Current morphological theory

A comment on Robert Beard & Mark Volpe: *Lexeme-Morpheme Base Morphology*

Richard Sproat

I present a computational model that allows for a clean formal account of a wide variety of morphological phenomena. Much of what i present is based on the large literature on finite-state approaches to morphology, but it is more coherent in that I propose that all morphological operations can be modeled using a single regular operation: composition. I define this operation and give several examples of its application. Two morphological models, 'Item and Arrangement' approaches (e.g. Lieber's), and 'Item and Process' approaches (e.g. Beard's) are compared with respect to this computational model. I will argue that there is not as much difference between these views as morphologists like to suppose.

1. Introduction

Some years ago I wrote a Ph.D. dissertation entitled *On Deriving the Lexicon* (Sproat 1985). It is not much cited nowadays, in part because the work was to some extent superseded by Lieber (1992), which developed a theory of morphology along much the same lines, and in part because there seems to be a collective memory span of perhaps ten to fifteen years in theoretical linguistics, and 1985 is just too long ago for anyone to remember.[1] Another reason, perhaps, that my earlier work is not much cited is that my main goal there was to attack the then dominant theory of **Lexical Phonology and Morphology**, by arguing that the constructs of stratum ordering, the tight coupling of phonological operations with morphological operations, so-called **bracketing paradoxes** and notions like **opacity** as implemented with **bracket erasure**, could all be derived from general syntactic and phonological assumptions, without recourse to a highly structured lexical component. Lexical Phonology has, as far as I can tell, largely gone the way of the dinosaurs, and so theories such as my own work, which depended upon it for their raison d'être, are no longer quite as relevant. Nonetheless, there is a sense in which my earlier approach is still relevant: While I no longer subscribe necessarily to the particular details of what I developed there, I am still sympathetic to the view that there is something wrong with highly articulated theories of morphology.

I have been asked to comment on one such theory, namely that of **Lexeme Morpheme Base Morphology** (LMBM), as summarized in Beard and Volpe's paper, and developed at length in (Beard, 1995).[2] There are a number of issues that one might comment on. One such important issue is the status of their **Unitary Grammatical Function Hypothesis**, by which they mean that there are but 44 universally available grammatical functions that play a role in both inflectional and derivational morphology. Among these functions are things such as *adession* ('on'), *inession* ('in'), *superession* ('over'), and *subession* ('under'). The problem with adducing universality to some of these functions is that, in light of the extensive comparative semantic work on spatial expressions by Steven Levinson and colleagues (e.g., Levinson et al., 2003), it is by no means clear what it means to claim that, say, 'on' is a universal function. Thus, while it is true that many languages do have an expression that may be translated with English 'on', in many instances, the particular way in which the 'on'-equivalent is used differs wildly from language to language. Languages have a habit of dividing up the range of spatial expressions along idiosyncratic lines, and it is thus at best misleading to suggest that there is a universal 'on' function since it is by no means clear what the semantics of such a function would be. At the very least this would need to be clarified.

In this reply, however, I will focus on a different issue, one that relates to LMBM's **Separation Hypothesis** in particular, and to classifications of morphological theories more generally. Despite the apparent differences between such theories as LMBM and, say, an **Item-and-Arrangement** approach such as Lieber's, is there in fact any difference between these theories, when considered at the right level? I shall suggest a heretical answer to this question. The heresy is born from my many years' work in computational linguistics, a field that has increasingly looked at the issue of wide coverage, and which has become increasingly impatient with complex mechanisms that explain relatively simple things. It is to a very brief review of computational theories of morphology that I now turn.

2. Computational morphology

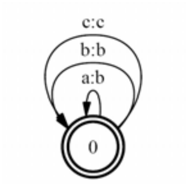
Computational approaches to morphology have been around for several decades but it is particularly within the last twenty years or so, with the work of Kaplan and Kay at Xerox PARC (eventually published as Kaplan and Kay, 1994), and Koskenniemi (1983), that computational morphology has come of age. This work derived in turn from an observation of C. Douglas Johnson (1972) that, under certain constraints, rewrite rule systems of the kind commonly used to express phonological rules, could be represented as **regular relations** and implemented computationally as **finite-state transducers**. Finite-state transducers compute string-to-string mappings, subject to the constraint that the relation expressed by the mapping can be constructed out of only the operations of concatenation, union and Kleene closure. Thus if R and S are regular relations, then so are R concatenated with S, R unioned with S, and R^*, the Kleene closure of R, that is zero or more instances of R concatenated with itself.

Regular relations are closed under a number of operations, the most significant of which is **composition**. This is to be understood in the algebraic sense of composition of two functions or relations. Thus, if R and S are regular relations then so is $T=R \circ S$, i.e. R composed with S. Furthermore, given an input x the result of applying T to x is the same as the result of applying first R and then S. Since finite-state transducers implement regular relations, the same operations that apply to regular relations apply also to transducers. Thus one can speak of, say, composing two transducers.

As an example of finite state transducers and their composition, consider a three-letter alphabet *{a,b,c}*, and a rule that changes any *a* in a string over that alphabet into a *b*. Such a rule would be written as:

(1) a -> b

A transducer that implements this rule over this alphabet looks as follows:

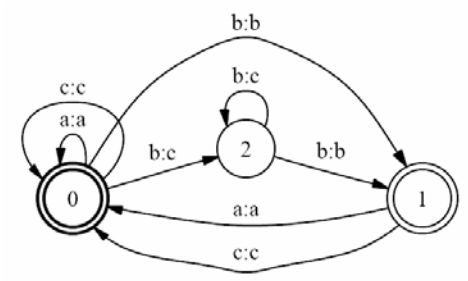


The interpretation is as follows: there is one state, labeled *0*, which in this case is both a starting state and a final state, the latter being indicated by the double circle. There are three arcs, each labeled with a colon-separated pair of symbols, representing the input and the output. Thus an arc labeled *a:b* maps input symbol *a* into output symbol *b.* To transduce a string, one starts in the initial state *0* and reads symbols from the string, matching them with labels on arcs on the input side and mapping them to the designated output label. Thus an input string *abbca* would map to *bbbcb.*

Consider now another rule that one might write as follows:

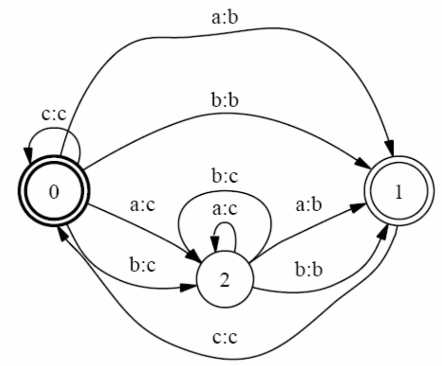
(2) b -> c / __ b

A transducer that implements this over the above alphabet is given below:



Here the structure is more complex – there are now three states, one initial (*0*, which is marked as such with bold circles), two final (*0*,*1*) and one that is neither initial nor final (*2*). But the mode of operation is the same as described above.

The two finite state transducers just described can be composed into a single transducer that implements the relation expressed by the two rules applied serially in the order given. Thus this transducer changes *a* into *b* and then furthermore changes any *b* (including those derived from *a*) into a *c* if it immediately precedes another *b* (including those derived from a). I leave it as an exercise for the reader to verify that the following transducer does in fact implement the described relation:



Composition has been used widely in implementations of phonological rewrite systems; see, again, Kaplan and Kay (1994). It also has a direct applicability in describing morphology. As we discuss in more detail in Roark and Sproat (forthcoming), composition is the most general computational operation that covers all morphological operations; the only exception is reduplication, which requires special machinery in a purely finite-state approach to morphology (since finite state techniques do not allow unbounded copying), but in fact requires special machinery in any approach.

Even vanilla affixation, which is more normally described in finite-state morphology using the operation of concatenation, can be subsumed under the operation of composition. Thus the more normal way of dealing with, say, suffixation would be to assume that a stem *A* is concatenated with a suffix B, to form a complex form Γ as follows (the center dot representing concatenation):

(3) $$\Gamma = A \cdot \beta$$

Yet one can also implement it as a composition operation where the intuition is that the relation involved is a relation that takes any stem, and returns that stem suffixed with the affix in question. Dispensing with drawing the transducers, we can define a new version of the affix as follows:

(4) $$\beta' = \Sigma^*[\epsilon : \beta]$$

Using standard notation, Σ denotes the whole alphabet of symbols, '*' denotes the Kleene closure (zero or more instances of members of Σ). Here we somewhat solecistically use this notation for a regular language to represent a regular relation where any string maps to itself. The symbol ε represents the empty string, so that the bracketed expression just means that the suffix β is inserted. Then one can redefine the operation of affixation as follows, using composition:

(5) $$\Gamma = A \circ \beta'$$

This may seem arcane, but note that affixes frequently have prosodic requirements on their base as in the case of English comparative affixation, or impose prosodic changes on their base as in Yowlumne (Archangeli 1984). Both of these can be handled by simple extensions of the above mechanisms, by replacing the no-op Σ^* with something that actually does some work, either imposing a restriction on the range of the relation, and thus the base, or else actually imposing a modification on the base. For Yowlumne, for example, consider alternations such as the following:

	Neutral Affixes		Template Affixes	
Root	*-al* 'dubitative'	*-t* 'passive aorist'	*-inay* 'gerundial' **CVC(C)**	*-ʃaa* 'durative' **CVCVV(C)**
caw 'shout'	caw-al	caw-t	caw-inay	cawaa-ʃaa-n
cuum 'destroy'	cuum-al	cuum-t	cum-inay	cumuu-ʃaa-n
hoyoo 'name'	hoyoo-al	hoyoo-t	hoy-inay	hoyoo-ʃaa-n
diiyl 'guard'	diiyl-al	diiyl-t	diyl-inay	diyiil-ʃaa-n
ʃilk 'sing'	ʃilk-al	ʃilk-t	ʃilk-inay	ʃiliik-ʃaa-n
hiwiit 'walk'	hiwiit-al	hiwiit-t	hiwt-inay	hiwiit-ʃaa-n

Thus for example the gerundial suffix *-inay* imposes a template **CVC(C)** on the base, no matter what the original shape of the base. This can be implemented as a transducer as follows:

(6) $$T_{cvc(c)} = CV[V:\epsilon]^*C[V:\epsilon]^*C?$$

This transducer works by deleting extraneous vowels and thus, in somewhat procrustean fashion, forces the base into the right shape. (See Roark and Sproat [forthcoming] for a treatment of the durative affix.) The full gerundive affix can then be represented as follows:

(7) $$\kappa_1 = T_{cvc(c)}[\epsilon : \text{inay}[+\text{GER}]]$$

Composing κ_1 with the stem then has the simultaneous effect of suffixing *-inay* plus the feature [+GER], and modifying the stem.

In Roark and Sproat (forthcoming), we show that regular relations, and thus finite-state transducers, can implement various kinds of morphology that fall under the general rubric of **prosodic circumscription** (McCarthy and Prince 1990), including Arabic broken plurals, infixation, and root-and-pattern morphology, as well as subtractive morphology and morphological alternations implemented using feature changes. We also argue that **morphemic** requirements as described in Aronoff (1994) are also readily handled by the identical mechanism. So for example, the Latin Third Stem, which Aronoff shows underlies a great many morphologically derived forms, can simply be treated as a prosodic requirement on the base, that can be handled in the same way as the prosodic requirements on the base of comparative affixes in English would be handled. Basically one just needs to reify the notion of **Third Stem**, which justifies creating a diacritic – e.g., [ThirdStem] – which marks all such instances; affixes that require this form merely specify that the presence of such a diacritic defines their bases.

So composition can model affixation, and indeed it is the most general finite-state device that can handle the full range of morphological operations.[3] And as we saw earlier, composition can handle rewrite rules of the kind that were classically used in generative phonology; indeed, as Karttunen (1998) shows, a simple variant on composition called **lenient composition** can also handle the cascade of violable constraints mooted by Optimality Theory. If composition can handle so many things and in particular if it can provide a mechanism for describing both affixation and rules, this should give one pause. Many morphological theories have staked their territory on the basis of whether they handle morphological derivation by rules or by lexemic affixes. The foregoing discussion suggests that perhaps there is no difference between the two. We now discuss this issue in more detail.

3. Differences between morphological theories

Traditional descriptions of morphological theory made a contrast between **Item-and-Process** and **Item-and-Arrangement** approaches to morphology. In the latter case, morphemes were all assumed to have entries in the lexicon, and complex words were constructed out of morphological pieceparts by concatenation or other combinatory operations. In Item-and-Process approaches, only major categories such as noun, verb or adjective have lexical entries; affixes and other 'functional' elements are introduced by rules that are sensitive to particular morphosyntactic features.

One of the useful contributions of Stump (2001) is the observation that there are really two dimensions to this issue. First of all, theories may be *lexical* or *inferential*.

Lexical theories are those theories in which all morphemes are given lexical entries, so that the third singular verbal affix *-s* in English has an entry associated with the features 3^{RD} SINGULAR, PRESENT and INDICATIVE. In contrast, *inferential* theories posit rules that are sensitive to such morphosyntactic features but, crucially, do not introduce them. Secondly, theories may be incremental or realizational. **Incremental theories** assume that rules or morphemes always add information when they are applied. Thus *likes* has its meaning by virtue of the addition of *-s* to *like*. Under **realizational theories**, in contrast, the introduction of form is licensed by particular morphosyntactic features. Stump shows that all four logically possible configurations are instantiated in the literature. To take just two examples, a typical lexical-incremental theory would be a classic Item-and-Arrangement model such as Lieber (1992). A typical inferential-realizational theory would be Stump's own approach.

Another inferential-realizational theory is Lexeme-Morpheme Base Morphology, as presented in Beard (1995) and in the Beard and Volpe paper in this volume. The **Separation Hypothesis***,* in particular, is a clear statement of an inferential theory, since affixes do not introduce features, but rather are sensitive to them. Similarly, the Separation Hypothesis is realizational since it is not the affixes that add information to bases, but rather morphosyntactic rules, which license the addition of particular affixes.

On the face of it the two most radically opposed models – namely lexical-incremental versus inferential-realizational – would appear to represent mechanically quite different approaches to morphological description. The problem is that it is easy to show that the theories are equivalent, both at the formal level and at the level of computational modeling. To see the first point consider Blevins' (2003) lexical-inferential analysis of West Germanic weak verb stem morphology.

Blevins observes that in all West Germanic languages – English, German, Frisian and Dutch – the past tense, past participle and perfect participle all share the same stem, which is formed with a dental. In English, this is exemplified by examples such as the following:

(8) 1. PAST: *John whacked the toadstool*
2. PERF: *John has whacked the toadstool*
3. PASS: *The toadstool was whacked*

The past form in 1, the perfect participle in 2 and the passive participle in 3 all share the same phonological property. In other West Germanic languages, some of these forms may have additional material. Thus in German:

(9) 1. PAST: *Er māh**t**et das Heu* 'He mowed the hay'
2. PERF: *Er hat das Heu gemäh**t*** 'He has mowed the hay'
3. PASS: *Das Heu wurde gemäh**t*** 'The hay was mowed'

The common feature of the forms is the dental **-t**, but there is additional material in each case: in the past there is a stem vowel *-e*, and in the perfect and passive participles, there is the prefix *ge-*. Blevins argues that one cannot view the dental suffix as being a single morpheme with a common semantics, and indeed this is also the conclusion reached by analyses based on morphemes, such as that of Pinker (1999), who argues for four distinct homophonous dental morphemes. This duplication is an embarrassment for lexical-incremental accounts to be sure, but it is important to bear in mind that the duplication is simply a fact of the data, and has to be incorporated somewhere in the model. In inferential-realizational accounts, such as the one Blevins provides, such duplications are handled by allowing many-to-one mappings between semantic features, and the morphological *exponents* of those features. Thus a realization function **R** is defined as follows, for English, where *Fd(X) = Xd* is a function that suffixes *-d* to the stem:

(10) **R** ([PAST]) = *Fd(X)*
R ([PERF]) = *Fd(X)*
R ([PASS]) = *Fd(X)*

Thus we have a many-to-one mapping between three semantic features, and a single exponent. Let us try to define these notions more formally as follows. First of all, we will use the abstract catenation operator '•' to represent the catenation of *-d* with the stem, and so we can redefine *Fd(X)* as follows:

(11) $Fd(X) = \lambda(X)[X \bullet d]$

Second, the realization expressions presumably do not just realize, say, [PAST], but realize it with respect to a certain base, the same base to which *-d* is ultimately attached. Let us assume an operator $\oplus$ to represent the addition of the relevant feature. Thus we would write:

(12) $\mathbf{R}\ (\lambda\ (X)[X \oplus PAST]) = \lambda\ (X)[X \bullet d]$

Now, one assumes that what it means to realize a particular feature or set of features on a stem by means of a particular morphological exponent is that one simultaneously adds the feature, and realizes the exponent of that feature. So we should be able to collapse the above into:

(13) $\lambda\ (X)[X \oplus PAST \wedge X \bullet d]$

Here, $\wedge$ simply denotes the fact that both the feature combination and the catenation operations take place. But, we can condense this expression further by collapsing the two combinatoric expressions into one:

(14) $\lambda (X)[X < \oplus; \bullet >< PAST; d >]$

Here $< PAST; d >$ is simply a pairing of the morphosyntactic/semantic feature with the phonological exponent. We use $< \oplus; \bullet >$ to represent a catenation pair which combines elements on the morphosyntactic/semantic side using $\oplus$ and elements on

the phonological side using •. In this formulation, we also need to consider *X* to be a morphosyntactic-phonological pairing, but we will leave this implicit in our notation.

At this point it should already be clear that the above expression is simply a singleplace curried version of the expression:

(15) $X < \oplus; \bullet >< PAST; d >$

This is clearly just a formulation of a lexical-incremental model.

One comes to the same conclusion if one considers the issue from a computational point of view. Recall that in Section 2 we argued that the most general operation that captures morphological operations is composition. Recall in particular that suffixation could be defined in terms of composition as follows:

(16) $\Gamma = A \circ \beta'$

where B′ represents a regular relation that inserts a suffix at the end of the base *A*. Consider now the agentive nominalizations discussed by Beard and Volpe. Thus we have forms like *read-er*, *stand-ee*, *correspond-ent*, and *record-ist,* involving the affixes *-er*, *-ee*, *-ent*, and *-ist,* as well as cases like *cook*, where there is apparently no affix. Since the particular affix chosen is dependent upon the base, any treatment of English morphology must treat the choice of affix as a lexical specification of the base. In many cases (though not apparently in this one) there may be lexical regularities such that bases with a particular general form predictably take a given affix. But no matter: the information must be lexically specified somewhere.

One computational implementation of this phenomenon would consist, straightforwardly, of a rule that introduces an affix at the end of a string as conditioned by a feature specification on the base. Thus for the case of *–ent* one might write the following rule:

(17) ε → ent / [+ent][+noun,+agentive] Σ^*__ $

Here we use *$* to represent the end of the string; as before, we use Σ^* to represent zero or more of any character from the alphabet of symbols. Thus the rule states that *–ent* is inserted at the end of a string if the base is marked with the features [+ent] and [+noun,+agentive]. A set of such rules would be needed to handle the range of affixes found for English agentive nominals. Call this set of rules **R**. Clearly we can define a new 'metarule' **R'**, defined as the union of all **R**:

(18) $R' = \bigcup R$

This new **R’**, which is simply an amalgamation of all of the individual affixes, can now be composed with any base to produce an agentive form appropriate for that base’s lexical specification.

We now need to ask the question of where the feature [+noun, +agentive] on the base comes from. Presuming that agentive forms are to be derived via a morphological process (and that does indeed seem to be the assumption that Beard and Volpe make), we must presume that there is a morphosyntactic-feature-introducing rule – let us call it **M** – that applies to nouns and introduces those feature specifications.[4] We assume that **M** is composed with a base **B** to produce a form that is marked with those features:

(19) **B○M**

This then can be composed with **R′** to produce an agentive form appropriate to **B**:

(20) **[B○M]○R’**

But now suppose we simply reassociate the brackets above (the operation of composition is associative) to yield:

(21) **B○[M○R’]**

Now observe that one can precompose **M○R’** into a single tranducer – call it **R”**, which has the following two properties:

- It introduces the morphosyntactic feature specification [+noun, +agentive].
- It spells out the morphology phonologically in a way that is appropriate to the base.

In short, **R”** encodes a lexical-incremental model of morphology.

There is still a question to be answered. One of the consequences of the Separation Hypothesis is that there are a limited number of affixes (or phonological exponents more generally) in a language, and these relate in a non-one-to-one fashion with the morphosyntactic operations of the language. Why should this in general be true? Presumably the fact that there are a limited number of phonological exponents is, in and of itself, of little interest: in any case such a set must surely be finite, and if we assume a bound on the length of added morphological material, then of necessity the set of possible affixes will be bounded and, depending upon the phonotactics of the language, potentially quite small. Of more interest is the non-one-to-oneness of the mapping between form and function, something that Beard and Volpe quite rightly point out is clearly true of functional morphemes, and barely if ever true of lexemes.

On the other hand, a look at cases like the English agentive suggests that at least some of the explanation must be purely historical. For example, it is presumably due to the Latin heritage of words like *correspond* that we have *correspondent*; similarly for *recordist.*[5] Presumably what needs to be said is that either someone at some point decided that a Latinate form should be used for the agentive of *correspond*; or else such forms took on the agentive function via lexical drift from a borrowed or constructed Latin form. Once they took on this function, **morphological blocking** (Aronoff 1976), itself probably derivable from the **Elsewhere Principle**, would presumably favor the more specialized form over the more general form. (Note however that such blocking is not absolute since alongside *recordist* one also finds *recorder.*)

This much anyone would have to agree on. Furthermore, there is clearly no requirement that the kind of one-to-many mappings one finds in English should necessarily hold in other languages. So one clearly finds languages like Mandarin, where agentives are formed completely regularly (via the affixation of *-zhe*). So it is simply a property of the history of English (and of course other languages) that one finds the particular array of facts that one does. The fact that speakers tolerate such a situation is presumably because there is no significant extra cognitive load to learning specialized forms of agentives, given that a significant amount of effort already goes into learning the lexical idiosyncracies of the language. In any event, it is not clear that once one has made the appropriate historical observations, one needs to say anything more to account for the kinds of observations that motivate the Separation Hypothesis.

Returning to the main point of this section, the Separation Hypothesis clearly suggests an inferential-realizational theory in Stump's (2001) terms. Such models are cast in direct opposition to lexical-incremental theories. But what the argument in this section suggests is that – as with Karttunen's (1998) computational analysis of Optimality Theory versus traditional Generative Phonology – there may not be much if any formal difference between the two approaches.

4. Conclusions

In claiming that there are no fundamental differences between morphological theories, I am well aware that I am treading on dangerous ground. Furthermore, since I have only sketched a few arguments in support of this view, I have no right to expect that most of the other contributors to this volume, or indeed many of its readers, will be convinced by what I have said. On the other hand, I am hoping that at least there will be some agreement that the views are worth considering.

The points I make here are part of a much larger program, one that I hope to develop further over the next few years. Theoretical linguistics has changed relatively little in its goals over the last five decades. One of those goals has been to develop theories of language, which are usually argued to be, or often merely presumed to be, reflections of an innate language faculty. The arguments for this position invariably take the following basic form:

- Linguistic phenomena are complex, and it is hard to see how a child could learn such complex phenomena if a fairly articulated model were not there to give him or her a leg up.
- Furthermore, there are certain observed universals that we need to account for so we need a constrained theory that accounts for these.

The first point clearly relates to assumptions about learnability. The second may relate to what humans are able to learn, but it may also equally well relate to models of historical development. With regard to the latter, it is noteworthy that there has been important recent work in phonology that has argued that most if not all of what we observe in the world's sound systems can be attributed to historical development (Blevins 2004). A core point here is that historical accounts are surely needed anyway, and once you have those it is not clear that synchronic accounts do any useful work.

With regard to the learnability, the assumptions of theoretical linguistics, as we have noted, have not changed much in the last half century. Meanwhile, the world around it has changed substantially. As Shalom Lappin and I argued in a recent posting to the Linguist List (Sproat and Lappin 2005), machine learning has made some significant advances over the last decade or so, and purely statistical learners have achieved some impressive results in the domain of unsupervised learning of syntax (e.g., Klein and Manning 2004) and also in morphology (e.g., Goldsmith 2001; Johnson and Martin 2003). Nobody thinks that unsupervised statistical methods have solved the problem of language acquisition. On the other hand, one can no longer fall back on one of Chomsky's favorite arguments that statistical methods are simply incapable of learning interesting generalizations about language. This point, plus the increasing evidence from child language acquisition showing the misguided nature of the old ideas about "poverty of stimulus" and lack of negative evidence (Chouinard and Clark 2003), suggest that it is high time we rethink the very foundations of linguistic theory. In particular, it is necessary, in my view, to question to what extent one needs highly articulated models that assume a complex innate language faculty.

And this in turn suggests some questions that bear more specifically on morphological theory. For example, in twenty years' time, will morphologists still be debating whether morphological theories should be incremental or realizational, or will such questions have become non-issues in favor of questions like what historical models might explain the distribution of morphological forms that we see, and which machine-learning methods most adequately account for how children acquire complex morphology? Were I a betting man I would be placing my bets on the latter.

Richard Sproat
Linguistics
Electrical and Computer Engineering
Computer Science
Beckman Institute
University of Illinois at Urbana-Champaign
Foreign Languages Building 4016D,
707 South Mathews Avenue, MC-168
Urbana, IL, 61801

rws@uiuc.edu

Notes

[1] This explains why Minimalists (Chomsky, 1995) can freely and sloppily recreate the devices of Generative Semantics, without admitting that this is what they are doing.

[2] To be fair, LMBM is much less complex than a theory such as that of Stump (2001).

[3] As we noted already, something special needs to be done for reduplication in anyone's model.

[4] Alternatively, it might change feature specifications of the base and also add information relating the denotation of the agentive to a particular argument of the verb. These details do not affect the point at hand.

[5] Alternations like *baker*, versus *standee*, might be explained on the basis of unaccusativity, if one can assume that the 'agent' of *stand* is actually an internal argument; thus on that analysis the reason we have *-ee* in this case is the same reason as we have *-ee* in *employee*, *trainee*, or (folk-etymologically) *mentee.*

References

Archangeli, Diana. 1984. *Underspecification in Yawelmani Phonology and Morphology.* Ph.D. dissertation, MIT.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. MIT Press: Cambridge, MA.

Aronoff, Mark. 1994. *Morphology by Itself: Stems and Inflectional Classes*. Linguistic Inquiry Monographs 22. MIT Press: Cambridge, MA.

Beard, Robert. 1995. *Lexeme Morpheme Base Morphology*. Albany: Suny Press.

Beard, Robert and Volpe, Mark. 2005. "Lexeme-Morpheme Base Morphology." In: P. Štekauer and R. Lieber (eds.), *Handbook of English Word-Formation*. Dordrecht: Springer, 189-205.

Blevins, James. 2003. "Stems and paradigms." *Language* 79: 737-767.

Blevins, Juliette. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns.* Oxford: Oxford University Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.

Chouinard, Mathias and Clark, Eve. 2003. "Adult reformulations of child errors as negative evidence." *Journal of Child Language* 30: 637-669.

Goldsmith, John. 2001. "Unsupervised Acquisition of the Morphology of a Natural Language." *Computational Linguistics* 27: 153-198.

Halle, Morris and Marantz, Alec. 1993. "Distributed Morphology and the pieces of inflection." In: K. Hale and S. J. Keyser (eds.), *The View from Building 20*. Cambridge: MIT Press, 111-176.

Kaplan, Ronald and Kay, Martin. 1994. "Regular models of phonological rule systems." *Computational Linguistics* 20: 331-378.

Karttunen, Lauri. 1998. "The proper treatment of optimality in computational phonology." In: *Proceedings of FSMNLP'98. InternationalWorkshop on Finite-State Methods in Natural Language Processing*. Bilkent University, 1-12.

Koskenniemi, Kimmo. 1983. *Two-Level Morphology: a General Computational Model for Word-Form Recognition and Production.* Ph.D. dissertation, University of Helsinki.

Johnson, C. Douglas. 1972. *Formal Aspects of Phonological Description.* The Hague: Mouton.

Johnson, Howard and Martin, Joel. 2003. "Unsupervised Learning of Morphology for English and Inuktitut." In: *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 43-45.

Lieber, Rochelle. 1992. *Deconstructing Morphology*. Chicago and London: University of Chicago Press.

Levinson, Stephen, Sergio Meira, and the Language and Cognition Group. 2003. "Natural Concepts in the Spatial Topological Domain – Adpositional Meanings in Crosslinguistic Perspective: an Exercise in Semantic Typology." *Language* 79: 485-516.

Klein, Dan and Manning, Christopher. 2004. "Corpus-Based Induction of Syntactic Structure: Models of Dependency and Constituency." In: *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004).*

Matthews, Peter. 1972. *Inflectional Morphology*. Cambridge: Cambridge University Press.

McCarthy, John and Prince, Alan. 1990. "Foot and word in prosodic morphology: The Arabic broken plural." *Natural Language and Linguistic Theory* 8: 209-284.

Pinker, Steven. 1999. *Words and Rules.* London: Weidenfeld and Nicholson.

Roark, Brian and Sproat, Richard. Forthcoming. *Computational Approaches to Syntax and Morphology*. Oxford: Oxford University Press.

Sproat, Richard. 1985. *On Deriving The Lexicon.* Ph.D. dissertation, MIT.

Sproat, Richard and Lappin, Shalom. 2005. "A Challenge to the Minimalist Community." Linguist List 16: 1156.

Stump, Gregory. 2001. *Inflectional Morphology: A Theory of Paradigm Structure.* Cambridge: Cambridge University Press.