scope of what they describe, as the island). Take, for example,

(33) Jane's report that B: please Ralph

It is presupposed that one knows

The other interjections are accented
is stopping to rephrase or encasement
of oh and well) or if they echo

report contains the word oh, well
(This last interpretation would say.)

Another area in which interjections
ways has to do with preposed ad
paper.) Interjections are accented
for example,

(34) In 1979, \( \{ \text{oh, we'll, men} \} \) say, \( \{ \text{--ah!} \} \)

(35) On that spot, \( \{ \text{oh, well, say} \} \) say, \( \{ \text{--ah!} \} \)

but they are odd after preposed
example,

(36) Quietly, \( \{ \text{*oh, well, say} \} \) *Jac

(37) For some reason, \( \{ \text{*oh, say} \} \) ??
ABSTRACT

This is a state-of-the-art review of word formation morphology. The paper surveys three loosely knit 'schools' of word formation: (1) the Generative school, (2) the Continental school, and (3) the Slavist school. It points out that much work in word formation is being duplicated because of a lack of coordination and communication between the respective schools. The paper concludes with a list of the 10 most perplexing problems facing word formationists today: (1) semantic-syntactic asymmetry, (2) derivational gaps, (3) lexicalization, (4) semiproductivity, (5) rule order, (6) derivational direction or 'motivation,' (7) the apparently ineluctable ad hocness in the derivation of relative adjectives and some noun compounds, (8) the location of the word formation component in the grammar, (9) confusion of nongenerable and nonoccurring forms, and (10) the problem of currency (how to determine which of several generable forms with appropriate meanings will in fact gain currency). The problems are discussed in the text and associated with appropriate references in the bibliography. (Author)
WORD-FORMATION: THE ANARCHY OF THE ART*

by

Robert E. Beard

When Syntactic Structures appeared in 1957, it produced very little effect among linguists dealing with problems of word-formation, even though it presented a theory which promised to solve an old dilemma for them: the prediction of neologisms.

Even after the appearance of Lees' The Grammar of English Nominalizations in 1960, word-formationalists took scant note of the generative potential of its model. There remained problems of establishing and assigning productivity gradients, of the super-power of derivational transformations and the design of proper constraints, and there were formidable semantic problems to overcome. Moreover, most linguists involved in word-formation investigations were dealing either with diachronic questions and/or strictly morphological ones. Even though many such linguists, especially the Slavists, depended upon crude semantic categories for their affixal taxonomy, these categories were consistently assumed and never analyzed in detail. Nor was the question of the relationship of

*This is a revised version of a paper delivered at the 19th annual meeting of The International Linguistic Association, March 9-10, 1974, New York.
word-formation to the rest of grammar seriously attacked. Instead, "word-formation" was generally defined as a field of linguistics rather than a linguistic process itself.

In the past five years, interest in word-formation has perceptibly increased. There have been advances in adjacent areas of linguistics which have influenced and will continue to influence that interest: two new semantic theories (interpretive and generative), a case theory, a theory of constraints. The sheer number of linguists involved in word-formation studies is indeed increasing but the questions they have taken up have as usual turned out to be more complex than had been anticipated. While concrete progress has been made in the description of individual languages and the isolation of critical universal questions, not even the starting point for an integral theory can be agreed upon today.

Instead of concerted progress toward a unified theory, a survey of the literature reveals rather quickly that in the 60's and early 70's, work in word-formation is being conducted in three different areas of the world and until quite recently the cross-fertilization has been negligible. Even presently, the influence too frequently is uni-directional. I have in mind now three more-or-less groups of scholars who to varying degrees can be differentiated into what I will call (1) the Generative school, still most active in this country, (2) the Continental school, which probably centers around the seminal work of H. Marchand and his students, and which developed from the Geneva school of C. Bally and the Copenhagen school of L. Helmseljev and O. Jespersen, and (3) the Slavist school of V. V. Vinogradov which also developed from
Bally, though more from the Prague school of N. Trubetskoi and R. Jakobson. I will not consider separately S. Lamb's school of stratificational grammar, since it is less known for its contribution to word-formation methodology, but at least one suggestion along stratificational linguistic lines will be discussed.

The purpose of this paper, it must be emphasized, is not to categorize research being done and certainly not the linguists conducting the research. The use of the term "school" in this instance is intended as a purely heuristic device. The definition of these schools is obviously clumsy, since individually the linguists in them are free to move back and forth among the assumptions of all three schools thus blurring boundaries, and because all of the schools are dynamic and are themselves in constant organic flux. I do hope, however, to demonstrate that there is evidence to indicate that unnecessary duplication is occurring, identical derivational processes and categories are receiving different names from different scholars working on the same problem. In sum total it appears that while schools are discernable, the full argumentational advantages of separate schools is not being realized and at the additional cost of the loss of the symbiotic advantage of collective cooperation. The result is a mild state of anarchy, but an anarchy nonetheless. This paper is intended as a research synapse, a nexus of directions in word-formation to convince word-formationists of each other and each other's work.

D. Worth (1968: 403) and H. Marchand (1969: 31), of the Slavicist and Continental schools respectively, have recently come independently to the conclusion that the description of a morpheme
must always be tri-partite, comprising phonological, syntactic, and semantic descriptions. Both have observed also that most of the literature until recently has been morphological in nature without reference to semantics and seldom to syntax. This has been due to what Chafe (1970: 60) has described as the "phonetic bias which underlies both structuralism and syntacticism." The deeper research has carried us into the structure of morpheme combinations and the processes by which they synchronically and diachronically obtain, the clearer it has become that semantic features are vitally involved. However, since semantic theories which integrate into a unified theory of grammar have only recently become available, it is not surprising that work on problems of word-formation have been predominately taxonomic, based on morphophonemic analysis to the virtual exclusion of serious semantic considerations.

As a matter of fact, it was only recently established that word derivation is impossible without recourse to syntactic features. Nida's (1949) classic six principles which included the definition of the morpheme as the "minimal meaningful unit of language," all rely heavily on parallel semantic and phonological information (the former left undefined) in analyzing word-formation classes. Nida does recognize a significance in "structural series" and "structural classes" for the identification of derivational morphological classes, but nowhere does he suggest that features restricting morphemes to these classes and other subclasses will be required in the lexical

---

1. To be precise, Marchand uses the terms "morphological semantic, grammatical" while Worth also uses "grammatical" in place of my "syntactic." I intend them no misrepresentation in adapting their terminology to that of this paper, although slight variations may exist between their terminologies.
description of the morphemes. Due in large part to Nida's, Bloomfield's, and others' impact on descriptive morphology in the 40's and 50's, most research in word-formation even to the present has centered on descriptions of the affixational phonological taxonomy of derivational morphemes, rather than words in phrases (Halle: 1963, Worth: 1966b, 1966c, 1968). Those writers who do refer to broad semantic categories (Belić: 1949, Vinogradov: 1960, Marchand: 1969), neither closely define them nor apply them systematically in their methodology.

There is already evidence that the two descriptivist schools are expanding upon their base of massive taxonomic data in two directions: (1) toward generative, i.e., predictive rules and (2) toward rules that are amenable to relatively precise semantic descriptions. In his most recent investigations, Worth (1972) has begun to raise semantic questions in connection with his generative rules for Russian word-formation. Marchand protests that his new interest in the syntactical origins of neologisms is unrelated to advances in transformational generative grammar, but is instead a reflex of Bally's (1944: 102) dictum that "a morphological syntagma is nothing but a reduced form of an explicit syntagma, the sentence" (Marchand: 1966, p. 133). Without a precise definition of what the process of "reduction" is in this approach, one must assume that it is some sort of transformation.

More interesting in the light of advances in the Continental school, however, are the recent works of Marchand's students, especially R. Brekle (1966, 1970) and L. Lipka (1971, 1972), who are attempting to adapt Marchand's insights to generative semantics.
Brekle's article is an attempt at a semantic description of the compound nouns whose structural description Marchand has so thoroughly perfected. In his book, Brekle attempts to develop a generative "sentence semantic" system to accompany a transformational grammar of compounds. The system, adapting the terminology and technique of symbolic logic, predicts the semantic classes of derived compounds on the basis of their component provenient forms. Lipka (1972) claims the same sort of generative semantic approach as Brekle. He posits 10 T-rules which purport to derive denominal and deverbal verb-particle constructions, but the rules are incomplete and without sufficient constraints. More interesting is his article (1971) which examines in detail the problems of compound and relative adjective derivation.

These several examples are not meant to imply that all word-formationalists of what I have been calling the Continental and Slavicist schools are expanding the scope of their rules in the direction of greater abstraction. Most indeed continue to restrict their aspirations to the description of a predetermined sampling of words, usually selected from one or more dictionaries, without reference to syntax or semantics. Since Marchand and Worth have been leaders in their field, however, one is justified in suspecting that their attitudes represent an avant-garde in their respective schools and a dissatisfaction with those strictures of the structuralist approach which have been discussed in detail elsewhere. I will simply re-enumerate them here.

The first of the strictures of structuralism results from the basically taxonomy-oriented methodology which does not provide
a means of explaining (predicting) how word-formation operates to generate neologisms. Nowhere in the grammar of a language is it more evident than in word-formation that there are more potential linguistic forms than there are forms already in usage. The classification of derived forms has in no case resulted in the predictability of those forms in their morphonemic, syntactic and semantic totality nor provided a method for doing so. Structuralism also fails to place word formation into any unified theory of language. Marchand (1969: 2) defines word-formation as a field of linguistics rather than a process of language. Dokulil (1962: 7) does the same, and although Vinogradov (1960: 15) claims that it is a process, his methodology differs little from that of other Slavicists. This defect is well-known and related to the wider criticism of structuralism's being an arbitrary organization of data from the smallest to the largest units rather than an attempt to reflect the actual processes of language, i.e. from meaning to sound and vice versa.

A third criticism is that structuralist approaches are incapable of defining the syntactic or semantic relationships between derived forms and the sentential context in which they occur. Certainly there is no way in which we can derive one adjective budgetary from a noun budget so that its relationships to all the possible nouns which it may determine are specified, e.g. budgetary item, budgetary director, budgetary year, budgetary crisis, budgetary matter, etc. Secondary lexical items must be derived in context not only to predict the semantic reading of relative adjectives like budgetary, but for a host of other reasons as well.
When R. B. Lees' *The Grammar of English Nominalizations* appeared in 1960 it promised a generative method of word-formation. With the emergence of the Fodor-Katz-Postal interpretive semantic model in 1963-1964, the theoretical solution to the semantic problem seemed just beyond the horizon to some. Taken together, these two methods offered a means to generate productive derived forms from underlying phrase structures which had already been assigned the correct semantic interpretation in their most extended form. The transformation of such an underlying phrase structure into a neologism would alter only the structure of the phrase, not the semantic interpretation. Thus the morphophonemics of the derived word could be generated and the meaning would already be projected onto it.

After applying this model for a decade, it has become increasingly evident that not only does it fall short of providing a vehicle for clear and precise syntactic-semantic predictions, the relation of word-formation to grammar as a whole remains unclear. Especially critical is the question of its relationship to the lexicon. After much criticism and discussion (cf. Bolinger: 1968; Dik: 1967; Marchand: 1965a, 1965b; Rohrer: 1966; Schacter: 1962; Winter: 1965), Chapin (1967) and Chomsky (1970) sharpened the criticism and Chomsky proposed a "new" model, the so-called "lexicalist" model.

The shortcomings of Lees' model as Chomsky sees them are as follows. First, the rules of the model are too strong since derivations are not as productive as, say, gerunds or participles. Word-formation via T-rules is faced with a complex system of constraints...
which must be accounted for in the lexical description of all
derivational morphemes and which will probably be local and ar-
bitrary rather than universal. Second, the generative powers of
transformations do not predict nor offer any real promise of being
capable of predicting lexicalized meanings of derived forms. Thus
words whose meanings have become lexicalized, certainly those with
idiomatized meanings, must be listed in the lexicon in any event.
There is reason to believe that their number is somewhere near the
number of derived words in usage.

One final problem which will face transformationalists (and
lexicalists) for some time to come is that of rule ordering. Chapin
(1970) deals with this problem as it pertains to English nouns and
adjectives. Generally speaking, the problem is one of cycling the
rules of derivation and/or suffixation so that derivations receiv-
ing more than one suffix will be derived with the proper set of
suffixes and meanings in the proper order. For example, in English,
"the -tion rule must apply to -ize derivatives, the -ize rule to
-al derivatives, and the -al to -tion derivatives, an untenable
situation to linear ordering" (Chapin: 1970, p. 59), e.g.
coeducationalization. To account for this Chapin offers the "epi-
cycle hypothesis" by which cycles would occur cyclically. Unfor-
tunately he leaves the hypothesis untested.

The lexicalist, on the other hand, would list all lexical items
in the lexicon, not marking them as to word class (N, V, Adj, ...).
Class would be assigned by the end nodes of the underived phrase
marker into which they are plugged by lexical rules. Schwartz (1970)
suggests special base rules of the sort A \(\rightarrow\) VP to accommodate even
derived words like *ocean-going* in the base component, with T-rules merely re-ordering the morphemes. Chomsky (1970) suggests this may be achieved via lexical rules. Crystal (1967) demonstrates how detailed typology may be avoided by using lexical features rather than word classes. All of these works taken together amount to a system of lexical insertion which could theoretically account for both lexical and derived stems within the base structure.

Still there are many problems which lexicalism as yet has not solved. One such problem falls between productive derivation and lexicalization: syntactic-semantic asymmetry.⁴ Botha (1968) deals with a variant of this problem in Afrikaans, Beard (1966) deals with it in Serbocroatian, and Makkai (1969) approaches it in Romance languages. Chomsky (1970) also raises it in mentioning the English denominals receiving -able (which is regularly assigned to deverbal derivations), e.g. *knowledgeable, honorable, reasonable*. While these adjectives seem semantically related to HAVING-type derivations (*grassy, furry, hairy*), their suffix relates them to ABLE-type suffixation. Botha suggests a separate "phonological dictionary" and "phonological matching rules" to solve the problem in Afrikaans, where the asymmetry is basically a phonological one. Beard proposes that the very separation of derivation from suffixation offers a resolution to the difficulty in Serbocroatian, where the problem is syntactic. Makkai would account for the morphological generalization of the occurrence of *venir* "to come" in words like *prevenir* "to warn" and *convenir* "to suit" without a symmetrically

⁴ The term is drawn from Bazell (1949; 1954).
concomitant semantic generalization, by mapping out structure and meaning separately. Since her model is stratificational, however, she feels no need to provide a set of mapping rules which relate meaning to sound, only an outline of relations.

Another problem mentioned by Chapin and Chomsky is that of ostensibly derived words with no obvious provenience, e.g. portable, legible, possible. All of the examples they offer are qualitative adjectives and are easily accounted for as lexical items. There are, on the other hand, underived relative adjectives such as civil (rights, war, defense) which present the same semantic problems as budgetary, but preclude any derivational provenience in context which might predict those senses. Still it is the lexicalist who faces the greater complex of problems in examining relative adjectives. Since the meaning of these adjectives vary with the noun they determine, the lexicalist must somehow cross-reference every lexical relative adjective and noun from which a relative adjective may be derived, with every other noun in the lexicon with which it may occur in order to capture even a general field of semantic readings. It may be the case, however, that the problem of relative adjectives is the same as the semantic problem of compound nouns, in which case the solution of one is a solution of the other—some if slight consolation.

The strict lexicalist would list even derivations in the lexicon along with lexical stems. Such an approach theoretically avoids the semantic problems of derivation, but only at the expense

---

1. I am grateful to Fred Householder for pointing out this difference and the difficulty it implies.
of losing wide generalizations not to mention generativity. The only work thus far that professes to present a strong lexical model is Stockwell, et al. (1973). Although P. Schacter, who designed the lexicon in this work, apologizes for its lack of sophistication on the grounds that it is experimental, one must admit in the end that the experiment fails insofar as it deals with word-formation. For example, one finds in this lexicon employ, employer, employee, employment as four separate entries. In addition to the redundancy of such a model, there seems to be a contradiction of observable fact inherent in it. For instance, Langacker's (1972: 41) admittedly exceptional example can be explained only in terms of a partial lexical retrieval:

"He is my employ...ER."
"Your employed or employee?"
"EE."

Such evidence would indicate that stem and suffixes are listed separately in the lexicon and, in exceptional cases, the rules of bound and free morphemes can be violated.

An alternative is outlined by Đurovič (1972). Đurovič describes a maximally determined lexical entry whose derivations are to be predictable despite "derivational gaps," i.e., derivations in a given family which do not occur or seem possible despite the absence of any definable constraints. He proposes that a "derivational microstructure" be determined "previously" (presumably in the lexicon) by "semantic differential features" in order to generate word derivations and predict gaps. Such features would be
extensions of Worth's (1966b) concept of a "derivational base stem" and would determine the full extent of a given stem's ability to receive suffixes. Unfortunately, he restricted his study to those forms available in dictionaries, that is to say, to occurrence rather than to generability.

A real contribution to lexicalist derivation has been made by Dokulil (1962). He places derivation partly in the lexicon and partly in morphology, though in a fashion not altogether clear. He has much advanced the science with his detailed development of "onomasiological categories" and his discussions of problems of determining the "motivating" or underlying form of a derivation. Similar in his placement of the derivational process among the other processes of grammar is Shaumyan (1965). Shaumyan posits a separate "word generator" in his applicative transformational model between his semantic "abstract generator" and the "sentence generator." The separation of word-formation from sentence-formation has also been suggested by Reibel (1963). His "feedback control" is designed to lexicalize derivations, but may also be the location of the derivational rules themselves.

Shaumyan, Apresyan, and Soboleva were the first to propose a workable generative semantic model. Although Shaumyan hedges considerably on his description of the semantic component of that model (his "abstract generator" is a "syntactic-semiotic system"), Apresyan (1967) has strengthened it considerably. Brekle's (1970) adaptation of logical positivist methodology parallels Apresyan's in many respects, yet it would appear the former was unaware of the latter's work.
Another major generative semanticist contribution in the struggle to develop a workable theory of word-formation was made by Chafe (1970: esp. 119ff). Chafe derives all forms via generative semantic rules such as this state-to-process rule:

\[(\text{S11-3}) \quad \text{state} \rightarrow \text{process} \quad \text{root} \rightarrow \text{root + inchoative}\]

The unit inchoative will appear in the surface structure in various forms, e.g. -en (wide\(\rightarrow\) widen), morphophonemic alternations (hot \(\rightarrow\) heat), sometimes as \(\emptyset\) (open \(\rightarrow\) open), sometimes as a "post-semantic development" (tired \(\rightarrow\) tire). The last example suggests that Chafe faces a problem of establishing motivation, but then his system remains in a very tentative state of development (see also Armstrong: 1973; van Schooneveld: 1974). A very promising direction for word-formation investigations is Chafe's generative semantic model combined with Dokulil's advances in sorting out onomasiological categories.

One final contribution should not be overlooked. The distinction between occurrence and generability is one all word-formationists have great difficulty establishing. Gleitman and Gleitman (1970) have made an initial step toward not only establishing a method for making that distinction but for measuring generability empirically. Although their objective is to demonstrate the extent paraphrases of derivations are recoverable, their greatest contribution may be their empirical model and the initial insights it provides into the problematics of separating judgements of grammaticality from secondary social and psychological interferences in the subjects.
An adequate theory of word-formation must explain how syntactic forms and semantic value come to be and come to be associated, as well as where the rules which generate them are located. At least three classes of sub-rebularity impede any semantic theory presently: (1) semantic-syntactic asymmetry, (2) derivational gaps and (3) lexicalization. Even generating the syntactic forms of neologisms founders on the demands of (4) constraining rule strength, (5) establishing rule order, (6) certifying the direction of "motivation," (7) avoiding ad hocness in the derivation of relative adjectives and noun compounds. Finally, there remain vaguenesses around the periphery of the word-formation component which prohibit its clear delineation and definition. (8) It is still unclear where this component is located a propos the grammar as a whole. There remains the question of "currency" with its two facets: (9) how does one empirically distinguish between a non-generable form and one which is simply non-occurring, and (10) how does one predict from the multiple, feasible forms generable at a given moment, the one which will actually gain currency in the society (be inserted into the lexicon of the language).

Despite the formidable problems we face, there remains three distinguishable schools of scholarship and the flow of information between them is erratic and uni-directional where it exists. The Slavicist Đurović works on a lexicalist assumption without reference to the generative grammarians struggling with the same problem. Chafe does not avail himself of Dokulil's onomasiological research even though it complements his theory. Makkai demonstrates the potential contribution of stratificational grammar which apparently
everyone has overlooked. Shaumyan continues to work pretty much unattended on a separate word generator. Brekle duplicates much of Apresyan's ground-breaking work in adapting symbolic logic to semantics. Of course, it must be kept in mind that most of the developments in the field have occurred swiftly—over the past 5-6 years. Nonetheless, it would seem that 20th century communications could serve us better.

One of the most telling areas in which the lack of cross-fertilization has resulted in excessive overlap and redundancy is one in which perhaps some accomplishments have recently been achieved: the study of relative adjectives and nominal compounds. Levi (1973) has produced an ingenious set of transformationalist arguments which distinguish "bona fide" adjectives from what "for want of a more inspired term" she calls "denominal non-predicate" adjectives. Although they are never defined, it appears that "bona fide" adjectives are the "qualitative" adjectives of C. Bally, which Slavicists have been studying for most of this century. Levi's "denominal non-predicate" adjectives are in fact relative adjectives. Levi goes on to credit Postal with several tests of adj ectivality which Vinogradov (1960: 301-0) applied to Russian.

The "anarchy" in word-formation study today is not simply explorations of basic questions in various directions. This is certainly to be expected so long as there is no consensus even as to where word-formation rules are located in the grammar. The anarchy of which I speak is a disassociativeness of these explorations which can and should be surmounted in favor of a more productive and beneficial mutuality.
SELECTED BIBLIOGRAPHY ON WORD-FORMATION


Bally, C. 1944. Linguistique générale et linguistique française, 2nd ed. Berne


1952. "Phoneme and morpheme analysis." Word 8: 33-8


Belić, A. 1949. Savremeni srpskohrvatski književni jezik, II. deo. Belgrade


Crystal, D. 1967. "Word classes in English." Lingua 17: 24-56


Henzen, W. 1957. *Deutsche Wortbildung,* 2nd ed. Tübingen


Jespersen, 0. 1909-1949. *A Modern English Grammar on Historical Principles,* VII. London/Copenhagen


Koziol, H. 1937. *Handbuch der englischen Wortbildunglehre.* Heidelberg


Lakoff, G. 1968. "Instrumental adverbs and the concept of deep structure." Foundations of Language 4: 4-29


Marchand, H. 1951. "Phonology, morphology and word-formation." Neuphilologische Zeitschrift 52: 87-95


----------- 1965b. "On the analysis of substantive compounds and suffixal derivatives not containing a verbal element." Indogermanische Forschungen 70: 117-45

----------- 1966. "On attributive and predicative derived adjectives and some problems related to the distinction." Anglia 84: 131-49

-----------


-----------


-----------


-----------


-----------


Nickel, G. 1968. "The contrast 'He is reading' — 'He is interesting' and related problems," in Brekle & Lipka (1968), pp. 150-9


-----------


Šanskij, N. M. 1953. *Osnovy slovoobrazovatel'nogo analiza.* Moscow

1968. *Očerki po russkom slovoobrazovaniju.* Moscow

1971. *V mire slov.* Moscow


Winter, W. 1965. "Transforms without kernels?" Language 41: 484-9


