CHAPTER 12

PARADIGM FUNCTION MORPHOLOGY AND THE MORPHOLOGY–SYNTAX INTERFACE

THOMAS STEWART AND
GREGORY STUMP

12.1 A Word-Based Interface Between Morphology and Syntax

Syntactic representations the phrases and sentences of a language accommodate the insertion of morpholexical (= morphological or lexical) expressions drawn or projected from its lexicon. Morphologists disagree about the types of morpholexical expression that is inserted and about the types of node into which insertion

takes place. Some argue that stems and affixes are inserted into separate nodes, so that the syntactic representation of a single word may involve several instances of morpholexical insertion; in such an approach, words have the status of syntactic complexes, so that a word's interaction with rules of syntax is mediated by the constellation of nodes ("morphemes") of which it is constituted.[1] Others argue that words are syntactic atoms instead—that they are inserted into syntactic structure as wholes and that their own internal morphological structure is unavailable to syntactic manipulation; in this approach, a word's interaction with rules of syntax is entirely determined by the unordered[2] set of morphosyntactic properties associated with the node that it occupies. Scrutiny of the empirical evidence reveals that the latter, word-based, conception of the morphology–syntax interface is more compatible with the range of behaviors exhibited by natural-language morphology than the former, morpheme-based, conception of this interface.

One pertinent sort of evidence is the fact that words which are completely alike in their external syntax may differ in their morphology. In the word-based approach, the syntactic behaviour of a word is associated with the unordered set of morphosyntactic properties situated at the node that it occupies; this approach therefore draws no connection between a word's syntactic behaviour and the exponence[3] of its properties. In the morpheme-based approach, by contrast, a word's syntactic behaviour is directly tied to the configuration of morphemes of which it is constituted; thus, the morpheme-based approach, unlike the word-based approach, predicts that words that are alike in their syntax should show similar exponence. This prediction is not borne out. English past-tense verb forms, for example, exhibit suffixal exponence (*tossed*), apophonic exponence (*threw*), extended exponence (*sold*), and null exponence (*hit*). In order to accommodate examples of the latter three types, proponents of the morpheme-based approach must assume that words that are alike in their syntax have similar morphology at some abstract level of representation but that this similarity is obscured by superficial operations; for instance, one might assume that *tossed*, *threw*, *sold*, and *hit* share an abstract structure of the form [V Tns] but that in the latter three cases, this structure is obscured by the transformational fusion of the V and Tns nodes or by the insertion of a zero suffix. There is no independent syntactic motivation for the postulation of such operations, whose sole rationale would be to get the morphology to fit the

[1] Indeed, the very notion "word" is epiphenomenal in a morpheme-based approach of this type.

[2] In this type of approach, a word's morphosyntactic properties are linearly unordered. Nevertheless, we assume that there may be at least two sorts of hierarchical relations among morphosyntactic properties. First, a morphosyntactic feature may be set-valued, so that its values include specifications for other features, e.g. AGR:{PER:1, NUM:sg}. Second, we leave open the possibility that grammatical principles might be sensitive to a ranking relation over morphosyntactic properties; see e.g. Stump (2001: 238 ff).

[3] In the morphology of a word w possessing a morphosyntactic property (or property set) p, the exponents of p are those morphological markings in w whose presence is associated with that of p. Exponence is the relation between a property (or property set) and its exponent(s).

syntax; nor is there a shred of independent evidence even given their common morphosyntactic content that *threw* has an underlyingly affixational structure like that of *tossed*. As this example suggests, the morpheme-based approach (unlike the word-based approach) is committed to the assumption that trees are as suitable for representing a word's morphology as for representing phrasal syntax; on this approach, all morphology is seen as fundamentally affixational, notwithstanding the prima facie counter-evidence of non-concatenative morphology of diverse sorts.

Another kind of evidence favouring the word-based conception of the morphology–syntax interface is the fact that words which differ in their external syntax do so in ways which correlate with their content, not with their form. This fact follows from the assumptions of the word-based approach, according to which syntax may be sensitive to a word's morphosyntactic properties but is in any event blind to its morphological form. In the morpheme-based approach, by contrast, the possibility is left open that words' morphological structure might correlate with differences in their external syntax that are not simply predictable from differences in their morphosyntactic content. The morpheme-based conception of the morphology–syntax interface is therefore permissive without motivation. For instance, the external syntax of the Fula verbs in Table 12.1 is fully determined by the unordered sets of morphosyntactic properties with which they are associated. The syntax is simply blind to whether these verbs inflect prefixally or suffixally; the third-person singular, first-person plural, and third-person plural forms do not function as a natural class with respect to any syntactic behaviour, nor do the complementary, suffixed forms. This is not an oddity of Fula; all human languages are like this.

A third type of evidence favouring the word-based conception of the morphology–syntax interface is the fact that languages differ morphologically in ways which cannot be attributed to independently motivated differences in their syntax. Thus, as Stump (2001: 25 ff) points out, the morphological expressions of tense and voice are in opposite orders in Albanian *lahesha* and Latin *lavābar* (both 'I was

Table 12.1 Some relative past active forms of the Fula verb LOOT 'wash'

	Singular	Plural
1	lootu-mi'	min-looti'
2	lootu-daa'	lootu-den' (inclusive) lootu-don' (exclusive)
3	'o-looti'	e-looti'

Source: Arnott 1970: 191

washed'), yet there is no independent justification for claiming that the operators of tense and voice participate in contrasting relations of c-command in these two languages; similarly, the morphological expressions of tense and subject agreement are in opposite orders in Latin *amābam* and Welsh Romany *kamávas* (both 'I loved'), yet there is no syntactic evidence for any difference in the nesting of TP and AgrP in these languages; and so on. The fact that the ordering the affixes of word often corresponds to an assumed nesting of functional categories has sometimes been used to argue for the morpheme-based approach to the morphology–syntax interface (cf. Embick and Noyer, in this volume). But this tendency can be seen simply as the effect of relevance on diachronic processes of morphologization (Bybee 1985: 38 ff); indeed, the latter explanation is easier to reconcile with the frequent incidence of "exceptions" such as *lavābar* or *kamávas*.

We conclude from this type of evidence that the interface between morphology and syntax is, in the terminology of Zwicky (1992: 356), a featural rather a formative interface—that the morphology and syntax of a language have only a limited shared vocabulary, which includes lexical categories and morphosyntactic properties but excludes such notions as affix or inflectional morpheme (Stump 2001: 18 ff). Accordingly, we believe that the adequacy of a morphological theory is, in part, a function of the extent to which it accommodates this conception of the morphology–syntax interface.

Here, we present an overview of Paradigm Function Morphology, a formally explicit morphological theory which presupposes a word-based interface between morphology and syntax. We begin by situating Paradigm Function Morphology within the general landscape of current morphological theories (sections 12.2 and 12.3), then proceed to a discussion of its central premises: the need to distinguish between content-paradigms and form-paradigms (sections 12.4 and 12.5), the need for both paradigm functions and realization rules in the definition of a language's morphology (sections 12.6 and 12.7), and the centrality of Pāṇini's principle (section 12.8). In section 12.9, we return to and elaborate on the word-based conception of the morphology–syntax interface afforded by PFM; we contrast this conception with the morpheme-based conception postulated by theories such as Distributed Morphology in order to highlight the significant empirical and descriptive advantages of the PFM approach (section 12.10). We summarize our conclusions in section 12.11.

12.2 What is PFM?

Paradigm Function Morphology (PFM) is an inferential–realizational theory of inflectional morphology which takes as its central premise the assumption that paradigms are essential to the very definition the inflectional system of a language.

It is REALIZATIONAL because it presumes that a word's inflectional markings are determined by the morphosyntactic properties which it carries; that is, it rejects the assumption, characteristic of INCREMENTAL theories, that words acquire their morphosyntactic properties only as an effect of acquiring the exponents of those properties. In addition, PFM is INFERENTIAL because it presumes that word forms are deduced from more basic forms (roots and stems) by means of rules associating particular morphological operations with particular morphosyntactic properties; that is, it rejects the assumption, characteristic of LEXICAL theories, that morphosyntactic properties are associated with inflectional markings just as lexico-semantic properties are associated with lexemes—in lexical entries or as "vocabulary items".[4]

The incremental–realizational distinction and the cross-cutting lexical–inferential distinction define theories of inflectional morphology of four logically possible types. As Stump (2001: 2f) shows, all four types are instantiated among current approaches to inflectional morphology: the lexical–incremental type is embodied by the theory advocated by Lieber (1992); the inferential–incremental type, by the theory of Articulated Morphology (Steele 1995); the lexical–realizational type, by Distributed Morphology (hereafter DM; Noyer 1992; Halle and Marantz 1993); and the inferential–realizational type, by the general approach of Word-and-Paradigm morphology (Matthews 1972; Zwicky 1985), A-morphous Morphology (Anderson 1992), Network Morphology (Corbett and Fraser 1993; Brown, Corbett, Fraser, Hippisley, and Timberlake 1996), as well as PFM.

12.3 Why an Inferential–Realizational Theory?

Semiotically oriented theories of morphology such as Natural Morphology (Dressler, Mayerthaler, Panagl, and Wurzel 1987) emphasize the cognitive value of isomorphism between units of content and units of form in morphological structure. From this perspective, an ideal system would have one and only one distinct, phonologically invariant, morpheme paired with each possible distinct morphosyntactic property. The ideal is not achieved in natural human language, however, not even in highly agglutinative language types. Languages commonly and successfully exploit all kinds of deviations from the canonical one-to-one pairing of form with content. This by no means undercuts the Natural Morphology

[4] Lexical theories of morphology may, however, maintain a distinction between stems and affixes, e.g. by assuming that affixes are inserted into syntactic structure later than stems (see Embick and Noyer, in this volume).

position; rather, it serves to point out that whatever the cognitive ideal, the morphological descriptive framework must be ready and able to allow for a range of morphological exponence beyond unifunctional, phonologically invariant, and consistently placed segmentable affixes. PFM, by insisting on the principled separation of content and exponence, allows for the range of observed morphological behaviours without necessitating structural zeroes, ad hoc hierarchical configurations, or treating some types of exponence as more or less "normal" in absolute terms. Incremental and lexical theories are less well suited to structures that are not built up in a monotonic increasing fashion out of discrete, intrinsically meaningful pieces.

Therefore, theoretical approaches that take affixation as basic and all other exponence as somehow deficient (e.g. segmentally underspecified reduplicants, floating mutation features), are hard-pressed to accommodate such non-canonical exponence in the morphological description. Canonical inflection is compatible with a variety of theoretical approaches; it is the noncanonical phenomena that provide the basis for choosing among them.

Incremental theories are based on the sometimes tacit assumption that inflectional markings are added to words in order to allow them to acquire their full set of morphosyntactic properties; accordingly, this type of theory implies that extended exponence (the appearance of more than one marking for the same property or property set) should never arise, for the simple reason that it is never motivated by the need to augment a word's morphosyntactic property set. Yet, extended exponence is widespread in inflectional morphology (Stump 2001: 3 ff). For instance, the default plural suffix *-où* appears twice in Breton *bagoùigoù* 'little boats': contrary to the basic premise of incremental theories, the addition of the second *-où* is not motivated by the need to supplement the word's morphosyntactic property set, nor is the stem *bagoùig-* (whose morphosyntactic property set is, if anything, already fully specified) acceptable, in itself, as a word for 'little boats' in Breton (**bagigoù* is likewise ungrammatical.) Under the assumptions of an inferential–realizational theory, there is no expectation that extended exponence should not arise, since there is no reason, a priori, why the morphology of a language should not contain two or more rules realizing the same property.

Because they portray inflectional markings as the source of a word's morphosyntactic properties, incremental theories imply that every one of a word's morphosyntactic properties should be interpretable as the contribution of a particular marking. But this, too, is an unsatisfied expectation in morphology: a word's morphological form may underdetermine morphosyntactic content. Consider an example from Sora (Austroasiatic; India). In Sora, the second-person plural affirmative non-past form of the verb DE 'get up' is *ə́deten* 'you (pl.) get up'; see Table 12.2. This form has an overt marking for tense (the nonpast suffix *-te*), a conjugation-class marker *-n*, and a default plural prefix *ə́-*; nowhere does it exhibit an overt exponent of second person. Yet, it is unmistakably the second-person

plural form: its first-person plural inclusive counterpart is *detenbe* (in which the appearance of the first-person plural inclusive suffix *-be* overrides that of *ə-*); its first-person plural exclusive counterpart is *ədetenay* (which contains the first-person exclusive suffix *-ay* also appearing in *detenay* 'I get up'); and its third-person plural counterpart is *detenji* (in which the appearance of the third-person plural suffix *-ji* overrides that of *ə-*).[5] Thus, not all morphosyntactic content of *ədeten* can be seen as the contribution of an inflectional marking; nor could one say that Sora verb forms receive a second-person interpretation by default, since both the second- and third-person singular counterparts of *ədeten* (*deten* 'you (sg.)/s/he gets up') lack any overt expression of person. At this juncture, proponents of incremental theories might propose that Sora possesses one or more phonologically empty person markers.

While the use of phonetically null affixes is not new, it is nevertheless a questionable formal device, if only because the putative distribution of such affixes is hard to demonstrate empirically. Often, zero affixes arise in a Structuralist implication on analogy with the distribution of one or more overt affixes with comparable but contrastive meaning. The incremental position in general implies that any content found in a word beyond the lexical meaning of the root is added either through a discrete operation with no phonological effect (Steele 1995) or through the concatenation of a phonetically null but contentful affix at some morpheme boundary. Taken to its logical conclusion, this move engenders either a large population of homophonous null affixes or a potentially long derivation of string-vacuous rule applications. In either case, the argument is developed theory-internally, and it is therefore unfalsifiable. In inferential–realizational theories, however, nothing so exotic as zero affixes is needed to account for the Sora facts; instead, one need only assume that in the inflection of Sora verbs of the DE type, there happens not to be any rule explicitly realizing the property "second person".[6]

Lexical theories also carry unwarranted implications about morphological form; in particular, they imply that inflectional markings are like lexically listed words in at least two ways. First, they imply that inflectional markings are inserted from the lexicon into phrase-structural nodes, and are therefore always linearly ordered with respect to the expressions with which they combine. Second, they imply that two types of relation may hold between an inflectional marking and a morphosyntactic property (or property set): an inflectional marking may express a particular morphosyntactic property (set), or it may be restricted to the context of a particular property (set). Neither of these implications is well motivated. First, inflectional

[5] The fact that the suffixation of *-be* or *-ji* overrides the prefixation of *ə-* shows that these affixes are members of an ambifixal position class (one whose members include both prefixes and suffixes); for discussion, see Stump (1993*b*, 2001: 133, 284f).

[6] For a detailed analysis of Sora verb morphology in an inferential–realization framework, see Stump (2005).

Table 12.2 Affirmative paradigms of the Sora verb DE 'get up'

		Nonpast	Past
Singular	1	de-te-n-ay	de-le-n-ay
	2	de-te-n	de-le-n
	3	de-te-n	de-le-n
Plural	1 incl	de-te-n-be	de-le-n-be
	1 excl	ə-de-te-n-ay	ə-de-le-n-ay
	2	ə-de-te-n	ə-de-le-n
	3	de-te-n-ji	de-le-n-ji

Source: Biligiri 1965: 232 ff

markings do not necessarily combine with other expressions in the same ways that words do. The pluralization of Somali *díbi* 'bull', for example, is effected by a prosodic inflectional marking (*dibí* 'bulls'); representing the morphology of *dibí* as an affixational structure is at fundamental odds with any observable evidence. Second, there is no empirical motivation for assuming that an inflectional marking must be seen as expressing one set of morphosyntactic properties but as selecting for some other set of such properties; instead, one may always simply assume that the only relation between an inflectional marking and a set of morphosyntactic properties is the relation of exponence. In the inflection of Swahili verbs, for instance, the default mark of negative polarity is a prefix *ha-*, as in *hatutataka* 'we will not want'. In the inflection of negative past-tense verb forms, the default past-tense prefix *li-* is overridden by a special suffix *ku-*: *tulitaka* 'we wanted', but *hatukutaka* 'we did not want'. Although one could certainly treat *ku-* as expressing past tense but selecting for a negative context, there is no evidence to favour this approach over the simpler approach treating *ku-* as an exponent of both past tense and negation. An inferential–realizational theory of morphology such as PFM is fully compatible with this simpler approach.

We conclude on the basis of these considerations that the most adequate theory of morphology is both inferential and realizational. Logically, a realizational theory requires an explicit account of the association of morphosyntactic properties with their exponents, and an inferential theory requires an explicit account of the principles regulating the ways in which morphological rules compete or combine in the definition of inflected forms; PFM furnishes both of these, as we show in sections 12.6–8. First, however, we discuss a third distinctive aspect of PFM, namely its theory of paradigms (sections 12.4 and 12.5).

12.4 Content Paradigms and Form Paradigms

Paradigms participate in the definition of two different grammatical domains (Stump 2002, to appear a; Ackerman and Stump 2004). On the one hand, a lexeme's paradigm distinguishes the various ways in which it can enter into the definition of phrase structure. In the syntax of Latin, noun phrases and their heads are specified for three morphosyntactic properties: a property of GENDER (masculine, feminine or neuter), which is lexically stipulated for each noun lexeme, and thus invariant within a given noun paradigm; one of six properties of CASE, as listed in (1) below; and a property of NUMBER (singular or plural). The paradigm of a Latin noun therefore canonically contains twelve cells, one for each of the twelve sorts of N nodes into which it might be inserted. The masculine nominal lexeme AMĪCUS 'friend', for example, provides the paradigm schematized in (1); each cell in this paradigm is schematized as the pairing of AMĪCUS with a different gender–case–number specification. (The paradigm itself need not, of course, be listed lexically; it need only be accessible by projection from information specified in the lexeme's entry.) Seen as a response to the needs of syntax, (1) constitutes a content paradigm.

(1) Content paradigm of the lexeme AMĪCUS 'friend'

a. ⟨ AMĪCUS, {MASC NOM SG} ⟩	*g.* ⟨ AMĪCUS, {MASC NOM PL} ⟩
b. ⟨ AMĪCUS, {MASC VOC SG} ⟩	*h.* ⟨ AMĪCUS, {MASC VOC PL} ⟩
c. ⟨ AMĪCUS, {MASC GEN SG} ⟩	*i.* ⟨ AMĪCUS, {MASC GEN PL} ⟩
d. ⟨ AMĪCUS, {MASC DAT SG} ⟩	*j.* ⟨ AMĪCUS, {MASC DAT PL} ⟩
e. ⟨ AMĪCUS, {MASC ACC SG} ⟩	*k.* ⟨ AMĪCUS, {MASC ACC PL} ⟩
f. ⟨ AMĪCUS, {MASC ABL SG} ⟩	*l.* ⟨ AMĪCUS, {MASC ABL PL} ⟩

Besides entering into the definition of phrase structure, paradigms participate in the definition of a language's morphological forms. In a realizational theory of morphology, rules of inflection apply to the pairing of a root[7] with a morphosyntactic property set; the paradigm of a Latin noun therefore provides an inventory of twelve such pairings, as in (2). Realization rules such as those in (4) apply to the pairings in (2) to determine the realizations listed in (3). Seen as a response to the needs of morphology, (2) constitutes a form-paradigm.

[7] Here and below, we adhere to the following terminological usage: a word form is a synthetic realization of a cell in a paradigm; a stem is a morphological form which undergoes one or more morphological rules in the realization of a cell in a paradigm; and a lexeme's root is its default stem.

(2) Form paradigm of the root *amīc* 'friend':

a. ⟨ amīc, {masc nom sg} ⟩
b. ⟨ amīc, {masc voc sg} ⟩
c. ⟨ amīc, {masc gen sg} ⟩
d. ⟨ amīc, {masc dat sg} ⟩
e. ⟨ amīc, {masc acc sg} ⟩
f. ⟨ amīc, {masc abl sg} ⟩
g. ⟨ amīc, {masc nom pl} ⟩
h. ⟨ amīc, {masc voc pl} ⟩
i. ⟨ amīc, {masc gen pl} ⟩
j. ⟨ amīc, {masc dat pl} ⟩
k. ⟨ amīc, {masc acc pl} ⟩
l. ⟨ amīc, {masc abl pl} ⟩

(3) Realizations of the cells in (1) and (2)

a. amīcus
b. amīce
c. amīcī
d. amīcō
e. amīcum
f. amīcō
g. amīcī
h. amīcī
i. amīcōrum
j. amīcīs
k. amīcōs
l. amīcīs

(4) Some Latin morphological rules

a. **Stem-formation rule**
Where root R is a second-declension nominal, R's thematized stem is R*u*.

b. **Realization rules**
Where X is the thematized stem of a second-declension root R and R is an adjective or masculine noun,
i. cell ⟨ R,{masc nom sg}⟩ is realized as X*s*;
ii. cell ⟨ R,{masc voc sg}⟩ is realized as R*e*;
...

There is, of course, a close connection between the content paradigms and its form paradigms of a language. In particular, each cell in a content-paradigm (i.e. each content cell) normally corresponds to a particular cell in a particular form paradigm (i.e. to a particular form cell); this form-cell is its form-correspondent. In general, the realization of a content cell is that of its form correspondent. Thus, because the content cell in (1*a*) has the form cell in (2*a*) as its form correspondent, they share the realization in (3*a*).

In the canonical case, there is an isomorphic relation between a language's content and form paradigms: a lexeme L has a single root R, and for each morphosyntactic property set σ with which L is paired in some cell ⟨L, σ⟩ of its content paradigm, the form correspondent of ⟨L, σ⟩ is ⟨R, σ⟩ (so that the realization of ⟨L, σ⟩ is that of ⟨R, σ⟩). This isomorphic relation might be formulated as the rule of paradigm linkage in (5), in which ⇒ is the form-correspondence operator.

(5) **The universal default rule of paradigm linkage**
Where R is L's root, ⟨L, σ⟩ ⇒ ⟨R, σ⟩

One might question whether it is actually critical to make a distinction between content and form paradigms. In those instances in which the default rule in (5) has effect, the distinction seems genuinely redundant. But as we show in the following section, (5) is sometimes overridden, and it is in precisely such instances that make it necessary to distinguish the two types of paradigm.

12.5 Why Two Types of Paradigm?

Deviations from the default relation of paradigm linkage in (5) are of diverse kinds; some of the more common types of deviation are listed in Table 12.3. We consider each of these in turn.

12.5.1 Deponency

In instances of DEPONENCY[8] (Table 12.3, row 1), the realization of a content-cell $\langle L, \sigma \rangle$ is that of a form cell $\langle X, \sigma' \rangle$, where $\sigma \neq \sigma'$. Latin furnishes the standard example of this phenomenon. In Latin, certain verbs—the deponents—are special

Table 12.3 Common deviations from (5)

Deviation	Paradigm linkage
1. Deponency	$\langle L,\sigma \rangle \Rightarrow \langle X,\sigma' \rangle$, where $\sigma \neq \sigma'$ and normally, $\langle L_1,\sigma \rangle \Rightarrow \langle X_1,\sigma \rangle$ and $\langle L_1,\sigma' \rangle \Rightarrow \langle X_1,\sigma' \rangle$
2. Syncretism	
directional:	$\langle L,\sigma \rangle$, $\langle L,\sigma' \rangle \Rightarrow \langle X,\sigma \rangle$
nondirectional:	$\langle L,\{\tau \dots\} \rangle$, $\langle L,\{\tau' \dots\} \rangle \Rightarrow \langle X,\{\tau \vee \tau' \dots\} \rangle$
3. A single content-paradigm's realization conditioned by multiple inflection classes ((a) principal parts phenomenon; (b) systematically associated inflection classes; (c) heteroclisis)	$\langle L,\sigma \rangle \Rightarrow \langle X,\sigma \rangle$ $\langle L,\sigma' \rangle \Rightarrow \langle Y,\sigma' \rangle$

[8] Although the term "deponency" is frequently associated with exceptional verb morphology in the classical Indo-European languages, we use this term in a more general way to refer to any instance in which a word's morphology is at odds with its morphosyntactic content. Construed in this way, deponency takes in a range of phenomena in a wide range of languages.

Table 12.4 Paradigm linkage in the inflection of nondeponent, deponent, and semideponent verbs in Latin

Lexemes	Content-cells	Form-correspondents	Realizations
MONĒRE	⟨ MONĒRE, {1 sg pres *act* indic} ⟩	⟨ mon, {1 sg pres *act* indic} ⟩	moneō
'advise'	⟨ MONĒRE, {1 sg perf *act* indic} ⟩	⟨ mon, {1 sg perf *act* indic} ⟩	monuī
	⟨ MONĒRE, {1 sg pres *pass* indic} ⟩	⟨ mon, {1 sg pres *pass* indic} ⟩	moneor
	⟨ MONĒRE, {1 sg perf *pass* indic} ⟩	⟨ mon, {1 sg perf *pass* indic} ⟩	monitus sum
FATĒRĪ	⟨ FATĒRĪ, {1 sg pres *act* indic} ⟩	⟨ fat, {1 sg pres *pass* indic} ⟩	fateor
'confess'	⟨ FATĒRĪ, {1 sg perf *act* indic} ⟩	⟨ fat, {1 sg perf *pass* indic} ⟩	fassus sum
AUDĒRE	⟨ AUDĒRE, {1 sg pres *act* indic} ⟩	⟨ aud, {1 sg pres *act* indic} ⟩	audeō
'dare'	⟨ AUDĒRE, {1 sg perf *act* indic} ⟩	⟨ aud, {1 sg perf *pass* indic} ⟩	ausus sum

in that their active forms exhibit the morphology typical of passive forms; certain other verbs—the semideponents—are special in that their perfect active forms exhibit the morphology typical of perfect passive forms. Verbs of both types involve a deviation from (5) in which active content cells have passive form cells as their form correspondents; Table 12.4 illustrates this with selected cells from the paradigms of the non-deponent verb MONĒRE 'advise', the deponent verb FATĒRĪ 'confess', and the semi-deponent verb AUDĒRE 'dare'.

To account for the inflection of deponents and semideponents, we assume that in Latin, the default rule of paradigm linkage in (5) is overridden by the more specific rules of paradigm linkage in (6).

(6) *a.* Where L is a deponent verb having root R, ⟨L,{active . . . }⟩ ⇒⟨R,{passive . . . }⟩
b. Where L is a semi-deponent verb having root R,
⟨L,{perfect active . . . }⟩ ⇒ ⟨R,{perfect passive . . . }⟩

12.5.2 Syncretism

In instances of syncretism (Table 12.3, row 2), two or more content cells share their form correspondent, hence also their realization. Because the pattern of paradigm linkage in such instances involves a many-to-one mapping from content cells to form cells, these are instances in which a content paradigm has more cells than the form paradigm by which it is realized. The Sanskrit paradigms in Table 12.5 illustrate. In Sanskrit, a neuter noun has identical forms in the nominative and accusative. This syncretism is directional, since the nominative singular form patterns after the accusative singular form; thus, the suffix *-m* shared by the accusative singular forms *aśvam* and *dānam* in Table 12.5 also appears in the

Table 12.5 Inflectional paradigms of two Sanskrit nouns

		AŚVA 'horse' (masc.)	DĀNA 'gift' (neut.)
Singular	Nom	aśvaḥ	dānam
	Voc	aśva	dāna
	Acc	aśvam	dānam
	Instr	aśvena	dānena
	Dat	aśvāya	dānāya
	Abl	aśvāt	dānāt
	Gen	aśvasya	dānasya
	Loc	aśve	dāne
Dual	Nom, Voc, Acc	aśvau	dāne
	Instr, Dat, Abl	aśvābhyām	dānābhyām
	Gen, Loc	aśvayoḥ	dānayoḥ
Plural	Nom, Voc	aśvāḥ	dānāni
	Acc	aśvān	dānāni
	Instr	aśvaiḥ	dānaiḥ
	Dat, Abl	aśvebhyaḥ	dānebhyaḥ
	Gen	aśvānām	dānānām
	Loc	aśveṣu	dāneṣu

nominative singular form *dānam*. A second instance of syncretism appearing in Table 12.5 (and indeed, in all nominal paradigms in Sanskrit) is that of the genitive and locative cases in the dual. Unlike the syncretism of the nominative and accusative cases of neuter nouns, the genitive–locative dual syncretism is non-directional: there is no good basis for saying that the realization of either case patterns after that of the other.

To account for the first of these syncretisms, we assume that in Sanskrit, a neuter noun does not have separate nominative and accusative cells in its form paradigm; instead, it has three form cells (one singular, one dual, and one plural) whose case property is represented as nom∨acc, and each of these cells is the form-correspondent of both a nominative and an accusative cell in the noun's content-paradigm (as in row (*c*) of Table 12.6). The operator ∨ is defined as combining with two properties τ, τ' to yield a third property $\tau \vee \tau'$ such that any rule applicable in the realization of τ or τ' is also applicable in the realization of $\tau \vee \tau'$; the existence of a rule realizing the accusative singular through the suffixation of *-m* therefore produces a directional effect in the inflection of neuter nouns. To account for the genitive–locative dual syncretism in the paradigms of Sanskrit nominals, we

Table 12.6 Paradigm linkage in the inflection of two Sanskrit nouns

Lexeme	Content-cell	Form-correspondent	Realization
a. AŚVA 'horse'	⟨ AŚVA, {masc nom sg} ⟩	⟨ aśva, {masc nom sg} ⟩	aśvaḥ
	⟨ AŚVA, {masc acc sg} ⟩	⟨ aśva, {masc acc sg} ⟩	aśvam
b.	⟨ AŚVA, {masc gen du} ⟩ ⟨ AŚVA, {masc loc du} ⟩	⟨ aśva, {masc gen∨loc du} ⟩	aśvayoḥ
c. DĀNA 'gift'	⟨ DĀNA, neut nom sg} ⟩ ⟨ DĀNA, neut acc sg} ⟩	⟨ dāna, neut nom∨acc sg} ⟩	dānam
d.	⟨ DĀNA, neut gen du ⟩ ⟨ DĀNA, {neut loc du} ⟩	⟨ dāna, {neut gen∨loc du} ⟩	dānayoh

assume that a given nominal lacks distinct genitive dual and locative dual cells in its form paradigm; instead, it has a single form cell whose case property is gen∨loc, and this cell is the form correspondent of both the genitive dual and the locative dual cells in its content paradigm (as in rows (*b*) and (*d*) of Table 12.6). Because all nominals participate in this pattern of paradigm linkage and case and number are always expressed cumulatively in Sanskrit, no directional effects can arise in instances of the genitive–locative dual syncretism. We assume that the patterns of paradigm linkage in rows (*b*)–(*d*) of Table 12.6 are produced by the rules of paradigm linkage in (7), both of which override the default rule of paradigm linkage in (5).[9]

(7) Where L has R as its root, α = sg, du, or pl, and β = masc, fem, or neut,
   a. $\langle$L,{neut nom α}$\rangle$, $\langle$L,{neut acc α}$\rangle$ $\Rightarrow$ $\langle$R,{neut nom$\vee$acc α}$\rangle$
   b. $\langle$L,{gen du β}$\rangle$, $\langle$L,{loc du β}$\rangle$ $\Rightarrow$ $\langle$R,{gen$\vee$loc du β}$\rangle$

12.5.3 A Single Content-Paradigm's Realization Conditioned by Multiple Inflection Classes

The formulation of the default rule of paradigm linkage in (5) entails that each lexeme L has a root R and that for any relevant property set σ, the realization of $\langle$L,$\sigma\rangle$ is that of $\langle$R,$\sigma\rangle$. But the inflectional realization of a lexeme L sometimes involves two or more distinct stems (Table 12.3, row 3): that is, the form correspondent of $\langle$L,$\sigma\rangle$ may be $\langle$X,$\sigma\rangle$ while that of $\langle$L,$\sigma'\rangle$ is instead $\langle$Y,$\sigma'\rangle$, where X $\neq$ Y. The clearest instances of this sort are those in which X and Y belong to distinct inflection classes. At least three classes of such instances can be distinguished; we discuss these in the next three sections.

[9] The account of syncretism presented here is based on that of Baerman (2004).

12.5.3.1 *Principal Parts*

In many languages, a lexeme L is realized through the inflection of distinct stems (belonging to distinct inflection classes) in different parts of its paradigm. Very often in such cases, the lexeme's stems are independent, in the sense that one stem's form and inflection-class membership may neither determine nor be determined by those of another stem. Traditionally, a lexeme of this type is said to have several principal parts, which simply have to be memorized as idiosyncratic lexical properties. In Sanskrit, for example, a verb's present-system inflection comprises its present indicative, optative, and imperative paradigms as well as its imperfect paradigm, and its aorist-system inflection comprises its aorist and injunctive paradigms; there are ten present-system conjugations and seven aorist-system conjugations. A verb's present-system conjugation is, in general, neither predicted by nor predictive of its aorist-system conjugation. Thus, although the verb PRACH 'ask' follows the sixth present-system conjugation and the *s*-aorist conjugation, membership in the sixth present-system conjugation class neither entails nor is entailed by membership in the *s*-aorist conjugation class, as the examples in Table 12.7 show. Thus, the lexicon of Sanskrit must typically stipulate at least two principal parts for a given verb: a present-system and an aorist-system form.

Traditionally, a word's principal parts are assumed to be fully inflected words, but this is not necessary; one could just as well assume that a principal part is simply a stem belonging to a particular inflection class. The inflection of the Sanskrit verb PRACH 'ask' might then be assumed to involve instances of paradigm linkage such as those in Table 12.8, in which the form-correspondent for a

Table 12.7 Present-system and aorist-system conjugations of ten Sanskrit verbs

Verbal lexeme		Conjugation class	
		Present-system	Aorist-system
RAKṢ	'protect'	first	*s*-aorist
DRĀ	'run'	second	*s*-aorist
HĀ	'go forth'	third	*s*-aorist
MṚC	'injure'	fourth	*s*-aorist
VLĪ	'crush'	ninth	*s*-aorist
GUR	'greet'	sixth	root-aorist
SIC	'pour out'	sixth	*a*-aorist
KṢIP	'throw'	sixth	reduplicating aorist
SPHṚ	'jerk'	sixth	*iṣ*-aorist
LIH	'lick'	sixth	*sa*-aorist

Table 12.8 Paradigm linkage in the inflection of Sanskrit PRACH 'ask'

Content cell	Form correspondent	Realization
⟨ PRACH, {3 sg pres indic act} ⟩	⟨ pṛccha$_{[Vl]}$, {3 sg pres indic act} ⟩	pṛcchati
⟨ PRACH, {3 sg aor indic act} ⟩	⟨ aprākṣ$_{[s\text{-}Aorist]}$, {3 sg aor indic act} ⟩	aprākṣīt

present-system content cell contains the principal part PṚCCHA (ϵ sixth conjugation), while the form-correspondent for an aorist-system content-cell contains the distinct principal part APRĀKṢ (ϵ s-aorist conjugation). On this analysis, the default rule of paradigm linkage in (5) is overridden by the more specific rule in (8).

(8) Given a verbal lexeme having X (a member of conjugation class [α]) as a principal part, ⟨L, σ⟩ ⇒ ⟨X, σ⟩ provided that
   a. present or imperfect ϵ σ and α = I, II, III, IV, V, VI, VII, VIII, IX, or X; or
   b. aorist ϵ σ and α = root-aorist, *a*-aorist, reduplicated aorist, *s*-aorist, *iṣ*-aorist, *sa*-aorist, or *siṣ*-aorist.

12.5.3.2 *Systematically Associated Inflection Classes*

In instances of the principal parts phenomenon, the realizations of a lexeme are built on two or more stems belonging to distinct and mutually unpredictable inflection classes. There are, however, instances in which a lexeme's realizations are based on multiple stems belonging to distinct but mutually predictable inflection classes. Consider, for example, the Sanskrit paradigms in Tables 12.5 and 12.9. Nouns belonging to the *a*-stem declension (exemplified in Table 12.5 by the nouns AŚVA 'horse' and DĀNA 'gift') are never feminine, while nouns belonging to the derivative *ā*-stem declension (exemplified by SENĀ 'army' in Table 12.9) are all feminine. Accordingly, the inflection of the adjective PĀPA 'evil' in Table 12.9 involves two stems, belonging to two distinct declension classes: the stem *pāpa-* 'evil' follows the *a*-stem declension and is used for masculine and neuter forms, while the stem *pāpā-* follows the derivative *ā*-stem declension and is used for feminine forms. Unlike the association of the sixth present-system conjugation and the *s*-aorist conjugation in the inflection of Sanskrit PRACH, the association of the *a*-stem and derivative *ā*-stem declensions in the inflection of Sanskrit adjectives is highly systematic (Whitney 1889: §332). We assume that this association is effected by the rule of paradigm linkage in (9), which—overriding the default rule (5)—entails that if an adjectival lexeme has *a*-stem form-correspondents, then

Table 12.9 Inflectional paradigms of two Sanskrit nominals

		SENĀ 'army'	PĀPA 'evil'		
		(fem.)	Masc	Neut	Fem
Singular	Nom	senā	pāpaḥ	pāpam	pāpā
	Voc	sene	pāpa		pāpe
	Acc	senām	pāpam		pāpām
	Instr	senayā	pāpena		pāpayā
	Dat	senāyai	pāpāya		pāpāyāi
	Abl	senāyāḥ	pāpāt		pāpāyāḥ
	Gen	senāyāḥ	pāpasya		pāpāyāḥ
	Loc	senāyām	pāpe		pāpāyām
Dual	Nom, Voc, Acc.	sene	pāpāu	pāpe	pāpe
	Instr, Dat, Abl.	senābhyām	pāpābhyām		pāpābhyām
	Gen, Loc.	senayoḥ	pāpayoḥ		pāpayoḥ
Plural	Nom, Voc.	senāḥ	pāpāḥ	pāpāni	pāpāḥ
	Acc	senāḥ	pāpān	pāpāni	pāpāḥ
	Instr	senābhiḥ	pāpāiḥ		pāpābhiḥ
	Dat, Abl.	senābhyaḥ	pāpebhyaḥ		pāpābhyaḥ
	Gen	senānām	pāpānām		pāpānām
	Loc	senāsu	pāpeu		pāpāsu

Table 12.10 Paradigm linkage in the inflection of Sanskrit PĀPA 'evil'

Content cell	Form correspondent	Realization
⟨ PĀPA, {masc nom sg} ⟩	⟨ pāpa, {masc nom sg} ⟩	pāpaḥ
⟨ PĀPA, {neut nom sg} ⟩	⟨ pāpa, {neut nom sg} ⟩	pāpam
⟨ PĀPA, {fem nom sg} ⟩	⟨ pāpā, {fem nom sg} ⟩	pāpā

the feminine cells in its content paradigm have *ā*-stem form-correspondents; accordingly, (9) licenses the instances of paradigm linkage in Table 10.

(9) Given an adjectival lexeme having X*a* (ϵ *a*-stem declension) as its root, then there is a stem X*ā* (ϵ *ā*-stem declension) such that $\langle L, \sigma \rangle \Rightarrow \langle X\bar{a}, \sigma \rangle$ if feminine ϵ σ.

12.5.3.3 *Heteroclisis*

In instances of both the principal parts phenomenon (e.g. the present- and aorist-system inflection of Sanskrit verbs) and that of systematically associated inflection classes (e.g. the gender inflection of Sanskrit adjectives), lexemes belonging to a particular category characteristically depend on two or more distinct stems for their inflection. In instances of heteroclisis, by contrast, the cells of a lexeme's content paradigm are exceptional precisely because they have form correspondents whose stems are distinct and indeed belong to distinct inflection classes. In Sanskrit, for example, the lexeme HṚD(AYA) 'heart' inflects as an *a*-stem nominal *hṛdaya-* in the direct (i.e. nominative, vocative, and accusative) cases, but as a consonant-stem nominal *hṛd-* in the remaining, oblique cases; see the paradigms in Table 12.11; that is, the cells in HṚD(AYA)'s content-paradigm have the divergent pattern of form-correspondence exemplified in Table 12. The heteroclisis of Sanskrit HṚD(AYA) follows from the assumption that HṚD(AYA) has *hṛd-* as its root

Table 12.11 The heteroclite inflection of Sanskrit HṚD(AYA) 'heart'

		ĀSYA 'mouth'	HṚD(AYA) 'heart'		TRIVṚT 'threefold' (neuter forms)
	Stem	āsya	hṛdaya	hṛd	trivṛt
	Declension	neuter *a*-stem		neuter C-stem	
Singular	Nom	āsyam	hṛdayam		trivṛt
	Voc	āsya	hṛdaya		trivṛt
	Acc	āsyam	hṛdayam		trivṛt
	Instr	āsyena		hṛdā	trivṛtā
	Dat	āsyāya		hṛde	trivṛte
	Abl	āsyāt		hṛdas	trivṛtas
	Gen	āsyasya		hṛdas	trivṛtas
	Loc	āsye		hṛdi	trivṛti
Dual	Nom, Voc, Acc	āsye	hṛdaye		trivṛtī
	Instr, Dat, Abl	āsyābhyām		hṛdbhyām	trivṛdbhyām
	Gen, Loc	āsyayos		hṛdos	trivṛtos
Plural	Nom, Voc, Acc	āsyāni	hṛdayāni		trivṛnti
	Instr	āsyāis		hṛdbhis	trivṛdbhis
	Dat, Abl	āsyebhyas		hṛdbhyas	trivṛdbhyas
	Gen	āsyānām		hṛdām	trivṛtām
	Loc	āsyeṣu		hṛtsu	trivṛtsu

Source: Whitney 1889: 149

Table 12.12 Paradigm linkage in the inflection of Sanskrit HṚD 'heart'

Content cell	Form correspondent	Realization
⟨ HṚD, {neut nom sg} ⟩	⟨ hṛdaya, {neut nom sg} ⟩	hṛdaya-m
⟨ HṚD, {neut loc sg} ⟩	⟨ hṛd, {neut loc sg} ⟩	hṛd-i

and that the rule of paradigm linkage in (10) overrides the default rule (5) in the definition of HṚD(AYA)'s direct-case forms; in the definition of its oblique-case forms, (5) remains unoverridden.

(10) For any direct case α, if α ϵ σ, then ⟨HṚD, σ⟩ ⇒ ⟨*hṛdaya*, σ⟩.

The notion of paradigm linkage is quite powerful, but there are numerous imaginable ways in which it might be systematically restricted. Pending the completion of thorough typological investigations of the phenomena of deponency, syncretism, and multiple inflection-class conditioning, we have no specific restrictions to propose at present, but hope eventually to do so in light of ongoing research; see for example Stump (to appear *a*) and Baerman, Brown, and Corbett (to appear).

12.6 Paradigm Functions and Realization Rules

Paradigm Function Morphology gets its name from a theoretical construct central to the realizational definition of a language's inflectional morphology. Intuitively, a paradigm function is a function from cells to realizations. More precisely, for any form cell ⟨X, σ⟩ having Y as its realization in some language ℓ, the paradigm function PF_ℓ of language ℓ is a function from ⟨X, σ⟩ to ⟨Y, σ⟩; and for any content cell ⟨L, σ⟩ such that ⟨L, σ⟩ ⇒ ⟨X, σ′⟩ in some language ℓ (where σ may or may not equal σ′), the value of PF_ℓ(⟨L, σ⟩) is that of PF_ℓ(⟨X, σ′⟩). Thus,

$PF_{English}$(⟨LIKE, {third-person singular present indicative}⟩)
= $PF_{English}$(⟨*like*, {third-person singular present indicative}⟩)
= ⟨*likes*, {third-person singular present indicative}⟩,

and so on. (We will see momentarily why a paradigm function's value is the pairing of a realization with a property set rather than simply the realization alone.)

Table 12.13 Future- and past-tense forms of Swahili TAKA 'want'

		Positive	Negative
Future tense	1sg	ni-ta-taka	si-ta-taka
	2sg	u-ta-taka	ha-u-ta-taka (→ hutataka)
	3sg (class 1)	a-ta-taka	ha-a-ta-taka (→ hatataka)
	1pl	tu-ta-taka	ha-tu-ta-taka
	2pl	m-ta-taka	ha-m-ta-taka
	3pl (class 2)	wa-ta-taka	ha-wa-ta-taka
Past tense	1sg	ni-li-taka	si-ku-taka
	2sg	u-li-taka	ha-u-ku-taka (→ hukutaka)
	3sg (class 1)	a-li-taka	ha-a-ku-taka (→ hakutaka)
	1pl	tu-li-taka	ha-tu-ku-taka
	2pl	m-li-taka	ha-m-ku-taka
	3pl (class 2)	wa-li-taka	ha-wa-ku-taka

The paradigm function's of a language is defined in terms of its more specific realization rules. Realization rules are of two types: rules of exponence associate specific morphological operations with specific morphosyntactic property sets; rules of referral specify instances in which the realization of some property set by one rule is systematically identical to that of some (possibly distinct) property set by some other rule or rules. In the notational system of Ackerman and Stump (2004), realization rules take the form in (11):

(11) $X_C, \sigma{:}\tau \rightarrow Y$

This should be read as follows. Where $\langle X, \sigma \rangle$ is a pairing such that $\tau \subseteq \sigma$ and X belongs to class C, $\langle X, \sigma \rangle$ is realized as $\langle Y, \sigma \rangle$. Thus, consider the fragment of Swahili inflectional morphology in (11).

In order to account for this set of forms in a realizational analysis of Swahili inflection, the twelve realization rules in (12) must be postulated. These rules are organized into four rule blocks; since the morphology of the forms in Table 12.13 is purely affixal, each of the four blocks in (12) can be seen as housing an affix-position class. The rules within a given block are disjunctive in their application: if one applies, the others do not.

(12) Some Swahili realization rules

Block A:	*a.* $X_V, \sigma{:}\{\text{TNS:fut}\}$	→ *ta*X
	b. $X_V, \sigma{:}\{\text{TNS:past}\}$	→ *li*X
	c. $X_V, \sigma{:}\{\text{POL:neg, TNS:past}\}$	→ *ku*X

Block B:	*d.* X_V,σ:{AGR(su):{PER:1, NUM:sg} }	→ *ni*X
	e. X_V,σ:{AGR(su):{PER:2, NUM:sg} }	→ *u*X
	f. X_V,σ:{AGR(su):{PER:3, NUM:sg, GEN:{1,2} } }	→ *a*X
	g. X_V,σ:{AGR(su):{PER:1, NUM:pl} }	→ *tu*X
	h. X_V,σ:{AGR(su):{PER:2, NUM:pl} }	→ *m*X
	i. X_V,σ:{AGR(su):{PER:3, NUM:pl, GEN:{1,2} } }	→ *wa*X
Block C:	*j.* X_V,σ:{POL:neg}	→ *ha*X
Block D:	*k.* X_V,σ:{POL:neg, AGR(su):{PER:1, NUM:sg} }	→ *si*X
	l. X_V,σ:{}	→ (⟨X, σ⟩ : B) : C

Pāṇini's principle determines which of a block's rules applies in the realization of a pairing ⟨X, σ⟩: without exception, it is the narrowest of the applicable rules in that block. Rule (13*a*) is narrower than rule (13*b*): (i) if class C is a proper subset of class C′; or (ii) if C = C′ and the morphosyntactic property set τ2 is a proper subset of τ1. (We discuss the centrality of Pāṇini's principle in PFM in Section 12.8.)

(13) *a.* X_C, σ:τ_1 → Y
b. $X_{C'}$, σ:τ_2 → Z

There is no intrinsic ordering among a language's blocks of realization rules; rather, the interaction among rule blocks is determined by a language's paradigm function. Thus, in order to account for the rule-block interactions embodied by the forms in Table 12.13, one might propose the following provisional definition of the Swahili paradigm function:

(14) Provisional definition of the Swahili paradigm function

$PF_{Swahili}(\langle r, \sigma\rangle) = (\ (\langle r, \sigma\rangle : A) : B) : C$

N.B.: The notation "⟨X, σ⟩ : Block *n*" means "the result of applying the narrowest applicable rule in Block *n* to the pairing ⟨X, σ⟩".

This definition accounts for most of the negative forms in Table 12.13, whose definitions each involve the application of a rule from Block A, a rule from Block B, and a rule from Block C; the definition of *hatutataka* 'we will not want', for example, involves the application of *ta*-prefixation (12*a*), *tu*-prefixation (12*g*), and *ha*-prefixation (12*j*). This definition shows why a paradigm function's value is the pairing of a realization with its property set rather than simply the realization alone: because a realization rule is defined as applying to a pairing of the type ⟨morphological expression, morphosyntactic property set⟩, it must yield a value of this same type if a subsequent realization rule is to apply directly to its output; and because a paradigm function's value is defined as the result of applying a particular succession of realization rules, this value must likewise be of the same type.

Although the definition in (14) suffices to account for most of the negative forms in Table 12.13, the first-person singular negative forms present a problem: these

involve the portmanteau rule of *si*-prefixation in (12k). To account for the fact that the *si-* rule is a portmanteau, we have situated it in a special block D; by virtue of its default rule (12l), Block D is paradigmatically opposed to Blocks B and C together. By virtue of (12l), the Swahili paradigm function may be reformulated as in (15).

(15) Improved definition of the Swahili paradigm function
$PF_{Swahili}(\langle r,\sigma\rangle) = (\langle r, \sigma\rangle : A) : D$

This definition accounts for all of the negative forms in Table 12.13, whose definitions each involve the application of a rule from Block A and a rule from Block D: the definition of a first-person singular negative form involves the Block D rule of *si*-prefixation (12k); the definition of all other negative forms instead involves the default Block D rule in (12l), whose application entails the application of a Block B rule and a Block C rule.

Although definition (15) of the Swahili paradigm function accounts for the negative forms in Table 12.13, the positive forms seem to present another problem, since Block C in (12) provides no rule for the realization of positive forms. Such instances fall within the compass of the Identity Function Default (16), according to which every rule block in every language has an identity function as its least narrow rule. Once this is assumed, the definition in (15) accounts for the full range of forms in Table 13.

(16) Identity Function Default
Universally, the following rule acts as the least narrow member of any rule block:
$X_{any}, \sigma{:}\{\} \rightarrow X$

Thus, suppose that we wish to know the realization of the form-cell $\langle$*taka*, $\sigma\rangle$, where σ is the property set {1 pl future affirmative}. By definition (15), this is the form defined by the evaluation of ($\langle$*taka*, $\sigma\rangle$: A) : D. By Pāṇini's principle, this value is the result of applying rule (12l) to $\langle$*tataka*, $\sigma\rangle$ (itself the result of applying rule (12a) to $\langle$*taka*, $\sigma\rangle$). The result of applying (12l) to $\langle$*tataka*, $\sigma\rangle$ is, by Pāṇini's principle, the result of applying (16) to the result of applying rule (12g) to $\langle$*tataka*, $\sigma\rangle$. The final value $\langle$*tutataka*, $\sigma\rangle$ identifies *tutataka* as the realization of $\langle$*taka*, $\sigma\rangle$. Similar proofs are possible for all of the forms in Table 12.13.

12.7 Why Paradigm Functions?

One of the clearest points of contrast between the formalism of PFM and that of other morphological theories is the extensive reference to paradigm functions in PFM. Despite the fact that they are unique to PFM, paradigm functions are, we

claim, essential to the definition of the morphology of a language. Here we present some of the central reasons for their postulation.

12.7.1 *Rule-Block Interactions*

One might suppose, naively, that all interactions among a language's rule blocks in the definition of its word forms are the effect of a strict linear ordering of its rule blocks. In actuality, rule blocks interact in complex ways, and this is one of the reasons for postulating paradigm functions. For instance, one rule block may be paradigmatically opposed to a combination of two or more other rule blocks, as we have already seen: in Swahili, the exponent *ni-* of first-person singular subject agreement and the exponent *ha-* of negation are introduced by blocks B and C, respectively; but in the definition of a verb's first-person singular negative forms, the application of *ni*-prefixation and that of *ha*-prefixation are excluded by the application of the rule introducing the first-person singular negative prefix *si-*. In view of the fact that *si-* pre-empts both *ha-* and *ni-*, the rule of *si*-prefixation cannot be plausibly situated in either Block B or Block C; instead, it must be assumed to occupy a distinct block D, as in (12k).[10] The relation of paradigmatic opposition between Block D and the combination of blocks B and C can then be attributed to the paradigm function in (15) together with the rule of referral in (12*l*): according to (15), a verb draws its inflectional markings from Blocks A and D; by rule (12*l*), a verb draws inflectional markings from Blocks B and C only in the absence of any applicable rule in Block D.

Paradigm functions are also necessary for specifying other types of complex rule-block interaction (Stump 1993*c*, 2001: ch. 5). For instance, a single rule block may participate in the instantiation of more than one of a word's affix positions; that is, it may define parallel position classes. Thus, in Lingala, the rules expressing subject agreement in the affix position labelled I in Table 12.14 are, with only a few exceptions, identical to the rules expressing object agreement in the affix position labelled III; this identity can be accounted for by defining the Lingala paradigm function as in (17) (where each Roman numeral names the rule block responsible for the corresponding affix position in Table 12.14) and postulating a single, additional rule block to which Blocks I and III both default (Stump 2001: 144 ff).

(17) Definition of the Lingala paradigm function
$PF_{Lingala}(\langle r,\sigma\rangle) = (((((\langle r, \sigma\rangle : IV) : V) : III) : II) : I)$

Rule blocks may also be reversible, applying in one sequence in the realization of some morphosyntactic property sets but in the opposite sequence in the realization

[10] By virtue of its default rule (12*l*), Block D is an example of what Stump (2001: 141) terms a portmanteau rule block.

Table 12.14 Position-class analysis of some Lingala verb forms

Affix position						Gloss
I	II	III	(root)	IV	V	
na-	ko-		sál	-ak	-a	'I always work' (2nd habitual present)
ba-		m-	bet	-ak	-í	'they hit me' (historical past)
na-	ko-	mí-	sukol	-ak	-a	'I often wash myself' (2nd habitual present)
to-	ko-		kɛnd		-ɛ	'we are leaving' (present continuative)

Note: cf. Dzokanga 1979: 232

of other property sets. Thus, in the realization of Fula verb forms in the relative tenses, the application of the rule block realizing subject agreement ordinarily precedes that of the rule block realizing object agreement, as in the definition of (18*a*) *mball-u-mi-*ɓ*e*' 'I helped them'; but in instances in which a first-person singular subject coincides with a second- or third-person singular (class 1) object, the application of these two rule blocks is reversed, as in (18*b*, *c*). This relationship between the two rule blocks can be accounted for by postulating a paradigm function whose evaluation involves applying the subject-agreement block before the object-agreement block in the default case but involves the opposite order of application in the realization of certain morphosyntactic property sets (Stump 2001: 149 ff).

(18) *a.* *mball-u-mi-*ɓ*e*'
help-REL:PAST:ACT-I-them:CLASS.2
'I helped them.'

b. *mball-u-maa-mi*'
help-REL:PAST:ACT-you:SG-I
'I helped you (sg.).'

c. *mball-u-moo-mi*'
help-REL:PAST:ACT-him:CLASS.1-I
'I helped him.' (Arnott 1970: Appendix 15)

Examples of this type embody one kind of motivation for the postulation of paradigm functions—namely, the need to specify the different ways in which rule blocks may interact in the definition of a language's inflected word forms. There is, however, an additional type of motivation: a paradigm function makes it possible to refer to the realization of a paradigm's cells independently of the particular morphological operations by which their realization is defined; for this reason, any rule that must refer to a cell's realization without referring to the specific morphology of this realization necessitates the postulation of a paradigm function. There are several examples of this type of motivation; we discuss three of these in sections 12.7.2 to 12.7.4.

12.7.2 Head Marking

A matter of recurring interest in the morphological literature is the question of whether morphological expressions, like phrases, have heads. This question has been answered in different ways by different people. Our view is that of Stump (1993*a*, 1995, 2001): that a morphological expression is headed if and only if it arises through the application of a category-preserving rule of word formation. A category-preserving rule of derivation or compounding is one which allows one or more morphosyntactic properties of a base to persist as properties of its derivative. Thus, the Sanskrit rule producing preverb–verb compounds is category-preserving since it applies to verbs to yield verbs: *vi* 'away' + *gam* 'go' (v.) → *vi-gam* 'go away' (v.).

Morphological head-marking is the inflection of a headed morphological expression on its head; thus, the Sanskrit preverb–verb compound *vi-gam* exhibits head-marking because it inflects on its head *gam*: *vy-a-gacchat* 's/he goes away'. Interestingly, not all headed morphological expressions exhibit head-marking; in Breton, for example, headed derivatives in *-ad* '-ful' inflect at their periphery: *tiad* 'houseful' (← *ti* 'house' + *-ad*), plural *tiadoù*, **tiezad* (cf. *tiez* 'houses'). Whether or not a headed morphological expression exhibits head-marking is determined by the category-preserving rule by which it arises. Word-to-word rules are category-preserving rules that give rise to headed expressions that exhibit head-marking; root-to-root rules are category-preserving rules that give rise to headed expressions that do not exhibit head-marking. Thus, the Sanskrit rule of preverb–verb compounding is a word-to-word rule; the Breton rule of *-ad* suffixation is a root-to-root rule.

In PFM, the relation between word-to-word rules and head-marking is captured by means of the universal principle in (19). As Stump (1995, 2001: ch. 4) shows, this principle entails both empirical generalizations in (20); for this reason, it is preferable to approaches that attribute head marking to feature percolation or to head operations (inflectional operations which are stipulated as applying to a headed expression's head), which fail to account for these generalizations.

(19) **Head-Application Principle**
If a headed morphological expression x_2 arises from a morphological expression x_1 through the application of a word-to-word rule r such that $x_2 = r(x_1)$, then for each cell $\langle x_1, \sigma \rangle$ in x_1's form-paradigm and its counterpart $\langle x_2, \sigma \rangle$ in x_2's form-paradigm, $PF(\langle x_1, \sigma \rangle) = \langle y, \sigma \rangle$ if and only if $PF(\langle x_2, \sigma \rangle) = \langle r(y), \sigma \rangle$.

(20) *a.* **The Coderivative Uniformity Generalization**
Headed roots arising through the application of the same category-preserving rule are alike in exhibiting or in failing to exhibit head marking.

b. **The Paradigm Uniformity Generalization**
Roots that exhibit head marking do so categorically, throughout their paradigm of inflected forms.

The definition of the Head-Application Principle must refer to the realization of a paradigm's cells independently of the particular morphological operations by which their realization is defined; it therefore provides additional motivation for the postulation of paradigm functions in morphological theory.

12.7.3 Periphrasis

Much recent work suggests that periphrasis is a kind of morphological exponence; see Börjars, Vincent, and Chapman (1997), Sadler and Spencer (2001), Ackerman and Stump (2004), and Stump (to appear *b*). Accordingly, we assume that in instances of periphrasis, the realization of a cell in some paradigm consists of two or more words. Consider, for example, the second-person past forms of the Eastern Mari verb KOL 'die' in Table 12.15: the affirmative forms are synthetic, but the negative forms are periphrastic, consisting of KOL's affirmative gerundial stem *kolen* together with a negative, present-tense form of the copula UL. Like many instances of periphrasis, those in Table 12.15 are non-compositional: for instance, the morphosyntactic properties associated with the periphrase *kolen omə̑l* 'I did not die' are not the sum of those associated with its individual parts; in particular, neither *kolen* nor *omə̑l* has a past-tense property. Ackerman and Stump (2004) argue that the realization of negative second-person past forms in Eastern Mari involves a rule of exponence that applies to a form-cell ⟨X,{TNS:2past, POL:negative, . . . }⟩ to produce a periphrase consisting of X's affirmative gerundial stem combined with PF_{Mari}(⟨UL, {TNS:present, POL:negative, . . . }⟩); because this rule must refer to the realization the auxiliary UL independently of the particular morphological operations by means of which its realization is defined, it provides additional motivation for the postulation of paradigm functions. The same is true of many rules of periphrastic realization.[11]

12.7.4 Paradigm Linkage

By assumption, a paradigm function yields the same value when applied to a content cell as when it applies to that cell's form correspondent (section 12.6). For this reason, the rules of paradigm linkage in a language can be formulated as

[11] The recognition that periphrasis is a form of morphological exponence gives rise to a new conception of morphology: a language's morphology can no longer be seen merely as a system defining that language's individual word forms; instead, it must be seen as defining that language's paradigms, some of whose cells are realized as synthetic expressions (as in the case of Latin *laudor* 'I am praised') but others of which might be realized as periphrases (as in the case of *laudātus sum* 'I have been praised'). An important issue for this conception of morphology is that of distinguishing true periphrases (those defined by a language's morphology) from ordinary, syntactically defined word combinations; see Ackerman and Stump (2004) for a discussion of some relevant criteria.

Table 12.15. Second-past realizations of the Mari *em*-conjugation verb KOL 'die' (Eastern dialects)

		Affirmative	Negative
Singular	1	kol-en-am 'I died'	kolen oml 'I didn't die'
	2	kol-en-at	kolen otl
	3	kol-en	kolen oɣl
Plural	1	kol-en-na	kolen onal
	2	kol-en-da	kolen oðal
	3	kol-en-t	kolen oɣtl

Source: Alhoniemi 1985: 110,116

clauses in the definition of that language's paradigm function—specifically, as those clauses defining its evaluation when it applies to content cells. For instance, the default rule of paradigm linkage in (5) might be reformulated as in (21); the Latin rule of paradigm linkage in (6*a*) might be reformulated as in (22); and so on. This way of formulating rules of paradigm linkage eliminates the need to define relations of form correpondence independently of the definition of a language's paradigm function. Because these formulations must make reference to the shared realization of form cells and content cells without referring to the specific morphology of this realization, they provide additional motivation for the postulation of paradigm functions in morphological theory.

(21) **The universal default rule of paradigm linkage**
Where R is L's root, $PF(\langle L,\sigma\rangle) = PF(\langle R,\sigma\rangle)$

(22) Where L is a deponent verb having R as its root,
$PF_{Latin}(\langle L,\{active\ldots\}\rangle) = PF_{Latin}(\langle R,\{passive\ldots\}\rangle)$

12.8 The Centrality of Pāṇini's Principle

A theory of rule interaction is not well served if the outcome of any given interaction is not predictable on principled grounds. That is, the more a theory relies on outcomes imposed by stipulation, rather than having outcomes follow from more general, predictive principles, the less highly valued that theory of rule interaction should be. Pāṇini's principle has a long history in linguistic theorizing, having played a central role in the metatheory of the pre-eminent grammarian of

Classical Sanskrit whose name it bears. According to this principle, competition between two applicable rules is resolved in favour of the narrower rule (i.e. that rule whose domain is a proper subset of the other rule's domain). Later instantiations of this principle have been invoked by Anderson (1969, but cf. 1986) and Kiparsky (1973) and popularized in modern times as the so-called Elsewhere Condition.

Paradigm Function Morphology relies on the original Pāṇinian insight, avoiding later redefinitions that introduce caveats and riders that weaken the overall predictiveness of the principle. In PFM, the claim is (as mentioned above) that if a rule block is invoked in the definition of a cell's realization, the narrowest applicable rule in that block will apply to the exclusion of all other rules in the same block. Just in case there is no explicit realization rule in the block that applies, the Identity Function Default (16) applies instead. PFM makes the strong claim that Pāṇini's principle is sufficient, in every case, to decide the evaluation of each block of realization rules; this claim is the Pāṇinian Determinism Hypothesis (Stump 2001: 62ff).

In order to identify the narrowest applicable rule in a block, it is necessary first to select those rules in the block that are indeed applicable in the realization of the cell in question, and then to determine which of those rules is the narrowest. (Inapplicable rules are naturally irrelevant to the definition of a cell's realization.) Both the applicability and the relative narrowness of a given realization rule are determined with reference to two fundamental aspects of its formulation: the rule's property-set index (represented as τ in schema (23)), which identifies the set of morphosyntactic properties realized by the application of the rule; and its class index (represented as C in schema (11)), which identifies the inflection class of the forms to which the rule applies.

Given a realization rule r in the format 'X_C, σ: $\tau \rightarrow$ Y', the applicability of r to a form–property-set pairing $\langle X, \sigma \rangle$ is determined (*a*) by comparing the indexed property set τ to the fully specified property set σ appearing in $\langle X, \sigma \rangle$, and (*b*) by establishing the relationship between the indexed inflection class C and the expression X appearing in $\langle X, \sigma \rangle$. In order for r to be applicable to $\langle X, \sigma \rangle$, σ must be an extension[12] of τ and X must be a member of C. If τ includes any property not included in σ or if C excludes X, then r is inapplicable to $\langle X, \sigma \rangle$. If two rules r1, r2 belonging to the same rule block are both applicable to $\langle X, \sigma \rangle$, then the relative narrowness of r1 and r2 is determined by comparing their property-set indices and their class indices: if r1 and r2 have the same class index but the property-set index of r1 is an extension of that of r2, then r1 is narrower than r2; if the class designated by r1's class index designates a subset of that designated by r2's class index, then r1 is narrower than r2. In either instance, the narrower of the two applicable rules applies to the exclusion of its competitor, in accordance with Pāṇini's principle.

[12] For present purposes, an extension can be thought of as a superset; for a more precise definition, see Stump (2001: 41).

Note that the "winning" rule in such instances is not necessarily narrower than its competitors with respect to both its property-set index and its class index; indeed, our formulation of narrowness allows the property-set index of a losing rule r2 to be an extension of that of the winning rule r1 if the class designated by r1's class index is a subset of that designated by r2's class index.

But what of the logical possibility that two rules belonging to the same rule block might both be applicable to $\langle X, \sigma \rangle$ but might be such that neither is narrower than the other? In that case, Pāṇini's principle would logically fail to resolve the competition between the two rules. Our assumption is that instances of this sort are excluded by a universal well-formedness condition on the constitution of inflectional rule blocks; see Stump (2001: 23f, 73ff). This assumption makes it possible to maintain the Pāṇinian Determinism Hypothesis, according to which a rule's applicability and its relative narrowness always suffice to determine whether it participates in the realization of a given cell; under this hypothesis, there is never any appeal to stipulated, unprincipled, or extrinsic ordering of realization rules.

The Pāṇinian Determinism Hypothesis is a more parsimonious conception of the resolution of rule competition than is assumed in other current morphological theories. In DM, the closest thing to a rule block is a list of vocabulary items that are in competition insofar as they realize specifications the same features; the vocabulary items in a "rule block" of this sort do not necessarily belong to the same position class, nor is their insertion necessarily disjunctive if they are semantically compatible. Thus, competition among vocabulary items in DM is regulated not only by the Subset Principle (essentially the equivalent of Pāṇini's principle) but also by extrinsic ordering relations among vocabulary items and a feature-discharge principle, according to which a feature which has been realized once cannot be realized again by a subsequent rule.[13]

12.9 The Morphology–Syntax Interface in PFM

We now return to the question of how morphology interfaces with syntax. PFM affords a word-based interface, allowing words to be inserted as units into terminal nodes. We assume that in the syntactic representations of a language, every node

[13] Embick and Noyer (this volume) make no mention of arguments against the feature-discharge hypothesis that have appeared in the literature (e.g. Stump 2001: 156–66), but tacitly acknowledge the force of such arguments by admitting that vocabulary insertion isn't always sensitive to whether a feature has been discharged; oddly, they don't seem to recognize that this admission reduces the feature-discharge hypothesis to vacuity.

belonging to a lexical category is fully specified for the morphosyntactic features accessible to it (in that language); we also assume that a terminal node X may be associated with a lexeme L provided that (*a*) L belongs to category X and (*b*) L's lexical properties (e.g. its argument structure) unify with those of node X. On that assumption, the simplest cases of lexical insertion satisfy the following condition:

(23) If a node of category X has the morphosyntactic property set σ and is associated with lexeme L, then the (synthetic) realization of the content-cell ⟨L, σ⟩ is inserted into X.

For instance, if node V has the morphosyntactic property set {1 pl future affirmative} and is associated with the Swahili verbal lexeme TAKA 'want', then (23) allows the realization of the content cell (namely the word form *tutataka*) to be inserted into V. In this conception of lexical insertion, the terminal node into which a word form is inserted determines the content cell which that word form realizes.[14] The condition in (23) suffices to account for canonical instances of lexical insertion in all languages. There are, however, certain phenomena that necessitate a somewhat more complicated conception of lexical insertion; nevertheless, even these phenomena are compatible with the assumption of a word-based interface between morphology and syntax.

One type of complication arises in instances in which the realization of a cell is a combination of two or more words. We assume that the words constituting a periphrase are inserted interdependently—that, for instance, the parts of the Latin periphrase *laudātus sum* 'I am praised' are inserted into two V nodes in such a way that *laudātus* heads *sum*'s complement. Numerous questions arise for the proper formulation of this sort of interdependent insertion. How do the two V nodes participate in determining the content-cell ⟨LAUDĀRE, {1 sg pres perf pass indic}⟩ which *laudātus sum* realizes?[15] Is the syntactic relation between the nodes into which the parts of a periphrase are inserted subject to universal constraints? How—if at all—does the syntactic relation between the parts of a periphrase enter into its semantic interpretation? Pending further research, we shall defer proposing specific answers to these questions; for the moment, we assume that the lexical insertion of periphrases is subject to language-specific restrictions.

A second complication for the canonical conception of lexical insertion in (23) is the incidence of shape alternations (Zwicky 1992). While there is little in the way of compelling justification for the claim that syntax is sensitive to phonological

[14] On the assumption (*a*) that the syntax of English situates the morphosyntactic property "genitive" on the final word (rather than on the head) of a genitive noun phrase, and (b) that the morphology of English defines a genitive realization for any word that may end a noun phrase, (23) suffices to account for the insertion of the word *else's* in a phrase such as *someone else's hat*. See Lapointe (1990), Miller (1991), Halpern (1992), and Stump (2001: 126–30) for relevant discussion.

[15] The difficulty of this question emerges particularly clearly in instances of non-compositional periphrasis.

constituents, there is abundant evidence that syntax sometimes determines a word's phonological shape; in particular, conditions on the phonological shape of word forms are sometimes imposed in the context of other specific words or in particular syntactic constructions. We assume that conditions of this kind are enforced through the distribution of SHAPE PROPERTIES, whose requirements must be satisfied, to the degree possible, by any word or constituent appearing in the relevant context. Word forms sometimes have an idiosyncratic shape set from which one or another shape is selected according to shape properties of the contexts in which they appear; the English indefinite article, for example, has the shape set {*a*, *an*}, whose members are used in preconsonantal and prevocalic contexts, respectively.[16] On the other hand, shape alternations sometimes operate more systematically, as a class behaviour. Certain (but not all) instances of Celtic initial consonant mutation may be analyzed as the spelling out of shape properties distributed by particular grammatical words or constructions (see Stewart 2004 for Scottish Gaelic in particular). In Scottish Gaelic, interrogative modality is marked by an interrogative particle at the left edge of the sentence, and the first word following the particle appears in its nasalized shape: *Am bris thu e*? 'Will you break it?' (Here the initial /b/ in *bris* is pronounced as a prenasalized [^{m}b] or as a plain nasal [m].) While the preceding nasal segment can be identified as the diachronic source of the nasalizing mutation, it is not the case that a preceding nasal is a sufficient condition for selecting a word form's nasalized shape; instead, interrogative modality must be seen as having two expressions in this example—the interrogative particle *am* and the shape property associated with the nasalizing mutation.

The word-based morphology–syntax interface afforded by PFM is fully consistent with the observed separation of morphological structure from syntactic structure (Section 12.1) and accommodates both canonical instances of lexical insertion subject to condition (23) and more complicated instances involving periphrasis and shape alternations. This is an important basis for preferring PFM to theories such as DM: because PFM accommodates a word-based interface, it is in that respect much more restrictive than a theory necessitating a morpheme-based interface, since a theory of this latter sort allows for much more extensive interaction between a word's morphological structure (as distinguished from its morphosyntactic content) and its syntactic context. In the section which follows, we show that this is very far from the only reason for preferring PFM as a theory of morphology.

[16] Although similar alternate shapes were formerly available for some possessive adjectives (*my* ~ *mine*, *thy* ~ *thine*), this alternation is restricted to the indefinite article in modern English; that is to say, a shape distinction has undergone a retreat in its generality over time, but its effects remain as a residue, limited to one word of the language (and are thus of questionable status as a synchronic rule, rather than as a stipulated fact about *a* ~ *an* alone).

12.10 The Nature of General and Specific Patterns in Morphology

Zwicky (1986) distinguishes two different ways in which linguists formulate the relation between a general pattern and a special deviation from that pattern. In what he calls the "General as Basic" (or BASI) approach, the deviation has the general pattern as its underlying form and derives from it through the application of one or more rules; in the "General as Default" (or DEFO) approach, the deviation simply excludes the general pattern, which therefore has no role at all in its formation.[17] Although the two approaches might appear to be notational variants, they are not:

> What is crucial is that a BASI analysis is *derivational*, while a DEFO analysis is *monostratal*: rules of the former type map representations into representations and so induce a series of representations, each of which is available as the locus for the statement of other generalizations (that is, as a stratum . . . at which conditions can be stated or to which rules can apply), while the rules of the latter specify a set of conditions, some of them overridden by others, but all holding for a single stratum of representations. (Zwicky 1986: 307)

PFM is resolutely a DEFO theory of morphological form—a declarative theory having neither "underlying structures" nor "derivations", whose rules are nothing more than definitions; as such, it affords a word-based interface between morphology and syntax. DM, by contrast, is just as resolutely BASI—a non-declarative theory in which words have "underlying structures" which undergo multi-stage syntactic derivations; as such, it necessitates a morpheme-based interface. Which approach is right for morphology? Zwicky observes that while

> BASI can be extended easily to give analyses for phenomena covered by DEFO, . . . no simple tinkering will extend DEFO to cover characteristically BASI phenomena, in particular feeding interactions. It follows that DEFO is the more constrained view of rule interaction within a component of grammar and so has a prior claim on our attention. That is, DEFO ought itself to be the default view of how rules work in a component of grammar, with the more powerful BASI view adopted only on the basis of evidence that DEFO is inadequate for that component. (p. 308)

No evidence has ever been presented which decisively excludes a DEFO theory of morphology such as PFM.[18] Nevertheless, proponents of DM argue that no

[17] See Pullum and Scholz (2001) for relevant discussion from a computational perspective.

[18] Without so much as a reference, Embick and Noyer (this volume) assert that "[a]rticulated Lexicalist approaches make a number of precise empirical predictions, predictions which we take to have been disconfirmed", then say nothing more on the subject. In place of this facile dismissal, we (and, we suspect, most readers) would have preferred to know which predictions are at issue and the evaluative criteria by which these predictions have been assessed; we have no doubt that the reports of lexicalism's demise are exaggerated.

justification is necessary for adopting a BASI approach to morphology; their rationale is that this move keeps the number of generative mechanisms in their theory to a desired minimum. This isn't a cogent justification, however. As a theory of inflectional morphology, there is no meaningful sense in which PFM is generative: it simply assigns morphological realizations to content cells, which we equate with terminal nodes in syntax; that is, PFM simply interprets terminal nodes as realizations. Moreover, this is true in one fashion or another of all realizational theories of inflection; only incremental theories of inflection (see section 12.2) can be seen as generative in the sense intended by Embick and Noyer.

The central argument advanced in favor of DM is the fact that in at least the simplest cases, the claim in (24) holds true.

(24) The morphological structure of a word is precisely what it is predicted to be by the operation of head-movement through a nested succession of functional categories.

But if we restrict our consideration to language's simplest cases, we miss much of the evidence that is most useful for distinguish adequate theories from inadequate ones. And the fact is that once we move beyond the simplest cases, the claim in (24) is wildly disconfirmed—by portmanteau affixes (sections 12.6 and 12.7.1), by reversible position classes (section 12.7.1), by head marking (section 12.7.2), by syncretism (section 12.5.2), and so on. Proponents of DM must therefore resort to the postulation of a range of PF operations (rules of fission, fusion, rebracketing, local dislocation, and impoverishment) whose sole motivation is the desire to equate affix positions in a word's morphology with functional heads in syntax. But every recourse to such rules casts doubt on the leading premise that morphology is just syntax. Indeed, many morphologists feel that rules of this sort effectively render the theory unfalsifiable: if morphological patterns consistent with (24) are taken as confirming the theory but patterns inconsistent with (24) are not seen as disconfirming it (being instead attributed to otherwise unmotivated PF movements), what could possibly disconfirm the theory? Proponents of DM counter this conclusion by insisting (in Embick and Noyer's words) "that the operations that apply at PF are minimal readjustments, motivated by language-particular requirements". But this is far from clear; through a creative combination of rebracketings and local dislocations, one can, for example, permute the members of a finite string at will, and it has never been shown how the principles of DM could manage to restrict this process. (See Williams, in this volume, for discussion.)

Moreover, not even minimal readjustments of the sort countenanced in DM are needed in a DEFO theory such as PFM. Thus, consider again the Swahili examples in Table 12.13. In DM, each of the negative verb forms in Table 12.13 would have the structure in (25); in the case of the first singular negative form *sitataka* 'I will not want', however, the POL and AGR nodes would have to undergo a fusion operation to accommodate the insertion of the portmanteau prefix *si-*, an exponent of both

first-person singular subject agreement and negative polarity. In PFM, by contrast, the definition of *sitataka* involves no fusion operation at all: *sitataka* doesn't arise from an abstract structure in which polarity and agreement appear as distinct "morphemes"; instead, it is directly defined as the realization of the form-cell ⟨*taka*, {1 sg neg fut}⟩. This is the better analysis, since there is no empirical evidence from either syntax or morpology that *si-ta-taka* arises from a three-prefix structure. Similar remarks hold true for a range of phenomena. In the analysis of (18*c*) (Fula *mball-u-moo-mi*' 'I helped him'), for example, DM would require a local dislocation of the object-agreement morpheme in an abstract structure such as (26); but as was seen above, the DEFO analysis of reversible position classes afforded by PFM depends neither upon any dislocation operation nor upon the empirically unmotivated assumption that *mball-u-moo-mi*' arises from an abstract representation in which object agreement follows subject agreement.[19]

(25) (POL * (AGR * (TNS * (V * $\sqrt{}$ROOT))))

(26) ((((V * $\sqrt{}$ROOT) * TNS) * SUBJ-AGR) * OBJ-AGR)

The assumption that all words have abstract morphological structures in which each of the relevant functional heads constitutes an affixal morpheme engenders a number of undesirable consequences in DM. Prominent among these is a heavy reliance on zero affixes—for plural number in *moose*-Ø, for past tense in *cut*-Ø, for past participial status in *come*-Ø, and so on. But the postulation of zero affixes in DM is suspicious. Not only are they without empiricial justification (a fact proven by the possibility of dispensing with them altogether in inferential theories of morphology); they also exhibit a distribution that is different from that of any real (empirically motivated) affix. In DM, non-concatenative operations do not constitute vocabulary items, but are instead treated as effects triggered by vocabulary items. Accordingly, the assumptions of DM entail that English morphology (and that of most languages) evinces two rather amazing coincidences. First, default suffixes such as plural *-s*, past-tense *-ed*, comparative *-er* (and so on) are overridden by suffixes that are identical in their phonological form: thus, *men*-Ø, *sang*-Ø and *worse*-Ø all have a suffix with zero phonology. Second, these zero suffixes all have the effect of triggering nonconcatenative operations of apophony or stem suppletion: *man* → *men*, *sing* → *sang*, *bad* → *worse*. Again and again, DM analyses involve sets of competing vocabulary items whose default member is overridden by an affix having zero phonology; again and

[19] Besides being questionable on theoretical grounds, some of Embick and Noyer's claims are dubious at a purely observational level. For instance, in arguing that the Huave reflexive suffix *-ay* exhibits local dislocation, they assert that it is subject to a morphological well-formedness condition which requires it to appear "directly before the final inflectional affix of a verb, if any". In support of this assertion, they allege the ungrammaticality of forms such as *t-e-kohč- ay-os-on* 'we cut ourselves', attributing their data to Stairs and Hollenbach (1981). But this source doesn't actually say that *t-e-kohč- ay-os-on* is ungrammatical; in fact, Stairs and Hollenbach (1969: 40) cite the incidence of forms such as *ta-kohc-ay- os-on* 'we cut [past] each other', in which *-ay* precedes two suffixes.

again, this zero affix triggers a nonconcatenative operation. DM portrays these coincidences as purely accidental, stipulating them over and over in one set of vocabulary items after another. In inferential theories of inflection such as PFM, by contrast, the affixation operations are in direct paradigmatic opposition to nonconcatenative operations. For this reason, neither of the suspicious coincidences arises: words like *men*, *sang* and *worse* are not alike in their suffixal phonology because they have no affixes, nor are their various non-concatenative operations conditioned by the presence of any affix.

The assumption that affixes basically correspond to functional heads also leads to the expectation that a word's affix positions and its morphosyntactic features will stand in a one-to-one correspondence. While this sort of correspondence can be found in some languages, it is far from the norm. In Kabyle Berber, for example, a finite verb expresses subject agreement in three affix positions, and it agrees with respect to three morphosyntactic features, those of person, number, and gender; see Table 12.16. Yet, there is no one-to-one correspondence between affix positions and features; on the contrary, the prefix *i-* in *i-wala* 'he saw' expresses third person, singular number, and masculine gender; the first suffix *-m* in *t-wala-m-t* 'you (fem.) saw' expresses second person and plural number; and the second suffix in *t-wala-m-t* expresses plural number and feminine gender.

The problems which such instances of extended and overlapping exponence present for the claim in (24) are widely recognized. But the counterevidence to (24) is in fact much more extensive than has been widely appreciated. The phenomenon of syncretism (section 12.5.2), for example, poses serious problems for maintaining the claim in (24). Proponents of DM have claimed that analyses exploiting underspecified vocabulary items and rules of impoverishment suffice to reconcile this claim with the incidence of syncretism, but they do not suffice: some types of syncretism (e.g. bidirectional syncretisms: Stump 1993b, 2001: 212ff; Baerman 2004) do not submit to either type of analysis. The phenomenon of deponency (section

Table 12.16 Completive paradigm of Kabyle Berber WALI 'see'

		Singular	Plural
1		wala-ɣ	n-wala
2	masc	t-wala-ḍ	t-wala-m
	fem	t-wala-ḍ	t-wala-m-t
3	masc	i-wala	wala-n
	fem	t-wala	wala-n-t

Sources: Hamouma n.d.: 79; Chaker 1983: 112

12.5.1) disconfirms the claim in (24) in the most dramatic way possible: a deponent word exhibits the morphology appropriate to a set of morphosyntactic properties which it manifestly does not possess. Non-compositional periphrasis (section 12.7.3) is similarly problematic for DM: here a periphrastic expression's morphosyntactic properties are distinct from the sum of those of its individual parts. It is perhaps not surprising that some of these phenomena have yet to figure with any prominence in the DM literature; all, however, have been shown to be compatible with the assumptions of PFM.

12.11 Conclusions

In the foregoing discussion, we have identified a range of criteria for evaluating theories of morphology and its interface with syntax. Empirically, the best-motivated theory of morphology must be both inferential and realizational (section 12.3); the best-motivated theory of the morphology–syntax interface must be word-based rather than morpheme-based (section 12.1); and all else being equal, a DEFO theory is preferable to a BASI theory because of the greater restrictiveness inherent in DEFO analyses (section 12.10). By these criteria, PFM must clearly be preferred to a number of existing alternatives.

Further reading

Ackerman, F. and Stump, G. (2004), "Paradigms and Periphrastic Expression: A Study in Realization-Based Lexicalism", in L. Sadler and A. Spencer (eds.), *Projecting Morphology*, Stanford, CA: CSLI Publications, 111–57.

Anderson, S. R. (1992), *A-Morphous Morphology*, Cambridge: Cambridge University Press.

Matthews, P. H. (1972), *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*, Cambridge: Cambridge University Press.

Stump, G. T. (1993*a*), "On Rules of Referral", *Language* 69: 449–79. Repr. in F. Katamba (ed.), *Morphology: Critical Concepts in Linguistics*, London: Routledge, 2003.

—— (1993*b*), "Position Classes and Morphological Theory", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1992*, Dordrecht: Kluwer, 129–80.

—— (1995). "The Uniformity of Head Marking in Inflectional Morphology", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1994*. Dordrecht: Kluwer, 245–96.

—— (2001). *Inflectional Morphology*, Cambridge: Cambridge University Press.

—— (2002). "Morphological and Syntactic Paradigms: Arguments for a Theory of Paradigm Linkage", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2001*, Dordrecht: Kluwer, 147–80.

—— (2005), "Referrals and Morphomes in Sora Verb Inflection"', in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2005*, Berlin: Springer, 227–51.
—— (2006), "Heteroclisis and Paradigm Linkage", *Language* 82: 279–322.
Zwicky, A. M. (1985), "How to Describe Inflection", in M. Niepokuj, M. Van Clay, V. Nikiforidou, and D. Feder (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, 372–86.
—— (1986), "The General Case: Basic Form Versus Default Form", in V. Nikiforidou, M. Van Clay, M. Niepokuj and D. Feder (eds.), *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, 305–14.

References

Ackerman, F. and Stump, G. T. (2004), "Paradigms and Periphrastic Expression: A Study in Realization-Based Lexicalism", in L. Sadler and A. Spencer (eds.), *Projecting Morphology*, Stanford: CSLI Publications, 111–57.
Alhoniemi, A. (1985), *Marin Kielioppi*, Helsinki: Suomalais-Ugrilainen Seura.
Anderson, S. R. (1969), *West Scandinavian vowel systems and the ordering of rules*, Doctoral dissertation, MIT.
—— (1986), "Disjunctive Ordering in Inflectional Morphology", *Natural Language and Linguistic Theory* 4: 1–31.
—— (1992), *A-Morphous Morphology*, Cambridge: Cambridge University Press.
Arnott, D. W. (1970), *The Nominal and Verbal Systems of Fula*, Oxford: Oxford University Press.
Baerman, M. (2004), "Directionality and (Un)Natural Classes in Syncretism", *Language* 80: 807–27.
Brown, D. and Corbett, G. (to appear), *The Syntax–Morphology Interface: A Study of Syncretism*, Cambridge: Cambridge University Press.
Biligiri, H. S. (1965), "The Sora Verb: A Restricted Study", in G. B. Milner and E. J. A. Henderson (eds.), *Indo-Pacific Studies, Part II: Descriptive Linguistics* [*Lingua* 15], Amsterdam: North-Holland, 231–50.
Börjars, K., Vincent, N. and Chapman, C. (1997), "Paradigms, Periphrases and Pronominal Inflection: A Feature-Based Account", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1996*, Dordrecht: Kluwer, 155–80.
Brown, D., Corbett, G., Fraser, N., Hippisley, A. and Timberlake, A. (1996), "Russian Noun Stress and Network Morphology", *Linguistics* 34: 53–107.
Bybee, J. (1985), *Morphology: A Study of the Relation Between Meaning and Form*, Amsterdam: Benjamins.
Chaker, S. (1983), *Un parler berbère d'Algerie (Kabylie)*, Doctoral dissertation, Université de Paris V.
Corbett, G. G. & Fraser, N. M. (1993), "Network Morphology: A DATR Account of Russian Nominal Inflection", *Journal of Linguistics* 29: 113–42.
Dressler, W. U., Mayerthaler, W. Panagl, O. and Wurzel, W. U. (1987), *Leitmotifs in Natural Morphology*, Amsterdam: John Benjamins.
Dzokanga, A. (1979), *Dictionnaire lingala–français, suivi d'une grammaire lingala*, Leipzig: Verlag Enzyklopädie Leipzig.

Halle, M., and Marantz, A. (1993), "Distributed Morphology and the Pieces of Inflection", in K. Hale and S. J. Keyser (eds.), *The View from Building 20: Linguistic Essays in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press., 111–76.

Halpern, A. L. (1992), *Topics in the Placement and Morphology of Clitics*, Doctoral dissertation, Stanford University.

Hamouma, H. (n.d.), *Manuel de grammaire berbère (kabyle)*, Paris: Association de Culture Berbère.

Kiparsky, P. (1973), " 'Elsewhere' in phonology", in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*, New York: Holt, Rinehart & Winston, 93–106.

Lapointe, S. G. (1990), "EDGE features in GPSG", in M. Ziolkowski, M. Noske, and K. Deaton (eds.), *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, i, Chicago: Chicago Linguistic Society, 221–35.

Lieber, R. (1992), *Deconstructing Morphology*, Chicago: University of Chicago Press.

Matthews, P. H. (1972), *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*, Cambridge: Cambridge University Press.

Miller, P. H. (1991), *Clitics and Constituents in Phrase Structure Grammar*, Doctoral dissertation, University of Utrecht.

Noyer, R. R. (1992), *Features, Positions and Affixes in Autonomous Morphological Structure*, Doctoral dissertation, MIT.

Pullum, G. K. and Scholz, B. (2001), "On the Distinction Between Model-Theoretic and Generative–Enumerative Syntactic Frameworks", *Lecture Notes in Computer Science* 2099, Berlin: Springer-Verlag, 17–43.

Sadler, L. and Spencer, A. (2001), "Syntax as an Exponent of Morphological Features", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2000*, Dordrecht: Kluwer, 71–96.

Stairs, E. F. and Hollenbach, B. E. (1969), "Huave verb morphology", *International Journal of American Linguistics* 35, 38–53.

—— (1981), "Gramática huave", in G. A. S. Kreger and E. F. Stairs (eds.), *Diccionario huave de San Mateo del Mar*, Mexico City: Summer Institute of Linguistics.

Steele, Susan, (1995), "Towards a Theory of Morphological Information", *Language* 71: 260–309.

Stewart, Jr., T. W. (2004), *Mutation as Morphology: Bases, Stems, and Shapes in Scottish Gaelic*, Doctoral dissertation, The Ohio State University.

Stump, G. T. (1993*a*), "How Peculiar is Evaluative Morphology?", *Journal of Linguistics* 29: 1–36.

—— (1993*b*), "On rules of Referral", *Language* 69: 449–79. [Repr. 2003 in Francis Katamba, (ed.), *Morphology: Critical concepts in linguistics*, London: Routledge.]

—— (1993*c*), "Position Classes and Morphological Theory", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1992*, Dordrecht: Kluwer, 129–80.

—— (1995), "The Uniformity of Head Marking in Inflectional Morphology", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 1994*, Dordrecht: Kluwer, 245–9.

—— (2001), *Inflectional Morphology*, Cambridge: Cambridge University Press.

—— (2002), "Morphological and Syntactic Paradigms: Arguments for a Theory of Paradigm Linkage", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2001*, Dordrecht: Kluwer, 147–80.

—— (2005), "Referrals and Morphomes in Sora Verb Inflection", in G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2005*. Berlin: Springer Verlag.

—— (To appear *a*), "Heteroclisis and Paradigm Linkage", *Language* 82.

—— (To appear *b*), "Morphological Blocking and Pāṇini's Principle", in F. Ackerman, J. Blevins, and G. Stump (eds.), *Paradigms and Periphrasis*, Stanford: CSLI Publications.

Whitney, W. D. (1889), *Sanskrit grammar* [2nd edition]. Cambridge, MA: Harvard University Press.

Zwicky, A. M. (1985), "How to Describe Inflection", in M. Niepokuj, M. Van Clay, V. Nikiforidou and D. Feder (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, 372–386.

—— (1986), "The General Case: Basic Form Versus Default Form", in V. Nikiforidou, M. Van Clay, M. Niepokuj, and D. Feder (eds.), *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Linguistics Society, 305–14.

—— (1987), "Morphological Rules, Operations, and Operation types", *ESCOL '87*, 318–34.

—— (1992), "Some Choices in the Theory of Morphology", in R. D. Levine (ed.), *Formal Grammar: Theory and Implementation*, 327–371. New York and Oxford: Oxford University Press.