Stem alternations and multiple exponence

Matthew Baerman and Greville G. Corbett

Abstract

In a canonical inflectional paradigm, inflectional affixes mark distinctions in morphosyntactic value, while the lexical stem remains invariant. But stems are known to alternate too, constituting a system of inflectional marking operating according to parameters which typically differ from those of the affixal system, and so represent a distinct object of inquiry. Cross-linguistically, we still lack a comprehensive picture of what patterns of stem alternation are found, and hence the theoretical status of stem alternations remains unclear. We propose a typological framework for classifying stem alternations, basing it on the paradigm-internal relationship between the features marked by stem alternations versus those marked by affixes. Stem alternations may mark completely different features from the affixes (§2), or the same features (§3). Within the latter, the values may match (§3.1) – a rare situation – or be conflated (§3.2). Conflation in turn may involve natural semantic/morphosyntactic classes (§3.2.1), or phonological conditioning (§3.2.2), or be morphologically stipulated (§3.2.3). These patterns typically reveal stems' continued allegiance to lexical as opposed to inflectional organizing principles.

1 Introduction

1.1 Canonical stems and odd ones

Consider the two verb paradigms below, from Daga, a language of the Trans-New Guinea phylum spoken in Milne Bay Province, Papua New Guinea. In (1a), there is a single invariant stem and five distinct suffixes (1SG and 1PL are identical). In (1b) there are six distinct stems but only two suffixes, one for 3PL and another for everything else.[1]

(1)   Daga, present continuous forms (Murane 1974: 51, 70)

	(a) 'get'	(b) 'go'
1SG	batnag–ivin	ang–ewan
2SG	batnag–ingi	ag–ewan
3SG	batnag–iwan	a–ewan
1PL	batnag–ivin	an–ewan
2PL	batnag–iwanin	ais–ewan
3PL	batnag–iwanum	amo–an

While subjective impressions are at best of dubious value, it is striking that (1a) looks normal and (1b) looks odd. In (1a) morphosyntactic values are realized by different suffixes, with the lexical stem remaining inert. This conforms to canonical inflection (Corbett 2007, presented below). The stem provides lexical information – which, for a single lexeme, 'ought' not to vary within the paradigm – and the affix provides the morphosyntactic information. But in (1b) the burden of morphosyntactic realization is almost entirely shunted to the stem, with the suffix contributing precious little information.

In the present example these subjective impressions fit with the general system of the language: (1a) represents the typical case, while the sort of pattern seen in (1b) is found with just a few lexemes. We would hardly expect to find the reverse distribution, so we can reassure ourselves that the canonical role of inflectional stems is to express lexical meaning, not grammatical. The stem alternation pattern in (1b) is lexically specified, and might be thought to be not really part of the inflectional system proper. On such a view stem alternations are a language-specific, indeed a lexeme-specific, peculiarity, of no real consequence to our understanding of inflectional morphology in general.

And yet stem alternations are a common feature of inflectional paradigms, so that no account of inflection would be complete without consideration of them. In order to give such an account, we need first to pose the simple question: What kind of stem alternations are there? This is not an easy question to ask, let alone answer, because it is not obvious what the appropriate terms of comparison are. For the present, we approach the question by looking at the extreme case (1b) and asking ourselves what exactly is odd about it. It is not the features being expressed; those are simply the normal ones expressed in the language. It is not the number of alternants; six is a lot, but plenty of languages have as many (e.g. French, as in Bonami & Boyé (2002: 55, fn 10)). What is strikingly unusual is the relationship between the stem alternants and affixation, which is the reverse of what we expect; and yet exactly what our expectations are is rarely explicitly formulated. In order to clarify this we propose a typology of the relationship between stem alternation patterns and affixal patterns.

1.2 Stems and affixes in canonical typology

As the paradigms in (1a) show, the distinction between stem and affix is fluid, presenting an obvious challenge to any attempt to oppose the two to each other. But

Table 1. Canonical inflection

	comparison across *cells* of a lexeme	comparison across *lexemes*
lexical material (≈ shape of stem)	same	different
grammatical material (≈ shape of affix)	different	same
outcome (≈ shape of inflected word)	different	different

there are, we contend, distributional regularities that these terms capture, which have proven to be indispensible in linguistic descriptions. Since the data show such variation, it would be valuable to have a fixed point from which we could measure. It is here that the canonical approach is of help. Consider the analogy of the cardinal vowels. We know that vowels vary according to how front and how close they are. Daniel Jones pushed both scales to their logical end point and proposed that the maximally front and maximally close vowel should be taken as a cardinal vowel, a standard from which other similar vowels can be measured. Note that this is a useful point to fix, irrespective of how many languages have such a vowel. Similarly we know that inflected forms combine grammatical and lexical information in varying ways. We would like a canonical standard from which to measure; this is proposed in Table 1, based on Corbett (2007: 9).[2]

Let us begin with a single lexeme (the middle column in Table 1). In the canonical situation, the lexical material, realized by the stem, is the same throughout the paradigm. On the other hand, the grammatical material, realized by the affix, is different in each cell. The outcome is that every morphosyntactic specification is realized differently for a given lexeme. If we now compare lexemes cell by cell, comparing across the lexicon (the right column in Table 1), then in the canonical situation the stem of each lexeme will be different from that of other lexemes (different lexical material is realized differently). The affixes can be the same across lexemes, and this produces the situation in which each form of each lexeme is unique.

This is not a pattern which we find often (though as a lexeme (1a) comes close to being canonical in this respect). The point is that the canonical situation is clearly defined and easily recognizable, so that we can use it as a fixed point in the theoretical space, from which we can measure non-canonical instances. To get a sense of the system, we should see how various non-canonical phenomena fit. Suppose the lexical material is not always realized in the same way for a given lexeme: then we have stem alternations, and in the extreme case we have suppletion. If the grammatical material is not always realized differently, within the lexeme, we are dealing with syncretism. Comparing across lexemes, if their lexical material is not realized differently, we have an instance of homophony. And finally, if grammatical material is expressed differently by different lexemes, we have allomorphy, and in more extreme cases we have inflectional classes. This schema gives us the basis for tackling the problem we have set.[3]

The schema in Table 1 has implications both for the distribution of inflectional features within the paradigm, and for their phonological/morphological expression.

Our focus in the present paper is the former, but we must also say a few words first about the latter. A canonical stem alternation should be suppletive and lexically idiosyncratic. In the following discussion we have used this criterion to identify stems, but it should be understood that the distinction we make between stems and affixes is relative to a particular paradigm; that is, where there are two parallel systems of alternation, one may have more lexical properties and the other more affixal properties. We have selected examples where this is fairly clear. Nevertheless, languages abound in indeterminate cases. Consider the distribution of suffixed -*gi* in the paradigms in from Gaagudju, a language of the Top End of Australia.

(2)   Gaagudju (Harvey 2002: 410, 429, 457)

	(a) 'hide (INTR)'	(b) 'swell'	(c) 'be blocked'
PP	–mardéedji-gi	–balbarée-gi	–djurríinjdji
PI/PIRR	–mardéedji-ri	–balbarra-gée-ni	–djurrinjdji-gée-ni
PR	–mardéedji-ri	–balbarrée-gi	–djurríinjdji-gi
FUT	–mardéedji	–balbarrée-gi	–djurréenjdji
EV	–mardéedji-gi	–balbarra-gée-ya	–djurrinjdji-gée-ya

NB: subject prefixes are omitted here

In (2a) -*gi* appears to be an affixal marker of the past perfective and evitative (and is so treated by Harvey (2002)). In (2b) it is found throughout the entire paradigm, and would appear to be the final syllable of the stem (the alternation of *i* and *a* to *e* under stress is an at least semi-regular morphophonological process). But in (2c), which appears to be a hybrid of the two patterns, it is unclear what is going on: is -*gi* part of an alternating stem or is it a tense-aspect-mood suffix? A definite answer as such is of little use by itself, but an understanding of the properties of clearer cases may help us to at least talk sensibly about what is going on in such paradigms.

1.3 Stems in morphological theory

The status of stems varies among different morphological theories. In some, e.g. Paradigm Function Morphology (Stump 2001) and Network Morphology (Brown & Hippisley forthcoming), stems are treated as a distinct object by the rules which construct the paradigm. Other approaches deny any special status to stems, either because all components of the inflected word form are construed as the same type of object, as in Distributed Morphology (Embick & Halle 2005), or because no status is accorded at all to anything below the level of the fully inflected word, as in word-based morphology (Blevins 2006). Our purpose is not to argue one way or another for the theoretical status of stems, but rather to highlight patterns which morphological theories ought to take into account. This task is quite independent of whether or not stems are accorded a distinct status by the theory. Indeed, even the stem-free approach outlined by Blevins (2006), although expressed in terms of the implicative relations between whole word forms within a paradigm, nevertheless assumes a division between different elements which could well be labelled 'stem' and 'affix', if only informally.

Consider his example from Sámi (Finno-Ugric) in (3). The forms illustrated show an alternation between a geminate and a short consonant, /tt/ ∼ /t/ with 'piece' and /gg/ ∼ /g/ with 'work'.

(3) First declension nouns in Sámi (Blevins 2006: 546, citing Bartens 1989)

	'piece'		'work'	
	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM	bihttá	bihtát	bargu	barggut
GEN/ACC	bihtá	bihtáid	barggu	bargguid
ILL	bihttái	bihtáide	bargui	bargguide
LOC	bihtás	bihtáin	barggus	bargguin
COM	bihtáin	bihtáguin	bargguin	bargguiguin
ESS		bihttán		bargun

The distribution of geminate versus short alternants in the two nouns are mirror images of each other, which allows the abstraction of the implicational structure in Table 2.

Table 2. Implicational structure of (3)

GENITIVE/ACCUSATIVE SG		NOMINATIVE SG	
↔	LOCATIVE SG	↔	ILLATIVE SG
↔	ALL PLURALS	↔	ESSIVE
↔	COMITATIVE SG		
	= LOCATIVE PL		

Even though the network in Table 2 makes no explicit reference to 'stems' versus 'affixes', segmentation into the equivalent components must be assumed in order to interpret the relationships. The interpredictability expressed by the double arrow '↔' translates into identity of the first two syllables, minus the final consonant of the second syllable, if there is any, and minus any third syllable. The implied division is given in (4a, b).[4]

(4) Elements referenced and omitted in Table 2

	a. referenced elements		b. omitted elements	
	'piece/work'			
	SINGULAR	PLURAL	SINGULAR	PLURAL
NOM	bihttá/bargu	bihtá/barggu		–t
GEN/ACC	bihtá/barggu	bihtá/barggu		–id
ILL	bihttá/bargu	bihtái/barggu	–i	–ide
LOC	bihtá/barggu	(= COM SG)	–s	
COM	bihtá/barggu	bihtá/barggu	–in	–iguin
ESS		bihttá/bargu		

In whatever way we construe the distinction between (4a) and (4b), there are distinct generalizations that apply to each, and any analysis will want to capture these. In particular, the elements in (4b) are identical for all nouns, while the implicational structures corresponding to Table 2 actually vary between different inflection classes (along with their phonological form, of course). As argued above in §1.2, in such cases where there are two or more cross-classifying systems within the paradigm, one of them will typically have more lexical properties and warrant being called a 'stem', if only contingently.[5]

Our strategy will be to look for examples that are relatively uncontroversial in terms of the stem–affix divide; that is, those that are sufficiently close to the canonical. We then examine the relations between the stem and affix, examining what information each provides. This will enable us to see why the Daga forms in (1b) appear so surprising. First we take the simpler case, where different features are involved (§2), and then look at those where the features are the same (§3).

2 Different features

We consider here instances where the stem and affix realize different features. At one extreme these may be in complementary distribution. For example, in Modern Greek, the verb stem marks aspect (imperfective vs perfective), while the affixes mark person and number of the subject, and tense. In (5), ðén- is the imperfective (traditionally *imperfect*) stem and ðés- the perfective (traditionally *aorist*).

(5)   Modern Greek 'tie' (Holton, Mackridge & Philippaki-Warburton 1997: 120f)

	imperfective	perfective
1SG NONPST	ðén-o	ðés-o
2SG NONPST	ðén-is	ðés-is
3SG NONPST	ðén-i	ðés-i
1PL NONPST	ðén-ume	ðés-ume
2PL NONPST	ðén-ete	ðés-ete
3PL NONPST	ðén-un	ðés-un
1SG PST	é-ðen-a	é-ðes-a
2SG PST	é-ðen-es	é-ðes-es
3SG PST	é-ðen-e	é-ðes-e
1PL PST	ðén-ame	ðés-ame
2PL PST	ðén-ate	ðés-ate
3PL PST	é-ðen-an	é-ðes-an

Such instances of a complete division of labour are hard to find, but it seems likely that the typical example of this type is one in which the feature marked on the stem is one with more lexical semantic consequence (*relevance* in Bybee's (1985) terms) than that marked affixally, as above. Indeed, a moderately plausible alternative view of the word in (5) would be that the aspectual pair was actually two distinct lexemes, with aspect a lexical rather than an inflectional feature.

A more common situation is to find some degree of overlap, such that some feature is shared by stem and and affix, and some other feature is the sole responsibility of one or the other. For example, in Yabem, an Oceanic language of Papua New Guinea, prefixes distinguish person, number and mood, while stems distinguish only number and mood. The full paradigm is given in (6), while (7) repeats the information with the prefixes and stems separated for clarity. (Note that 1st exclusive plural and 2nd person plural are always syncretic.)

(6)   Yabem 'go (there)' (Dempwolff 1939 [2005]: 17)[6]

	realis	irrealis
1SG	gà–jà	jà–nòm
2SG	gò–jà	ò–nòm
3SG	gè–jà	è–nòm
1INCL	tá–já	tá–ná
1/2PL	á–já	á–ná
3PL	sé–já	sé–ná

(7)   Example (6) decomposed

	prefix		stem		
	realis	irrealis	realis		irrealis
1SG	gà–	jà–		SG	–nòm
2SG	gò–	ò–			
3SG	gè–	è–			
1INCL	tá–		–já		
1/2PL	á–		+	PL	–ná
3PL	sé–				

Both prefixes and stems conflate values. Even so, the prefix paradigm still requires reference to all three features of person, number and mood, while the stem paradigm requires reference only to number and mood. Thus prefix and stem partly coincide in the features they mark, but the prefix marks an 'extra' feature. On the assumption that the affix is the canonical locus of inflection, this represents a more canonical distribution than if it were the stem hosting the extra feature.

3 Same features

This brings us to the more complex situation where stem and affix distinguish the same features, as is the case above with mood and subject number in the Yabem example. For expository clarity we illustrate the typological parameters below with instances of complete overlap, i.e. where the feature inventory is exactly the same for both. In this situation what is relevant is the relationship of the feature values to each other.

3.1 Matching values

Logically the simplest relationship is one where the values match. In the noun paradigm in (8a), from the West Nilotic language Nuer, stem and affix alternate for the same case-number values: nominative singular, genitive/locative singular and plural.

(8)   Nuer noun paradigms (Frank 1999: 84 f., 87)

	'bark (of dog)'	'ear'	'meat'
NOM SG	gua	jith	riŋ
GEN/LOC SG	gui̠-kä	jith-kä	riän̠
PL	guiä̠-ni̠	jith-ni̠	riiŋ̠

To the extent that such perfect matching occurs it is restricted to a small inventory of values. Even in the Nuer example it is probably a lucky accident, as we also find the same affixation pattern with no stem alternation (8b), and the same stem alternation pattern with no affixation (8c).

3.2 Conflation

Much more commonly we find that when the same feature is marked on both stem and affix, at least one of these components conflates some of the values. There are three possibilities here, which we discuss in turn.

3.2.1 Natural classes of features

Sometimes the forms resulting from conflated values can be attributed to the meaning of the features themselves. Consider the past tense transitive verbal paradigm in (9a) from Koyi Rai, a language of the Kiranti group of the Himalayan branch of Tibeto-Burman, spoken in Eastern Nepal. There are eleven person-number values which map onto two stems: jɔmd- is found in the singular (of all persons) and in the 3rd person (all numbers), while jɔmts- is found in the dual and plural of 1st person (inclusive and exclusive) and 2nd person. In (9b) the order of the values has been shifted from that used by Lahaussois in order to consolidate the stems.

(9)   Past tense of Koyi Rai 'hit' (with implied 3SG object) (Lahaussois 2009: 12)

	a. full paradigm		b. stem only
1SG	jɔmd-uŋa	1SG	
1INCL DU	jɔmts-asi	2SG	
1EXCL DU	jɔmts-asu	3SG	jɔmd-
1INCL PL	jɔmts-aki	3DU	
1EXCL PL	jɔmts-akɔ	3PL	

2SG	jɔmd-una	1INCL DU	
2DU	jɔmts-asina	1EXCL DU	
2PL	jɔmts-ani	1INCL PL	
3SG	jɔmd-u	1EXCL PL	jɔmts-
3DU	jɔmd-usi	2DU	
3PL	jɔmd-uni	2PL	

This two-way stem opposition is found in a number of verbs, and may be realized in various ways; for the moment we adopt the labels 'stem A' and 'stem B':

(10) Past tense stem oppositions in Koyi Rai (Lahaussois 2009: 14–16)

	'hit'	'throw'	'shoot'	'bring'	'eat'
stem A	jɔmd-	ward-	ɔpʰ-	hoʔ-	pa-
stem B	jɔmts-	war-	ɔpts-	huʔ-	pja-

At first glance the distribution of the stems appears to be disjunctive: stem A (*jɔmd-*) is found with the singular (of any person) or 3rd person, while stem B (*jɔmts-*) is found with the dual and plural of 1st and 2nd person. However, Lahaussois (2009) interprets the alternation as realizing a kind of inverse marking, involving person and number hierarchies. In the person hierarchy 1st and 2nd person outranks 3rd person, and within 1st and 2nd person the singular outranks dual and plural. The 'direct' form (stem A) is used where the subject matches or outranks the object on the compound hierarchy, while the 'inverse' form is used where the object outranks the subject. (Lahaussois does not flesh out the hierarchy in enough detail to verify all the combinations.) On this interpretation the stems reflect a natural class grouping of person values (1st and 2nd person combined as 'discourse participant' or 'real' person) and number values (dual and plural combined as 'non-singular').

3.2.2 Phonological conditioning

The Koyi Rai stem alternation discussed above lends itself to another interpretion (pointed out by an anonymous reviewer, and partly suggested by Lahaussois herself): the phonological environments in which the two stems appear are contrastive. This will be even clearer if we look at it alongside the non-past paradigm (11), where a third stem *jɔm-* appears in the environments corresponding to the 'direct' stem A. This third stem appears before consonant-initial suffixes, stem A appears before *u*-initial suffixes, and stem B appears elsewhere. Although there are no purely phonological rules that would account for this, there is a phonological correlation.

(11) Koyi Rai 'hit' (with implied 3SG object) (Lahaussois: 2009: 12)[7]

	a. full paradigm		b. stem only	
	NON-PAST	PAST	NON-PAST	PAST
1SG	jɔm-dɔ	jɔmd-uŋa		
2SG	jɔm-dana	jɔmd-una		
3SG	jɔm-da	jɔmd-u	jɔm-	jɔmd-
3DU	jɔm-dasi	jɔmd-usi		
3PL	jɔm-dani	jɔmd-uni		
1INCL DU	jɔmts-isi	jɔmts-asi		
1EXCL DU	jɔmts-isu	jɔmts-asu		
1INCL PL	jɔmts-iki	jɔmts-aki		
1EXCL PL	jɔmts-ekɔ	jɔmts-akɔ	jɔmts-	
2DU	jɔmts-isina	jɔmts-asina		
2PL	jɔmts-ini	jɔmts-ani		

The direct∼inverse contrast would then be encoded in the suffixes themselves, and only indirectly in the stem.[8] In as much as the shape of the suffixes must be accounted for in any case, this would be the more parsimonous approach. Nevertheless it should be borne in mind that not all stem alternations in Koyi Rai lend themselves to such an interpretation, since there is a class of verbs which display a non–past ∼ past stem alternation quite independent of the phonology of the suffix, e.g. the 3rd singular non–past *re-di* 'laugh' versus the 3rd singular past *rja-di* (Lahaussois 2009: 11).

A more transparent example of phonological conditioning comes from Gumbaynggir, a Pama-Nyungan language of New South Wales, Australia. In the sample verb paradigm in (12) there are 18 different suffixes that map onto four stems, *njaig-*, *njai-*, *nja:g-* and *nja:-*.

(12) Gumbaynggir 'see' (Smythe 1948–49: 181)

	present tense			non-present tense
IND	PRS-INDEF	njaig-i	IMP	SG OBJ	nja:g-a
IND	PRS-FUT	njaig-iw	IMP	PL OBJ	nja:g-ili
IND	PRS-PST	njaig-iŋ	SBJV		nja:g-ala
PTCP	PRS-INDEF	njaig-indi	INF		nja:g-eigu
PTCP	PRS-DEF	njai-djindi	VBL NOUN	ACTION	nja:g-igam
IND	PRS-DEF	njai-dji	VBL NOUN	AGENT	nja:g-igir
			PTCP	FUT IMM	nja:g-undi
			IND	FUT IMM	nja:g-u
			IND	FUT REM	nja:-jiw
			IND	PST	nja:-waŋ
			PTCP	FUT REM	nja:-jiwundi
			PTCP	PST	nja:-waŋandi

In part the distribution of the four stems follows 'natural class' conflation as described above (§3.2.1): forms with the stem vowel *-ai-* all share a present meaning (as reflected in the glosses), opposed to forms in *-a:-*. But the alternation of forms with or without a stem-final *-g-* has no featural correspondence. Nevertheless, a brief inspection of the paradigm shows that it has instead a phonological correspondence: forms with *-g-* occur before vowel-initial suffixes, while forms in which the *-g-* has been deleted occur before consonant-initial suffixes.

3.2.3 Morphomic patterns

Other examples show that the conflation may also be arbitrary, corresponding to nothing either in the feature system nor, at least synchronically, to anything phonological. Rather, the stems are morphomic (Aronoff 1994). Indeed, stem alternations constitute the most often cited examples of morphomes, with those of Romance verbs being particularly familiar (Bonami & Boyé 2002, Maiden 2005); these are striking not just for the morphotactic diversity and morphosyntactic arbitrariness of the patterns, but for their diachronic persistence, having been maintained and propagated throughout the whole family. A similar situation is found in the Chinantecan languages, of the Otomanguean family, spoken in Central Mexico. Various patterns of stem alternation have been reconstructed for the proto-language (Rensch 1989: 21f), and continue to be found throughout the family. One example is shown in (13), from Lealao Chinantec.

(13) Lealao Chinantec 'listen' (Rupp 1996: 424, 427–429)

	PRS	FUT	PRT
1SG	nuu–y^4	hi^2-nuu–y^{42}	ma^3-nuú–y^4
1PL INCL	nuu^{32}-a^2	hi^2-niuu^2-a^2	ma^3-niuu^2-a^2
1PL EXCL	nuu^{32}-ah^1	hi^2-niuu^2-ah^1	ma^3-niuu^2-ah^1
2SG	nuu–y^3	hi^2-nuu–y^3	ma^3-niuu–y^3
2PL	nuu^3-ah^3	hi^2-nuu^3-ah^3	ma^3-niuu^3-ah^3
3	nuu^3	hí^4-nuú^4	ma^3-nuú^3

Note: superscript numerals indicate tone height, with '1' as the highest and '4' as the lowest. The acute accent indicates so-called ballistic stress, a prosodic feature which we subsume here under the general rubric 'tone'.

Inflection for subject person–number and tense involves three intersecting elements:

- segmental affixation (suffixes for person–number and prefixes for tense)
- tonal alternation
- stem alternation (palatalization of the initial consonant)

Let us set aside the tonal alternations on the stem for the moment, and concentrate on the affixes and segmental alternations of the stem. The affixes delimit 18 distinct cells, as shown in (14a). The two stems, palatalized *niuu* and plain *nuu*, divide the paradigm into two parts, according to the 'B' pattern (the designation is due to Merrifield 1968), as shown in (14b), where the palatalized stem is indicated by shading. The stem *niuu* is found in the 1st plural (inclusive and exclusive) of the future, and in the 1st plural and 2nd person (singular and plural) of the preterite, while stem *nuu* is found in the remainder of the paradigm.

(14) Affixal and stem patterns extracted from the verb 'listen' in (13)

	a. affixes			b. stem		
	PRS	FUT	PRT	PRS	FUT	PRT
1SG	Ø-...-á⁴	hi²-...-á⁴	ma³-...-á⁴	nuu	nuu	nuu
1PL INCL	Ø-...-a²	hi²-...-a²	ma³-...-a²	nuu	niuu	niuu
1PL EXCL	Ø-...-ah¹	hi²-...-ah¹	ma³-...-ah¹	nuu	niuu	niuu
2SG	Ø-...-u³	hi²-...-u³	ma³-...-u³	nuu	nuu	niuu
2PL	Ø-...-ah³	hi²-...-ah³	ma³-...-ah³	nuu	nuu	niuu
3	Ø-...-Ø	hí⁴-...-Ø	ma³-...-Ø	nuu	nuu	nuu

This distribution of stem alternants cannot be derived from the distribution of affixes in (14a), nor is it morphosyntactically coherent in any obvious way. We assume it must simply be morphologically specified.

Other stem alternation patterns are possible too. For example, alongside the 'B' pattern seen in (14b), some verbs such as 'spill' (15) display the 'C' pattern alternation, in which the palatalized stem includes the 1PL present and all 3rd person forms as well.

(15) Stem alternation pattern C, illustrated with stem alternants of the verb 'spill' (Rupp 1996: 458)

	PRS	FUT	PRT
1SG	tũũ	tũũ	tũũ
1PL INCL	tiũũ	tiũũ	tiũũ
1PL EXCL	tiũũ	tiũũ	tiũũ
2SG	tũũ	tũũ	tiũũ
2PL	tũũ	tũũ	tiũũ
3	tiũũ	tiũũ	tiũũ

Both patterns can in turn occur nested within the same paradigm, given the right morphological means. This is seen in (16), in a verb that has both suppletion and palatalization (here realized as a vowel alternation).

(16) Verb with two nested patterns

'take' (Rupp 1996: 460)

	PRS	FUT	PRT
1SG	uuyh^4	hi^2-uuyh^{42}	ma^3-uúyh^4
1PL INCL	hɨ̃h^{32}-a^2	hi^2-hĩh^2-a^2	ma^3-hĩh^2-a^2
1PL EXCL	hɨ̃h^{32}-ah^1	hi^2-hĩh^2-ah^1	ma^3-hĩh^2-ah^1
2SG	uúyh^3	hi^2-uúyh^3	ma^3-hĩh^3-u^3
2PL	uúyh^3-ah^3	hi^2-uúyh^3-ah^3	ma^3-hĩh^3-ah^3
3	hɨ̃h^{32}	hí^4-hɨ̃h^2	ma^3-hɨ̃h^4

The suppletive stem *hɨ̃h* is opposed to *uuyh* according to the 'C' pattern, and nested within that the palatalized variant *hĩh* (i.e. with the stem vowel /ĩ/ in place of /ɨ̃/) within it according to pattern 'B'. Unsurprisingly, it is the more substantial alternation (suppletion) which defines the context for the more superficial one (vowel alternation).

This leaves still the vexing question of how to treat the tonal alternations. From a purely morphotactic point of view the status of tone is unclear: is it a property of the stem, or is it a discrete autosegment? Giving it a label is itself of little value, but if we adopt the canonical approach as outlined above (§1), we can characterize its behaviour, and in particular, the nature of its alternation pattern. Abstracting away from the actual tone values, there are 57 distinct paradigms in terms of the set of oppositions between cells, all of which show some conflation of the full inventory of 18 paradigmatic cells. Some of the conflations line up with the natural class pattern described above (§3.2.1); thus, 1st person plural inclusive and exclusive are always identical, as are 2nd person singular and plural. But many of the patterns combine disjunctive person–number and tense values in a way similar to the behaviour of the stem alternations that we have just seen, though the patterns never coincide.

On the other hand, the sheer number of distinctions brought about by tonal alternations gives them the flavour of affixal patterns, with paradigms having up to seven distinct tones, as in (17); recall that '1' is the highest and '4' the lowest tone.

(17) Tone class I–2.3 (Rupp 1996: 464)

	PRS	FUT	PRT
1SG	32	1	2
1PL INCL		42	
1PL EXCL			
2SG		3	4
2PL			
3	4	4	

The tonal alternations here approach the pattern of affixal alternations, suggesting that the intermediate status of tone in this language is not just morphotactic, but also distributional (recall that the acute indicates 'ballistic' stress). That is, the lexical specificity of tone patterns makes them look like a property of the stem, while the sheer number of feature values they mark make them look like affixes.

4 Conclusion

Let us return to the problem posed by the Daga data in (1). We suggested that in the canonical situation, the lexical information is realized by the stem, and the grammatical information by the affix. If we were to find the reverse situation, systematically, this would mean that we had labelled stem and affix incorrectly. Note, however, that we do find instances of stems marking all the information, as in one of our Nuer nouns in (8). Against the background of the full system of the language, however, we can accept our initial impression that the lexemes with the inert stem (as in 1a) are expected, and those like (1b), with more grammatical information realized on the stem than on the affix, are indeed unusual and surprising.

    Apart from the cleanest type of system, where all grammatical information is expressed by the affix, we find other possible divisions of labour between stem and affix. In these instances it is equally important to be clear about whether we are discussing a system or individual lexemes within it. The most minor deviation from the canonical type is one in which both stem and affix realize grammatical information, but they mark different features. Here we expect that stem alternations will mark the features with greater semantic relevance (in Bybee's 1985 terms). So if the features are aspect and person, being marked on verbs, we expect aspect (having greater semantic relevance to the lexical semantics of the verb) to be realized on the stem, and person to be realized by affixation. This expectation is often met, both at the level of the system and of individual lexemes within it. However, if there are multiple layers of formants we may struggle to determine whether we have a complex stem or a complex affix. In the extreme case, where information is conveyed regularly and primarily through the stem, there can be a case for saying that we are dealing with different, derivationally related, lexemes (an issue raised by the Greek data in §2).

More interesting are the cases where stem and affix display multiple exponence. We provided an initial typology here. The conceptually simplest system is that in which not only the available features match, but also the values match between stem and affix. This situation is surprisingly rare; we illustrated it with Nuer nouns, and then only for a subset of the lexical items. What we find much more frequently is some sort of conflation between the values expressed by the stem and the affix. In some instances the conflation may reflect a natural class, as in Koyi Rai verbs. We should be careful not to over-interpret such instances, because there are many systems with no such justification. Sometimes we can identify morphophonological conditioning, as was part of the pattern in Gumbaynggir. Or we may find purely morphological patterns, as in Lealao Chinantec. Indeed, both types of pattern most likely have their origin as the by-product of affixation, with the particulars of the phonological interaction between stem and affix determining whether one ends up with a morphosyntactically coherent pattern or a purely morphological one. The robustness of stem alternation patterns such as those seen in Chinantec shows that being morphosyntactically incoherent is no barrier to morphologization.

Notes

1. This work has been funded by the European Research Council (grant ERC-2008-AdG-230268 MORPHOLOGY); whose support is gratefully acknowledged. We would also like to thank audiences at the Workshop "Stems in Inflection and Word Formation" at the 14th International Morphology Meeting, Budapest, May 2010, and at the University of Kentucky, November 2010, who made helpful comments on the material discussed in this paper. We appreciate the useful suggestions of the editors and two anonymous referees.
2. Here we abstract away from fundamental differences in the composition or structure of the paradigm, as found for instance when a part of the paradigm is periphrastic or deponent.
3. For recent work in this tradition, see Brown, Chumakina and Corbett (forthcoming).
4. Alternatively, '↔' might be construed as referring specifically to the CC ∼ C alternation. But in that case we would still have to account for the invariant portion of the paradigm, which would give us a three-way division along the lines of 'stem', 'stem final consonant' and 'suffix'.
5. In the abstractive approach described by Blevins (2006), as in other exemplar-based approaches, the implicational structures that describe the paradigm need not be segregated, meaning that a word form can accommodate multiple conflicting segmentations. This is fully in keeping with the canonical approach described in the present paper.
6. The practical orthography employed by Dempwolff has been modified to match the more standard one used by Ross (1993). Acute accent indicates high tone, grave accent low tone.
7. The reader may note a certain complementarity between the non-past forms where the suffix begins with *d* and the stem A past tense forms where the stem ends in *d* (e.g. 3SG non–past *jɔm-da* versus past *jɔmd-u*) and wonder if the segmentation is correct. But as can be seen in the last three examples in (10), a verb's stem A does not always end in *-d*: corresponding to *jɔmd-* other verbs may have stem-final consonant deletion ('shoot', 'eat') or ablaut ('bring'). So the complementarity in (11) is not systematic.

8. Lahaussois (2009) interprets the *jɔm-* ~ *jɔmts-* alternation as phonologically conditioned, and the *jɔmd-* ~ *jɔmts-* alternation as determined by the subject and object feature values.

References

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

Bartens, Hans-Hermann. 1989. *Lehrbuch der saamischen (lappischen) Sprache*. Hamburg: Helmut Buske Verlag.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42. 531–573.

Bonami, Oliver & Gilles Boyé. 2002. Suppletion and dependency in inflectional morphology. In: Frank van Eynde, Lars Hellan & Dorothee Beermann (eds), *Proceedings of the 8th International HPSG Conference*, 51–70. Stanford: CSLI.

Brown, Dunstan, Marina Chumakina & Greville G. Corbett (eds). Forthcoming. *Canonical morphology and syntax*. Oxford: Oxford University Press.

Brown, Dunstan & Andrew Hippisley. Forthcoming. *Network morphology*. Cambridge: Cambridge University Press.

Bybee, Joan. 1985. *Morphology*. Amsterdam/Philadelphia: John Benjamins.

Corbett, Greville G. 2007. Canonical typology, suppletion and possible words. *Language* 83. 8–42.

Dempwolff, Otto, Joel Bradshaw & Francisc Czobor. 2005. *Otto Dempwolff's 'Grammar of the Jabêm Language in New Guinea'* [translated and annotated edition of *Grammatik der Jabêm Sprache auf Neuguinea* (1939)]. Honolulu: University of Hawai'i Press.

Embick, David & Morris Halle. 2005. On the status of stems in morphological theory. In Twan Geerts, Ivo van Ginneken & Haike Jacobs (eds), *Proceedings of Going Romance 2003*, 59–88. Amsterdam: John Benjamins.

Frank, Wright J. 1999. *Nuer noun morphology*. M.A. thesis, State University of New York, Buffalo.

Harvey, Mark. 2002. *A grammar of Gaagudju*. Berlin: Mouton de Gruyter.

Holton, David B, Peter Mackridge & Irene Philippaki-Warburton. 1997. *Greek: a comprehensive grammar of the modern language*. London: Routledge.

Lahaussois, Aimée. 2009. Koyi Rai: An initial grammatical sketch. *Himalayan Linguistics Archive* 4. 1–33. Available at <http://www.linguistics.ucsb.edu/HimalayanLinguistics/grammars/HLA04.html>.

Maiden, Martin. 2005. Morphological autonomy and diachrony. In: Geert Booij & Jaap van Marle (eds). *Yearbook of Morphology 2004*, 137–175. Dordrecht: Foris.

Merrifield, William R. 1968. *Palantla Chinantec grammar* (Papeles de la Chinantla 5, Serie Científica 9). Mexico: Museo Nacional de Antropología.

Murane, Elizabeth. 1974. *Daga grammar. From morpheme to discourse*. Norman, Oklahoma: SIL.

Rensch, Calvin R. 1989. *An etymological dictionary of the Chinantec languages* (Studies in Chinantec languages 1). Dallas: SIL and University of Texas, Arlington.

Ross, Malcolm D. 1993. Tonogenesis in the North Huon Gulf Chain. In Jerold A. Edmondson & Kenneth J. Gregerson (eds), *Tonality in Austronesian languages* (Oceanic Linguistics Special Publications 24), 133–153. Honolulu: University of Hawai'i Press.

Rupp, James. 1996. *Diccionario chinanteco de San Juan Lealao, Oaxaca*. Tucson: SIL.

Smythe, W. E. 1948–49. Elementary grammar of the Gumbáiŋgar language. *Oceania* 19/2. 130–191.

Stump, Gregory T. 2001. *Inflectional morphology: a theory of paradigm structure*. Cambridge: Cambridge University Press.

Authors' addresses: (Matthew Baerman, Greville Corbett)
Surrey Morphology Group
Department of English
Faculty of Arts and Human Sciences
AC Building
University of Surrey
Guildford, Surrey GU2 7XH
UK
Email: m.baerman@surrey.ac.uk, g.corbett@surrey.ac.uk