

Final Vowel Agreement in the Ekegusii Verb

I. Introduction

Bantu languages have been the source of a great deal of discussion in Lexical Functional Grammar, due to their agglutinating nature as well as the myriad of important things they can tell about how to theoretically treat many syntactic concepts. Bresnan & Mchombo (1995) used the Bantu language Chichewa to establish how the concept of lexical integrity functions in agglutinating languages, and the Bantu language noun class system in particular. Earlier, Bresnan & Mchombo (1987) used Chichewa to describe how the f-descriptions contained in specific morphemes within the verb can greatly impact the resulting f-structure to the point that multiple incorporated pronouns can be found verb-internally. The Bantu verb becomes even more the focal point of the f-structure in this way, as many languages in this family produce quite complex f-structures that only consist of a single verb.

In this paper, I will discuss how the many verbal morphemes in the Bantu language known as Ekegusii interact with one another and create combinatory constraints via their separate f-descriptions, focusing on the agreement found in the final vowel morpheme, which appears to be central to these phenomena. By doing so, I will be implementing in theory what I have already implemented computationally using finite-state methods and constraints within the xfst formalism (Karttunen, 2003). Therefore, a secondary goal of this paper is to theoretically motivate a previously designed application.

II. Typology and Description

Ekegusii is an E.10 Bantu language spoken in the Kisii Highlands of Western Kenya. Like most Bantu languages, it is highly agglutinating and displays pro-drop to the extent that a single

verb can represent a semantically rich sentence with subject and object PRED values encoded, as well as varying tense, aspect, and mood. The verb contains a high level of sophistication such that this lexical category will be the singular focus of this paper. Because of the verb's ability to articulate a complete f-structure monolexically, the following c-structure rule will be considered, rendering LFG nodal c-structure representations unnecessary for this paper, given the economy of expression. (Bresnan, 2001) Note, however, that some depictions of the verb's morphemic constituents will be displayed.

(1)

$$S \rightarrow V$$

This is clearly not the only rule in Ekegusii, but is the pertinent rule for this topic of discussion. A normal Ekegusii sentence is also capable of the regular $S \rightarrow NP VP$ rule seen in other Bantu languages as well, and is probably the more regular construction. The language also is one of a typologically small number of languages that distinguish the recent and distant past solely through tones. (Nurse & Philipson, 2006) Further research into the language has revealed a habitual aspect that is also exclusively tonal. (Elwell, 2005) Published descriptive work includes work by Whitely (1960, 1965) and Cammenga (2002). Unfortunately, the earlier work did not have available to it phonological manners of derivation to properly locate high tone linkage, and the later work ignores tone all together, which renders description less than accurate for a language that distinguishes tonal morphemes so acutely. Furthermore, neither of these texts consider the language in the LFG formalism, meaning there has been no real work in the language using it. Because of this, the topic at hand is indeed ripe for discussion, at least in this language.

Ekegusii utilizes morphological skeletons for verbs that are as follows:

(2)

Finite: (IN) – SM – TM – (LM) – (OM) – ROOT – (EXT) – FV

Infinitive: {(PPFX)/(IN)} – PFX – NEG – OM – ROOT – (EXT) – FV

In these skeletons, the verbal root is underlined. The following table describes each morphological constituent and their contribution to the f-structure.

(3)

<i>Morphological Category</i>	<i>Name</i>	<i>F-S Contributions</i>
IN	Initial Nasal	(↑ ASP), Discourse Functions
SM	Subject Marker	(↑ SUBJ), (↑ NEG)
TM	Tense Marker	(↑ TENSE)
LM	Limitative Marker	Discourse, Aspect
OM	Object Marker	(↑ OBJ)
ROOT	Root	(↑ PRED)
FV	Final Vowel	(↑ TENSE), (↑ ASP), (↑ XCOMP)
NEG	Negation	(↑ MOOD)
PPFX+PFX	Pre-prefix and Prefix	(↑ TENSE) [non-finite]

III. Combinatory Issues and Analysis

Ekegusii has a large inventory of closed-class morphemes for each morphological section other than the root, which maintains categorization as open-class. Several large points will be discussed in this paper regarding the combinatory restrictions placed on the verbal morphemes due to their f-structure, mainly focusing on the final vowel. Each pertinent issue will be described and then analyzed. First, however, there are two points that should be discussed in brief that exist as combinatory issues but can be handled without considering f-description interaction in this theory.

A. Argument Structure

While verbal roots are intrinsically transitive or intransitive, the argument structure of a verb can be altered by morphemes specifically designed to add or remove arguments. There are

many advantages to considering lexical mapping theory when viewing the verbal stem (including the root, extensions, and the final vowel) as the ultimate source of the PRED feature.

Approaches for capturing this using f-descriptions will not be entertained as a possible supplement to the concept of lexical mapping theory for incorporated pronoun licensing, as it could prove to be theoretically redundant.

In finite-state application, featural constraints should be and are made to assure that a transitivizing or passivizing morpheme does not occur with the improper number or type of arguments. This may be a fundamental difference between computational application and theory, however, and therefore does not license further discussion in this paper.

B. Infinitive Alternation

There are no attested forms of the infinitive marker /o-ko-/ co-occurring with the many SM morphemes. Following this, there are varying arguments to consider with regards to the placement of the infinitive. To treat it as a gerundive class marker, as we also see in Bantu languages, it may be prudent to consider it an SM, but to treat its impression upon tense, putting it in the TM position could be motivated. Doubt can be cast on both considerations, though, when it is considered that often infinitives will find their SUBJ values in higher f-structures, and TM morphemes cannot co-occur with an infinitive morpheme because of the violation of uniqueness that would ensue in the (\uparrow TENSE) feature. Due to these considerations, two separate rules would be prudent to describe the morphemes comprising the verb in LFG, as seen below:

- (4a) $V \rightarrow V_{IN} V_{SM} V_{TM} V_{LM} V_{OM} V_{Root} V_{EXT} V_{FV}$
 (4b) $V \rightarrow V_{PPFX} V_{PFX} V_{NEG} V_{OM} V_{Root} V_{EXT} V_{FV}$

This is as opposed to a single regular expression using xfst, which ultimately relies on the featural diacritics in order to prevent co-occurrence of subject and tense morphemes with the

infinitive. Note that in (4b), the initial nasal/pre-prefix differentiation is absorbed into a single morphemic category. This is, at this point, for the sake of economy, and will ultimately have no bearing on the actual analysis in this paper.

C. Final Vowel Agreement

Agreement with the final vowel will be the main topic of discussion, because it is dependent upon so many morphemes. This section will be subdivided by more basic final vowel variations, ending with sections which take more than one feature into account. Here, the problem will be discussed and evidence will be given, with an analysis offered at each proper juncture. To preface this discussion, the indicative final vowel morpheme /-a/ exists as the most unmarked final vowel. However, because much of what will be accounted for here in the f-descriptions of verbal morphemes exist to constrain, it will be mentioned last when more marked morphemes are described first. This will show the number of features that must be disallowed in this final vowel in order for this unmarked morpheme to emerge.

A. Past Tense

The past tense displays final vowel morphemic agreement. Regardless of whether the tense is recent or distant past, a single morpheme is used to agree with the tense markers for both:

(5)

n-yugor-a	n- n- áá- yugór -été	n- n- áa - yugór -été
1sg-OPEN-ind	IN-1sg-rpast-OPEN-past	IN-1sg-dpast-OPEN-past
'I open'	'I opened (recently)'	'I opened (long ago)'

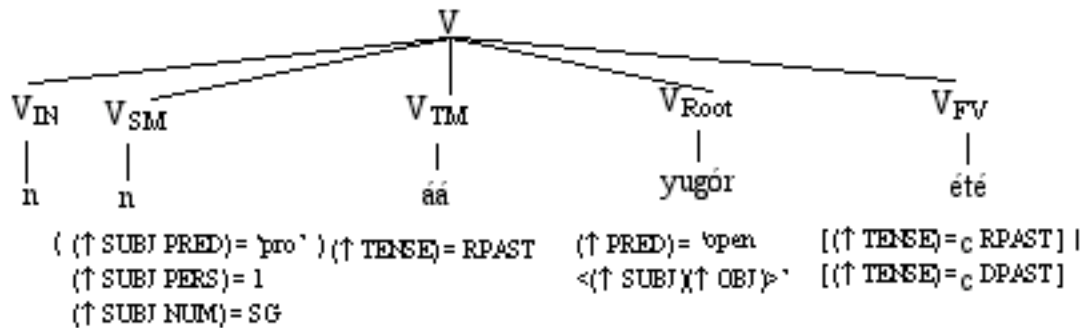
Because of this alternation, the f-description of the past tense morphemes and the final vowel must agree. Furthermore, the final vowel should agree such that it will co-occur with both the recent and distant past. Rather than using an underspecifying PAST feature for tense and

further describing the recent and distant past features through other features, a disjunctive f-description will be preferred for the final vowel:

- (6)
- áá- V_{TM} (↑ TENSE) = RPAST
 - áa- V_{TM} (↑ TENSE) = DPAST
 - ete V_{FV} [(↑ TENSE) =_c RPAST] | [(↑ TENSE) =_c DPAST]

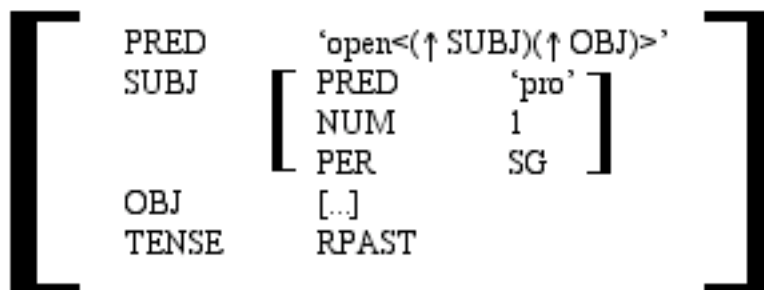
This will result in the following morphemic structure for the second example in (5): open

(7)



Note that there is no f-description for the initial nasal. More work will be needed to properly assign an f-description to this morpheme, because its shape in the language is as amorphous as its meaning. As this paper's focus is on combinatory constraints and not the nuances of aspect and discourse, which the initial nasal influences, it is perhaps more prudent to waive treating it than proposing something which would take much time and discussion to motivate. The structure in (7) will yield the following incomplete f-structure, necessitating only an OBJ to be found in the VP (alternately, if there were an OM, it could be found in V):

(8)



Because of the disjunction we see in the f-description for the final vowel, as long as there is a TM morpheme associated with the recent or distant past, the proper final vowel will be licensed. However, further considerations will show that an extra feature must be considered in the f-description.

B. Negation

Negation occurs in Ekegusii in two ways. The most common is through a set of portmanteau SM morphemes derived from their respective unmarked morphemes and the NEG marker /ta/. Shown here are both sets of SM morphemes:

(9)

	sg	pl
1	ng'-	to-
2	o-	mo-
3	a-	ba-

Non-negative

	sg	pl
1	tíín-	ntó-
2	tó-	mó-
3	tá-	mbá-

Negative

/-ta-/ can still be found in the infinitive, where it is the only indicator of negation:

(10)

go	-sar-	i-	a	go	-ta	-sar-	i-	a
inf-	SPOIL	-pass-	ind	inf-	neg-	SPOIL	-pass-	ind
	'to spoil'				'not to spoil'			

In the following examples, it can be seen that while the present tense prefers insertion of a pleonastic auxiliary for negation (similar to DO-insertion in English), the SM morphemes are the same for both the compound present tense and the non-compound past tense:

(11)

n	-teeng	-a	tíín	-d-	í	go	-teeng-	a
1sg-	DANCE	-ind	1sg.neg-	BE-pres.neg	inf-	DANCE-	ind	
	'I dance'			'I do not dance'				

perfective only exists in the present tense, as perfective constructions in the past rely on the recent or distant past differentiation or compound verbs. Because of this, the TM morpheme must be treated as specifically perfective in nature.

(16)

n	-aa-	bún	-íre	n	-aa-	yugó	-ire
1sg-pft-	BREAK-	pft		1sg-pft-	OPEN-	pft	
			'I have broken'				'I have opened'

It is not clear what morpheme has the f-description that contributes to the ASP feature of the f-structure, and what morpheme is simply constrained to the PERF attribute of ASP.

Considering the probable historical basis of the TM morpheme as being past tense, I will assume that ASP comes from the final vowel:

(17)

-ire	(↑ ASP) = PERF
-aa-	(↑ ASP) =c PERF

One advantage to making /-ire/ the source of the ASP feature valuation is that any other TM morpheme than /-aa-/ will cause a violation of uniqueness in the f-structure. Conversely, the morpheme /-aa-/ will not appear because of its constraining equation without /-ire/ providing proper ASP valuation. This f-description reversal between the FV and TM will therefore account for any case of undervaluation that may accidentally license an unattested combination.

D. The Subjunctive and Future

Another final vowel /-e/ is found in cases of subjunctive and future constructions. Because it is so multi-faceted, it will have a complex f-description. The following are examples of where it is found:

(18)

Future:

ní- n-dáa- bún -é
IN-1sg-fut-BREAK-subj
'I will break'

ní -n-dáa- yugór -é
IN-1sg-fut-OPEN-subj
'I will open'

Subjunctive:

n- táget -e o-mw- áak -e
1sg-WANT-subj 2sg-3sg-HIT-subj
'I want you to hit him'

ní-n- ch -é n- soom -e
IN-1sg-COME-subj 1sg-READ-subj
'I will read soon (I am coming reading?)'

Note that this final vowel occurs not only with a specific tense, the future, but also with both verbs involved in examples of subordinate clauses. Because this morpheme is multi-faceted the f-description must again utilize disjunctions to properly account for the various agreements:

(19)

-e $V_{FV} [(\uparrow \text{TENSE}) =c \text{FUTURE}] | (\uparrow \text{XCOMP}) | (\text{GF } \alpha \uparrow \text{XCOMP})$

The first item in the disjunction requires the future morpheme /-ráa-/ with an f-description that evaluates $(\uparrow \text{TENSE})$ with FUTURE. The next item is an existential constraint which looks for an XCOMP in the immediate level of f-structure. This accounts for the fact that matrix clause verbs display the /-e/ final vowel. By using inside-out functional uncertainty, the final item tests for the presence of an XCOMP within the XCOMP function itself, reflecting the presence of this final vowel morpheme in the subordinate clausal verb.

E. The Indicative

Given the broad characteristic of the occurrences of non-indicative final vowels in the language, the f-description for /-a/ will necessitate a great deal of constraining equations. Furthermore, it can be seen in the following data that there is no tense marker that co-occurs with the indicative:

(20)

m- bún -á
1sg-BREAK-ind

n- teeng -a
1sg-DANCE-ind

‘I break’

‘I dance’

While it has been asserted by many Bantuists that the tense marker is null in the present, null lexical items are dispreferred over explanations that could account for featural items without positing an empty category to provide this f-description. Because of this, attempts will be made to explain the mutual exclusivity of the final vowel /-a/ and the tense marker. Looking back, it was asserted that there was no TM morpheme found in the infinitive. Further, it can be seen that infinitive constructions use the indicative /-a/ as the final vowel morpheme. This can clearly all be considered in the f-description for /-a/.

A proper f-description will prevent instances of subjunctivity, perfectivity, negativity, and non-present tense. Keeping this in mind, the final vowel must be properly constrained, but also must account for the lack of a TM despite the interpretation of present tense. These considerations are accounted for by the following f-description:

(21)

$$\begin{array}{l} \text{-a} \quad V_{FV} \quad [(\uparrow \text{TENSE}) = \text{PRES}] \mid [(\uparrow \text{TENSE}) =_c \text{INF}] \\ \quad \quad \quad \sim(\uparrow \text{XCOMP}) \\ \quad \quad \quad \sim(\text{GF} \alpha \uparrow \text{XCOMP}) \\ \quad \quad \quad (\uparrow \text{ASP}) =_c \sim\text{PERF} \\ \quad \quad \quad (\uparrow \text{TENSE}) = \text{PRES} \rightarrow (\uparrow \text{NEG}) =_c \sim+ \end{array}$$

This f-description accomplishes all of these through constraining equations, a negative existential, and the optionality of either being the source of present tense or necessitating infinitive tense from another source. The tense disjunction maintains that there must be a TENSE value. A morpheme in the TM will generate position will contain an f-description that assigns tense. TM morphemes are blocked from combining with /-a/ by necessitating either the infinitive or assigning PRES to $(\uparrow \text{TENSE})$. By having a TM and /-a/, we would find either a violation of completeness or the lack of a necessitated feature attribute, INF. The negations of the two XCOMP-related equations featured in (19) serve the opposite function in this case—to prevent the verb from being in a subordinate clause or a matrix clause taking a subordinate clause.

The next two constraining equations bar perfective and negative constructions. While the perfective is essentially impossible with TM /-aa-/, which contains a constraining equation, it is prudent to insure that no perfective aspect can be found. However, this f-descriptor may be considered superfluous. Finally, we see a conditional equation that maintains that if the tense is present, then it cannot also be negative. By stating this conditionally, we avoid a constraining equation that may result in the prevention of attested infinitive forms with /-ta-/ and /-a/ co-occurring. Note that the previous account of tense stemming from the negative final vowel will further bolster the analysis for /-a/ in a negative infinitive construction, given these agreement conditions.

IV. Conclusion

At this point, a large series of combinatory possibilities for the final vowel has been accounted for. This paper has established that there are a great amount of dependencies between the many morphemes that constitute the verb in Ekegusii and, surely, many other Bantu languages as well. Furthermore, many combinatory restrictions can be accounted for by considering the source of a feature valuation; the concept of uniqueness will prevent many unattested forms. The final vowel, being the major hub of agreement in Ekegusii, is not only key to the f-structure, but a complex sign of the many features at work in it. This series of constraining equations is proof to that, and concretizes the importance of properly capturing the f-description found in the lexicon.

While the formalisms are fundamentally different, LFG and xfst both necessitate feature agreement to properly constrain the combinatory capabilities of the Ekegusii verb. This motivates the origination of different features in different morphological positions and their dependencies which ultimately influence all attested forms. In both theories, lexical entries require featural constraints and featural valuations. With two slightly disparate frameworks—

finite-state methods, which use no memory, and LFG, which necessitates some amount of memory as a kind of unification grammar—proving this assertion, it is clear that the Bantu verb is featurally constrained in its combinatory possibilities.

V. References

Bresnan, Joan (2001). *Lexical Functional Syntax*. Blackwell Publishers.

Bresnan, Joan & Sam Mchombo (1987). “Topic, Pronoun, and Agreement in Chichewa.” *Language* 63(4):741-782.

(1995). The lexical integrity principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13: 181-254.

Cammenga, Jelle (2002) Phonology and morphology of Ekegusii: a Bantu language of Kenya. Köln: Rüdiger Köppe Verlag.

Elwell, Robert. (2005). “A Morphosyntactic Analysis of the Ekegusii Verb.” Honors Thesis Presentation, University at Albany, May 2005.

Karttunen, Lauri & Kenneth Beesley (2003). *Finite State Morphology*. CLSI Publications.

Nurse, D. and G. Phillipson. (2006). “Common Tense Aspect Markers in Bantu” forthcoming in *Journal of African Languages and Linguistics*.

Whiteley, W. H. (1965). *A Practical Introduction to Gusii*. Nairobi: East African Literature Bureau.

(1960). *The Tense System of Gusii*. Kampala: East African Institute of Research.