

Real-Time Morphology: Symbolic Rules or Analogical Networks?

Author(s): Bruce L. Derwing and Royal Skousen

Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society (1989), pp. 48-62

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

Real-Time Morphology: Symbolic Rules or Analogical Networks?

Bruce L. Derwing
University of Alberta

and

Royal Skousen
Brigham Young University

1. Introduction

Some time ago, in describing to one of us the main theme of H.G. Wells' novel, 'The War of the Worlds', a friend explained that a group of powerful alien beings had invaded the earth and found ordinary earthlings to be relatively easy pickings; luckily for us, however, the microbes of our planet were a different kettle of fish and, having no suitable defences against them, these alien monsters '*succame* to our diseases' and the planet was saved.

As is well known, the child language literature is replete with anecdotes of this kind. To cite just three such examples (all coincidentally involving numbers): if four airplanes can be said to fly in *formation*, then two must fly in *twomation* (Sturtevant 1947); if it can be *too hot* in one place, why not *three hot* in another? (Jespersen 1922, cited by Hockett 1970:89); and if *forty-four*, *sixty-six*, and *seventy-seven* are all good number names, why not *fivety-five* or even *onety-one* (for 11) (Derwing 1976).

Surely no linguist would want to attribute such isolated examples as these to knowledge of 'rules', i.e., one for each individual case. In fact, as Householder (1971:63) has pointed out, the 'only candidate so far proposed for this job is analogy', the identification of a sameness of similarities (or differences) with other forms in the lexicon: *become:became::succumb:succame*, etc. Interestingly, such giants of earlier linguistic eras as Hermann Paul and Leonard Bloomfield (who, of course, were both acutely aware that 'analogical change' was also a well-documented historical process) linked all of what is now called 'linguistic creativity' to this notion: the process of 'freely creating' novel forms, says Paul, 'we call formation by analogy' (1891:97), while Bloomfield argues that 'a regular analogy permits a speaker to utter speech-forms which he has not heard; we say that he utters them on the analogy of similar forms which he has heard' (1933:275).

For all of this, few generative grammarians have ever paused in their own rush to theory to take the notion of analogy very seriously; one might say, in fact, that one of the fundamental tenets of the 'Chomskian revolution' was the implicit (or not so implicit) rejection of such a simple and relatively straightforward notion in favor of more abstract and convoluted approaches to the problem of making 'infinite use of finite means' (cf. Chomsky 1965:8). Kiparsky's brief (two-page) dismissal is a fairly typical one, which ends with the

conclusion that 'at the point at which ... analogies begin to make the right generalizations, they are indistinguishable from rules' (1975:189).

Needless to say, the recent emergence of 'connectionist' models from psychology is gradually forcing upon the linguistic community a re-assessment of this position, a renewal of interest in the very old idea that there is more to analogy than the mere cataloguing of a few anecdotal examples of the kind that appeared in our introduction above (see, for example, Bybee 1988, as good evidence of this). This is not a paper about 'connectionism', however, at least not about the specific model that most people currently associate with this term, viz., the theory of 'parallel distributed processing' [PDP] currently under development by David Rumelhart, Jay McClelland, and their research associates (see Rumelhart & McClelland 1986 [R&M]). This is, however, a paper about a whole class of possible models (including PDP) that can all reasonably be referred to as 'analogical' and that can be sharply differentiated from the kind of 'symbolic' or rule-based approaches that are exemplified by generative grammars in linguistics (cf. Pinker & Prince 1988).

In brief, therefore, this paper will attempt to (1) elucidate some of the key features that differentiate the 'analogical' approach from the generative or rule-based one and then (2) proceed to consider some evidence, based on psycholinguistic research carried out over quite a number of years, which lends, as we see it, sufficient *prima facie* support in favor of the analogical approach to (re-)establish it as a candidate worthy of a much more careful hearing by linguists than has been afforded in recent years.

2. Rule-based vs. Analogy-based Theories

How, then, can analogy-based theories be distinguished from the kind of rule-based theories we now find almost exclusively represented in linguistics? Some important hints can be found in a rather musty old (unpublished) paper by John Ohala (1972), written way back in those medieval times when duplication was still being performed by means of 'ditto machines'. In the context of trying to find some more satisfactory explanation for certain morphological regularities in English than that prescribed in classical generative phonology [CGP] (as codified especially in Chomsky & Halle 1968), Ohala lists several important differences in the requirements of the kind of 'independent phonological rules' that characterized CGP and those of a theoretical alternative based on what he called 'analogical rules', but which we shall simply characterize as the 'analogical approach'.¹ With his permission, we summarize a few of these below, with some re-organization and other minor modifications imposed:

Independent-Rules Approach

1. Requires fewer items in basic lexical store, e.g., abstract underlying forms for individual morphemes.
2. Less need for rapid search of lexicon.
3. Requires considerable long-term storage of rules.
4. Computation of derived forms is rather complex, depending on particular derivation.
5. Inductive learning of such rules is extremely difficult, especially if highly abstract underlying forms are necessary.
6. Requires keeping all the lexical and phonological machinery carefully hidden from inspection by speaker's conscious mind.³

Analogical Approach

- 1a. Requires much greater number of items to be stored in lexicon, e.g., words.
- 2a. More need for rapid search of lexicon.
- 3a. No storage of rules. New 'rules' can be created on spur of the moment by reference to existing words, then forgotten.
- 4a. Computation is relatively simple and roughly equal in complexity for all derivations.²
- 5a. No rules to acquire *per se*, but rather building up a network of lexical connections.
- 6a. No such secrecy required; phonological contents of lexicon subject to introspection.

One thing that is readily apparent from the list of characteristics summarized here is that some rather fundamental trade-offs are involved. In just those areas (items 1 and 2 above) where the analogical approach is seemingly cumbersome and complex (a very large lexicon and a consequent need for a very rapid lexical search mechanism), the rule-based approach is relatively 'simple' by comparison (a much smaller lexicon to be stored and searched through); in other areas (items 3 and 4), however, it is the rule-based approach which is encumbered with the excess machinery (rules to store and complex derivations to run through before words become available for use) in comparison with the analogical approach (no rules *per se* to worry about, with direct access to words possible in retrieval). Clearly, both classes of theories are burdened with formidable learning problems (item 5): a large set of intricately inter-related rules and (at least in CGP) an abstract, bare-bones lexicon in one case vs. an immense lexicon of systematically interconnected

whole words in the other), i.e., neither side takes language acquisition to be a piece of cake. Finally (in 6), though the independent-rules approach seems to be the evident loser, we dare not make too much of the contrast at this point, knowing as little as we do about the mechanisms or limits of introspection in general. (Nonetheless, we are faced with at least a minor intellectual puzzle of sorts: if the trappings of the grammar are indeed as varied and complex as contemporary theories suggest, why is it that no one but a few professional linguists seems to have even the slightest inkling of any of it?)

On balance, though, based, at least, on the present discussion, we can find no compelling reason for choosing between these two fundamentally different alternative approaches on *a priori* grounds. Both have their advantages and both have their disadvantages. For lack of attention in recent decades, of course, the analogical approach is at a very large disadvantage from the standpoint of detailed theoretical development (a shortcoming that all of the major figures involved would, no doubt, readily admit). But no such short-term deficit is any clear indicator of long-term benefits or survivability. In any event, as in all areas of serious scientific inquiry, theoretical choices must always be made in the light of available empirical evidence, and in this connection we can report on a fairly considerable body of psycholinguistic data that seems to bear on the main issue at hand.

3. Some Psycholinguistic Indicators

Through a series of ingenious experiments performed about a decade ago, Robert Stanners and his collaborators at Oklahoma State University made extensive headway in clarifying some key questions concerning the nature of the lexicon and the place of morphology in it. These studies all exploited the phenomenon of repetition or identity priming, an effect noted some years earlier in connection with some independently motivated studies involving a yes-no lexical decision task (Forback, Stanners & Hochhaus 1974). In the lexical decision task itself, subjects had merely to decide as quickly as possible whether a CRT-presented letter string was or was not a real word, and the priming effect might be informally described as a kind of sled-greasing phenomenon: if the same word was presented twice on the same series of trials, decision times were substantially reduced on the second presentation, even if the two presentations were separated by as long as 10 minutes or by as many as 36 intervening items. In the later work of Stanners *et al.* (1979b), this effect was adapted to the investigation of whether or not morphologically complex words (such as *discomfort* or *unaware*) were represented in the mental lexicon as unitary wholes, or whether their morphological constituents were represented separately and the full

words synthesized by rule, as needed. The authors reasoned that if the latter were the case for the word *discomfort*, for example, the prior presentation of other words containing its presumed parts (i.e., the base word *comfort* and some other word containing the prefix *dis-*, such as *disarm* or *disfavor*) ought to function just as effectively as a prime as the prior presentation of the whole word *discomfort* itself. They found, however, that this was not the case: though the word 'fragments' yielded a significant priming effect in comparison with the unprimed response, the effect was also significantly weaker than when the whole word prime was used. (Moreover, something the authors do not specifically note, the response patterns for clear-cut 'free root' cases, such as the two already illustrated, were very much the same as for the more problematic 'bound root' cases that they also investigated, as illustrated by such words as *retrieve* and *progress*.) Such evidence supports what Butterworth (1983) has called the 'Full Listing Hypothesis' (FLH), that is, while not all word-forms necessarily appear in the mental lexicons of speakers (see Stemberger & MacWhinney 1988 for some evidence that bears on this point), the forms that do appear there are represented as whole words, not as disembodied parts.

By much the same token, in other studies (Stanners *et al.* 1979a, Kempley & Morton 1982; but cf. Fowler, Napps & Feldman 1985), some indications were found of significant differences in the strength of these secondary priming effects, depending on the type of morphological relationship involved: best for inflectional variants (as when priming the base word *sing* with *sings* or *singing*), intermediate for derivatives (as when priming with *singer*), and weakest for any irregular variants that the base word might have (as when priming with *sung*). Therefore, as Cutler sums up in her own survey of this and other evidence (1983:58), a picture of the mental lexicon has emerged in which, at least for English,⁴ (a) words are represented in their full, 'undecomposed' forms and (b) the representations for morphologically related words are 'connected in some way'. Clearly, all of this is highly compatible with a general 'analogical network' type of framework, particularly one where morphological variants are connected, with varying degrees of strength, with a central basic or 'root' word.

It must now be pointed out that this quite considerable evidence in favor of a massive word-store network contrasts sharply with the total absence of evidence for anything like the generative notion of the morpheme-invariant underlying form (see Linell 1979 for an independent critique). It thus flies directly in the face of one of the fundamental assumptions of much of modern phonological theory (and certainly CGP), viz., that the chief function of rules is to simplify the lexicon. (As Chomsky & Halle put it, 'the lexicon specifies only

idiosyncratic features of lexical entries, omitting all those that can be determined by general rule' [1968:166]). Ironically, though this practice became so widespread as to constitute virtually the conventional wisdom in the field for two decades or more, no good argument, to our knowledge, has ever been given for adopting it in the first place. This is surprising, first of all, because it has long been recognized that the economies so achieved in the phonological domain 'cannot be extended to the semantic system, since the meaning of derived words frequently cannot be recovered from the meaning of their constituent parts' (Henderson 1985:223). Moreover, in the context of the 'computation' issue raised in item 4, it is clear that to choose lexical simplification as the be-all and end-all of linguistic methodological practice is to put a quite arbitrary premium on storage at the expense of ease of retrieval, since the less is stored, the more reconstructive computation has to be done prior to retrieval (Derwing 1973:154, n. 2 and 1988); in fact, as Aitchison emphasizes in her excellent survey (1987:9ff), there are very good reasons for thinking that it is considerations of retrieval that ought to be given priority concern, in view of the extremely rapid rates at which the processes of speech production and (especially) comprehension are normally carried out. (And, indeed, the general psycholinguistic finding is, consistent with all that has been said above, that 'morphologically complex words are ... no more difficult to access from the lexicon than morphologically simple words' [Cutler 1983:73].) All these things considered, therefore, the notion of 'rule' under discussion here seems to be scarcely tenable.

We cannot dismiss the entire rule-based approach in quite so quick and off-hand a manner, however, as the ill-begotten lexical simplification function is certainly not the only one that can be (or has been) conceived for rules. One somewhat more promising alternative was proposed, for example, by Vennemann (1974), who suggested that rules be viewed as learned generalizations about the lexicon and thus might serve not to simplify the lexical word-store, but rather help to organize it or give it structure — which is obviously something that is going to have to be done in one way or another, in any event, if efficient retrieval is ever to be achieved. Under this conception, the notion of rule is perfectly compatible, in principle, with the FLH and with all the other findings about the mental lexicon so far discussed (though no mechanism is as yet provided within this framework for the kind of 'network of morphological connections' that the data suggest is also required). The learning problem (item 5 above) would also appear to be considerably ameliorated under a conception of this sort, which lends itself quite naturally to what has come to be known as the 'true generalization condition' on the learnability of rules (Vennemann 1974; Hooper 1976). For contrary to many assertions (e.g., Anderson 1987:342), such a constraint is not

arbitrary but is instead consistent with the kind of general capacities that human beings are already known to possess, notably, the ability 'to extract regularity from the environment' (see Derwing 1973 for an extended argument, especially pp. 66, 200-201, 310). What is not so consistent — and which is why there is a 'serious learning problem' associated with the more abstract conception of rule — is the conception of language it entails that must presume an immense (but otherwise unmotivated) innate schematism simply to 'make possible' its acquisition by the child (Chomsky 1969:67).⁵

There are, however, other potentially serious difficulties associated with the notion of rule, even under the kind of radical reconception as the one just outlined. One of these is the following ontological problem: what kind of reality can be ascribed to a notion of rule whose mental existence is not open to introspection (item 6 above) and whose operations are ordered in non-real time (cf. Cook 1974). Linguistic models, we will recall, are not conceived as models of what speakers and hearers actually do (so-called 'performance models', often pejoratively described), but rather as models of what they (in some obscure 'implicit' sense) are presumed to 'know' (so-called 'competence models', which are characterized as somehow 'neutral' with respect to the processes of speech production or comprehension [Chomsky 1965 and elsewhere]).

And there is more. For in addition to the contrasts (independently) formulated by Ohala (1972), as already noted, Skousen (*Analogical modeling of language* [AML], in press) lists some others of even more potentially serious import:

Independent-Rules Approach

7. Contextual space is partitioned into well-defined rule contexts.
8. Transitions in behavior are sharp and precise.
9. Usage: find the correct rule that applies to the given context.
10. Usage is a function of the description.

Analogical Approach

- 7a. Contextual space remains atomistic.
- 8a. Transitions in behavior are gradual and fuzzy.
- 9a. Usage: find an appropriate example to model behavior after.
- 10a. Usage is the description.

As noted in 7 and 8 above, the kind of rules that appear in formal grammars explicitly and sharply demarcate the conceptual space. Though some attempts have been made to temper this rather harsh property (as through the *ad hoc* device of the so-called 'variable rule' [Labov 1970; Cedergren & Sankoff 1974]), in its classic form a rule either applies to a representation or it does not; a particular

morphological construction is either regular or irregular; a sentence is either grammatical or ungrammatical; etc. As Hockett (1970), in particular, has emphasized, this has led to a conception of the grammar of a language (and hence of the language that such a grammar generates) as a 'well-defined system', i.e., one that can be 'completely and exactly characterized by deterministic functions' (p. 45). Hockett then proceeds to spend most of the rest of his book arguing, on the basis of a host of examples, that natural (i.e., human) language is just not like that, but is rather a much more flexible, open and ill-defined thing, with analogy, once again, seen as the primary creative mechanism.⁶ By much the same token, Bolinger, in a particularly delightful as well as compelling article (1974), argues not only that the edifice of language 'has more patching and gluing about it than it has architectonics' (p. 1), but also (as an implicit commentary on what we have already said about the nature of the mental lexicon) that 'the human mind is [perhaps] less remarkable for its creativity than for the fact that it remembers everything' (p. 2) — and that, even in syntax, 'idiomaticity is a vastly more pervasive phenomenon than we ever imagined' (p. 3).

4. On Formalizing the Notion of Analogy

We could cite all of the examples provided in these and numerous other sources, however, and still fail utterly to convince — the main reason, of course, being that, until quite recently, at least, the principle of analogy suffered from a single, but ultimately fatal, fault of its own: it was too flexible, too open-ended and too ill-defined to serve as a suitable explanatory vehicle for what was genuinely productive in language, as it failed to exclude endless other creative productions that occurred only very rarely, if at all, as Kiparsky (1975:188) has properly emphasized.

The main problem with the traditional notion of analogy, in other words, was that there was no limit to its use: almost any form could be used to explain the behavior of another form, provided there was some similarity, however meager, between the two forms. But the problem is not 'inherent' in the notion in principle, as Kiparsky thought (*loc. cit.*), but is rather a consequence of the informality with which the notion has tended in the past to be characterized. By contrast, Skousen (AML) has recently presented a detailed and explicit definition of analogy which seeks to overcome this problem in a principled and well-motivated way; specifically, the range of search is sharply delimited through explicit characterization of a notion of *analogical set* for a given context. Though technical details are too complex to be treated in a satisfactory way here, the basic principles involved are those of *supracontextual homogeneity* and *random selection*, meaning that, if the given context does not lead to single, definitive solution in the lexicon, a range of surrounding

supracontexts is explored until a point of supracontextual *heterogeneity*, explicitly defined, is reached; a random choice is then made from among the set of possible analogical examples made available by the search.

To illustrate these principles with concrete examples, Skousen presents a number of detailed cases, including (1) leakage in the direction from *an* to *a* in children's use of the English indefinite article, (2) the spelling of /h/-initial words by English-speaking adults, (3) VOT crossover data for the English bilabial stop phonemes /p/ and /b/, (4) terms of address in colloquial Egyptian Arabic, and (5), the most extensive example, dialectal variation (and historical changes) involving the Finnish past tense forms. In all these examples, speaker variability is the rule rather than the exception, and the model in each case makes a set of explicit quantitative predictions about expected outcomes. To take just one specific illustration here, his model predicts three possible outcomes for the past tense forms of the three Finnish verbs shown in Table 1, viz., either to replace the stem-final vowel by /-i/ (=V-i in the table), or by /-oi/ (=a-oi), or else to replace the stem-final dental stop+vowel sequence by the suffix /-si/ (=tV-si). For speakers who have not fixed a past tense for such low frequency verbs (e.g., children), or who may have forgotten them (cf. the English *succumb* example used at the start of this paper), the predicted relative probabilities of the expected outcomes are also shown (as percentages).

Table 1. Outcomes for Three Infrequent Finnish Verbs

1. Verbs in -'aV[+son]ta		<u>Pr(V-i)</u>	<u>Pr(a-oi)</u>	<u>Pr(tV-si)</u>
kaarta-	'swerve'	0	48.6	51.4
saarta-	'surround'	0.1	41.9	57.9
	Average	0.1	45.3	54.7
2. Verb in -'aVta		<u>Pr(V-i)</u>	<u>Pr(a-oi)</u>	<u>Pr(tV-si)</u>
raata-	'toil'	0	99.6	0.4

Note especially that even though all these verbs can, in theory, take all three outcomes, Skousen's model predicts a clear difference in behavior between those verbs that have a sonorant immediately preceding the /t/ and those that don't.⁷ Quantitative predictions of this kind have the great advantage, of course, that they can, in principle, be tested against the behavior of real speakers and learners of the language, so we need not wait forever to see what, if anything, this particular model has going for it.

One brief comparison might usefully also be made at this point between Skousen's model and the competing 'connectionist' or

'interactive activation model' of the PDP group, which can also be characterized as a variant of the general analogical approach, in a broad sense (e.g., both share the key property of dispensing with independent rules in favor of some notion of a network of connections among forms). An important difference between the two sub-approaches, however, is that a connectionist model does not make available an alternative set of outcomes to choose from, but predicts behavior by having the various possibilities compete with one another until stability is reached and a single, preferred outcome emerges. One empirical difficulty with this design feature (see Skousen's AML for others) is that it cannot readily mirror the ability that speakers have to choose an alternative outcome when additional input indicates that the original choice of outcome is wrong. Thus, for instance, when confronted with the nonce spelling YEAD, speakers typically suggest a pronunciation (usually /yid/) with no difficulty. If told that their first pronunciation is incorrect, these speakers also have no difficulty in producing an alternative (such as /yed/). There is also evidence that speakers can quickly alternate from one outcome to another, especially when one or more of the resulting choices is considered strange by the speaker. (Skousen gives an example of a child, 5 years and 10 months of age, looking at a picture of the Grand Canyon and making a rapid-fire series of attempts to pluralize the word *cliff*: /klɪftɪz/, /klɪfs/, /klɪvz/, /klɪfs/....'). In Skousen's approach the rules of usage can be readily extended to find alternative outcomes when the first outcome chosen is rejected for some reason, as all possible outcomes are readily available for inspection in the analogical set defined for the context in question. In order to model this kind of ability using a connectionist approach, however, the connections pointing to the first outcome chosen would presumably have to be deactivated or momentarily disconnected and the entire system allowed to re-establish stability on an entirely new basis. This seems both cumbersome and implausible as a practical procedure.

But to return to our main theme, we can press on to the usage issue (items 9-10 above), where it becomes immediately clear that the analogy model-builders are all involved in the construction of models of real-time mental activities, i.e., performance models of real language users. This has two immediate and important advantages. For one, an explanation of the mystery of the 'missing intuitions' is readily available, as nothing more than a specific example of the general, classical distinction between procedural knowledge ('knowledge how') and declarative knowledge ('knowledge that'). Rumelhart has put it this way:

The knowledge that we have about language seems to be largely embedded in the procedures involved in the production and comprehension of linguistic utterances.

This is evidenced by the relative ease with which we perform these tasks when compared with our ability to explicate the knowledge involved in them (1979:2; see also Derwing 1973:251-258).

The second (and by far most important) advantage of modeling usage instead of merely describing forms, as we have already seen, is that the kind of quantitative claims that usage models make about linguistic behavior are straightforward and can be empirically tested. As has been widely recognized, this is not yet so for rule-based systems, which still largely lack a critical 'heuristic' component which might convert them into testable real-time models of linguistic performance (see Derwing 1973:259-296 for an extended critique). Since rule-based systems have been traditionally conceived as descriptions of forms, not of activities, they are not readily interpreted as real-time models of performance, and are interpreted only with great difficulty and uncertainty in any other psychologically relevant sense, as well (see Skousen 1979 and Derwing 1980). Until the requisite supplemental machinery is all in place, therefore, the place of so-called 'competence models' in general cognitive theory will remain very uncertain, at best, and little in the way of serious quantitative evaluation is likely to ensue. We thus expect that theoretical development of analogical models, now at last seriously underway, will be rapid, substantial and largely forward-moving, while progress with rule-based systems continues to advance mostly laterally, in response to what Ohala has aptly described as 'a kind of Brownian motion through the possible theoretical space' (1988:2).

5. Conclusions

To sum up: we have seen at least ten important empirical and conceptual differences between the rule-based and analogical approaches; these give the lie to the gratuitous suggestion by Pinker & Prince that the analogical approach — or at least the particular R&M version thereof — may, in the end, turn out to be 'nothing more than an implementation of a symbolic rule-based account' (1988:182). The differences between the two approaches are both substantial and fundamental; they also, incidentally, provide a potentially vast empirical ground on which the relative merits of the two approaches may, over time, be carefully and systematically weighed. While the evidence is not yet all in, by any means, we should not be blind to the fact that what weight of evidence we do have supports, quite overwhelmingly, the analogical approach, confirming Antilla's speculation that 'Memory or brain storage is on a much more extravagant scale than we would like to think; even the most "obvious" cases can be stored separately' (1972:349, cited in Bolinger 1974:2).⁸

To be sure, given the current 'mental set' of our discipline, news of

this sort is not likely to bring much joy to the hearts of a great many linguists. Even so staunch a 'live and let live' type as Fred Householder once wrote, in fact, that 'A linguist who could not devise a better grammar than is present in any speaker's brain ought to try another trade' (1966:100). But why? Whose grammars are learned or acquired by speakers, after all — linguists' grammars, or the grammars (if we may call them that) that are actually present in individual brains? And whose grammars, after all, actually play a role in the very real processes of language production and comprehension that real speakers and hearers actually engage in? And whose grammars go wrong in any of the various aphasic disorders? Surely, if linguists are content to write arbitrary grammars, based on arbitrary theories and judged by arbitrary decision criteria, they remove themselves by default from those very areas of explanation that most of them, we gather, have hoped might lift their field out of the backwater of its taxonomic past.

There is one final irony in all this that is perhaps also worth mentioning. Specifically, though the focus of this paper has been on the great chasm of contrast that separates rule-based theories from analogical ones, it seems that this very notion of a formal (generative) rule was itself based on the formulation of an analogy, namely, the analogy between the notion of a formal language (as known within automata theory in mathematics) and a so-called 'natural' language (as learned and manipulated by real language users; see Derwing 1973:284ff. for discussion). In other words, the rejection of the principle of analogy in general by nearly an entire generation of linguists has been based on the acceptance of one very specific analogy — and a false one, at that, if the evidence outlined above is anywhere near the mark. Once we have managed to get all of our analogies (as well as our priorities) straight, what we suspect is that some version of the analogical approach is likely to turn out to be the real answer to most of our problems, after all.

Notes

¹ Skousen (in press) uses the terms 'structural' vs. 'analogical' to distinguish these two fundamentally different approaches to language description.

² But see below for the problem of limiting the range of the search.

³ It is for this reason that R&M refer to the rule-based approach as 'the explicit inaccessible rule view' (1986:217).

⁴ This qualification is necessary and potentially important, as the vast bulk of research on which the FLH is based has been research on English, or at least on a small set of languages typologically very close to English; clearly, the situation is a deplorable one and needs to be corrected without delay (see Derwing 1988 for further discussion).

⁵ This situation is a major topic of concern in Derwing (1973), where it is discussed at length.

⁶ Ironically, one of the chief criticisms that Pinker & Prince (1988) raise against R&M's approach is that it fails to distinguish sharply between regular and irregular cases of the English past tense (p. 137), whereas the weight of evidence, of course, is that no such sharp boundary exists for actual learners or even adult speakers (see Derwing 1988).

⁷ Note also that the *Nykysoomen sanakirja* (Sademiemi 1973) list both *a-oi* and *tV-si* as possible variants for the first two verbs, but only *a-oi* for the third, *raata-*.

⁸ Cf. also Hunt (1982), who sees analogy as the fundamental cognitive strategy.

References

- Aitchison, J. 1987. *Words in the mind*. Oxford: Basil Blackwell.
- Anderson, S.R. 1987. *Phonology in the twentieth century: theories of rules and theories of representations*. Chicago: University of Chicago Press.
- Antilla, R. 1972. *An introduction to historical and comparative linguistics*. New York: Macmillan.
- Bloomfield, L. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bolinger, D. 1974. Meaning and memory. *Forum Linguisticum* 1: 1-14.
- Butterworth, B. 1983. Lexical representation. In B. Butterworth (ed.), *Language production*. Vol. 2. New York: Academic Press.
- Bybee, J.L. 1988. Morphology as lexical organization. In M. Hammond & M. Noonan (eds.), *Theoretical morphology: approaches in modern linguistics*, 119-141. San Diego: Academic Press.
- Cedergren, H.J. & D. Sankoff. 1974. Variable rules: performance as a statistical reflection of competence. In D. Sankoff (Ed.), *Linguistic variation: models and methods*, 85-95. New York: Academic Press.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1969. Linguistics and philosophy. In S. Hook (ed.), *Language and Philosophy*, 51-94. New York: New York University Press.
- Chomsky, N. & M. Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Cook, V.J. 1974. Is explanatory adequacy adequate? *Linguistics* 133: 21-31.
- Cutler, A. 1983. Lexical complexity and sentence processing. In G.B. Flores d'Arcais & R.J. Jarvella (eds.), *The process of language*

- understanding*, 43-79. New York: Wiley.
- Derwing, B. L. 1973. *Transformational grammar as a theory of language acquisition: a study in the empirical, conceptual and methodological foundations of contemporary linguistics*. Cambridge: Cambridge University Press.
- Derwing, B. L. 1976. What kind of rules can children learn? In W. von Raffler-Engel, & Y. Lebrun, eds., *Baby talk and infant speech*, 68-78. Amsterdam: Swets & Zeitlinger.
- Derwing B. L. 1980. English pluralization: a testing ground for rule evaluation. In G. D. Prideaux, B. L. Derwing & W. J. Baker (eds.), *Experimental linguistics: integration of theories and applications*, 81-112. Ghent: E. Story-Scientia.
- Derwing, B.L. 1988. Some characteristics of the mental lexicon: psycholinguistic evidence. Paper presented at the 3rd International Morphology Meeting, Krems, Austria. To appear in W. Dressler (ed.), *Morphologica 1988*.
- Forback, G.B, R.F. Stanners & L. Hochhaus. 1974. Repetition and practice effects in a lexical decision task. *Memory and Cognition* 2:337-339.
- Fowler, C. A., S. E. Napps & L. Feldman. 1985. Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition* 13: 241-255.
- Henderson, L. 1985. Issues in the modelling of pronunciation assembly in normal reading. In K.E. Patterson, J.C. Marshall & M. Coltheart (eds.), *Surface dyslexia*, 459-508. London: Erlbaum.
- Hockett, C.F. 1970. *The state of the art*. The Hague: Mouton.
- Hooper, J. Bybee. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
- Householder, F.W. 1966. Phonological theory: a brief comment. *Journal of Linguistics* 2: 99-100.
- Householder, F.W. 1971. *Linguistic speculations*. Cambridge: Cambridge University Press.
- Hunt, M. 1982. *The universe within: a new science explores the human mind*. New York: Simon & Schuster.
- Jespersen, J.O.H. 1922. *Language: its nature, development, and origin*. London: G Allen & Unwin.
- Karlsson, F. 1986. Frequency considerations in morphology. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung (ZPSK)* 39: 19-28.
- Kempler, S.T. & J. Morton. 1982. The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology* 73: 441-454.
- Kiparsky, P. 1975. What are phonological theories about? In D. Cohen & J.R. Wirth (eds.), *Testing linguistic hypotheses*, 187-209. New York: Wiley.

- Labov, W. 1970. The study of language in its social context. *Studium Generale* 23:30-87.
- Linell, P. 1979. *Psychological reality in phonology: a theoretical study*. Cambridge: Cambridge University Press.
- Ohala, J.J. 1972. On the design of phonological experiments. Paper read at the annual meeting of the Linguistic Society of America, Atlanta.
- Ohala, J.J. 1988. The cognitive cost of phonological analysis. *Berkeley Cognitive Science Report* #53.
- Paul, H. 1891. *Principles of the history of language*. Translated by A.H. Strong. London: Longmans, Green.
- Pinker, S. & A. Prince. 1988. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28:73-193.
- Rumelhart, D.E. 1979. Analogical processes and procedural representations. *CHIP* 81. LaJolla, CA: Center for Human Information Processing, University of California at San Diego.
- Rumelhart, D.E., J.L. McClelland & the PDP Research Group. 1986. On learning the past tenses of English verbs. In *Explorations in the microstructures of cognition, vol II, psychological and biological models*, 216-271. Cambridge, MA: MIT Press.
- Sadeniemi, M. (ed.). 1973. *Nykysuomen sanakirja (Dictionary of modern Finnish)*. Porvoo: Werner Söderström.
- Skousen R. 1979. Empirical interpretations of psychological reality. In E. Fischer-Jørgensen, J. Rischel & N. Thorson (eds.), *Proceedings of the Ninth International Congress of Phonetic Sciences*, Vol. 2, 121-128. Copenhagen: Institute of Phonetics, University of Copenhagen.
- Skousen, R. (in press). *Analogical modeling of language*. Dordrecht: Kluwer.
- Stanners, R.F., J.J. Neiser, W.P. Herson & R. Hall. 1979a. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18:399-412.
- Stanners, R.F., J.J. Neiser & S. Painton. 1979b. Memory representation for prefixed words. *Journal of Verbal Learning and Verbal Behavior* 18: 733-743.
- Stemberger, J.P. & B. MacWhinney. 1988. Are inflected forms stored in the lexicon? In M. Hammond & M. Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics*, 101-116. San Diego: Academic Press.
- Sturtevant, E.H. 1947. *An introduction to linguistic science*. New Haven: Yale University Press.
- Vennemann, T. 1974. Words and syllables in natural generative phonology. In *Papers from the parasession on natural phonology*, 346-374.