An Approach to Describing Afroasiatic Templatic Morphologies

Gene Gragg - University of Chicago (U.S.A.)

1. Introduction: Describing Paradigms

The object of this article is to discuss the organization of a complex Afroasiatic paradigm with extensive internal vowel alternation, and to investigate a possible general framework suitable for the description and analysis of a large database of morphological paradigms, and in particular one which proposes to deal with important amounts of data from non-concatenative morphologies.[1] Apart from and preliminary to (as a preparation for) any typological, diachronic, or, most important, synchronic linguistic uses which will be made of this information, we need a consistent framework for the exploration of the most basic layers of the information content of complex paradigms Although that is not the main focus in this article, the amount and complexity of the data involved naturally imply an important eventual computational dimension to the project.

The data to be analyzed in this preliminary exploration will be the complex stem morphology of the Beja "prefixing", stem-changing verb. [2] After an enumeration of the attribute-value dimensions of the paradigm (**Sect. 2**), I will present some important general descriptive parameters and outline what appears to me to be a promising formal framework for paradigm description and analysis (**Sec. 3**), and explore its fuller implementation in Beja (**Sec. 4**).

2. The Attributes and Values of the Beja Prefixing Verb-Stem Paradigm

The exemplary paradigm of Appendix A illustrates the dimensions and range of stem-forms of the Beja prefixing (stem-changing) verbs. The paradigm involves five dimensions, i.e., morphosyntactic categories which find lexical expression in the Beja prefixing verb. These are, in the format *Category: val$_1$ val$_2$ val$_3$...*:

1. *RootClass: ccc ccy cc*. The tense-derived paradigm is traditionally subdivided by three root classes as in (**2.1**). The normalized radicals of each class are assigned to the five occurring stem templates

---

2. Beja is a Cushitic language spoken along the Red Sea coast in Sudan and adjacent areas in Eritrea. As opposed to what is the case in Semitic, "Prefixing Verb" in Beja is an inflection class. Verbs in Beja belong either completely to the prefixing verb class (with stem-internal vowel change) or completely to the suffixing verb class, with invariant stem. Somewhat less than half the verbal vocabulary. belongs to the prefixing class. The data here is basically that of Hudson 1973a,b 1974, 1976. This data has been conveniently collected also in Appleyard 2007. The Bedja stem-changing verb taken here as uniformly triconsonantal, which in fact seems to be the case to a surprising degree. Reduplicating (frequentative-intensive) and multiply derived (causative of passive, middle of causative, etc.) verb stems are not considered. In the interest of morphophonemic simplification, the accentual system is not taken into account. An initial exploration of much of this data was done in Gragg 2005, 2006 – the emphasis here is on the descriptive framework.

as outlined in **2.2** (the numerical subscript on the stems of the paradigm in Appendix A indicate which stem is associated with each form; cf. below for paradigm distribution of stem templates).

(2.1) Root Classes:

Root Class	Normalized	Paradigm Exemplar
ccc	$C_1C_2C_3$	dbl "collect"
ccy	C_1C_2Y2	dgy "bring back"
cc	$(W2)C_2C_3$	lw "burn"

(2.2) Stem Template Structure:

Template	Structure
TEMPLATE1	$C_1V_1 C_2 V_2C_3.$
TEMPLATE2	$V_1 C_1C_2 V_2 C_3.$
TEMPLATE3	$C_1 C_2 V_2 C_3.$
TEMPLATE4	$C_1 V_2C_2 (C_3= Y2)$
TEMPLATE5	$(C_1= W2) C_2 V_1 C_3$

Note in particular, as an inspection of the forms will reveal, that V_1 in these templates are not prefix vowels, but always initial templatic stem vowels. In addition, a certain number of paradigm stems are augmented with an infix (e.g., the -n- *B pres sg* marker mentioned above) or prefix (e.g., the derived stem markers for T S M) consonant.

The root classes *ccc* and *ccy*, as might be suggested by their designations, do in fact correspond fairly well to Semitic "strong" and "final weak" root. For templates 1-3, *ccc* and *ccy* forms differ only by the effects of some fairly straightfoward morphophonemics for Y2 and W2 outlined in (**Sect. 4C**). For reasons discussed at length in (Gragg 2005), and which should become clear as we examine the data, the two consonants of the biconsonantal roots correspond in every way to the second and third consonants of the other classes (as opposed to, say, the first and second, or first and third). In addition the "missing" root consonant arguably has a rounding effect on the first stem vowel in certain forms. Because of this the class is normalized as a kind of "initial weak" root class, with a missing C_1, perhaps best represented simply by "W2" in template formulae.

Templates 4 and 5 however require more comment, as they represent the approach chosen here (among other conceivable approaches) to "weak" verb morphology, when this is not reducible to simple combinatory morphophonemics. [3] Both of these B-stem templates attest to the greater formal diversity within this stem (as opposed to the large-scale leveling of the T S M stems), and to a greater differentiation between the "strong" and "weak" verb morphology. It can be seen from inspection of the subscripts on the forms of Appendix A that *ccy* verbs differ from *ccc* in using, according to our analysis, TEMPLATE1 in the simple base present, with infixed, rather than stem-initial position for the V_1 which undergoes the important *ee => ii* aorist ablaut ("Ablaut1" below): *-deeg-, -diig-* (as opposed to *-eedgi-, -iidgi-* of the base intensive). Whatever the ultimate (historical) morphological or phonological grounds

---

3. There is, and always has been for centuries of Semitic descriptive tradition, a trade-off between a somewhat more abstract root analysis (sometimes conceived of in historical, sometimes in phonological terms) with a relatively more modest stem-template inventory but a certain amount of morphophonemics, and a very concrete root analysis (for Beja, say, an analysis in terms of CCC, CCi, and CC roots) with a usually larger stem-template inventory but sometimes little or no morphophonemics. It should be noted that the formal framework explored here is neutral between these descriptive poles, and is perfectly compatible with either.

for this, it may well represent a Cushitic archaism.[4] By the same token, the unique $C_1V_2C_2$ structure of TEMPLATE4 (limited to the *ccy* jussive) might represent a transfer into this CVC structure of the important V_2 *i => a* jussive ablaut ("Ablaut2", below) present only in this root class: *-diig-, -daag-* (as opposed to *-iidgi-, -iidga-* of the intensive). In a similar fashion TEMPLATE5 provides for the vowel quality of the exceptional *B cc intense past* and *negative*. What may be going on here is that, although the Wcc verbs are remarkably consistent in surfacing with a pattern *c2v2c3*, the analogy of *ccc* and *ccy past* and *negative* forms beginning -CaaC… (with V_1 = *aa*: *-daabi(i)l-, -daag(i)Y2-*) have induced *-laaw-, -laaw-*, with a paradigmatic V_1 in what would normally be V_2 position.

2. *DerivedStem: base mid caus pass refl*. The dimension of derived stem category has a typical Afroasiatic inventory: B – base, T – middle, S – causative, and M – passive. In addition there is an interesting R – reflexive stem which is in some tenses partially identical with the T stem, in others has no special affix and a characteristic a-vocalization, and in still others (in the biconsonantal root-class) has a unique t-infix.

3. *AugmentedStem: smpl intns*. As a sub-dimension, each derived stem, including the base but not the R-stem, can have one or more "augmented" counterparts involving lengthening or reduplication of a template element or element sequence. These augmented stems have a high degree of lexical, dialectal, and idiolectal variability, and have never been thoroughly documented for any variety of Beja. However, the best documented and most paradigmatically consistent augmented stem, here termed the "intensive", is widely distributed in the lexicon (but does not occur with the T-, S-, or M-stem in the biconsonantal rootclass, or with the R-stem in any rootclass), and is basically signalled by a long first stem vowel. [5]

4. Tense: *pres past aor juss neg* . Along a dimension which can be labeled "Tense" there is a "present", two past tenses (with mutually contradictory and extremely confusing nomenclature in the literature), here termed "past" and "aorist"; a modal form here labeled "jussive" for comparative purposes, and another form underlying several negative tenses and here labeled simply "negative". Not registered in this paradigm is a non-prefixing participial form which also underlies a number of suffixing secondary tenses, including a future.

5. SubjectNumber: *sg pl*. The most restricted of the dimensions involves only the present tense of the B-Stem, which has a -n- infixed before the "middle" radical for verbs with singular subject. The data is

4. Note, for example, the middle-inflexing CCV stems with the -Ca(a)Caa- stem structures in Sasse's (1980) reconstructed Proto-East Cushitic *CC*V-root:

	CCC		CCV	CV(V)C
present	*a-CCaC-	*a-CCu/iC-	*u/i-Ca(a)Caa-	*a-Ca(a)C-
past	*u/i-CCu/iC-		*u/i-CCu/i-	*u/i-Cu(u)/i(i)C-

5. Notice the puzzling non-attestation (cf. **4.4**) of a representation for *T intns past* in all rootclasses, and of *T, S, M intns* for *cc* verbs. The non-occurrence of the form in question, which would apparently have to be *-t-daabaal-/-t-daagaay-* by TSTEM:4 and VERB10, may possibly have something to do with a constraint against a sequence of two long *identical* vowels (note the B intns neg -daabiil-). There is not enough information on Beja usage to say whether the *T smpl past, -t-dabaal-,* can systematically subsitute for its intensive counterpart. For the *cc* verbs one would have to speculate what an "underlying" -{t, s, m}-"W"-aaCaaC- (as opposed to a hypothetical -{t, s, m}-"W"-aCaaC-) might have been. In the present context these facts are simply stipulated (in TSTEM:6-7, MSTEM:6, SSTEM:5 below)

given in (**3.10**) below. The fact that the consonant in question is the first surface C in *cc* verbs is one good reason to consider this a systematically/ paradigmatically initial-weak root class**.**

3. Parameters and Framework of the Description

A. Parameters: Stem, Root, Template

The description which follows makes intensive use of the traditional morphological constructs of stem, root, and template. Linguistically the present enterprise is in the spirit of recent renewed interest in a "word and paradigm" (WP) approach to morphology, and in the structure and role of paradigms in linguistic analysis[6]. Not the least attractive feature of what we may call "neoparadigm approach" is precisely that it aims at the "broad descriptive coverage" that, as Blevins (2006) observes, is not well served by many contemporary theoretical approaches.[7] However, even though, in the tradition of recent WP, words and their paradigmatic relations are central to this undertaking; it will be observed that the actual paradigm examined in detail here is of the verb *stem* – the word shorn of its person-number-gender (PNG) subject-agreement prefixes and suffixes. Thus in all instances the full-word specification for the paradigm cell in question would have to include subject agreement PNG features, which would mean adding to the stem appropriate affixes from (**3.1**)[8].

But in fact in Beja, as frequently in Afroasiatic, PNG affixes occupy a relatively independent morphological slot, which interacts only minimally with peripheral consonantal or vocalic stem elements.[9]

(**3.1**), showing the familiar Afroasiatic "ʔ -t-y-t-n" pattern, gives the relevant formatives for Beja and, for comparison, Afar, Berber (Prasse's [1973 vol III, p. 9] *système normal*), and Ge'ez - all of which have a comparable independence of verb stem and PNG affix: The very stability of the verbal prefixing PNG paradigm over the millennia of Afroasiatic linguistic history demonstrates the long-term paradigmatic independence of these elements**.**

---

6. This tradition can be characterized perhaps most generally as envisaging the paradigm as a whole as a property of the lexeme, and of its cells as realizations of a set of morphosyntactic properties associated inherently with the lexeme, or with its syntactic, semantic, or pragmatic contexts. The term "WP" refers to an exchange of articles between Hocket 1954 and Robbins 1959. The renewed interest in the paradigm as a linguistic unit, starting perhaps with Matthews (1972), was brought into contact with the generative tradition in, e.g., Zwicky (1985) and Aronoff (1994). More recently one could cite Stump (2001, 2005), Blevins (2003, 2006), Hay & Baayen (2005). Beyond the boundary of strictly "WP" approaches to paradigm, note the Optimal Paradigm approach of McCarthy (2003).

7. "Contemporary theoretical approaches [to morphology] are, for the most part, indifferent to the task of providing broad descriptive coverage. The narrow focus of theoretical studies is often justified by a desire to isolate aspects of a grammatical system that are taken to be particulary interesting or revealing. There is usually an implicit assumption that the less interesting parts of a system would be amenable to a similar, if somewhat more tedious, analysis. Yet it is far from clear that an assumption of this sort is warranted in the domain of morphology." (Blevins 2006, p. 541).

8. Specifically, the 138-cell stem paradigm of Appendix A, in which each cell represents a value for a set of values <rootclass, stemclass, tense> becomes a 1104-cell paradigm with cell values for <rootclass, stemclass, tense, subj-num, subj-pers, subj-gender>.

9. Note Hay & Baayen 2005: "WPM [Word and Paradigm Morphology]'s claim that only full words have representations in the lexicon is arguably too strong. Stems and affixes may well develop their own reprsentations. Even so, such representations probably depend for their continuing existence on the graded support they receive from paradigmatic analogy" (p. 343). Note also the observation further on that "The generalization turns out to be that as we move forwards from the stem in a multiply suffixed word, the suffixes encouted are progressively less 'fused' with their host in terms of relative frequency and junctural phonotactics. In other words, morphological structure fades as we move in towards the stem" (p. 345 with a reference to Bybee 1985).

(3.1) Subject PNG Affixes

			Beja	Afar	Ge'ez	Berber
sg	**1**		ʔ a-	Ø-	ʔ ə-	Ø-…-äγ
	2	**m**	(ti-)…-a	t-	tə-	t- …-äd
		f	(ti-)…-i	--	tə-…-i	t- …-äd
	3	**m**	ʔ i-	y-	yə-	y-
		f	ti-	t-	tə-	t-
pl	**1**		ni-	n-	nə-	n-
	2	**m**	ti-…-na	t-…-n-	tə-…-u	t-…-äm-
		f	--	--	tə-…-a	t- …-mät
	3	**m**	ʔ i-…-na	y-…-n-	yə-…-u	Ø-…-än
		f	--	--	yə-…-a	Ø-…-nät

In addition to stem, unabashed and extensive use will be made of two pervasive traditional constructs of morphological description in the domain of Afroasiatic (principally, of course Semitic): the stem *template* and the *root*. The latter of course has been the subject of more than a thousand years of discussion in the grammatical literature (cf. among recent surveys del Olmo Lete 2003), and still underlies countless unanswered questions, synchronic and diachronic, linguistic and psycholinguistic. In the present context I merely take the root to be a set of phonological constraints, systematically relatable to a stem-template and shared by a set of lexical items – often, but not always, with some concomitant communality in semantics, morphology, syntax, and perhaps even pragmatics. The template of course has been in recent decades a renewed object of study for the multi-tiered, overlapping, simultaneous phenomena in language, beginning in suprasegmentals and continuing on to influence in one way or another most recent treatments of discontinuous morphologies.[10] Whatever its theoretical implications, it always been at least implicitly present in Semitic and non-concatenative Afroasiatic grammatical literature, and clearly remains an indispensable descriptive device for non-concatenative morphologies**.**

B. A Descriptive Framework: Default Inheritance Hierarchies

As will be demonstrated, the process of accounting for the forms in the stem paradigms of Appendix A consists essentially in repeated combinations of a restricted inventory of statement types:

1. Most basically, a set/string, or path, of morphosyntactic properties, for example the *c2* of a *simple present singular* or the *v2* of an *aorist-jussive* is realized, i.e., associated with a value, for example *c2 is preceded by n* or *v1 is +high*.

---

10. Cf., in addition to the long modern literature going back to Goldsmith 1979 and McCarthy 1979, in an Ethiopian context Hayward 1988.

2. Furthermore it is clear that these generalizations (or realization rules, as they are sometimes called) differ in scope: n-infixation happens only in the Base sub-paradigm, whereas aorist-jussive ablaut ("Ablaut1") is valid across the whole verb paradigm.

3. Morever many, if not most, of the most general or simple of the realization rules have specific exceptions – instances where a more specific set or path of morphosyntactic properties are realized differently. For example, in the face of the sweeping generalization made by the Ablaut1 realization rule, it remains the case that in the *T sub-paradigm*, the *v1 of the intensive aorist-jussive ccy* is identical with the present (/aa/) and not raised.

The claim implicitly being made then is that a paradigm can be specified by a set of path-value associations over a given domain or node in a hierarchy of such nodes, but that any general statement can be preempted by a more specifically determined association. This latter, very well known, principle governing the interaction of grammatical generalizations, is frequently referred to as the "Otherwise" or "Elsewhere" condition – and in a more general logic context as "default" or "nonmonotonic" inheritance. It has been used explicitly or implicitly in linguistic description for more than two millennia, since the time of the Indian grammarian, Panini. In accordance with this principle, wherever a number of generalizations ("rules") are competing for possible application, the most specific applicable principle will apply, and *otherwise* the more general principle (the one with fewer specifications). [11] It is important to keep in mind the difference between this Paninian principle of rule application and the familiar device of rule ordering. The realization rules informally specified in this section, and more formally in the next are not extrinsically ordered in any sense. Rather the ensemble of rules constitutes a network of simultaneous constraints on possible word shapes. This corresponds to the computational implementation of these rules in word recognition/production as an iterative pattern matching in which the most detailed possible specification for a form is chosen as the generalization to be applied**.**

It was in fact recognized by some linguists of the "neoparadigm tradition" that the elements just discussed in fact amount to a outline of a simple, but robust approach to paradigm and morphology. That is, that:

"inflectional paradigms within a given language typically resemble one another in all but a few characteristics. What is needed is a natural way of saying 'this paradigm is just like that paradigm, except for this property'. A promising approach is to use nonmonotonicity and inheritance machinery to capture such matters of family resemblance." (Cahill & Gazdar 1997, p. 211)

The basic step was the creation by Gerald Gazdar and Roger Evens in the late 80's of a clean, minimalist "Lexical knowledge description language"[12] which uses default inheritance in the simultaneous satisfaction of multiple interacting constraints. Morphologies in this vein, which emerged in the mid- to late-nineties tended to be conceived as networks – hence "network morphology" [13]. An attractive feature

---

11. On "Paninian Determinism" cf. Stump 2001, pp. 21-25. One might paraphrase the basic idea behind nonmonotonic inheritance as follows: "Monotonic inheritance – if your ancestor has a feature, you've got it.; non-monotonic inheritance – if your ancestor's got it, that's the default, and you've got it, unless you have inherited something more specific from a separate source".

12. The language was termed DATR. There is a large literature, but cf. especially Evens & Gazdar 1996. For default networks in general, also Daelemans etal 1992, and more generally Briscoe et al 1993.

13. On morphology in this context: Cahill & Gazdar 1997, 1999;Corbett & Fraser 1993; Fraser & Corbett 1994; Brown etal 1996.

of these morphologies for the purposes of a large automated database is that they are not only computationally tractable, but have in fact been computationally implemented.

This vein has continued in the "paradigm function" work of Gregory Stump, who tends to use function terminology rather than network terminology. For our purposes the morphological analyses are perfectly comparable. A synthesis of morphological theory from this point of view is given in Stump 2001[14]. Very importantly for our purposes, Stump and his Computer Science colleague at University of Kentucky, Raphael Finkel, have developed an implementation and extension of DATR which they call KATR (Finkel et al. 2002), and demonstrated its applicability to a fragment of Hebrew in Finkel & Stump 2002. The present investigation follows closely the approach outlined by Finkel & Sump 2002, and its computational facet was done with an implementation of KATR[15].

The view of morphology being developed here, starts from the three observations/claims made above concerning the generalizations one needs to make as one goes though a paradigm of the form (Appendix A).

1. The first and perhaps most obvious of these is that there are groups of paradigm cells which are characterized by a common value realized at a location of the structural description (in this context read "template element") by a common set of morphosyntatic values. For example: "the template element *c2* in the *smpl pres sg* is realized as n+c2" or "*v1 aor* is realized as +high". These are the basic facts of morphology; we will call them path/value associations, and note them as follows:

(3.2) Basic notation for morphological information

$<$c2 smpl pres sg$>$ == n c2
$<$v1 aor$>$ == +high

2. The second observation is that these generalizations often come in blocks that can be more or less restricted in scope. Thus the first generalization holds only in the B stem, the second is valid throughout the verbal paradigm. These blocks will be referred to as "nodes", the loci at which the "facts" of morphology are situated. They will be noted with a name (conventionally in caps[16]) followed by a colon, followed by a set of of related facts, followed by a period. For the two facts noted above the notation would be:

(3.3) Nodes: Loci for collection of morphological facts

BSTEM:
    $<$c2 smpl pres sg$>$ == n c2**.**
VERB:
    $<$v1 aor$>$ == +high**.**

Moreover "locus of related morphological facts" fits precisely what we mean by "lexical item" – i.e., a collection of the "givens", the inherent properties of the "base items" in any morphological schema. Thus for one of the exemplary lexical items of the paradigm of Appendix A: [17]

(3.4) A Lexical Node:
COLLECT-B:

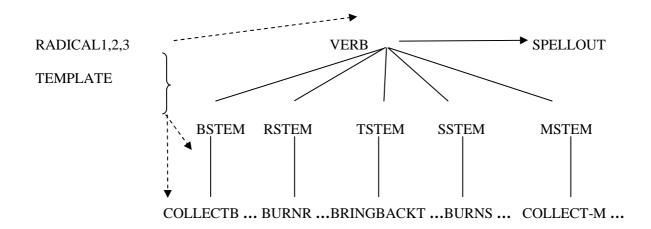|	
<rootclass> == ccc	% 1
<root> == d b l	% 2
<semantics> == "collect"	% 3
< > == BSTEM.	% 4

The irreducible properties of a lexical item include of course much more syntactic and semantic information, which may or may not be usefully noted in <path> == value form – but that is not of concern in this context. In this lexical node (in fact it will be the one we work with) for ease of reference the "facts" are numbered sequentially after the "%" ("comment" mark of the DATR/KATR formalism); a morphological fact will be referred to with the notation NODE: *SeqNum*. COLLECT-B: %4 is especially to be noted. "< >", the empty path marker (the "catch-all" leading subpath of every path), is the "otherwise" notation. *Any* further property-value determinations not yet specified at the present node will be inherited from associations specified at the target node (in this case the BSTEM).

3. The third claim is that blocks of morphological facts can be related hierarchically in such a way that higher nodes in the hierarchy define default properties shared by the nodes they dominate, ultimately the terminal, lexical, nodes**.**

Observations 2 and 3 imply that a sequence of such "parent-child" inheritance nodes should naturally yield an inheritance hierarchy. And for the Beja data of Appendix A, the simple inheritance hierarchy of the nodes in **3.6** appears to be adequate:

---

17. Note that "derived stem" is taken here as an *inherent* property of the lexeme, as opposed to "tense", "number" etc., where are variable properties of the words realizing the lexeme. This has to do with the relatively lower degree of productivity, and lesser semantic predictability of derived verb forms. This is a relatively conservative stance, and obviously there are more and less productive ways of deriving lexemes from other lexemes, and in Beja these lexeme-to-lexeme process *are* relatively productive. Computationally, for the generation of the paradigms of Appendix A, it makes little difference whether there are lexems COLLECT-B, COLLECT-R, etc., or only one lexeme COLLECT, with derived stem as an additional variable property of the lexeme/root.

3.6 Beja Morphological Hierarchy

This hierarchy follows of course a fairly traditional schema of Afroasiatic (and especially Semitic) grammars, whereby generalizations which are true of the verb as a whole are separated from generalizations specific to individual stems. [18] Moreover, as we shall see, there are cross-generalizations shared between specific stem classes, which can be readily stated in this framework. On the other hand, there is nothing apriori or ironclad about this specific hierarchy. For example, not used here is a general hierarchical level corresponding to the frequently all-important division of morphological generalizations according to root class ("strong" versus "weak", with the special sub-classes of "weak" root morphologies). Although, as we shall see, root-class distinctions enter into statements of morphological facts ("realization rules"), no separate level seems necessary here – in part perhaps because of the "morphophonemic" approach taken here to the ccy and cc rootclasses (see **Sect. 4C**).[19]

Orthogonal to the hierarchy are the nodes TEMPLATE and RADICAL. The former is the set of realization rules which assign stem templates to word realizations of lexemes by virtue of information at verb, derived stem, and lexeme level, and thus give new c and v paths to be given values by further pattern matching at the relevant nodes. At the end of the iterative process a *path* (not a *set*) of values are handed off to the morphophonemic component labeled SPELLOUT

To recapitulate, our starting point is the fairly complex data array we call a paradigm, the set of word forms which realize all possible combinations of morphosyntactic properties associated inherently with the lexeme, i.e., the set of word forms which in fact constitute a lexical entry. If we agree to represent a word, i.e., a paradigm cell, as Lexeme:<α β γ . . .>, where "Lexeme", as above, is an arbitrary name of the lexeme and :<α β γ . . .> is the set of morphosyntactic values determining a specific word-form of that lexeme, then we might envision the abstract schema for a word (as opposed to a lexeme) as something like:

---

18. It should be noted also that the general picture in (4.1) is not unlike the view of Hebrew lexemes, and the morphological role of Hebrew verbal inflectional classes ('binyanim') suggested in Aronoff 1994 (Ch. 5). This general flow also seems to underly the fragment of Hebrew given in Finkel and Stump 2002, which greatly influenced, and inspired, the analyses in this paper.

19. Parallel work in Ge'ez such a layer might be useful – and there are some common generalizations, for example, those pointed out in footnote 3 , not accounted for succinctly here.

3.8 Word Schema

COLLECT-B:
    $<$rootclass$>$ == ccc   % 1
    $<$root$>$ == d b l   % 2
    $<$semantics$>$ == "collect"   % 3
    $<$ $>$ == BSTEM.   % 4
    $<$aug$>$ == α
    $<$tens$>$ == β
    $<$num$>$ == γ
    $<$pers$>$ == δ
    $<$gen$>$ == ϵ
    . . .

where α β γ **. . .** refer to values corresponding to attributes which must or can be specified in the possible realizations of the lexeme as a word (or word-part, as in the paradigm of Appendix A)**.**

Viewed thus as an array of possible realizations of Lexeme:$<$ α β $\gamma>$, the cells of Appendix A might be schematically represented as:

(**3.9**)

COLLECT-B:$<$smpl sg pres$>$	COLLECT-B:$<$intns pres$>$	**. . .**	COLLECT-M:$<$smpl pres$>$	COLLECT-M:$<$intns pres$>$
.	.	.	.	.
BURN-B:$<$smpl neg$>$	BURN-B:$<$intns neg$>$	**. . .**	BURN-M:$<$smpl neg$>$	[BURN-M:$<$intns neg$>$]

where each cell of (**3.9**) is in fact a query directed to the nodes of the morphological default inheritance hierarchy.[20] A complete paradigm would be specified by generating in some mechanical fashion a set of queries for a given lexical node covering all allowed feature combinations as in the paradigm schematically outlined in (**3.9**).[21]

C. Application to Beja

Considering globally the paradigm of Appendix A, one can see that all the forms involve one or two template vowels (v1, v2) intercalated among three template consonants in (according to the analysis given here) five template patterns. The identity of the templatic c's is already implicit in a lexical item's root. and, following Finkel & Stump 2002, will be derived here by the function-like, single-path/value nodes RADICAL1/2/3:

3.7 The "RADICAL" Node/Function

RADICAL1: $<$c1 c2 c3$>$ == c1 **.**
RADICAL2: $<$c1 c2 c3$>$ == c2 **.**

---

20. Cf. "KATR generates output based on queries directed to nodes representing individual words" (Finkel & Stump 2002, p. 3).

21. Such a paradigm-generating capacity is in fact built into the DATR/KATR systems.

RADICAL3: <c1 c2 c3> == c3 .

The identification of the "interdigitated" vowels, however, and their pattern of distribution is considerably more complex. If in Beja, as often in non-concatenative morphologies, it is hard to assign consistent morphemic function to individual vowels as such, or even to a discontinuous "vocalic melody", nevertheless it is clearly not the case that we are reduced to simply stipulating the identity of *v1* and *v2* for each individual cell of the stem paradigm. Generalizations over paradigm cells *are* possible. Moreover, even if these generalizations do not generally permit a neat partitioning of the morophological grid into contiguous blocks of subparadigms, they do permit a networked hierarchy of vowel assignments affecting various combinations of paradigm categories and values which are represented in the rows and columns of the paradigm.

To begin with a phenomenon already alluded to, the *B simpl pres sg* is realized by the stem template *c1v1c2v2c3*, where *v1* is realized as /a/ and *c2* is preceded by /n/. (**3.10** extracts the relevant rows and columns from the Beja paradigm of Appendix A)

(3.10) Infix -n- present sg:

rootclass	tense	num	B-smple
ccc	pres	sg	**-danbiil-**
ccy	pres	sg	**-dangi-**
cc	pres	sg	**-nliiw-**

Otherwise in the *B present, aorist, or jussive, simple or intensive* (paradigm rows and columns extracted in **3.11**), the template is *v1 c1c2 v2 c3* and *v1* is realized as a non-low, non-back vowel -- high where so stipulated by an aorist-jussive ablaut rule to be formulated below, but otherwise non-high /ee/ (for the boxed *ccy* B smple juss cf. the special generalization about ccy jussives taken up below in connection with (**4.4**)).

(3.11) *v1* in the unmarked present, aorist, and jussive

rootclass	tense:	num: B	B-smple	B-intns
ccc	present	[sg	-danbiil-]	**-eedbil-**
		pl	**-eedbil-**	
	aorist		**-iidbil-**	**-iidbal-**
	jussive		**-iidbil-**	**-iidbal-**
ccy	present	[sg	-dangi-]	**-eedgi-**
		pl	**-deeg-**	
	aorist		**-diig-**	**-iidgii**
	jussive		**-daag-**	**-iidga**
cc	present	[sg	-nliiw-]	**-eeliw-**
		pl	**-eeliw-**	
	aorist		**-iiliw-**	**-iiliw-**
	jussive		**-iiliw-**	**-iiliw-**

With the exception of the boxed cells (see below), the data of **(3.12)** show that *B pres intense or simple pl v2* is short, but that otherwise throughout the paradigm *Verb pres or neg v2* is long.[22] Inspection of this extract, and as well of the Appendix A paradigm as a whole reveals another regularity – other things being equal, the unmarked values of v1 and v2 are *a* and *i* respectively.

(3.12) Long /ii/ in pres and neg ("Ablaut 1")

root	Tens	B	B-intns	Refl	T	T-intns	S	S-intns	M	M-intns
ccc	Sg	-danbiil-	-eedbil-	-t-dabiil-	-t-dabiil-	-t-daabiil-	-s-dabiil-	-s-daabiil-	-m-dabiil-	-m-eedbiil-
	Pl	-eedbil-								
		-dabiil-	-daabiil-	-dbaal-	-t-dabiil-	-t-daabiil-	-s-dabiil-	-s-daabiil-	-m-dabiil-	-m-eedbiil-
ccy	Sg	-dangi-	-eedgi-	-t-dagi-	-t-dagi-	-t-dagi-	-s-dagi-	-s-dagi-	-m-dagi-	-m-eedgi-
	Pl	-eedgi-								
		-dagi-	-dagi-	-dgaay-	-t-dagi-	-t-dagi-	-s-dagi	-s-dagi	-m-dagi-	-m-eedgi-
cc	Sg	-nliiw-	-eeliw-	-eet-liiw-	-t-ooliiw-	- - -	-s-ooliiw-	- - -	-m-ooliiw-	- - -
	Pl	-eeliw-								
		-liiw-	-laaw-	-laaw-	-t-ooliiw-	- - -	-s-ooliiw-	- - -	-m-ooliiw-	- - -

Finally, notice that in a number of cases a realization rule applies for more than one morphosyntactic category, i.e., to disjunctions of categories such as: *present or negative*, or *present, aorist, or jussive*. To handle these disjunctions DATR defines variables with a notation *"#vars $nonpast: pres aor juss."* meaning *"$nonpast* is defined as a variable covering *present, aorist,* or *jussive*.[23]

Limiting ourselves for the moment to the simple present, the regularities pointed out in connection with the paradigm excerpts **(3.10-12)** could be formulated in a default inheritance morphological hierarchy **(3.13)**[24]:

(3.13) A Morphological Hierarchy

#vars $pres: pres neg**.**
#vars $nonpast: pres aor juss**.**

BSTEM:
    < > == VERB.   % 1
    {stemclass} == base   % 2
    {stempref} ==   % 3
    {c2 pres simple sg} == n RADICAL2:<"<root>"> % 4
    {v1 pres simple sg} == a   % 5
    {v1 $nonpast} == ee   % 6

---

22. This will be referred to henceforth as Ablaut1. Note that in *ccy*, as explained below in **Sect. 4C** both VCiiY (e.g. -dangiiY-) and CCiY (e.g., -eedgiY) yield final Ci (e.g., -dangi-, -eedgi-)

23. In Appendix B it will be observed that the "#vars" notation is useful for defining phonological as well as morphological classes, and that classes so defined can be used as a basis for more, or less, inclusive classes by "+", or "-", operations.

24. Displayed in **3.13** is one of the important differences between KATR and DATR. In the latter, in path/value statements are always of the form <path> == value, and <path> is always an *ordered sequence* of morphosyntactic atoms. KATR admits also path/value statements of the form {path} == value, where {path} is simply a *unordered set* of atoms, i.e., where order is irrelevant. On the linguistic rationale for this see Stump 2001, pp. 270-273.

VERB:
    < > == SPELLOUT:< "<stempref>" TEMPLATE > % 1
    <c1> == RADICAL1:<"<root>">   % 2
    <c2> == RADICAL2:<"<root>">   % 3
    <c3> == RADICAL3:<"<root>">   % 4
    {v1} == a   % 5
    {v2} == i   % 6
    {vf1} ==   % 7
    {vf2} ==   % 8
    {vf2 $pres} == +long   % 9
TEMPLATE:
    <smpl pres sg> == "<c1>" "<v1>" "<vf1>" "<c2>" "<v2>" "<vf2>" "<c3>" % 1
    <pres> == "<v1 >" "<vf1>" "<c1>" "<c2>" "<v2 >" "<vf2>" "<c3>"   %2

As a concrete context, let us suppose that we are interrogating this system with the query COLLECT-B:<pres smpl sg>. The BSTEM node/block, identifies the stemclass (BSTEM:2) and says that the stemprefix is #null# (BSTEM:3). It handles the infixation of "n" by realizing c_2 as an "n c_2", where the second radical is derived from the lexeme's *root* value by way of RADICAL2 (BSTEM:4). Note the very important use of the quoted path ' "<root>" '. In this system, a quoted value is an instruction to find the value in question in the original query node. Thus in our case RADICAL2 will take as its "argument" the value of COLLECT-B:<root>, which of course (by COLLECT-B:3) is "d b l", from which RADICAL2 returns "b". Note also the Paninian aspect of this RR: Since more of the query path <pres smpl sg> is matched/consumed in BSTEM:4 than in VERB:4, which gives a competing value for c_2, BSTEM:4 "wins out". BSTEM:5 gives a value for *v1 pres smpl sg*, which is again in Paninian competition with BSTEM:6, which gives an "otherwise" value for *v1 pres*. BSTEM:1 says that further specifications are inherited from VERB.

VERB, for its part, assigns values for otherwise unassigned *c1, c2, c3* (VERB:2-4), and then states in VERB:5,6 that the "unmarked" (otherwise unspecified) *v1* is *a* and *v2* is *i*. Also, throughout the paradigm, as noted above, *v2*, whatever its quality, is +long in the *pres* and *neg*. Note here the way in which, in the absence of a serious investigation of the way in which complex feature matrices should be handled in a *<path> == value* context, vowel features have been linearized in a rather ad hoc manner for the purposes of this investigation, as a "*vf1*" or "*vf2*" atom, which in the template follows "*v1*" and "*v2*" respectively. Various ablaut-like specifications such as "*+long*", "*+high*" are assigned to these atoms, which are then spelled out in the obvious manner in the SPELLOUT node, where, e.g., the sequence "*a +long*" is replaced by "*aa*", or "*ee +high*" by "*ii*". By default (VERB:7, 8) these atoms are not realized. Finally VERB:1 a) asks the original query (COLLECT-B) to find what value, if any, should be given to an atom *stempref* (in our example identified as #null# by BSTEM:4); b) sends the query string, inherited from the original query, to the TEMPLATE node for template assignment (again done by interrogating the original query for each element of the template); and c) bundles the resulting ordered sequence of paths to SPELLOUT, which extracts the phonological information from each path, making the required morphophonemic replacements (cf. below).

Informally it is not difficult to see, once the node-linking and default information formalisms are grasped, how the individual *<path> == value* statements correspond to obvious generalizations about paradigm communalities, and how individual paradigm cells are "derived." Thus returning to the example query COLLECT-B:<pres smpl sg>, it is obvious that this form will have a *c1v1vf1c2v2vf2c3* template (by TEMPLATE:1), where stempref is #null#, *v1* is *a*, and *c2* is *nb* (by BSTEM:3, 4, 5), c1 and c3 are *d* and *l* (by VERB:2, 4), v2 is *i*, vf1 #null#, and vf2 *+long* (VERB:6, 7, 9), yielding:

SPELLOUT:<(#null#) d a (#null#) nb i +long l> => danbiil

To see in detail how this process might work step-by-step (as might be done in a computational implementation), **(3.14)** shows a sort of "trace" of such an explicit "derivation"[25], from an initial query to a final result**.**

(3.14) COLLECT-B:<smpl pres sg> = danbiil

QUERY	YIELDS	BY MATCH
COLLECT-B:<smpl pres sg>	BSTEM:<smpl pres sg>	COLLECT-B:4
BSTEM:<smpl pres sg>	VERB:<smpl pres sg>	BSTEM:1
VERB:<smpl pres sg>	SPELLOUT:< "<stempref>" TEMPLATE: <smpl pres sg> >	VERB:1
"<stempref>"	COLLECT-B:<stempref pres smpl sg>	*[original query node]*
COLLECT-B:<stempref pres smpl sg>	(#null#)	BSTEM:3
TEMPLATE:<smpl pres sg>	"<c1>""<v1>""<vf1>""<c2>" "<v2>""<vf2>""<c3>"	TEMPLATE:1
"<c1>""<v1>""<vf1>""<c2>" "<v2>""<vf2>""<c3>"	*[successively back to query node]*	
COLLECT-B:<c1 smpl pres sg>	d	VERB:2
COLLECT-B:<v1 smpl pres sg>	a	BSTEM:5
COLLECT-B:<vf1 smpl pres sg>	(#null#)	VERB:7
COLLECT-B:<c2 smpl pres sg>	n b	BSTEM:4
COLLECT-B:<v2 smpl pres sg>	i	VERB:6
COLLECT-B:<vf2 smpl pres sg>	+long	VERB:9
COLLECT-B:<c3 smpl pres sg>	l	VERB:4

Finally, "SPELLOUT:<#null# d a #null# n b i +long l smpl pres sg>" yields "danbiil" by combining "i +long" into "ii", extracting in linear order the phonological material, and ignoring the non-phonological**.**

4. The Templatic Structure of Beja Stems and a Descriptive Framework

From this point we will continue through Appendix A, noting the regularities – the explicit formalisms however will be found in Appendix B**.**
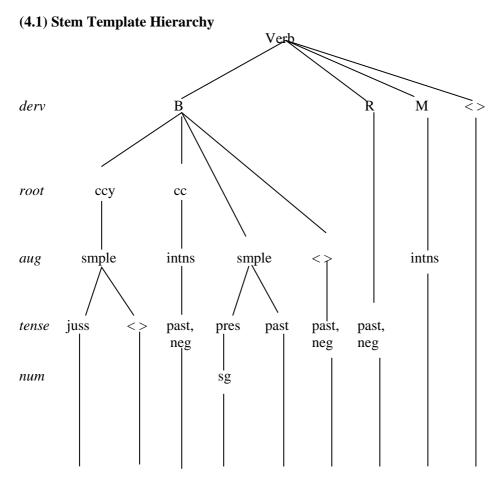
A. The Stem-Templates

Up to now we have talked about the stem template and specification of its v/c values together,. but the generalizations turn out to be independent of one another, and better made separately. Examined more in detail, it becomes aparent that Beja stem-template assignment proceeds, not form-by-form, but in a more hierarchical fashion, whereby, within a class of stems, a certain more specific stem or stem-subclass receives one stem assignment, while the rest, not so specified (i.e., the *otherwise* case), receive another.

25. Note that this is not so much a "derivation" in the usual linguistic sense of a step-by-step transversal of a set of ordered rules, but an iterative, left-to-right, depth-first pattern matching, on the basis of a set of simultaneous constraints, where the longest matching pattern "wins". Establishing a derivation in a system of this kind is more like proving a theorem on the basis of a set of axioms. Morphological descriptions in the DATR tradition are often in fact termed "theories", and it is no accident that the first computational implementation was in the logic-programming language Prolog, to which recent KATR development has returned

The pattern is that of a cascading sequence of specific cases, followed by more general *otherwise* generalizations. Thus, *B ccy simple jussives* are assigned Template4, whereas all other tense forms stipulated *B ccy simple* get Template1. Similarly *B cc smple {juss, neg}* get assigned Template5. *Otherwise*, for all other rootclasses (the "< >" in **4.1**), *B simple present singlular* takes Template1, and Template3 in *B smpl past*; *otherwise* in all other B (simple or not, and including *ccy*) take Template2 in the present, aorist, and jussive. All *R {past, neg}* take Template3 in the past and negative, and all *M intensive* take Template2. *Otherwise* all forms of all roots are Template1. This complex of relationships can perhaps be made clearer in the hierarchical diagram of (**4.1**), where "<>" means "all otherwise/default" for the attribute in question**.**

This is in fact a default inheritance hierarchy, in which, for example, *ccy* B simple verbs are assigned the stem template value Template4 if they are also jussive (the more specific case) but otherwise Template1 (the default case). This complex of relationships can be expressed as a set of *path == value* statements at the level of TEMPLATE in Appendix B**.**

(4.1) Stem Template Hierarchy

				Verb					
derv			B			R	M	< >	
root		ccy	cc						
aug		smple	intns	smple	< >		intns		
tense	juss	< >	past, neg	pres	past	past, neg	past, neg		
num				sg					
stem	TMPL4	TMPL1	TMPL5	TMPL1	TMPL3	TMPL2	TMPL3	TMPL2	TMPL1

B. The Templatic C/V Constituents

With the template assignments complete, the *c* and *c* assignments for the rest of the BSTEM, for the other derived stems, and for the VERB can be made following the same steps as in **Sect 3**. The complete formal description is displayed in Appendix B, here we will simply note the regularities that need to be accounted for. We will start by noting three sweeping ablaut-like generalizations having to do with the formation of the aorist, jussive and intensive**.**

In the first place it should be clear from **4.2**, apart from the boxed exception which will be handled by the more specific TSTEM:5 (a relic of an older V => *a* aorist ablaut?), there is a very general across-the-board rule according to which throughout the paradigm the *aorist or jussive v1* is high (VERB:10), whatever the template or corresponding present-tense v1. This will change /ee/ to /ii/, /a/ to /i/, and /aa/ to /ii/, but also /oo/ to /uu/.[26] (We will refer to this vowel change as Ablaut2)**.**

(4.2) Present > Aorist/Jussive: Low > High ("Ablaut2") a,e>i + o > u

		B	B-intns	R	T	T-intns	S	S-intns	M	M-intns
ccc	pres	**-ee**dbil-	**-ee**dbil-	-t-d**a**biil-	-t-d**a**biil-	-t-d**aa**bil-	-s-d**a**biil	-s-aabiil-	-m-d**a**biil-	-m-**e**dbiil-
	Aor	**-ii**dbil-	**-ii**dbal-	-t-d**i**bil-	-t-d**i**bil-	-t-d**ii**bil-	-s-d**i**bil-	-s-d**ii**bal-	-m-d**i**bil-	-m-**ii**dbil-
	Juss	**-ii**dbil-	**-ii**dbal-	-t-d**i**bil-	-t-d**i**bil-	-t-d**ii**bil-	-s-d**i**bil-	-s-d**ii**bal-	-m-d**i**bil-	-m-**ii**dbil-
ccy	Pres	-d**ee**g-	**-ee**dgi-	-t-d**a**gi-	-t-d**a**gi-	-t-d**aa**gi-	-s-d**a**gi-	-s-daagi-	-m-d**a**gi-	-m-**ee**dgi-
	Aor	-d**ii**g-	**-ii**dgii	-t-d**i**g-	-t-d**i**g-	-t-d**aa**gi-	-s-d**i**g-	-s-d**ii**g-	-m-d**i**g-	-m-**ii**dgi-
	Juss	-d**aa**g-	**-ii**dga-	-t-d**i**ga-	-t-d**i**ga--t-di	-t-d**aa**ga-	-s-d**i**ga-	-s-d**ii**ga-	-m-d**i**ga-	-m-iidga-
cc	Pres	**-ee**liw-	**-ee**liw-	**-ee**t-liiw-	-t-**oo**lii-	- - -	-s-**o**liiw-	- - -	-m-**o**liiw-	- - -
	Aor	**-ii**liw-	**-ii**liw-	**-ii**t-liw-	-t-**uu**liw-	- - -	-s-**u**liw-	- - -	-m-**uu**liw-	- - -
	Juss	**-ii**liw-	**-ii**liw-	**-ii**t-liw-	-t-**uu**liw-	- - -	-s-**u**liw-	- - -	-m-**uu**liw-	- - -

In the second place, as can be seen from the data of Appendix A (extracted in **4.3**), the aorist has the same form as the jussive in *ccc* and *cc*. In *ccy* verbs, however, it is marked by a v2 ablaut to /a/ (for clarity, the presupposed morphophonemic representation is added below to the surface representation). Surprisingly, this *v2* is actually infixed in *v1* position in the B simple present, giving to the *ccy B simple aorist* and *jussive* the same CVC surface form as the *ccy B present singular*. As discussed above, this infixation is handled via template assignment in TEMPLATE:1,2. (This is the vowel change we will refer to as "Ablaut3", formulated in VERB:11.) As suggested by the linked boxed forms, perhaps the occurrence of *v2 = a* in <u>both</u> the aorist and the jussive of the B and S intensive (BSTEM:10, SSTEM:4) is a generalization of the *a* member of the ablaut pair in transitive intensives – possibly an indication that, as we would expect, the ablaut was not originally limited to the *ccy* root class**.**

---

26. Cf. below **3C** in "Morphophonemics" for the roundedness of v1 in the T, S, M cc verbs.

(4.3) Aorist vs jussive -- "Ablaut 3: i > a"

root	tense	B	B-intns	R	T	T-intns	S	S-intns	M	M-intns
ccc	aorist	-iidbil-	-iidbal-	-t-dibil-	-t-dibil-	-t-diibl-	-s-dibil-	-s-iibal-	-m-dibil-	-m-iidbil-
	juss.	-iidbil-	-iidbal-	-t-dibil-	-t-dibil-	-t-diibil-	-s-dibil-	-s-iibal-	-m-dibil-	-m-iidbil-
ccy	aorist	-d**ii**g-	-iidg**i**-	-t-dig- -t-dig-	-t-dig-	-t-daag-	-s-dig-	-s-diig-	-m-dig-	-m-iidg**i**-
	juss.	-d**aa**g-	-iidg**a**-	-t-dig**a**- -t-diga	-t-dig**a**- -t-diga	-t-daag**a**-	-s-dig**a**-	-s-diig**a**-	-m-dig**a**-	-m-iidg**a**-
ccy morph	aorist	-**iig**i**Y**-	-iidg**i**Y-	-t-ig**i**Y- -t-dig-	-t-dig**i**Y-	-t-aag**i**Y	-s-dig**i**Y-	-s-iig**i**Y-	-m-ig**i**Y-	-m-iidg**i**Y-
	juss.	-**aa**gY-	-iidg**a**Y-	-t-dig**a**Y- -t-diga	-t-dig**a**Y- -t-diga	-t-ag**a**Y-	-s-ig**a**Y-	-s-iig**a**Y-	-m-ig**a**Y-	-m-idg**a**Y-

Finally, as is clear from (**4.4**), throughout the paradigm the *intensive v1* is long (VERB:12).

(4.4) Long *v1* in intensive ("Ablaut4")

rootclass	tense:	B-ntns	T-intns	S-intns	M-intns
ccc	present	**-edbil-**	**-t-daabiil-**	**-s-daabiil-**	**-m-eedbiil-**
	past	**daabil-**	**- - -**	**-s-daabil-**	**-m-eedbaal-**
	aorist	**iidbal-**	**-t-diibil-**	**-s-diibal-**	**-m-iidbil-**
	jussive	**iidbal-**	**-t-diibil-**	**-s-diibal-**	**-m-iidbil-**
	negative	**daabiil-**	**-t-daabiil-**	**-s-daabiil-**	**-m-eedbiil-**
ccy	present	**-eedgi-**	**-t-daagi-**	**-s-daagi-**	**-m-eedgi-**
	past	**-daag-**	**- - -**	**-s-daag-**	**-m-eedgaay-**
	aorist	**-iidgii**	**-t-daagi-**	**-s-diig-**	**-m-iidgi-**
	jussive	**-iidga**	**-t-daaga-**	**-s-diiga-**	**-m-iidga-**
	negative	**daagi-**	**-t-daagi-**	**-s-daagi**	**-m-eedgi-**
cc	present	**-eeliw-**	**- - -**	**- - -**	**- - -**
	past	**-laaw-**	**- - -**	**- - -**	**- - -**
	aorist	**-iiliw-**	**- - -**	**- - -**	**- - -**
	jussive	**-iiliw-**	**- - -**	**- - -**	**- - -**
	negative	**-laaw-**	**- - -**	**- - -**	**- - -**

Of the intermediate TEMPLATE nodes, much of BSTEM has already been covered. The boxed exceptions of (**3.12**) noted above show that in B *v2 pres intense or simple pl* is short; this is expressed in BSTEM:7, 8, thus providing an exception for the very general default VERB:7. The T, M, and SSTEM nodes are much simpler than B, and quite straightforward. Apart from spelling out the stem prefix (TSTEM:3, SSTEM:3, MSTEM:3), only a couple of particular-case vowel specifications are necessary (note SSTEM:4 and TSTEM:5, mentioned above). The +low quality of the *aa* vowel common to the past of the low- or non-transitive T and M stems (TSTEM:4, MSTEM:5; as opposed to the *i* of the high-transitive B and S) is possibly to be related to the *a* of equally low-/non-transitive R (RSTEM:5).

The situation is somewhat different, however, with the R stem, which is at least as complex as the B stem. This interesting stem requires three sets of generalizations. First, as can be seen in (**4.5**), where the corresponding B and T forms are given for contrast, this stem has a past-negative with the "archaic"

CCVC form (TEMPLATE:8), and *v2* = /a/ (RSTEM:5, as just mentioned). (The *ccy* change –dgay2- => -dge- will be part of that rootclass's morphophonemics.) Second, the identity of the R *present, aorist,* and *jussive* with the respective tenses in the T stem for the "strong verb" (*ccc* and *ccy*), is an instance of paradigm syncretism, so-called parasitic formations, wherein two sets of paradigm forms collapse. In terms of our generalization format, this can be indicated by simply saying that the paradigm categories at a certain node are simply to be treated as the corresponding categories at another node. Capturing this generalization involves a slight extension of the formalism to allow generalizations of the form $NODE_1$: path == $NODE_2$, meaning that the configuration $NODE_1$:<path> will be interpreted as $NODE_2$:<path>**.**

Finally, the *cc* verbs take infix -t- before the second consonant from the end (i.e., initial consonant; RSTEM:8), and share the shape and vocalism of the B stem $nonpast (RSTEM:7, cf. BSTEM:6), where *v1* is likewise subject to the generalization about jussive ablaut (i.e., follows the B pattern). All of this makes the R stem look less like the other "derived" stems, and more like a B stem, with a -t- infix in its core non-past forms (*present, aorist, jussive*)**.**

(4.5) Reflexive

Rootcl	tense:	num:	B	R	T	M	M-intns
Ccc	present	sg	-anbiil-	-t-dabiil-	-t-dabiil-	-m-dabiil-	-m-eedbiil-
	past		-dbil-	-db**a**l-	-t-dab**aa**l-	-m-dab**aa**l-	-m-eedb**aa**l-
	aorist		-iidbil-	-t-dibil-	-t-dibil-	-m-dibil-	-m-iidbil-
	jussive		-iidbil-	-t-dibil-	-t-dibil-	-m-dibil-	-m-iidbil-
	negativ		-dabiil-	-db**aa**l-	-t-dabiil-	-m-dabiil-	-m-eedbiil-
Ccy	present	sg	-dangi-	-t-dagi-	-t-dagi-	-m-dagi-	-m-eedgi-
	past		-dgi-	-dg**e**-	-t-dag**aa**y-	-m-dag**aa**y-	-m-eedg**aa**y-
	aorist		-diig-	-t-dig-	-t-dig-	-m-dig-	-m-iidgi-
	jussive		-daag-	-t-diga	-t-diga	-m-diga-	-m-iidga-
	negativ		-dagi-	-dg**aa**y-	-t-dagi-	-m-dagi-	-m-eedgi-
Cc	present	pl	-eeliw-	-eet-liiw-	-t-ooliiw-	-m-ooliiw-	- - -
	past		-liw-	-l**a**w-	-t-ool**aa**w-	-m-ool**aa**w-	- - -
	aorist		-iiliw-	iit-liw-	-t-uuliw-	-m-uuliw-	- - -
	jussive		-iiliw-	-iit-liw-	-t-uuliw-	-m-uuliw-	- - -
	negativ		-liiw-	-l**aa**w-	-t-ooliiw-	-m-ooliiw-	- - -

C. Spellout: "Weak"-Rootclass Morphophonemics

As mentioned, although *ccy* and *cc* rootclasses could in principle be treated completely in their own right, both on its own merits and following Semitist tradition the decision has been taken here to handle these rootclasses as "weak" versions of the *ccc* rootclass, sharing, to the extent possible the stem-template and V/C assignments of the "strong" rootclass. This entails a modest amount of morphophonemics. Without extensive comment, I give below in **(4.6)** a selection of thirteen forms illustrating all the "deviant behavior" of these rootclasses. These classes are handled in a few instances (MP 5, 11-13) by special template assignment, and in the rest by some obvious morphophonemic generalizations, given informally in **(4.7)**.

(4.6) "Derivation" of weak forms

	Pdgm Cell	Template	*v1*	*v2*	ccc	Morpho Ponme.	Surface	Rules
1	M smpl pres	m-*c1v1 c2 v2c3*	a	ii	-m-dabiil	m-w2aliiw	m-ooliiw	MP1,3
2	M smpl aor	m-*c1v1 c2 v2c3*	i	i	-m-dibil	m- w2iliw	m-uuliw	MP1,3
3	B smpl pres sg	*c1v1* n-*c2 v2c3*	a	ii	-dan-biil	- w2an-liiw	n-liiw	MP2
4	B smpl aor	*v1 c1c2 v2 c3*	ii	i	-iidbil	-ii w2liw	-iiliw	MP3
5	B intns neg	*c1 c2 v1 c3*	aa	--	NA	- w2laaw	-laaw	MP3
6	B smpl past	*c1 c2 v2 c3*	--	a	-dbal	-dgay2	-dge	MP4
7	R neg	*c1 c2 v2 c3*	--	aa	-dbaal	-dgaay2	-dgaay	MP5
8	T smpl pres	t-*c1v1 c2 v2c3*	a	ii	-t-dabiil	-t-dagiiy2	-t-dagi	MP6
9	B smpl aor	*v1 c1c2 v2 c3*	ii	i	-iidbil	-iidgiy2	-iidgi	MP7
10	T smpl aor	t-*c1v1 c2 v2c3*	i	i	-t-dibil	-t-digiy2	-t-dig	MP8
11	B smpl pres pl	*c1v1 c2 v2c3*	ee	i	NA	-deegiy2	-deeg	MP8
12	B intns juss	*v1 c1c2 v2 c3*	i	a	NA	-iidgay2	-iidga	MP9
13	B smpl juss	*c1 v2c2 (c3)*	--	aa	NA	-daagy2	-daag	MP9

(4.7) Weak Radical Morphophonemics

	Rules	Exx. from (4.7)
MP1	+vow $\rightarrow$ +rnd / {s,m,t} – w2 __	## 1, 2
MP2	a $\rightarrow$ Ø / w2 __	# 3
MP3	w2 $\rightarrow$ Ø	## 1, 2, 4, 5
MP4	a y2 $\rightarrow$ e /- *cc* __	# 6
MP5	aa y2 $\rightarrow$ aay	# 7
MP6	ii y2 $\rightarrow$ i	# 8
MP7	i y2 $\rightarrow$ i / *cc* __	# 9
MP8	i y2 $\rightarrow$ Ø	## 10, 11
MP9	y2 $\rightarrow$ Ø	## 12, 13

The *ad hoc* formalism of **(4.7)** is obviously that of the raditional ordered rule, and other formulations are possible. While it is not clear that it is ultimately the most appropriate, a DATR formulation of SPELLOUT is given in Appendix B, since that what actually lies behind the formal system that actually produces the stem paradigm of Appendix A**.**

5. Conclusion

The object of this report has been to explore how one might usefully describe morphological paradigms involving complex internal vowel changes. It has posed thus the question: Beyond the exemplary-paradigm-cum-principal-parts knowledge basis which underlies traditional practice and much modern WP theory, what informational and structural properties does a complex examplary paradigm have that might form the basis of its extension to a targeted population of (principal-part marked) lexical items, and how does one describe this? While spell-out and realization rules may not capture the full scope of analogical extension, if all the information needed for the production of words from lexemes is

indeed contained in the paradigm and principal parts, then it should be useful and even necessary to state in some consistent formalism exactly what the information is that is analogically extended.[27]

Moreover, on an initial, data-processing level, I submit that there is a real value for the intellectual discipline and rigor of a description for the linguist to have a simple and robust formalism within which one can and must state explicitly everything which is going on (every "realization") in a paradigm -- a formalism which is adequate to the task, but which makes the fewest possible assumptions about the data being described, or about the criteria for "best" description. In addition, to the extent that this formalism is computationally implemented, one obtains the kind of control over what was sometimes referred to as "mere" descriptive adequacy (a poor sister to "explanatory adequacy"). Such a description and implementation can assure both the investigator and the reader that the morphological "facts" (path/value statements) assembled *do* in fact account for everything. One can guarantee that the description presented is *complete* and that alternative descriptions, of the same data or parallel data, if they are made within the same framework, are *comparable*.**[28]**

For non-concatenative (and thus Afroasiatic) morphologies in general, I hope to have demonstrated, following the lead of Finkel and Stump, that the traditional and/or widely used constructs such as stem, root, and template fit very naturally into a simple declarative formalism for morphological description. What do we learn specifically about Beja? Although we have not been able to report on comparable data from related languages in this context, work done in this framework on Afar and Ge'ez show an interesting distribution of paradigm structuring generalizations. East Cushitic Afar emerges as a complex, but highly lexicalized variant/version of this kind of non-concatenative morphology, and Semitic Ge`ez as a highly regularized one. Beja comes out as a language with relatively simple lexical structure, but highly complex in the cross-cutting generalizations that, as has been seen, need to be made at every level. The complexity of the B stem, as opposed to the general uniformity of the T, S, and M stems, shows the former as a locus of possibly deeply entrenched development, with possible reflexes of a once more differentiated conjugation system based on root-class. The R stem also, which clusters nearly as many specific morphological facts as the B, with its -t- infixes and its very different treatment of the bicnsonantal roots, also looks like the result of long-term development. Thus far what description might suggest – for Beja the independent historical question then imposes itself: Whence this complexity? What is original and what is home-grown?[29]

On a more practical level, apart from its potentially salutary role as a preliminary test of the adequacy and completeness of an explicit description, a very useful "public" function of a DATR-like description of a language could be in a morphological database of a number of languages in a family or area. For such a database, rather than being limited to inert displays of data, what might it be like to be able to:

(1) generate the paradigm of any arbitrary verb;

(2) establish correlations (via a link, for example) between a cell of the paradigm, and the set of generalizations (path-value statements) that govern that cell – both the path-value statements themselves, but also, more helpful I suspect, human-readable expansions (incorporable into the formal description through "comment" elaborations);

---

27. Cf. Blevins (2006, pp. 359 f.) on proportional analogies, and his reservations concerning the ability of spell-out and realization rules to capture the full scope of analogical extension. On the linguistic problem of analogical extension in general, cf. Skousen 2002.

28. From a computational point of view there are of course other usable models, with their strengths and weaknesses, and they should be explored in this context. Certainly relevant is the work of Karttunen on Finite State Morphology and paradigm as output of a finite state device: Beesley & Karttunen 2003; Karttunen 2003.

29. Recall for example the question raised in **Sect. 2** (footnote 2) concerning the comparison between the middle-inflexing *ccy* stems and the -Ca(a)Caa- stem structures of Sasse's (1980) reconstructed Proto-East Cushitic *CC*V-root.

(3) conversely, correlations between a statement, and all the cells in the exemplary paradigm which are partially defined by that statement.

The ultimate test of course would be to see how insightful or helpful the lining up of (potentially) homologous or comparable generalizations are in a set of languages being contrasted in some perspective, areal, typological, or genetic.

Finally, and with great trepidation, the question can be asked: Could there be a linguistically interesting point to all of this? So far the point of the discussion has been, not that a DATR-like description is the Truth, but that this is potentially a pretty useful instrument. This has been, as it were, my constant disclaimer. But underneath it all, could the fact that morphologies can apparently be described in such a computationally simple way (as syntax arguably cannot) have somehow a deeper significance? Could this imply that a declarative, hierarchical, default-inheritance could be, at least in part, what morphology is? The question clearly goes beyond this report – but it does capture one aspect of the complex, and not always serene relationship between computation and the study of natural language.[30]

BIBLIOGRAPHY

Almqvist, H.1881. *Die Bischari-Sprache Tu:-Bed'a:wiye in Nordost-Afrika (beschreibend und verglei-chend dargestellet)*. Upsalla, Kaiserliche Akademie der Wissenschaften.

Appleyard, D. 2007. "Beja morphology", in Kaye, ed., *Morphologies of Africa and Asia*, Winona Lake, IN, Eisenbrauns, pp. 447-480.

Aronoff, M. 1994. *Morphology by itself: stems and inflectional classes*, Cambridge MA, MIT Press.

Briscoe, T. – de Paiva, V. – Copestake, A., eds. 1993. *Inheritance, Defaults, and the Lexicon.*, Cambridge, Cambridge University Press.

Beesley, K.R. – Karttunen, L. 2003. *Finite State Morphology*, Stanford, CSLI Publications.

Blevins, J. 2003. "Stems and paradigms", *Language* 79/4:737-767.

Brown, D. – Corbett, Gr. – Fraser, N. – Hippisley, An. – Timberlake, A. 1996. "Russian noun stress and network morphology", *Linguistics* 34:53-107.

Cahill, L. – Gazdar, G. 1997. "The inflectional phonology of German adjectives, determiners and pronouns", *Linguistics* 35:211-245.

Cahill, L. – Gazdar, G. 1999. "German noun inflection", *Journal of Linguistics* 35:1- 42.

Carstairs-McCarthy, An. 2006. "Affixes, stems and allomorphic conditioning in paradigm function morphology", *Yearbook of Morphology* 2005:253-281.

Corbett, Gr. – Fraser N. 1993. "Network morphology: a DATR account of Russian nominal inflection", *Journal of Linguistics* 29:113-142.

Daelemans, W. – Gazdar, G. – de Smedt, K. 1992. "Inheritance in natural language processing", *Computational Linguistics* 18:205-218.

Daelemans, W. – De Smet, K. 1994. ·"Default inheritance in an object-oriented representation of linguistic categories",. *International Journal of Human-Computer Studies* 41:149-177.

Evans, R. – Gazdar, G. 1996. "DATR: A language for lexical knowledge representation", *Computational Linguistics* 22:167-216.

---

30. Note for example Karttunen's lament over computational morphology's "curious non-relationship with 'real' linguistics extending back to more than three decades" (Karttunen 2003, p. 214).

Finkel, R. V – Shen, L. – Stump, Gr. – Thesayi, S. 2002. "KATR: A set-based extension of DATR", http://cs.engr.uky.edu/~gstump/katrreport.pdf:1-29.

Finkel, R. – Stump, Gr. 2002. "Generating Hebrew verb morphology by default inheritance hierarchies". Paper presented at *ACL-02 Workshop on Computational Approaches to Semitic Languages* (http://acl.ldc.upenn.edu/acl2002/SEMITIC/pdfs/Semitic022.pdf).

Fraser, N.M. – Corbett, Gr.G. 1994. "Gender, animacy, and declensional class assignment: a unified account for Russian", *Yearbook of Morphology* 1994:123-150**.**

Goldsmith, J. 1979. *Autosegmental Phonology*, MIT Ph.D. diss. 1976, New York.

Gragg, G. 2005. "Morphology and root structure: A Beja Perspective", *Aula Orientalis* 23:23-34.

Gragg, G. 2006 "The 'weak' verb: Akkadian and Beja", in Deutscher, G. G. Deutscher and Kouwenberg N.J.C., eds, *The Akkadian language in its semitic context: studies in the Akkadian of the third and second millennium BC* , Leiden, pp. 19-29.

Hayward, R. 1978. "The Prefix Conjugation in Afar", in Fronzaroli, P., ed., *Atti del secondo congresso internazionale di linguistica camito-semitica. Firenze, 16-19 aprile 1974*, Firenze, pp. 355-368.

Hayward, R. 1988. "In defence of the skeletal tier", *Studies in African Linguistics* 19/2:131-172**.**

Hayward, R. – Orwin, M. 1991. "The prefix conjugation in Qafar-Saho: the survival and revival of a paradigm – Part I", *African Languages and Cultures* 4/2:157-176.

Hippisley, An. 2001. "Word formation rules in a default inheritance framework: a Network Morphology account of Russian personal nouns", *Yearbook of Morphology* 1999:221-262**.**

Hockett, Ch. 1954. "Two Models of Grammatical Description". *Word* 10:210- 231

Hudson, R.A. 1973. "Syllables, moras and accents in Beja", *Journal of Linguistics* 9:53-63.

Hudson, R.A. 1973. "An Item-and-paradigm approach to Beja syntax and morphology", *Foundations of Language* 9:504-548.

Hudson, R.A. 1974. "A structural sketch of Beja", *African Language Studies* 15:111- 142.

Hudson, R.A. 1976. "The Bedja Language", in Bender, L., ed., *The Non-Semitic Languages of Ethiopia*. East Lansing, MI,: pp. 139-154.

Karttunen, L. 2003. "Computing with Realizational Morphology. In Computational Linguistics and Intelligent Text Processing", in Gelbukh, A., ed., *Computational Linguistics and Intelligent Text Processing* (Lecture Notes in Computer Science 2588), Heidelberg pp. 205-216.

Matthews, P. 1972. *Inflectional Morphology: a theoretical study based on aspects of Latin verb conjugation.*. Cambridge.

Matthews, P. 1991. *Morphology (2nd edition)*.. Cambridge.

McCarthy, J. 1979. *Formal Problems in Semitic Phonology and Morphology*, MIT Ph.D. diss. 979**.**

McCarthy, J. 1981. "A prosodic theory of nonconcatenative morphology", *Linguistic Inquiry* 12:373-418.

McCarthy, J. 2003. "Optimal Paradigms", *Rutgers Optimality Archive:485-1201* (http://roa.rutgers.edu /files/485-1201/485-1201-M*cc*ARTHY-0-5.PDF):1-75.

Parker, E.M. – Hayward, R.J. 1985. *An Afar-English-French Dictionary*, London.

Prasse, K.-G. 1973. *Manuel de grammaire touarègue (tăhăggart)*, Copenhagen**.**

Reinish, L. 1893-1894. *Die Bed'auye-Sprache in Nordost-Afrika. I, II, III, IV* (Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. Philosophisch-historische Classe128/III, 128/VII, 130/VII, 131/III), Wien.

Robins, R.H. 1959. "In defense of WP", *Transactions of the Philological Society* 1959:116- 144.

Roper, E.M. 1928. *Tu Bed'awie: An elementary handbook for the use of Sudan government officials*, Hertford.

Sasse, H.-J. 1980. "Ostkuschitische und semitische Verbalklassen", in *Diem-Wild 1980*:153-174.

Spencer, An. 2003. "Review Article: Putting some order into morphology: reflections on Rice (2000) and Stump (2001)", *Journal of Linguistics* 39:621-646.

Spencer, A. 2004. "Generalized Paradigm Function Morphology" (http://privatewww.essex.ac.uk/~spena/papers/GPFM.pdf):1-46.

Stump, Gr. 2001. *Inflectional morphology: a theory of paradigm structure* (Cambridge studies in Linguistics 93), Cambridge.

Stump, Gr. 2005. "Rules about paradigms", in Orgun, C. – Sells, P., eds, *Morphology and the Web of Grammar: Essays in Memory of Steven G. Lapointe* (Stanford Studies in Morphology and the Lexicon), Stanford CA, pp. 49-82.

Stump, Gr. 2006. "Some criticisms of Carstairs-McCarthy's conclusions", *Yearbook of Morphology 2005*: 283-303.

Voigt, R. 1986. "Die Verbalklassen des Grundstammes im Bedauye", *Africana Marburgensia* 19/2:26- 37.

Voigt, R. 1988. "Zur Bildung des Präsens im Bed'auye", in *Bechhaus-Gerst-Serzisko 1988*, pp. 379- 407.

Zaborski, An. 1988. "Remarks on the Verb in Beja", in Arbeitman, Y.A., ed., *Fucus - a Semitic/Afrasian gathering in remembrance of Albert Ehrman*, Amsterdam, pp. 491- 498.

Zaborski, An. 1997. "Problems of the Beja present seven years later", *Lingua Posnaniensis* 39:145-153.

APPENDIX A: BEJA STEM FORMS

CCC Stem Forms

tense:	num:	B-smple	B-intns	R	T-smple	T-intns	S-smple	S-intns	M-smple	M-intns
Present	sg	-danbiil$_1$	-eedbil$_2$-	-t-dabiil$_1$-	-t-dabiil$_1$-	-t-daabiil$_1$-	-s-dabiil$_1$-	-s-daabiil$_1$-	-m-dabiil$_1$-	-m-eedbiil$_2$-
	pl	-eedbil$_2$-								
Past		-dbil$_3$-	-daabil$_1$-	-dbal$_3$-	-t-dabaal$_1$-	- - -	-s-dabil$_1$-	-s-daabil$_1$-	-mdabaal$_1$-	-m-eedbaal$_2$-
Aorist		-iidbil$_2$-	-iidbal$_2$-	-t-dibil$_1$-	-t-dibil$_1$-	-t-diibil$_1$-	-s-dibil$_1$-	-s-diibal$_1$-	-m-dibil$_1$-	-m-iidbil$_2$-
Jussive		-iidbil$_2$-	-iidbal$_2$-	-t-dibil$_1$-	-t-dibil$_1$-	-t-diibil$_1$-	-s-dibil$_1$-	-s-diibal$_1$-	-m-dibil$_1$-	-m-iidbil$_2$-
Negative		-dabiil$_1$-	-daabiil$_1$-	-dbaal$_3$-	-t-dabiil$_1$-	-t-daabiil$_1$-	-s-dabiil$_1$-	-s-daabiil$_1$-	-m-dabiil$_1$-	-m-eedbiil$_2$-

CCY Stem Forms

tense:	num:	B-smple	B-intns	R	T-smple	T-intns	S-smple	S-intns	M-smple	M-intns
present	sg	-dangi$_1$-	-eedgi$_2$-	-t-dagi$_1$-	-t-dagi$_1$-	-t-daagi$_1$-	-s-dagi$_1$-	-s-daagi$_1$-	-m-dagi$_1$-	-m-eedgi$_2$-
	pl	-deeg$_1$-								
past		-dgi$_3$-	-	-dge$_3$-	-t-dagaay$_1$-	- - -	-s-dag$_1$-	-s-daag$_1$-	-m-dagaay$_1$-	-m-eedgaay$_2$-
aorist		-diig$_{1/4}$-	-iidgi$_2$-	-t-dig$_1$-	-t-dig$_1$-	-t-daag$_1$-	-s-dig$_1$-	-s-diig$_1$-	-m-dig$_1$-	-m-iidgi$_2$-
jussive		-daag$_4$-	-iidga$_2$-	-t-diga$_1$-	-t-diga$_1$-	-t-daaga$_1$-	-s-diga$_1$-	-s-diiga$_1$-	-m-diga$_1$-	-m-iidga$_2$-
negative		-dagi$_1$-	-daagi$_1$-	-dgaay$_3$-	-t-dagi$_1$-	-t-daagi$_1$-	-s-dagi$_1$-	-s-daagi$_1$-	-m-dagi$_1$-	-m-eedgi$_2$-

CC Stem Forms

tense:	num:	B-smple	B-intns	R	T-smple	T-intns	S-smple	S-intns	M-smple	M-intns
present	sg	-nliiw$_1$-	-eeliw$_2$-	-eet-liiw$_1$-	-t-ooliiw$_1$-	- - -	-s-ooliiw$_1$-	- - -	-m-ooliiw$_1$-	- - -
	pl	-eeliw$_2$-								
past		-liw$_3$-	-laaw$_5$-	-law$_3$-	-t-oolaaw$_1$-	- - -	-s-ooliw$_1$-	- - -	-m-oolaaw$_1$-	- - -
aorist		-iiliw$_2$-	-iiliw$_2$-	-iit-liw$_1$-	-t-uuliw$_1$-	- - -	-s-uuliw$_1$-	- - -	-m-uuliw$_1$-	- - -
jussive		-iiliw$_2$-	-iiliw$_2$-	-iit-liw$_1$-	-t-uuliw$_1$-	- - -	-s-uuliw$_1$-	- - -	-m-uuliw$_1$-	- - -
negative		-liiw$_1$-	-laaw$_5$-	-laaw$_3$-	-t-ooliiw$_1$-	- - -	-s-ooliiw$_1$-	- - -	-m-ooliiw$_1$-	- - -

84

Appendix B: A Morphological Description of Beja[1]

%% 1. VARIABLE DECLARATIONS & RADICAL DEFINITIONS

#vars $cons: b f t d s š k g x h l r m n w y w2 y2**.**
#vars $weak: w2 y2**.**
#vars $cons2: $cons - $weak **.**
#vars $vow: i e a o u ii ee aa oo uu**.**
#vars $vnonrnd: i e a ii ee aa **.**
#vars $vnhghlong: ee aa oo **.**
#vars $vnhghshrt: e a o **.**
#vars $vlow: a aa **.**
#vars $boundary: &**.**
#vars $phonseg: $cons2 + $vow + $boundary**.**

#vars $strong: ccc ccy**.**
#vars $aor: aor juss**.**
#vars $pres: pres neg**.**
#vars $nonpast: pres aor juss**.**
#vars $past: past neg**.**
#vars $deriv: s t m **.**

RADICAL1: <$cons#1 $cons#2 $cons#3> == $cons#1 **.**
RADICAL2: <$cons#1 $cons#2 $cons#3> == $cons#2 **.**
RADICAL3: <$cons#1 $cons#2 $cons#3> == $cons#3 **.**

%% 2. LEXICON

COLLECT-B:
    <> == BSTEM:<<rootclass>>
    <rootclass> == ccc
    <root> == d b l.

COLLECT-R:
    <> == RSTEM:<<rootclass>>
    <rootclass> == ccc
    <root> == d b l.

COLLECT-T:
    <> == TSTEM:<<rootclass>>
    <rootclass> == ccc
    <root> == d b l.

---

1. Appendix B contains all the morphological statements necessary to generate all the forms of Appendix A, except for some implementation-dependent redundancies. The most recent fully-runnable version can be downloaded from URL: <oi.uchicago.edu/pdf/graggbejamorphology.pdf>. The major feature not commented on in the text are the "<rootclass>" and "<stemclass>" path specifications on lexeme and stem nodes respectively, which have the effect of explicitly passing this information "up" the morphological hierarchy.

COLLECT-S:
    <> == SSTEM:<<rootclass>>
    <rootclass> == ccc
    <root> == d b l.

COLLECT-M:
    <> == MSTEM:<<rootclass>>
    <rootclass> == ccc
    <root> == d b l.

BRINGBACK-B:
    <> == BSTEM:<<rootclass>>
    <rootclass> == ccy
    <root> == d g y2.

BRINGBACK-R:
    <> == RSTEM:<<rootclass>>
    <rootclass> == ccy
    <root> == d g y2.

BRINGBACK-T:
    <> == TSTEM:<<rootclass>>
    <rootclass> == ccy
    <root> == d g y2.

BRINGBACK-S:
    <> == SSTEM:<<rootclass>>
    <rootclass> == ccy
    <root> == d g y2.

BRINGBACK-M:
    <> == MSTEM:<<rootclass>>
    <rootclass> == ccy
    <root> == d g y2.

BURN-B:
    <> == BSTEM:<<rootclass>>
    <rootclass> == cc
    <root> == w2 l w.

BURN-R:
    <> == RSTEM:<<rootclass>>
    <rootclass> == cc
    <root> == w2 l w.

BURN-T:
    <> == TSTEM:<<rootclass>>
    <rootclass> == cc
    <root> == w2 l w.

BURN-S:
    <> == SSTEM:<<rootclass>>
    <rootclass> == cc
    <root> == w2 l w.

BURN-M:
    <> == MSTEM:<<rootclass>>
    <rootclass> == cc
    <root> == w2 l w.

%% 3. TEMPLATE AND VERB NODES

BSTEM:
< > == VERB:<<stemclass>>	% 1
{stemclass} == base	% 2
{stempref} ==	% 3
{c2 pres smple sg} == n RADICAL2:<"<root>">	% 4
{v1 pres smple sg} == a	% 5
{v1 $nonpast} == ee	% 6
{vf2 intens pres} ==	% 7
{vf2 pl pres} ==	% 8
{v2 ccc intens $aor} == a	% 9

    .

RSTEM:
< > == VERB:<<stemclass>>	% 1
{stemclass} == refl	% 2
{stempref} ==	% 3
{intens} ==	% 4
{v2 $past } == a	% 5
{$nonpast $strong smple} == TSTEM	% 6
{v1 cc $nonpast} == ee	% 7
{c2 cc $nonpast } == t RADICAL2:<"<root>">	% 8

    .

TSTEM:
< > == VERB	% 1
{stemclass} == mid	% 2
{stempref} == t &	% 3
{v2 past} == aa	% 4
{vf1 ccy intens $aor} == long	% 5
{past intens} ==	% 6
{cc intens} ==	% 7

    .

MSTEM:
< > == VERB:<<stemclass>>	% 1
{stemclass} == pass	% 2
{stempref} == m &	% 3
{v1 intens} == ee	% 4
{v2 past} == aa	% 5
{cc intens} ==	% 6

SSTEM:
    < > == VERB   % 1
    {stemclass} == caus   % 2
    {stempref} == s   &   % 3
    {v2 ccc intens $aor} == a   % 4
    {cc intens} ==   % 5

VERB:
    < > == SPELLOUT:< "<stempref>" TEMPLATE > % 1
    <c1> == RADICAL1:<"<root>">   % 2
    <c2> == RADICAL2:<"<root>">   % 3
    <c3> == RADICAL3:<"<root>">   % 4
    {v1} == a   % 5
    {v2} == i   % 6
    {vf1} ==   % 7
    {vf2} ==   % 8
    {vf2 $pres} == long   % 9 Ablaut1
    {vf1 $aor} == high   % 10 Ablaut2
    {vf2 ccy juss} == low   % 11 Ablaut3
    {vf1 intens} == long   % 12 Ablaut4

%% 4. TEMPLATE ASSIGNMENT

TEMPLATE:
    {base ccy smple juss} == TEMPLATE4   % 1
    {base ccy smple} == TEMPLATE1   % 2
    {base cc intens $past} == TEMPLATE5 % 3
    {base smple pres sg} == TEMPLATE1 % 4
    {base smple past} == TEMPLATE3   % 5
    {base $nonpast} ==   TEMPLATE2   % 6
    {pass intens} == TEMPLATE2   % 7
    {refl $past} == TEMPLATE3   % 8
    < > == TEMPLATE1   % 9

TEMPLATE1: <> == "<c1>" "<v1>" "<vf1>" "<c2>" "<v2>" "<vf2>" "<c3>".
TEMPLATE2: <> == "<v1 >" "<vf1>" "<c1>" "<c2>" "<v2 >" "<vf2>" "<c3>".
TEMPLATE3: <> == "<c1>" "<c2>" "<v2>" "<vf2>" "<c3>".
TEMPLATE4: <> == "<c1>" "<v2>" "<vf2>" long "<c2>" .
TEMPLATE5: <> == "<c2>" "<v1 >" "<vf1>" "<c3>".

%% 5. Spellout

SPELLOUT:
    % Lengthening (Ablaut1)
    <a long> == aa <>
    <i long> == ii <>

    % Ablaut2 (aor-juss hightening)
    <$vnhghlong high> == ii <>
    <$vnhghshrt high> == i <>

% Ablaut3 (Final y2 only)
<i low y2> == a <>
<ii low> == aa <>

% Initial w2 morphophonemics
<$deriv & w2 a> == $deriv & oo <>   % MP1
<$deriv & w2 a high> == $deriv & uu <>   % MP1 Part of Ablaut1
<w2 $vlow> == <> % MP2
<w2> == <> % MP3

% Third y2 morphophonemics
<$cons#1 $cons#2 a y2> == $cons#1 $cons#2 e <> % MP4
<a long y2> == aa y <> % MP5
<i long y2> == i <> % MP6
<$cons#1 $cons#2 i y2> == $cons#1 $cons#2 i <> % MP7
<i y2 > == <> % MP8

% general spell-out
<$phonseg> == $phonseg <>
<> ==