

# **Twentieth-century structural linguistics in America\***

**D. Terence Langendoen, Brooklyn College & CUNY Graduate Center**

**Presentation at the Columbia University Seminar on the History and Philosophy of Science**

**April 1983**

I don't need to remind you that the theory and practice of linguistics has undergone a series of radical and convulsive changes over the past eighty years. In the United States the first important development was the creation of a discipline devoted to the problem of describing the indigenous languages of North America, a feat that was accomplished largely by Franz Boas, Alfred Kroeber, Edward Sapir and a few others. Out of this work, together with work done by Leonard Bloomfield and others trained in the history and structure of the major languages of Europe and Asia, there began to grow a body of literature concerned with developing models for the description of languages in general, and for codifying procedures to obtain such descriptions. One of the earliest and most influential models was outlined in Bloomfield's "A set of postulates for the science of language", published in 1926. As the title suggests, Bloomfield proposed a hypothetico-deductive model in which all of the elements of a language, defined as "the totality of the utterances that can be made in a speech-community" are analyzed in terms of a hierarchy of smaller elements: phrases, words and morphemes. The elements, called 'forms', are all said to have 'meanings'. The most elementary meanings of a language are those associated with morphemes. On the other hand, morphemes can be analyzed into smaller phonological units, called "phonemes", which admit of no further analysis. Thus: phonemes combined to form morphemes. Associated with each morpheme is a unit of meaning. Morphemes combine phonologically and semantically to form words; words combine phonologically and semantically to form phrases. Phrases in turn combine to make larger forms; the maximal phrases or a language are called sentences. The number of forms in a language is assumed to be finite, but from this it does not follow that a language is a finite set of utterances, since utterances themselves are not defined as forms.

Bloomfield was not himself particularly interested in methodology, but many of his students and followers were, and procedures were soon developed to find or "discover" the various elements which combine to make up the forms of a language. These procedures all had in common the division of the stream of speech into successively smaller and smaller units (segmentation) and the grouping together of the units that occur at a given level of analysis (classification). The attempts to codify these procedures and the disputes that raged over those attempts dominated the discussions and publications of the period from roughly 1935 to 1955, and gave rise to the perception that linguistics, as practiced by the Bloomfieldian structuralists is a 'taxonomic' science (Joos 1957).

The publication of Noam Chomsky's *Syntactic Structures* in 1957 and the nearly simultaneous appearance of Robert Lees's review in *Language* put an abrupt end to such methodological concerns. Chomsky wrote:

As I interpret most of the more careful proposals for the development of linguistic theory, ... they attempt to state methods of analysis that an investigator might actually use, if he had the time, to construct a grammar of a language directly from the raw data. I think that it is very questionable that this goal is attainable in any interesting way, and I suspect that any attempt to meet it will lead into a maze of more and more elaborate and complex analytic procedures that will fail to provide answers for many important questions about the nature of linguistic structure.

Moreover Lees argued, based in part on work by Morris Halle, that any grammar that emerges from such procedures will of necessity not be the simplest grammar of the language under description nor be capable of describing adequately the full formal structure of the language. The capitulation of post-Bloomfieldian structural linguistics to Chomsky's theory of transformational- generative grammar, which he developed as an alternative theory of linguistic form, was precipitous and nearly total. Contemporary challenges to the hegemony of generative grammar largely had not so much to do with issues involving the analysis of linguistic forms, but rather with the question of whether linguistics should be limited to the investigation of the formal properties of language.

Certainly if Bloomfieldian structuralism is, as Chomsky wrote in 1964, "a kind of pre-Darwinian taxonomy concerned solely with the collection and classification of countless specimens", it has little to recommend it. But this characterization of the work of that period is not accurate. Indeed, it has some features that, if anything, make it as attractive a theory of linguistic form as generative grammar as currently conceived. What is required is not a disparaging dismissal of Bloomfieldian structuralism, but a sympathetic reconstruction of it using the more sophisticated formal apparatus of logico-mathematical linguistics that has recently been developed as an adjunct to work in generative grammar.

In 1961 Charles Hockett provided a demonstration that Bloomfieldian structuralist theory is not a taxonomic one. He wrote:

The simplest and earliest assumption about the relation between morphemes and phonemes was that a morpheme is composed [emph. original] of phonemes: the morpheme *cat* is composed of the phonemes /k/, /æ/, /t/ in that arrangement.... This conception is satisfactory as long as we confine our attention to certain parts of the material in any language, but it breaks down when we are confronted with such phenomena as English *knife* : *kniv(s)* or *(dog)s* : *(cat)s* : *ros(es)*. The very criteria that lead us to reject a morphemic identification of *cat* and *cat(alog)* ... also lead us to want to call *knife* and *knive-*, or /z/, /s/, and /iz/, the same morpheme. Such identifications are incompatible with an acceptance of the 'composed of' relation between morphemes and phonemes.

As Hockett went on to point out, two different lines of attack on the problem had been developed within Bloomfieldian structuralism. According to the first approach, in addition to the phonemic system of a language, another more abstract system was posited, known as its morphophonemic system. For example, the English phonemic system contains /f/ and /v/ and no other labiodental fricatives, whereas the English morphophonemic system contains also the morphophoneme //F//, which occurs in the morpheme *knife/knive-*. That morpheme can be said to be composed of the

morphophonemes //n//, //a//, //y// and //F// in that arrangement, i.e. as //nayF//. According to the second approach, rather than adding a phonemic layer more abstract than the original one, one adds a morphemic layer less abstract than the original one, which may be called a system of morpheme alternants, or 'allomorphs'. For example, the English morpheme /nayf/ is represented by the two allomorphs |nayf| and |nayv|, each of which is composed of phonemes in the usual way.

Either way, however, a morpheme is not composed of an arrangement of phonemes. On the first approach, a morpheme is composed of an arrangement of morphophonemes, but a morphophoneme is not composed of an arrangement of phonemes; instead it is represented by particular phonemes, e.g. //f// is represented by /f/, //v// by /v/ and //F// by both /f/ and /v/. On the second approach, a morpheme alternant is composed of an arrangement of phonemes, but a morpheme is not composed of an arrangement of allomorphs; again it is represented by particular allomorphs, e.g. /fayf/ by |fayf| alone, but /nayf/ by both |nayf| and |nayv|.

Hockett (1961) diagrammed the relations between morphemes and phonemes in Bloomfieldian structuralism as in Figure 1, where *C* is the 'is composed of an arrangement of' and *R* is 'is represented by'.

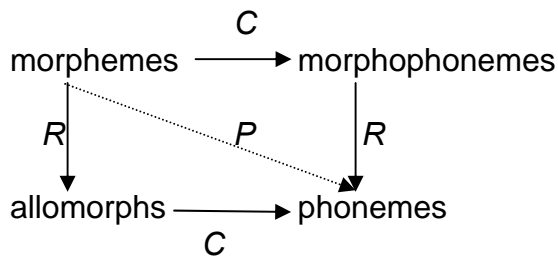


Figure 1

The direct (diagonal) relation between morphemes and phonemes is labeled *P* for 'is programmed into', 'is mapped into' or 'is realized as'. Hockett wrote: "In dealing with a particular language, it is the details of relation *P* that we are obliged to describe." Figure 1 shows two possible ways of decomposing *P* into the two relations *C* and *R*. It may be decomposed in other ways as well, as in an analysis that posits both allomorphs and morphophonemes, as in Figure 2, or not decomposed at all, as in an analysis that makes use of neither allomorphs nor morphophonemes, as in Figure 3.

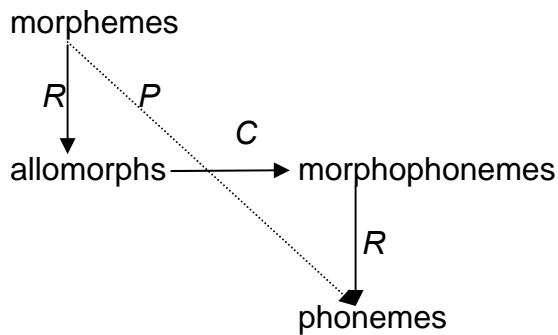


Figure 2

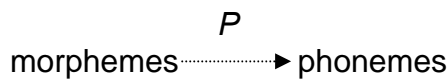


Figure 3

Of the three relations in Hockett's diagrams, only *C* is taxonomic. Since the overall relation *P* between morphemes and phonemes must be analyzed in terms of at least one non-taxonomic relation, it follows that Bloomfieldian structuralist theory cannot be viewed as totally taxonomic. But it is not only between morphemes and phonemes that a non-taxonomic relation obtains; in general, the mapping between any pair of contiguous levels in the hierarchy of Bloomfieldian structuralist theory is non-taxonomic in character. I conclude therefore that the characterization of that theory as taxonomic is mistaken.

The non-taxonomic character of the relation between morphemes and phonemes in Bloomfieldian structuralism is further revealed by its concern to further classify the phonological alternations of morphs into automatic and non-automatic, regular and irregular, word-internal and word-external, etc. For example, an alternation of phonemes is defined as automatic if it is conditioned solely by the phonemes in its environment. Thus the alternation between the allomorphs  $|nok|$  and  $|nokt|$  of the Latin morpheme  $/nokt/$ , exemplified in the words *nox* 'night' (nom. sg.) and *noctis* 'night' (gen. sg.) is automatic, as is the alternation between the allomorphs  $|rēk|$  and  $|rēg|$  of the Latin morpheme  $/rēg/$ , exemplified in the words *rēx* 'king' (nom. sg.) and *rēgis* 'king' (gen. sg.). However, to be able to analyze the alternation as automatic, we must posit  $//nokt//$  and  $//rēg//$  as the morphophonemic forms of these morphemes, since given the morphophonemic arrangements  $//nokt//+//s//$  and  $//rēg//+//s//$ , we automatically obtain the observed word forms (represented as arrangements of allomorphs)  $|nok|+|s|$  and  $|rēg|+|s|$ . On the other hand, given the morphophonemic arrangements  $//nok//+//is//$  and  $//rēk//+//is//$ , we do not automatically obtain  $|nokt|+|is|$  and  $|rēg|+|is|$ .

The need to distinguish a basic morphophonemic form of a morpheme, which need not even be the form of one of its alternants, as in the case of  $//nigw//$  in Latin, underlying the allomorphs  $|nik|$  and  $|niw|$ , exemplified in the words *nix* 'snow' (nom. sg.) and *nivis* 'snow' (gen. sg.), is dramatically illustrated by the following examples discussed by Rulon Wells in 1947. In Sanskrit, the past-participial suffix has the allomorphs  $|ta|$  and

|*na*|, among others. While the alternation is non-automatic, we may nevertheless take *//ta//* to be the basic form of the suffix, if only because of the vastly greater relative frequency of *|ta|*. Now the verb root *//ac//* 'bend' has the allomorphs *|ac|* and *|ak|*. Since *//c//* is automatically mapped onto */k/* before *//t//* in Sanskrit, we can view the choice of *|ak|* in *akna* 'bent' (= *|ak|+|na|*) as automatic, but only if *//ta//* is the basic form of the past-participial suffix. That is, given *//ac//+//ta//*, we obtain *|ak|+|na|* in two steps: first *//c//* automatically becomes */k/* before *//t//*; second *//t//* non-automatically becomes */n/* after *//ac//*. On the other hand, given *//ac//+//na//*, the result *|ak|+|na|* would not only be non-automatic, but also irregular.

Clearly the relation between the basic morphemic analysis *//ac//+//ta//* of the Sanskrit word *akna* can under no stretch of the imagination be called taxonomic, yet this is precisely the optimal analysis that emerges from the Bloomfieldian structuralist concern to determine which phonological alternations in a language are automatic and which are not.

If Bloomfieldian structuralist theory is not taxonomic, how does it differ from Chomskyan generative theory? For one thing, the formulation of structuralist descriptions tends to be algebraic -- Zellig Harris's structuralist theory of transformational grammar, for example, is explicitly algebraic in nature (Harris 1957, sec. 32) -- while that of generative theory is proof-theoretic. More importantly, the two theories make use of different levels of linguistic analysis, largely as the result of a fundamental difference in how the components of grammar that define those levels are individuated. For Bloomfieldian structuralists, the levels of analysis are antecedently given, and each component is designed to characterize the entities at a particular level. For Chomskyan generativists, components of grammar are individuated by the types of rules they contain, so that the levels of analysis can only be determined as the result of ongoing theorizing. One of the earliest disputes between Bloomfieldian structuralists and Chomskyan generativists in fact concerned the existence of the phonemic level. To Bloomfieldians, the phonemic level could hardly fail to exist, since it is that level at which linguistic forms are individuated: a phonemic representation in a language records just those phonological distinctions that native speakers of that language are capable of making and no others. To Chomskyan, the postulation of such a level stands in the way of formulating the simplest mapping from morphemic representations onto phonetic ones, and accordingly must be dispensed with.

The doctrine that each component of grammar must make use of entirely different sorts of rules is a generalization of Chomsky's more familiar doctrine of the 'autonomy of syntax' according to which the rules in each component of the grammar are stated solely in terms of concepts particular to that component. Ultimately the justification for maintaining the general doctrine of 'autonomy of components' has to do with simplicity or, more precisely, the elimination of duplication. Clearly within each component, no statement or rule will duplicate another. Then if the components are autonomous, no statement or rule of one component will duplicate that of another. On the other hand, within Bloomfieldian structuralism, duplication of statements across components is

rampant. For example in the morphemic component of the grammar of English, there will be a statement to the effect that *bil* 'demand for payment' is a morpheme, and in the lexical component of the grammar of English there will be another statement to the effect that *bil* 'demand for payment' is a word. Clearly these statements violate the doctrine of the autonomy of components since they belong to different components while at the same time sharing elements of vocabulary.

Given that the set of monomorphemic words consists of elements that are internally the same as morphemes and as words, the question arises how they are to be described in a grammar that conforms to the doctrine of the autonomy of components. One conclusion seems clear: Such a grammar cannot contain both a morphemic and a lexical component, one for describing the morphemes of a language and the other for describing its words.

However, duplication of statements also occurs in the lexical and syntactic components of Bloomfieldian structuralist grammars. As Bloomfield observed in his 1926 paper, the possessive morpheme /z/ in English attaches to phrases as well as to words. From this it follows that there are statements in both the lexical and syntactic component of a structuralist grammar of English that make use of this element, again contrary to the doctrine of the autonomy of components. Further, it would appear that strict adherence to that doctrine would force us to conclude that there is but a single component of grammar (morphology? lexicography? syntax?) that maps between morphemes and phrases.

In the interest of maintaining a priori, with the Bloomfieldian structuralists, that distinct components of grammar are involved in the analysis of morphemes, words and phrases, the doctrine of the autonomy of components must be abandoned. Rather than evaluating the statements or rules of grammar as a whole for simplicity, thus requiring that they appear at most once in a grammatical analysis, let the analysis decide which components each of its statements or rules are in, for example by a specification indicating for example that *bil* 'demand for payment' occurs both as a morpheme and as a word in the grammar of English, and that the possessive morpheme /z/ is both a lexical and phrasal affix in that grammar. Of course, a theory of 'permitted duplication' is required, to rule out, for example, the wholesale duplication of the syntax of phrases in the lexical component, a constraint which is needed to block the analysis of *the man I saw yesterday's* as a word in the phrase *the man I saw yesterday's hat*.

Much more can be said on this topic, but given the demands of time, I will conclude. The Bloomfieldian structuralist theory of language, when stripped of certain of its methodological assumptions, expressed in the format of contemporary logico-mathematical linguistics, and equipped with a proper account of semantics, remains a viable theory of language.

## References

- Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2:153-164.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.
- Harris, Zellig. 1957. Cooccurrence and transformation in linguistic structure. *Language* 33:283-340.
- Hockett, Charles F. 1961. Linguistic elements and their relations. *Language* 37:29-53.
- Joos, Martin. 1957. Introduction. *Readings in Linguistics*, ed. by Martin Joos. New York: American Council of Learned Societies.
- Lees, Robert B. 1957. Review of *Syntactic Structures*. *Language* 33:375-408.
- Wells, Rulon S. 1947. Automatic alternation. *Language* 25:99-116.

---

\* This version was prepared in July 2010 from handwritten notes that I read from during the presentation. It was edited to remove a few references that I couldn't track down. I also removed a paragraph describing the use of a hypergrammar (a grammar that generates another grammar) to account for the appearance of the same rule or statement in different components of the grammar.