Hieber, Daniel W. 2014. The processing of multi-categorial roots in Swahili. Term paper, 'Cognitive Foundations of Language', Prof. Fermín Moscoso del Prado, Fall 2014, UC Santa Barbara.

The processing of multi-categorial roots in Swahili

Daniel W. Hieber

University of California, Santa Barbara

1 Introduction

The nature of linguistic categories is one of the oldest (and thorniest) issues in linguistics: What parts of speech are there? Are they language-specific or universal? In classical antiquity, Dionysius Thrax attempted to answer these questions for Greek in his *Art of Grammar* by borrowing Aristotle's concept of necessary and sufficient features, resulting in the eight traditional parts of speech still in use today. This is often called the *featural approach* to determining linguistic categories.

As scholars became more aware of the world's more 'exotic' languages in the 1800s, and especially those of the Americas, it became more and more difficult to capture the great range of crosslinguistic diversity in linguistic categories using a featural approach. For any given proposed category, exceptions were always to be found. As Geeraerts explains:

> [T]he flexibility that is inherent in the absence of clear boundaries prevents the formulation of an essence that is common to all the members of the category. Because peripheral members may not be identical with central cases but may only share some characteristics with them, it is difficult to define a set of attributes that is common to all members of a category and that is sufficient to distinguish that category from all others. (Geeraerts 2006:149)

In response to this problem, in the mid-1970s there arose alternative, gradient approaches to lexical categorization (Geeraerts 2006; Hopper & Thompson 1984; Lakoff 1987; Rosch 1973; Ross 1972), in particular the influential prototype theory of Eleanor Rosch, popularized in linguistics by John Ross, which views categories as circumscribing particularly salient exemplars called prototypes. This more cognitively-based approach to lexical categories was an attempt to understand linguistic categorization as actually realized in the minds of speakers. But this raised the obvious question: What exactly *is* in the minds of speakers when it comes to linguistic categories?

This issue rose to particular prominence when discussing the issue of ‘conversion’ or ‘zero-derivation’, where a given lexeme was seen as functioning in multiple different parts of speech without any overt marking of its change in function. An example of this in English is the form *police*, which can function referentially as a noun (*I called the police*), modificationally as an adjective (*police brutality*), or even predicationally as a verb (*the teacher policed her students’ grammar*). Such lexemes are often called *flexible*, following Hengeveld (1992). There are generally three different positions linguists have adopted regarding the cognitive status of these different forms, represented schematically in Figure 1, Figure 2, and Figure 3 below. For each stance, lexemes are represented as having different kinds of relationships to the *pragmatic functions* of reference, modification, and predication, following Croft (2003) – or what are generally thought of as the functions coded by ‘Nouns’, ‘Verbs’, and ‘Adjectives’.

Figure 1. Precategorial approach to conversion

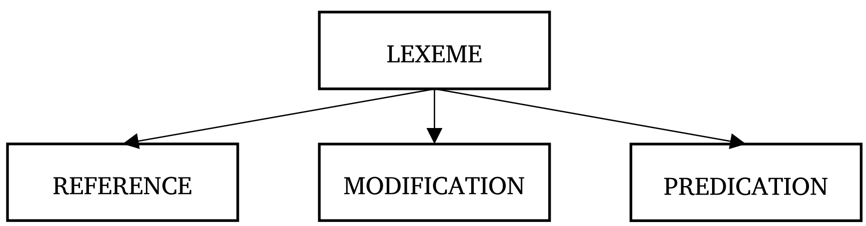


Figure 2. Zero-derivation approach to conversion

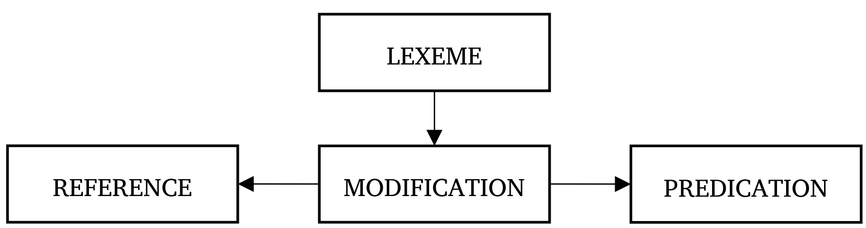
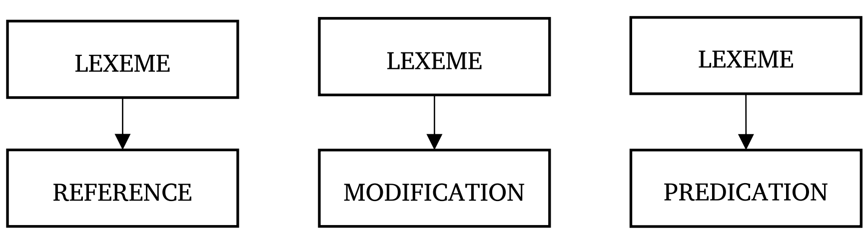


Figure 3. Multicategorial approach to conversion



In Figure 1, the three uses of *police* would be viewed as a single multifunctional or ‘precategorial’ lexeme, that only receives its particular interpretation as being used for reference, modification, or predication by virtue of its context within the utterance – including both morphosyntactic and discourse context. In Figure 2, the different uses of the lexeme are still viewed as related, but with one use standing out as primary, from which the other uses are derived. This could be either grammatically derived, in the sense of taking a null derivational morpheme ('zero-derivation'), or cognitively derived, in the sense that one function of the lexeme serves as the prototype, and the others are simply non-prototypical uses that are processed online and require pragmatic implicatures to interpret. Finally, Figure 3 is the approach which claims that seemingly related uses of a form like *police* are merely a historical accident, and that cognitively the three uses are entirely distinct lexemes in the mind.

1.1 The current study

The current study aims to shed some light on the question of which of these models is a better representation of the cognitive status of flexible lexemes, through an experimental study of response latencies in the recognition of Swahili word roots of varying lexical flexibility. Swahili is a Bantu (Niger-Congo) language spoken as a lingua franca throughout East Africa. It is an ideal candidate for a study of lexical flexibility because roots in Swahili often fulfill multiple pragmatic functions with no overt derivational morphology marking the change from one function to the other. For example, the root *-refu* appears as a noun in *marefu* 'a length', a verb in *kurefusha* 'to lengthen', and an adjective in *kirefu* 'long (inanimate)'. Even when the root appears with the inflectional morphology of whatever class it happens to be functioning in at the moment, it is still not always possible to determine the pragmatic function of the word without further context to disambiguate. For example, the word *mweusi*, derived from the root *-eusi*, is ambiguous between 'a black person' or the adjective 'black' (modifying an animate thing). The inflectional prefix *mw-* by itself is not sufficient to determine the pragmatic function of the word; the complete discourse context is required. Moreover, even *within* the Noun category, Swahili roots exhibit some lexical flexibility. Swahili has 9 noun class prefixes that can attach to a root, resulting in various nominal meanings. For example, the same root *-eusi* can form nouns like *weusi* 'the color black', *kweusi* 'darkness', or *mweusi* 'a black person'. Even though all of these are Nouns in the traditional categorization schema, linguists face the same problem here as they do with the word *police*: are these three words to be considered instances of the same lexeme or different lexemes, or is one of these forms more basic than the others? The same problem holds for adjectives, which utilize nearly the exact same set of prefixes as nouns do (giving rise to the ambiguity of function for words like *mweusi*, mentioned above). The complete list of nominal/adjectival prefixes and their associated meanings are given in Table 1.

Table 1. Swahili noun classes and their meanings

Class Prefix (Singular/Plural)	**Meanings**	**Example**	
m-/wa-	• humans	*mtu*	‘person’
m-/mi-	• plants	*mti*	‘tree’
	• body parts	*mdomo*	‘mouth’
	• geographic features	*mto*	‘river’
	• round things	*mkate*	‘bread (loaf)’
ji-/ma-	• manufactured products	*godoro*	‘mattress’
∅-/ma-	• natural or built places	*shamba*	‘farm’
ma-	• abstract/concrete concepts	*jambo*	‘issue’
	• body parts	*jicho*	‘eye’
	• fruits	*dafu*	‘coconut’
	• mass nouns (pl. form only)	*maji*	‘water’
ki-/vi-	• manmade things	*kiti*	‘chair’
	• geographic features	*kilima*	‘hill’
	• body parts	*kichwa*	‘head’
	• languages	*Kiswahili*	‘Swahili’
	• inanimate things	*kitu*	‘thing’
n-/n-	• manufactured objects	*nguo*	‘clothes’
∅-/∅-	• borrowings	*kalamu* (< Arabic)	‘pen’
	• natural phenomena	*baridi*	‘cold’
	• abstract concepts	*nguvu*	‘strength’
	• food	*ndizi*	‘banana’
u-/n-	• objects that appear in groups	*unyasi*	‘blade of grass’
u-	• uncountable nouns (no pl.)	*udongo*	‘soil, ground’
	• abstract nouns	*utu*	‘personhood’
	• names of countries	*Uganda*	‘Uganda’
ku-	• verbal nouns	*kusoma*	‘to read, reading’
	• indefinite locations	*kule*	‘thereabouts’
pa-	• definite locations	*pale*	‘there’
m(u)-	• interior locations	*mle*	‘in there’

Table 1 shows a few instances of noun-internal lexical flexibility. Compare the following sets of nouns, for example: *mtu* ‘person, *kitu* ‘thing’ and *utu* ‘personhood’; *mti* ‘tree’ and *kiti* ‘chair’.

Verbs, too, show a range of lexical flexibility. Swahili has a number of derivational suffixes that attach to a root to create a new verb, as in *kula* ‘to eat’ > *kulisha* ‘feed’ (lit. ‘cause to eat’). And these derivational suffixes can be stacked up to three long, yielding derivatives of derivatives such as *kulishana* ‘feed one another’. Again the same question arises: Are *kula*, *kulisha*, and *kulishana* to be considered instances of the same lexeme or different ones?

At the same time, not all roots exhibit the same degree of lexical flexibility. So while it is possible to form a verb from the root *-refu* ‘long’, it is not possible to do so (or at least, it is unattested) with the root *-kubwa* ‘large’, despite the fact that it would be easy to imagine a parallel to *kurefusha* ‘lengthen’ in **kukubwisha* ‘enlarge’. Likewise, not all roots can take every one of the 9 noun class prefixes. For example, while the root *-eusi* ‘black’ can be used to form the noun *kweusi* ‘darkness’, the root *-eupe* ‘white’ is not attested as the form **kweupe* ‘light/lightness’.

Both the high degree of lexical fluidity in Swahili and the variation in fluidity from one lexeme to the next make Swahili an ideal candidate for the present study, because it allows us to compare the response latencies in the recognition of words that have varying degrees of intra- and cross-categorial lexical flexibility. In order to make this comparison, however, it is necessary to first quantify the degree of lexical flexibility for the roots under consideration. This is discussed in the following section.

1.2 Quantifying lexical flexibility

To measure the lexical flexibility of different stems, I will make use of an information-theoretic measure of *inflectional entropy* of a root provided in Moscoso del Prado Martín, Kostić & Baayen (2004). This is a measure of the informational entropy calculated using the relative frequencies of the inflectional variants of a given root. It can be thought of as the “amount of information, that is, the minimum number of bits that would be necessary to encode the word in an optimal binary coding of all the words in the lexicon” (Moscoso del Prado Martín, Kostić & Baayen 2004:5). Put differently, it is a measure of the complexity of a root’s inflectional paradigm. For a given root *r* with *n* inflected variants each occurring with a frequency $f(r_i)$, the inflectional entropy of the word is defined as:

$$HI(r) = -\sum_{i=1}^{n} p(r_i) \log_2 p(r_i)$$
$$= -\sum_{i=1}^{n} \frac{f(r_i)}{f(r)} \log_2 \frac{f(r_i)}{f(r)}$$

where $f(r)$ is the cumulative root frequency of $r(f(r) = \sum_{i=1}^{n} f(r_i))$.

This measure is known to correlate negatively with response latencies for lexical decision tasks in English and Dutch (Moscoso del Prado Martín, Kostić & Baayen 2004; Baayen & Moscoso del Prado Martín 2005). That is, roots which show a higher degree of inflectional entropy take more time for speakers to process. However, these studies only looked at inflectional entropy within a single part of speech (nouns). The present study seeks to apply the same measure to examine the informational entropy of roots regardless of their lexical category. Therefore the inflectional entropy of each word would be calculated not just in relation to other instances of that root in the same lexical category, but in relation to all instances of that root, whether functioning nominally, verbally, or adjectivally.

A measure of inflectional entropy can therefore function as a heuristic for determining the lexical flexibility of a given root. The greater the lexical flexibility, the greater the inflectional entropy of a root. Based on this, one might predict that words which contain roots with a greater degree of lexical flexibility will elicit greater response latencies in a lexical decision task than words with roots of low lexical flexibility. However, in contrast to earlier studies, Milin, Filipović Đurđević & Moscoso del Prado Martín (2008) found this not to be the case for Serbian: the stem frequency of the word (analogous to root frequency in the present study) showed no significant correlation to response latencies in a lexical decision task. The present study will therefore contribute new data to the question of whether, and how strongly, root/stem frequency contributes to the lexical processing of words.

Beyond measuring just the lexical flexibility of a root, however, information theory also gives us a means of measuring how *canonical* any given use of a particular lexeme is. That is, we can calculate the degree to which a certain surface form diverges from typical uses of a root in a given category. If we know the frequency distributions of all the various inflectional possibilities for, say, adjectival roots, and calculate their average inflectional entropy as outlined above, it then becomes possible to compare the frequency distributions of a particular root against this average. Conceptually, this amounts to quantifying just how non-canonical a given root is for its class. Does this root behave in ways that deviate from the average member of its class? And how great is the deviation?

We measure this by calculating the *relative entropy* for each root, also known as the Kullback-Leibler divergence between two frequency probabilities. This measure, as applied here, takes the average frequency distribution for a class of words (in this case, adjectival roots), and measures the divergence between that distribution and the frequency distribution of a single root within that category. The resulting divergence measure therefore serves as a useful approximation of the canonicity of any given root, i.e. how much it is like the other members of its class. The Kullback-Leibler divergence $D(p\|q)$ is measured as follows:

$$D(p \parallel q) = \sum_{i \in \mathcal{P}} p(i) \log_2 \frac{p(i)}{q(i)}$$

$$D(p \parallel q) = \sum_{i \in \mathcal{P}} f(r_i)/f(r) \log_2 \frac{f(r_i)/f(r)}{f(r_i)/f(e)}$$

Here, the i variables are the possible inflected variants contained within the all the possible inflected variants of a root $\mathcal{P}$, and $p(i)$ and $q(i)$ are the probabilities of the inflected form i according to the distributions of p and q. $f(r)$ is the cumulative root frequency of a root r, $f(e_i)$ is the summed frequency of all the words within the same lexical category of r in the same inflected variant (i.e., the *inflectional exponent frequency*), and $f(e)$ is the cumulative frequency of all inflected roots within the lexical category (i.e., the cumulative inflectional exponent frequency, $f(e) = \sum_{i=1}^{n} f(e_i)$).

The measure of the relative entropy or canonicity will become useful in deciding which of the above models of the mental lexicon is most accurate. If the precategorial model (Figure 1) is correct, we should expect to see little effect for relative entropy on response latencies, because non-canonical uses of a root should be considered just as ‘basic’ as canonical ones; no use of a root is derived from some other use of the root.

2 Method

2.1 *Participants*

The participants for this study will consist of 100 students of psychology at the University of Dar es Salaam, whose participation fulfills course requirements. Participants will be screened for native fluency in Swahili and normal or corrected-to-normal vision. Each participant will take part in a single experimental session.

2.2 Materials

Swahili has only a small class of what are generally considered 'true' adjectives, i.e. adjectives that take the noun class agreement prefixes shown in Table 1. Other terms that English translates as adjectives are actually Swahili nouns used with the associative particle *-a* (e.g. *shati la udongo* 'muddy shirt' (adj.) < *udongo* 'mud, soil, earth' (n.)). The complete list of these agreeing adjectives is given in Table 2.

Table 2. Swahili agreeing adjectives

-aminifu	'honest, faithful'	*-geni*	'strange, foreign'	*-nono*	'big, fat (animal)'
-angavu	'bright'	*-gumu*	'hard'	*-ovu*	'bad'
-baya	'bad'	*-ingi*	'much, many, a lot'	*-pana*	'large, wide'
-bichi	'raw, green, unripe'	*-ingine*	'other, another'	*-pya*	'new'
-bivu	'ripe'	*-janja*	'cunning, crafty'	*-refu*	'long, high'
-bovu	'rotten'	*-kali*	'severe, sharp, cruel'	*-shupavu*	'brave, firm'
-chache	'few'	*-kavu*	'dry'	*-tamu*	'delicious, sweet'
-chungu	'bitter'	*-ke*	'female'	*-tupu*	'empty, naked'
-dogo	'small, little'	*-korofi*	'savage, brutal'	*-ume*	'male'
-ekundu	'red'	*-kubwa*	'big'	*-vivu*	'lazy'
-ema	'good, nice'	*-kuu*	'chief, main'	*-wivu*	'jealous'
-embamba	'thin'	*-kuukuu*	'used, old'	*-zee*	'old'
-epesi	'light, easy'	*-nene*	'fat'	*-zima*	'whole, adult'
-erevu	'malignant, cunning'			*-zito*	'heavy'
-eupe	'white, clear'			*-zuri*	'pretty, beautiful'
-eusi	'black, dark'				
-fupi	'short'				

Despite belonging to such a small class and being the only 'true adjectives' in the language, the forms in Table 2 actually show quite a bit of divergence in their distributional potential. Some of the forms occur more frequently in reference than modification despite being ‘adjectives’. And as mentioned earlier, some are attested as verbs while others are not. And those that are attested as verbs may or may not show a range of other verb forms derived from them. It is these adjectival roots and all of their attested derived and inflected forms that will be used in this experiment.

For each root, the inflectional entropy of the root will be calculated as laid out in the previous section. Since the frequency of a given inflectional form is known to influence lexical decision times (Kostić, Marković & Baucal 2003), and said frequency likely correlates with inflectional entropy, this variable will be accounted for by designing four experimental lists which contain roots that all have the same inflectional form (one of all nominals with the prefix *mw-*, one of all nominals with the prefix *u-*, one of all verbs with the infinitival prefix *ku-* and the derivational suffix *-isha* CAUSATIVE, and one of all verbs with the agreement prefix *i-* ‘it’ and the derivational suffix *-wa* PASSIVE). Average word frequencies, orthographic neighborhoods, and/or bigram frequencies will not be controlled for. Instead, these will be examined in the resulting regression analyses, since Baayen, Davidson & Bates () have shown this to be more appropriate for these kinds of variables.

Finally, Swahili pseudo-words will be added to each experimental list, ones which follow the phonotactic rules for Swahili but are not actually attested, and would not be meaningful as possible words. The end result is four experimental lists of varying lengths (since for any given root only certain combinations of inflectional affixes are possible, so that not all roots can appear in the *i-*+*-wa* construction, for example).

2.3 Procedure

Each participant will be randomly assigned to one experimental list, and the stimuli from that list presented to them in a random order (different for each participant). Visual stimuli will be presented in Andika font (designed specifically for readability among African populations following handwriting norms in the region), size 48, in black lettering using the standard Swahili alphabet. For each stimulus, the participant will perform a lexical decision task, which they have been verbally instructed to do before the experiment begins. They will be instructed to press the ‘yes’ button whenever the word is a Swahili word, and the ‘no’ button whenever it is not, and told to react as quickly and accurately as possible. The response times for each lexical decision will be recorded.

Each trial will begin with the visual presentation of a fixation dot for 1500ms, followed by the target word, which will remain on the screen for 1500ms or until a button response is given by the participant. When a response is given or the target word has been on screen for 1500ms, the stimulus disappears and the next trial begins. Each experiment will be preceded by 10 practice trials, after which the participant can ask questions, and which will not be considered in the final analysis. The experiment will be conducted using the experimental software *OpenSesame*.

3 Results

3.1 Preliminary data processing and normalization

The results will first be culled for outliers, so that participants who have response times that deviate significantly from the overall average response time are excluded, and test stimuli that produce abnormally-high (more than 20%) response errors are also excluded. Response latencies, surface frequencies, and root frequencies will then be transformed to a logarithmic scale to approximate normality. Surface and root frequency counts will then be orthogonalized to avoid problems of collinearity, by fitting a least-squares regression with the log surface frequency set as the predictor and the log root frequency set as the dependent variable. In the analyses of the resulting data, the result of this least-squares regression will then be used instead of the root frequency count.

3.2 Response latencies

The log-transformed reaction times will then be subjected to a multilevel regression analysis, using the linear mixed-effect model technique (Baayen, Davidson & Bates 2008). Independent (fixed effect) predictors will include: surface frequency (in log scale), root frequency (the residuals of the log scale orthogonalization), log number of orthographic neighbors (Coltheart et al. 1977), log bigram frequency (using the geometric mean of the bigram frequencies for each word, following Baayen et al. (2006)), and inflectional entropy. In order to obtain comparable magnitudes of effect estimates, all counts will be standardized to a z-score (after applying the log transformation where applicable). This transformation is innocuous for regression analyses in terms of evaluating the significance of effects, but puts all the β coefficients in the same scale. The random effects to be considered include: participants, experimental block, gender, and the stimulus item itself. Non-linear effects will also be investigated by including cubic and quadratic terms in the regression analyses.

For each independent variable, one of three possible results could emerge: positive correlation to response latencies, negative correlation to response latencies, or no correlation at all. It is safe to assume that surface frequency will correlate significantly with decreased response latencies, as highly frequent words are more likely to be quickly recognized. This effect has been found in a number of studies (Baayen, Dijkstra & Schreuder 1997; Milin, Filipović Đurđević & Moscoso del Prado Martín 2008; New et al. 2004; Baayen, Burani & Schreuder 1997). In contrast to surface frequency, root frequency will be assumed to show only a minor facilitatory effect on response latencies. Because the various surface forms of roots in the present study show drastically different distributions in terms of their lexical category and subcategory, it cannot be assumed that the frequency of nominal uses of a root will correlate to the frequency of, say, its verbal uses. A root that occurs frequently as a noun may occur quite infrequently as a verb, and vice versa. That said, the root *is* still the most significant meaning-bearing part of the word. And while Milin, Filipović Đurđević & Moscoso del Prado Martín (2008) find an effect for stem frequency (analogous to root frequency in the present study), the effect is noticeably smaller and at a lower level of significance than the effect for surface frequency. Therefore in discussing the results of this experiment, I assume a small to insignificant effect for root frequency. For the same reasons, similar effects will be assumed for orthographic neighborhood size. Since Milin, Filipović Đurđević & Moscoso del Prado Martín (2008) found no significant effects for mean bigram frequency, no correlation will be assumed here.

While Milin, Filipović Đurđević & Moscoso del Prado Martín (2008) found no significant effects for inflectional entropy / lexical flexibility on response latencies, earlier studies have (Moscoso del Prado Martín, Kostić & Baayen 2004; Baayen & Moscoso del Prado Martín 2005), and so it is reasonable to assume that some effect – perhaps small – will be found for inflectional entropy / lexical flexibility here as well.

No random effects will be assumed to be found in the results.

On the basis of the above assumptions, we can now consider what our hypothetical results tell us about the processing of flexible lexemes in Swahili.

4 Discussion

The hypothetical results presented here suggest a fairly straightforward interpretation: the inflectional entropy of a root influences of the recognition of words not just within a given part of speech, but also across parts of speech. Moreover, non-canonical uses of a root (in the sense of having a higher relative entropy) negatively influence the recognition of words more than canonical uses do. What does this tell us regarding the best model of the mental lexicon when it comes to 'conversion' and flexible lexemes? Looking back at Figures 1–3, we can immediately say that the multicategorial approach to conversion (Figure 1) is the least compatible with the present results. If it were true that the, say, nominal and verbal uses of a seemingly similar root were unrelated to each other, stored as separate lexemes in the mental lexicon, and simply homophonous, then we should not expect the cross-categorial uses of these homophonous roots to have any effect on lexical recognition. That is, inflectional entropy should have no effect on response time.

That leaves us with just the precategorial model (Figure 1) and the zero-derivation model (Figure 2). The main distinguishing feature between these two models is directionality. If directional effects are observed between canonical and non-canonical uses of a root, such that non-canonical uses of a root exhibit greater response latencies than canonical uses of a root, this supports a model of the mental lexicon where one use of a root is primary, and other uses are derived online – the zero-derivation model of Figure 2. But if no directional effects are observed, i.e. if the inflectional entropy of the root is a better predictor than the divergence of a given surface form from the average frequency distribution for that root, this supports the precategorial model of Figure 1.

The results presented here suggest that the former interpretation is the more appropriate, and that the zero-derivation model of the lexicon comes closest to capturing what happens cognitively when speakers use flexible lexemes. At the same time, it should be acknowledged that there is a high degree of diversity in the relative entropy for any given lexeme. For some roots, the non-canonical uses diverge hardly at all from the overall inflectional entropy of the root. For these roots, a precategorial model *is* more appropriate. This fact should serve as a cautionary note to linguists who want to say that an entire language is either flexible or rigid in its lexical categories. What the present study shows is that lexical flexibility is something that must be determined on a word-by-word basis, and that it is clearly a matter of degree and not necessary and sufficient conditions. It is only in the aggregate that a language begins to look more or less flexible on the whole, just as the present study suggests that Swahili is on net more rigid than flexible in its categories, while still exhibiting a high degree of flexibility in the case of many individual lexemes.

References

Baayen, R. H., D. J. Davidson & D. M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4). Elsevier Inc. 390–412. doi:10.1016/j.jml.2007.12.005. http://linkinghub.elsevier.com/retrieve/pii/S0749596X07001398 (10 July, 2014).

Baayen, R. H., L. B. Feldman & R. Schreuder. 2006. Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language* 55(2). 290–313. doi:10.1016/j.jml.2006.03.008. http://linkinghub.elsevier.com/retrieve/pii/S0749596X06000441 (8 December, 2014).

Baayen, R. Harald, C. Burani & Robert Schreuder. 1997. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. *Yearbook of morphology 1996*, 13–34.

Baayen, R. Harald, Ton Dijkstra & Robert Schreuder. 1997. Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language* 37(1). 94–117. doi:10.1006/jmla.1997.2509. http://linkinghub.elsevier.com/retrieve/pii/S0749596X97925092.

Baayen, R. Harald & Fermín Moscoso del Prado Martín. 2005. Semantic density and past-tense formation in three Germanic languages. *Language* 81(3). 666–698.

Coltheart, M., E. Davelaar, J. T. Jonasson & D. Besner. 1977. Access to the internal lexicon. *Attention and performance*, vol. VI, 535–556. Erlbaum.

Croft, William. 2003. *Typology and universals*. 2nd ed. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Geeraerts, Dirk (ed.). 2006. *Cognitive linguistics: Basic readings*. (Cognitive Linguistics Research 34). Berlin: De Gruyter.

Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. (Functional Grammar 15). Berlin: Mouton de Gruyter.

Hopper, Paul J. & Sandra A. Thompson. 1984. The discourse basis for lexical categories in Universal Grammar. *Language* 60(4). 703–752.

Kostić, Aleksandar, T. Marković & A. Baucal. 2003. Inflectional morphology and word meaning: Orthogonal or co-implicative domains? In R. H. Baayen & R. Schreuder (eds.), *Morphological structure in language processing*, 1–44. Mouton de Gruyter.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226471013.001.0001. http://bibliovault.org/BV.landing.epl?ISBN=9780226468044.

Milin, Petar, Dušica Filipović Đurđević & Fermín Moscoso del Prado Martín. 2008. The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian.

Moscoso del Prado Martín, Fermín, Aleksandar Kostić & R. Harald Baayen. 2004. Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94(1). 1–18. doi:10.1016/j.cognition.2003.10.015. http://www.ncbi.nlm.nih.gov/pubmed/15302325 (10 November, 2014).

New, Boris, Marc Brysbaert, Juan Segui, Ludovic Ferrand & Kathleen Rastle. 2004. The processing of singular and plural nouns in French and English. *Journal of Memory and Language* 51(4). 568–585. doi:10.1016/j.jml.2004.06.010. http://linkinghub.elsevier.com/retrieve/pii/S0749596X04000798 (19 December, 2014).

Rosch, Eleanor. 1973. On the internal structure of perceptual and semantic categories. In T. E. Moore (ed.), *Cognitive development and the acquisition of language*, 111–144. New York: Academic Press.

Ross, John Robert. 1972. The category squish: Endstation Hauptwort. *Papers from the Eighth Regional Meeting of the Chicago Linguistics Society*, 316–329. Chicago: University of Chicago Press.