, Kitumbatu in (4) below); for some others, it is given as optional to the full vowel variant (e.g. Kipemba, Kivumba). Recall that in Standard Swahili only *twende* 'let's go' shows this kind of reduction. It is also common to reduce the 1st person marker *ni* to *n* before some consonants, as seen in the examples in (4), where Standard Swahili forms (or underlying forms, where there are no comparable Standard forms) are given in parentheses for comparison.

(4) Reduction (coalescence) of the subject prefix before a verbal root

a. Kimakunduchi:
tuze (St. Sw. *tuuze*) 'let's sell', *nyuze* (St. Sw. *niuze*) 'let me sell', *nyone* (St. Sw. *nione*) 'let me see', *nyimbi* (< *niimba*) 'I sang', *nyono* (< *niona*) 'I saw', *nvata* (< *nivata*) 'I got'

b. Kitumbatu:
nyone (St. Sw. *nione*) 'let me see', *simbi* (St. Sw. *siimbi*) 'I don't sing', *suzu* (< *siuzi*) 'I don't sell'

c. Kipemba:
niezeke~nyezeke ‘let me thatch’, *nipate~mpate* (St. Sw. *nipate*) ‘let me get’, *ninunue~nnunue* (St. Sw. *ninunue*) ‘let me buy’

d. Kivumba:
wera (St. Sw. *waita*) ‘they call’, *nuze* (St. Sw. *niuze*) ‘let me sell’, *nambe~niambe* ‘let me say’, *senende~sienende* ‘I am not going’

Similarly, the reduction of the infinitive prefix *ku* is more advanced in non-Standard dialects than in Standard Swahili and occurs not only with high frequency lexical items but with many other verbs, too, as shown in (5a) and (5b) by Kimakunduchi and Chichifundi/Kivumba examples, respectively. The allomorphy is completely regularized in Chimwiini, where each allomorph is phonollogically conditioned with the *k*-allomorph before a vowel-initial stem, *x*-allomorph before voiceless consonants, and the full form *ku* in the remaining contexts (Kisseberth and Abasheikh 1976), as illustrated in (5c) below.

(5) Reduction of cl. 15 *ku* in non-Standard dialects
a. Kimakunduchi: *kona* ‘to see’ (St. Sw. *kuona*), *kwambizana* ‘to speak’ (St. Sw. *kuambiana*)

b. Chichifundi/Kivumba: *koga* ‘bathe’ (St. Sw. *kuoga*), *kosa* ‘to wash’ (St. Sw. *kuosha*)

c. Chimwiini: *kambiła* ‘to say’, *ki:mba* ‘to sing’, *xpika* ‘to cook’, *xfana* ‘to do’, *kugafa* ‘to make a mistake’, *kubu:sa* ‘to kiss’

Unlike in Standard Swahili, reductions of object markers are quite common in dialectal data, as illustrated in (6). All examples of the earlier mentioned subject markers as well as object markers in (6) are personal and I could not find comparable data containing non-personal markers of other classes. Consequently, no conclusive claim can be made as to the scope of these reductions: whether they are

limited to widely used personal markers or extend to class markers which generally have much lower text frequencies.

(6) Reduction of the object prefix

a. Kimakunduchi:
atone (St. Sw. *atuone*) 'let him see us', *akone* (St. Sw. *akuone*) 'let him see you', *anyuze* (St. Sw. *aniuze*) 'let him sell me'

b. Kipemba:
waniepuke~wanyepuke (St. Sw. *waniepuke*) 'let them avoid me', *wanifiche~wamfiche* (St. Sw. *wanifiche*) 'let them hide me'

c. Chichifundi:
were (St. Sw. *waite*) 'call them', *ruchiwajiza* (St. Sw. *tuliwaagiza*) 'we ordered them', *ulienira* (St. Sw. *aliyeniita*) '(one) who called me'

Another context for reduction is the juncture between a tense/aspect marker and a verbal root. As the examples in (7) illustrate, the vowel of the grammatical morpheme undergoes coalescence with that of the following verbal root.

(7) Reduction (coalescence) of the tense/aspect marker before a verbal root

a. Kimakunduchi:
tukemba (< *tukaimba*) 'we sang', *hatujemba* (< *hatujaimba*) 'we haven't sung', *tukoza* (< *tukauza*) 'we sold', *hatujoza* (< *hatujauza*) 'we haven't sold'

b. Chichifundi:
rukeba (St. Sw. *tukaiba*) 'we stole'

Subject prefixes may also undergo idiosyncratic reduction before a tense/aspect marker starting with a consonant which does not take place in Standard Swahili, except for the optional shortening of the 1st person marker *ni*, cf. *ninakuja ~ nnakuja ~ nakuja* 'I am coming'.

The deletion of the prefix's vowel and, subsequently, the deletion or fusion of the prefix's consonant leads to other changes, most conspicuous in the case of the 1st person marker *ni*, as shown in the examples in (8) below.

(8) Reduction (fusion) of the subject prefix before a tense/aspect marker

a. Kimvita:
t̲ʰakwambia (St. Sw. *nitakuambia*) 'I will tell you', *nnakuja* (St. Sw. *ninakuja*) 'I am coming'

b. Kipemba, Kimakunduchi, Kitumbatu:
tʰaimba (< *nitaimba*) 'I will sing', *tʰaaŋka* (< *nitaimba*) 'I will wake up'

c. Kitumbatu:
hapata (< *nikapata*) 'I got', *hona* (< *nikaona*) 'I saw', *sambili* (< *nisiambili*) 'I don't tell'

The final example in this section includes the data from Makunduchi, which show idiosyncratic reduction in the negative form of the frequent verb *jua* 'know'.

(9) Kimakunduchi idiosyncratic reduction in the negative form of the verb *jua* 'know'

a. *heji* (St. Sw. *hajui*) 'he does not know'

b. *sikwiji* (St. Sw. *sikujui*) 'I do not know you'

With the exception of the last example in (9) which involves a case of a lexically specified idiosyncratic change presumably limited to this one verb, the dialectal data demonstrate a large corpus of evidence for a widely extended diffusion of reduction processes in various morphophonemic contexts.

5. Innovative reduction in informal registers in Standard Swahili

The Standard Swahili data as previously presented in section 3 are to some extent idealized and representative of the textbooks, grammars and formal registers. In real language usage, the pronunciation is not so homogenous and variation between unreduced and respected reduced forms is bound to occur, especially in the spoken language and in colloquial or fast speech. Reduced innovations occasionally infiltrate into the written language, too, as shown in (10) by examples coming from Shafi Adam Shafi's novel *Vuta n'kuvute* (Dar es Salaam 1999). In the book, the reductions like these occur only in the dialog parts and never in the narrator's passages and they seem to represent "colloquial Swahili" rather than specific dialectal forms. As seen from these examples, reductions take place in multiple contexts and involve various morphemes: subject personal markers in a number of different environments, as in (10a), object prefixes in (10b) and other morphemes, as in (10c). But in the majority of cases the reduction affects high frequency expressions; it occurs with verbs, such as *kwenda* 'go' and *kwisha* 'end', but not with low frequency verbs, as for example *kuimba* 'to sing' which would always occur in its full form. The only exception is the behavior of the first person marker *ni* which is reduced with other verbs as well, as in *ntatangulia* 'I will go first', *n'kuvute* 'let me pull you' or *unan'tazama* 'you are looking at me'. All examples in (10) have formal equivalents included in parentheses; both variants of reduced/unreduced morphemes are underlined for convenience.

(10) Examples of reduction and coalescence in Shafi's novel *Vuta n'kuvute* (1999)

a. subject prefix reduction/coalescence:

ntatangulia (nitatangulia)	(SP+tense *ta*)
'I will go first'	
n'kuvute (nikuvute)	(SP+OP *ku*)
'let me pull you'	
yeshe (yaishe)	(SP+ verb *isha*)
'let them (cl. 6) end'	

haweshi (hawaishi) (neg+SP+ verb *isha*)
'they do not end'
haishi (haiishi) (neg+SP+verb *isha*)
'it (cl. 9) does not end'

b. object prefix reduction:
hakunambia (hakuniambia) (OP *ni*+ verb *ambia*)
'(s)he did not tell you'
nambiye (niambie) (OP *ni*+ verb *ambia*)
'tell me'
unan'tazama (unanitazama) (OP *ni*+ verb *tazama*)
'you are looking at me'
alikwabia (alikuambia) (OP *ku*+verb *ambia*)
'(s)he told you'

c. coalescence of other morphemes in pre-verbal contexts:
hendi (haendi) (neg *ha*+verb *enda*)
'(s)he is not going'
nikenda (nikienda) (AM *ki*+verb *enda*)
'if I go'
yasishe (yasiishe) (neg *si*+verb *isha*)
'let it (cl. 6) not end'

A more detailed study, based on the spoken language needs to be done to determine, how representative Shafi's dialog forms are for the colloquial Swahili and how much the variation is determined by the frequency criteria. I leave this for the future research.

Conclusion

The data discussed in this paper demonstrate that a number of the Standard Swahili morphophonological reductions are restricted to the high frequency lexemes. Because the affixes undergoing the weakening are found in many other lexical contexts, there is a potential of lexical diffusion and applying the same process in larger domains. This is especially encouraged by the fact that an internalized allomorphy rule can be easily formulated in phonological terms (a hiatus

resolution) and similar allomorphy characterizes a number of other Swahili affixes on a regular basis. Such extended reductions are commonly found in many non-standardized dialects. They also appear in informal varieties in Standard Swahili. This could indicate a case of inter-dialectal borrowing, but it could be as well treated as an independent process, especially since the more advanced reductions seem to be limited to the high frequency contexts.

References

Bertoncini-Zúbková, E., 1999, *Dialetti Swahili*, Naples: Istituto Universitario Orientale.

Bybee, J., 1985, *Morphology: A Study of the Relation between Meaning and Form*, Amsterdam – Philadelphia: John Benjamins.

Id., 2000, "The phonology of the lexicon: Evidence from lexical diffusion", in: Michael Barlow and Suzanne Kemmer (eds.), *Usage-Based Models of Language*, Stanford: CSLI, 65-85. [Reprinted in Bybee 2007: 199-215.]

Id., 2001, *Phonology and Language Use*, Cambridge: Cambridge University Press.

Id., 2002, "Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change", *Language Variation and Change*, 14, 261-290. [Reprinted in Bybee 2007: 235-264.]

Id., 2007, *Frequency of Use and the Organization of Language*, Oxford: Oxford University Press.

Id. and P. Hopper (eds.), 2001, *Frequency and the Emergence of Linguistic Structure*, Amsterdam – Philadelphia: John Benjamins.

Id., R. Perkins and W. Pagliuca, 1994, *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*, Chicago – London: University of Chicago.

Greenberg, J. H., 1966, *Language Universals*, The Hague – Paris: Mouton de Gruyter.

Hay, J., 2003, *Causes and Consequences of Word Structure*, New York – London: Routledge.

Hay, J., 2007, "The phonetics of 'Un", in: Judith Munat (ed.), *Lexical Creativity, Texts and Contexts*, Amsterdam: John Benjamins, 39-57.

Heine, B., 1993, *Auxiliaries: Cognitive Forces and Grammaticalization*, Oxford: Oxford University Press.

Hock, H.H., 1986, *Principles of Historical Linguistics*, Berlin – NY – Amsterdam: Mouton de Gruyter.

Hooper [Bybee], J., 1976, "Word frequency in lexical diffusion and the source of morphophonological change", in: William M. Christie, Jr., *Current Progress in Historical Linguistics*, Amsterdam: North-Holland, 95-105. [Reprinted in Bybee 2007: 23-34.]

Kisseberth, C. and Mohammad Abasheikh, 1976, "Chimwiini prefix morphophonemics", *Studies in the Linguistic Sciences*, 6, 2, 142-173.

Kraska-Szlenk, I., 2007, *Analogy: The Relation between Lexicon and Grammar*, (*LINCOM Studies in Theoretical Linguistics* 38), München: LINCOM Europa.

Id., 2009, "Size vis-à-vis frequency: minimality and maximality constraints in Swahili", *Language Sciences*, 31, 2-3. Special issue: Michael J. Kenstowicz (ed.), *Data and Theory: Papers in Phonology in Celebration of Charles W. Kisseberth*, 271-284.

Labov, W., 1972, "On the mechanism of linguistic change", in: Allan R. Keiler (ed.), *A Reader in Historical and Comparative Linguistics*, New York: Holt, Rinehart and Winston, 267-288. [Reprinted from: Georgetown University Monograph Series on Languages and Linguistics: Monograph No. 18, 1965, 91-114.]

Id., 1994, *Principles of Linguistic Change*, Vol. 1: *Internal Factors*, Oxford: Blackwell.

Maganga, C., 1991, *A Study of Morphophonology of Standard Kiswahili, Kipemba, Kitumbatu and Kimakunduchi*, Ph. D. dissertation, University of Dar-es-Salaam.

Mańczak, W., 1965, *Polska fonetyka i morfologia historyczna* [Historical Phonetics and Morphology of Polish], Warszawa: Państwowe Wydawnictwo Naukowe.

Id., 1969, *Le Développement phonétique des langues romanes et la fréquence*, Kraków: Uniwersytet Jagielloński.

Id., 1977, *Słowiańska fonetyka historyczna a frekwencja* [Slavic Historical Phonetics and Frequency], (*Prace Językoznawcze* 55), Kraków: Uniwersytet Jagielloński.

Id., 1978, "Irregular sound change due to frequency in German", in: Jacek Fisiak (ed.), *Recent Developments in Historical Phonology*, The Hague: Mouton, 309-319.

Id., 2008, *Linguistique générale et linguistique indo-européenne*, Kraków: Uniwersytet Jagielloński.

Nurse, D. and Thomas J. Hinnebusch, 1993, *Swahili and Sabaki: A Linguistic History*, Berkley: University of California Press.

Pierrehumbert, J.B., 2001, "Exemplar dynamics: Word frequency, lenition and contrast", in: Bybee and P. Hopper (eds.), 2001, *Frequency and the Emergence of Linguistic Structure*, Amsterdam – Philadelphia: John Benjamins, 137-157.

Traugott, E.C. and Bernd Heine (eds.), 1991, *Approaches to Grammaticalization*, Amsterdam: Benjamins, 2 vols.

Wang, W.S.-Y. (ed.), 1977, *The Lexicon in Phonological Change*, The Hague: Mouton.

Zipf, G., 1965 [1935], *The Psycho-Biology of Language: An Introduction to Dynamic Philology*, Cambridge, Mass.: MIT Press.