Inflectional morphology: A theory of paradigm structure

By Gregory T. Stump


Reviewed by Michael Maxwell
Linguistic Data Consortium

One of the dominant approaches to inflectional morphology in recent years is known as “realizational morphology.” The idea is that a word in a language which has inflectional morphology starts out as a stem (with derivational affixation already having taken place) plus a set of morphosyntactic features which describe the cell of the paradigm to which the inflected word belongs. For example, the cell of a paradigm which holds the third person singular present indicative verb might have a morphosyntactic feature structure like the following:

(1) SubjAgr: \[
\begin{array}{c}
\text{Person} \\
\text{Number} \\
\end{array}
\]

Tense: Present
Mood: Indicative

Then for each position (roughly speaking, a “slot”) in which an affix may occur, an affix is chosen which matches (or more precisely, unifies with) those features. If no affix belonging to a given position matches the features, that position is bypassed, giving the effect of a zero affix (but without the theoretical problems associated with zero morphemes).

Realizational morphology has well-motivated explanations not only for zero affixes, but also for multiple exponence (the situation where more than two affixes on a single word may mark the same morphosyntactic features), as well as some kinds of syncretism (the common tendency for more than one cell in a paradigm to be marked by the same affix).

Gregory Stump is one of the leading proponents of Realizational Morphology and is well qualified to present the theory. (Other advocates of Realizational Morphology include P. H. Matthews, Arnold Zwicky, and Stephen Anderson.) The first two chapters of this book could be used as an introduction to the theory of Realizational Morphology in a graduate level course; the remaining chapters are largely an argument for Stump’s particular version of the theory. It is these extensions to the core of Realizational Morphology that I would consider controversial, and the remainder of this review will be concerned with these extensions.
Chomsky 1965 introduced the idea of three levels of adequacy: observational, descriptive, and explanatory. It is generally quite clear when a theory fails the test of observational adequacy. But in morphology (unlike syntax), it would at least in principle be possible for a theory to achieve this level of adequacy by simply listing all the forms of words. In practice, such a ‘theory’ would be laughed down, since for at least some languages, the morphology is too complex for it to be plausible that a learner could encounter and memorize all the forms of every word. That is, listing every form fails on grounds of explanatory adequacy. Therefore, at least some morphology must be rule based.

Nevertheless, it is not a priori obvious at what point rule-based explanations of individual facts should end and listing begin. I will in fact suggest that Stump goes too far in proposing rule-based generalizations, and that it may be possible to eliminate whole classes of rules. It is simply not clear that humans make all the generalizations that one could make about morphological systems, or make those generalizations in the way Stump suggests. In short, it may be that morphology—as it is represented in the minds of native speakers—is simpler than Stump makes it out to be.

As an example of a complication that does not appear to be well motivated, consider the interaction of (morpho-)phonological rules with individual affixes, in Stump’s analysis of Bulgarian in chapter two. Stump follows Zwicky (in an unpublished paper) in claiming that rules are “indexed” for morphological processes. That is, each phonological rule refers to a certain set of affixes, and that it may be possible to eliminate whole classes of rules. It is simply not clear that humans make all the generalizations that one could make about morphological systems, or make those generalizations in the way Stump suggests. In short, it may be that morphology—as it is represented in the minds of native speakers—is simpler than Stump makes it out to be.

Moreover, the set of phonological rules that Stump gives are more complicated than necessary, one of the rules (17c, p. 48) actually being redundant. The complete analysis is given in the appendix to this review.

I hasten to add that particular analyses need to be tested in rigorous fashion. When I attempted to reanalyze the Bulgarian data, it quickly became apparent that a much simpler analysis than Stump’s (without indexing) was probably possible. The important word here is ‘probably’: because of the complexity of both the paradigms and the rules, I did not have confidence in my analysis until I had implemented it computationally. And in fact the redundancy of rule (17c) did not become apparent until the implementation was in place, and I could experiment with variations. I strongly urge anyone proposing a nontrivial morphological and phonological analysis to implement that analysis computationally; you are almost certain to find errors.

As another example of a possibly unmotivated mechanism, consider the discussion of syncretisms in Stump’s chapter seven. A syncretism occurs when a particular cell (or set of cells constituting a natural class) in a paradigm are identical in form with some other cell (or set of cells). For example, in Spanish the second person polite forms of verbs are always the same as the third person forms of the same plurality (a fact which is due to the historical origin of the second person polite forms).
Some syncretisms can be captured as natural classes, using morphosyntactic features, or as the default affix in some slot. Some forms of syncretism can be accounted for by saying that the syncretistic affix (or often, the zero affix) is the “elsewhere” case. For example, in English there is no need for a rule of referral to explain why all person/number combinations of the verb for a particular tense—except for the third person singular present—are identical. The explanation is simply that the only affix which distinguishes a person/number set is the third person singular present; all other forms (apart from the irregular forms of be), including the unaffixed present tense forms, are simply the “elsewhere” cases of their respective tenses.

But it is not always possible to capture syncretism by means of an “elsewhere” case; Stump refers to those instances of syncretism which cannot be captured in this way as being “stipulated” syncretisms, accounting for most of them by means of Rules of Referral. These are rules which stipulate that the form associated with a specified set of morphosyntactic features is identical to the form associated with some other set of features. But some stipulated syncretisms are accounted for by means of a distinct kind of rule, a “Symmetrical Syncretism Metarule.” The question which I would raise is whether these syncretisms should be stipulated in the grammar at all.

Notice that the question here concerns the grammar as represented in the native speaker’s mind. There is no doubt that pointing out syncretisms in a pedagogical grammar would be helpful to a second language learner. Syncretisms should also be pointed out in descriptive linguistic grammars, since they are fodder for the theory mills. And finally, the historical development of syncretisms is of interest. But my question is: do the grammars of native speakers contain stipulated syncretisms?

An example from Stump’s book may make the issue clearer: “In Rumanian, verbs belonging to any but the first conjugation have present indicative paradigms in which the 1sg form is identical to the 3pl form…” [p. 213]. The table of the present indicative verbs given by Stump shows that this syncretism has three effects:

- if the 3PL suffix is null, so is the 1SG;
- if the 3PL suffix has the form –u, so does the 1SG suffix (but none of the other suffixes take this form); and
- if there is a different stem form for the 3PL, the 1SG shares this stem.

But is this sufficient evidence for saying that the native speaker’s grammar must capture these facts? The syncretism is limited to a particular tense/mood combination, and even then it fails for verbs of the first conjugation class. It in fact seems quite possible that this comparatively minor generalization is simply missed by the native speaker. What we need to decide the question is empirical evidence. Historical change might supply this evidence, for instance if it could be shown that verbs of the first conjugation class were becoming like those of the other conjugation classes in obedience to this generalization.

The evidence for other stipulated syncretisms is stronger, in the sense that they have fewer exceptions than the above Romanian example, or cover more cases. One such example which Stump presents (pp. 216–7) is Hua, of Papua New Guinea (the data is from Haiman 1980). In this language, the 2SG and 1PL forms of verbs are marked by the same suffixes, of which there
are about a dozen for different moods, while the other person/number forms of verbs take a
different “default” suffix. I know of no morphosyntactic feature system in which second person
singular and first person plural would form a natural class, nor would the complementary set of
person/number combinations; hence, this would have to be a stipulated syncretism.

Still, it seems possible that the syncretism is simply not represented in the synchronic grammar.
One possibility is that native speakers have disjunctive feature sets for each suffix for the 2SG
and 1PL forms, with the non-2SG/1PL forms being the “elsewhere” suffix. Alternatively, native
speakers might simply have homophonous suffixes for these two morphosyntactic feature
combinations.

If we accept this analysis, we would be faced with the question of how such a striking situation
arose. I do not have any clear idea, although I will note that if we did not know the historical
origin of the Spanish syncretism between second person polite and third person verb forms, we
might puzzle over the origin of that syncretism, too. However, there is a strong resemblance
between the 2SG/1PL suffixes on the one hand, and the default suffixes on the other: in each
case, the two suffixes differ at their left end by a single consonant, as shown in the following
table (taken from Stump’s table 7.4, p. 216; I omit the meaning of the suffixes):

<table>
<thead>
<tr>
<th>Default</th>
<th>2sg/1pl</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>-e</td>
<td>-ne</td>
<td>0 → n</td>
</tr>
<tr>
<td>-hipana</td>
<td>-sipana</td>
<td>h → s</td>
</tr>
<tr>
<td>-hine</td>
<td>-sine</td>
<td></td>
</tr>
<tr>
<td>-ve</td>
<td>-pe</td>
<td>v → p</td>
</tr>
<tr>
<td>-va</td>
<td>-pa</td>
<td></td>
</tr>
<tr>
<td>-ma</td>
<td>-pa</td>
<td>m → p</td>
</tr>
<tr>
<td>-ma’</td>
<td>-pa’</td>
<td></td>
</tr>
<tr>
<td>-mi’</td>
<td>-pi’</td>
<td></td>
</tr>
<tr>
<td>-mana</td>
<td>-pana</td>
<td></td>
</tr>
<tr>
<td>-mane</td>
<td>-pane</td>
<td></td>
</tr>
<tr>
<td>-mae</td>
<td>-pae</td>
<td></td>
</tr>
<tr>
<td>-ga</td>
<td>-na</td>
<td>g → n</td>
</tr>
</tbody>
</table>

It seems plausible that something phonological is happening here—or more likely, something did
happen in the past. For instance, there may have been some single suffix marking 2SG and 1PL,
which would be odd, but not a syncretism, since it only involves a single affix); this hypothetical
suffix subsequently phonologically merged with the suffix to its right, and the phonological
process became opaque through later developments. Admittedly, the nature of this suffix (and
the phonological process of merger) is unclear. But the point is that there is no hard evidence
that the syncretism is currently represented in the grammars of native speakers.
A third example motivating Rules of Referral appears in Stump’s discussion of Lingala (pp. 144-149). In this language, transitive verbs have distinct prefix positions (slots) for subject and object agreement. A large number of prefixes in the two slots are the same, but a few affixes (mostly first and second person) differ. Stump’s analysis uses both Rules of Exponence and Rules of Referral: the Rules of Exponence in each slot constitute a block of affixes which are unique to each slot, while the Rules of Referral point to a block of affixes which are common to the two slots.

A simpler analysis, eliminating the Rules of Referral, would use a single set of agreement marking prefixes, which are shared between the two slots. (In essence, the prefixes would not be “owned” by the slots at all, only referred to.) Affixes which mark only subject agreement would appear only in the subject agreement slot by virtue of the fact that their features would be compatible only with subject agreement (cf. the feature structure at the beginning of this review), and similarly for object agreement affixes. Affixes which mark either subject or object agreement would have features compatible with either sort of agreement. Stump already uses such a feature structure, and my proposal removes a redundancy in his analysis: under his analysis, the fact that a prefix marks subject agreement, object agreement, or both, is stipulated twice: once by the features of the prefix and once by virtue of its membership in a particular affix block. Under my reanalysis, this information need only be given once, in the feature structure of the affix.

A similar situation occurs in Tzeltal: the prefixes on verbs, marking the ergative person, are identical (including their phonologically conditioned allomorphy) with the possessive prefixes on nouns. I do not believe this could be treated by Rules of Referral, which operate within paradigms, rather than across different parts of speech. But it can be explained by the alternative analysis I suggested, in which the affixes belong to some other component of the grammar, and can be referred to by the ergative slot on verbs and by the possessive slot on nouns.

What would be gained, by eliminating stipulated syncretisms and reanalyzing cases where affixes appear to belong to two or more slots, is the elimination of two classes of rules (Rules of Referral and Symmetrical Syncretism Metarules), with no loss to observational adequacy. In summary, I believe that inflectional morphology can plausibly be considered as simpler than Stump would have it.

I have talked at length in this review about areas of disagreement which I have with Stump’s theory. But that should not obscure the fact that this book is the most detailed and comprehensive exposition of inflectional morphology in many years. Regardless of your theoretical proclivities, if you are interested in inflectional morphology, you should read this book. The analyses are formal and detailed enough to be testable, the data is interesting, and the exposition is clear. And if you are trying to understand the inflectional morphology of a particular language, the ideas here may well help.

Appendix

This appendix contains a reanalysis of the phonological portion of Stump’s Bulgarian verb paradigm. The “morphophonological rules” which Stump proposes (his (17), p. 48) are shown
here for reference; I have translated his verbal descriptions into a more or less standard rule notation.

\[\begin{align*}
\text{a.} & \quad V \rightarrow 0 / \text{Stem, CONJ: [-T –C]} \quad \_
\text{b.} & \quad V \rightarrow 0 / \_
\text{c.} & \quad V \rightarrow 0 / V + \\
\text{d.} & \quad V \rightarrow [+stress] / \text{unstressed word} \\
\text{e.} & \quad [C +back] \rightarrow [+high] / V –back \\
\text{f.} & \quad Á \rightarrow e \\
\text{g.} & \quad Á \rightarrow é / C [V –back] \\
\text{h.} & \quad Á \rightarrow á
\end{align*}\]

Some explanations: Rule (a) applies to the [-T –C] class. This is a lexically marked set of verbs; the [-T] feature stands for “non-truncating,” while the [-C] feature indicates that the stem ends in a nonconsonant, and is actually redundant, given that rule (a) will not apply to a consonant-final stem.

As a side effect of the application of rule (b), the stress on the deleted vowel, if any, shifts to the remaining vowel. This fact is assumed here to follow by ordinary conventions of rule application. (Stump makes a similar assumption, apparently.)

Rule (c) should be self-explanatory. Rule (d) applies only in the absence of lexically marked stress, as indicated by the informal “unstressed word” on the right-hand side; again, this fact must be assumed to follow by convention. Rule (e) is a rule of palatalization (and would be captured by a rule of spreading under an autosegmental analysis).

The symbols “A” in rules (f)-(h) represent a morphophoneme whose surface reflex undergoes an alternation between [e] and [a]; the “Á” in rule (f) stands for the unstressed version, while the “Á” in rules (g) and (h) stands for the stressed version. I will not focus on the theoretical issues caused by morphophonemes, again assuming that there is a straightforward theoretical explanation. (One such explanation would be absolute neutralization.)

The crucial problem, in my view, is that Stump indexes each of these rules to apply following the attachment of a subset of the verbal affixes. Thus, whereas rule (b) would always bleed rule (c) if there were no other conditions, the two rules in fact are intended to apply to a disjunctive set of affixes. Indexing rules to affixes results in a powerful grammar—more powerful, I contend, than is necessary. I intend to show that it is in fact not necessary.

In the following table, I have indicated Stump’s indexing between affixes and rules, with the rules appearing across the top, and the suffixes down the left-hand side (the codes B1 through D9 are Stump’s). The “B,” “C,” and “D” refer to a slot containing a set of affixes, only one of which can attach. Thus, suffixes B1, B2, and B3 are mutually exclusive (and also mutually exclusive with the zero affix B4).

An “x” at the intersection of a row and column indicates that Stump stipulates that the rule applies following the attachment of the given affix, while the lack of an “x” indicates that the
Rule (a) always applies (if only vacuously), and thus it need not be indexed.

Rule (b) is conditioned morphosyntactically: it applies only (and always) when the tense is “present.” I will return later to how this fact should be indicated.

As mentioned above, rules (b) and (c) apply in a complementary set of positions. That is, on the assumption that rule (b) is ordered before (c), rule (c) applies unless rule (b) bleeds it. Thus, rule (c) does not need to be indexed, if rule ordering is allowed (or if, as I will suggest, rule (b) is actually a morphological process, not a phonological rule, and therefore intrinsically ordered before rule (c)).

Rule (d) applies only after the B suffixes. It does not seem obvious that the phonological conditions on its application can be revised so that it applies after the C and D suffixes, i.e., at the word level. In particular, a moraic solution does not seem plausible. One solution, then, would be to use a distinct morpheme boundary between the stem and the B suffixes. A more
modern version of this solution would assume a stratal distinction between the B suffixes on the one hand and the C and D suffixes on the other, with rule (d) (and rules (a) through (c)) applying on the earlier stratum. While the use of strata introduces a certain power to the grammar, it is certainly a constrained power when compared with the power resulting from arbitrarily indexing rules to affixes.

Rule (e) always applies (if only vacuously), thus its indexing is moot.

Rules (f) through (h) can apply at the word level. Of these, rule (g) is crucially ordered before (and bleeds) rule (h). The order of rule (f) with respect to these is not at issue.

Thus, apart from rule (b), which is morphosyntactically conditioned, the rules need not be indexed. What of rule (b), then? As alluded to above, rule (b) applies only to present tense verbs. It is, I would claim, not a phonological rule, but a morphological process: an affix of truncation. At any rate, its form would be:

\[(b') V \rightarrow 0 / __ \] [Pres Tns] + V

Eliminating the indexing is possible using Stump’s rules, modified slightly. But there is another analysis which also avoids indexing, and which may be preferable. Rules (a) through (c) are replaced by the following rules:

\[ (1) V \rightarrow 0 / __ \] [Pres Tns, +T] + V

\[ (2) V \rightarrow 0 / a __ \]

\[ (3) V \rightarrow 0 / __ ə \]

Rule (1) corresponds to part of the original rule (b) (the rest of (b) being covered by rule (3)), and is morphosyntactically conditioned. Again, it may be better treated as an affix of truncation (applying only to the [+Truncating] class of verbs); note that unlike Stump’s analysis, this morphosyntactically conditioned rule now applies before all the “ordinary” phonological rules, a fact which would fall out from its being a morphological rule (under the normal assumption that morphological process apply before phonological ones, at least in the absence of cyclic rule application, and abstracting away from the effects of strata).

Rule (2) corresponds to Stump’s rules (a) and (c). The reason rules (a) and (c) can be collapsed in this way follows from facts of Bulgarian morphology. First, consider Stump’s rule (a), repeated here for convenience:

\[ (a) V \rightarrow 0 / ] [\text{Stem, CONJ: [–T –C]}] __ \]

There are four classes of Bulgarian verbs with respect to the stem-final phoneme, captured by Stump’s classes [+T +C], [+T –C], [–T +C], and [–T –C]. The two [+C] classes end in consonants, hence stem-final vowel deletion can never apply to them, even if the exception feature [–T –C] is replaced by [–T].
Of the remaining two classes, the \([+T–C]\) verbs will fall under my rule (1), deleting the stem vowel, and thereby correctly bleeding my rule (2). (No stems end in a sequence of two vowels, thus deletion of the stem vowel by rule (1) always leaves a consonant-final stem, which does not meet the structural description of rule (2).)

The final class is the \([-T–C]\) verbs, \textit{all of which end in the vowel /a/} (Aronson 1968:67). This is the vowel—in fact, the only vowel—which triggers deletion by Stump’s rule (a). Thus, my rule (2) covers Stump’s rule (a) without further stipulation, since this class of verbs will be the only source of a stem-final /a/ (all other stem-final /a/ vowels having been deleted by rule (1)). Moreover, there are no other sources of an /a/ vowel preceding another vowel. In particular, the imperfect suffix –A (labeled B2 in the above chart) is not a source of /a/ vowels in the relevant environment, because it is always immediately followed by the non-present suffix C2 –x.

Now consider Stump’s rule (c), repeated here:

\[(c) V \rightarrow 0 / V + _\]

Under his formulation, this rule deletes a vowel after any other vowel; it is bled by his rules (a) and (b). But in fact this means that rule (c) is \textit{always} bled, because the only two sources of vowels are vowel-final stems (his \([-C]\) class) or a preceding vowel-final suffix. Vowel-final stems of the \([-T–C]\) class cause a following vowel to be deleted by Stump’s rule (a), thereby deleting the vowel which would otherwise undergo rule (c) (recall that Stump indexes rule (a) so that it applies everywhere); vowel-final stems of the \([+T–C]\) class delete the stem-final vowel by rule (b), thereby deleting the vowel which would otherwise serve as the environment of rule (c). Likewise, a suffix-final vowel preceding a suffix initial vowel deletes by rule (b) in all the relevant cases, although this is more difficult to show; it did not become apparent to me until I implemented the rules computationally.

Stump also uses indexing for Potawatomi (his A.14 in chapter three); again, it does not appear to be necessary, since of the four phonological rules he proposes, two rules apply after every affix, and two at the word level. This would be easily accommodated by stratal ordering. Further simplification may be possible, but the morphological analysis is already so complex that it becomes difficult to determine the exact effects of phonological rule application.

This leads to a more general issue: deciding when a given analysis meets the standard of observational adequacy. Simply put, it is almost impossible to work out all the repercussions of a complex analysis in one’s head, or even with pencil and paper. The only reliable way to test an analysis is with a computational implementation, a point which Stump makes (p. 262). I can only agree heartily with him on this.

\textbf{Notes}

[1] Of course, it is possible that phonology does not involve rules at all, but only constraints, as in Optimality Theory. While it may be possible to recast phonological (and morphological) rules as constraints, I will not attempt that in this review.
[2] The implementation, which uses Xerox’s xfst program, is downloadable from http://www.ldc.upenn.edu/About/Staff/maxwell [apparently no longer accessible. –ed.]. For information about the xfst program itself, see http://www.stanford.edu/~laurik/fsmbk/home.html.

[3] There is considerable dialectal variation in the use of the polite/informal distinction, with many present-day dialects virtually lacking plural informal forms. However, the focus in the text is on the syncretistic forms, not on their sociolinguistic use.

[4] Stump distinguishes two cases of identity of form: where the entire wordform matches, and where the affixes chosen from a particular block (“slot”) match. He later suggests that the former may be subsumed under the latter.

[5] Stump lists four conjugation classes in Rumanian. He accounts for the exceptionality of the first paradigm by making the Rule of Referral apply to all paradigms, but having a special Rule of Exponence (affixation rule) which applies only to the first paradigm. By the assumption that Rules of principle, this special Rule of Exponence overrides the more general Rule of Referral. The issue I am raising, however, is not how the rules interact, but whether Rules of Referral exist at all.

[6] The Spanish second person polite forms Usted and Ustedes are shortened forms of archaic vuestra merced and vuestras mercedes ‘your mercy (mercies)’. Since the archaic forms were common nouns, they took third person agreement, hence the present-day syncretism between second person polite and third person plural endings.

[7] Haiman considers the possibility that there is a synchronic phonological process at work, but does not reach any conclusion.

[8] The exact nature of a feature structure which is compatible with either a subject or object agreement is dependent on the particular theory of features.

References

