Mark Aronoff
English verbs in *Syntactic Structures*

Beauty is truth, truth beauty—that is all
Ye know on earth, and all ye need to know.

John Keats, Ode on a Grecian Urn (1820)

1 Introduction

The formally inclined take comfort in the last two lines of Keats's Ode on a Grecian Urn. We judge an analysis by its beauty, or elegance, or simplicity. In linguistics, there is no clearer demonstration of Keats's maxim than the analysis of English verbs in *Syntactic Structures* (henceforth SS). I am not alone in my belief that the beauty of this analysis played a large part in leading the field to accept the truth of Chomsky's claim that the description of human language calls for the use of transformations.

The system of English verbs provides one of the core pieces of evidence in SS for the value of transformations in grammatical theory and description. In this article, I will review Chomsky's analysis of English verbs and use it to partially reconstruct the (largely implicit) view of morphology that lies behind it—most centrally, the role that the morpheme played in SS and just what the term *morpheme* meant in that work. I will explore the roots of this view in structuralist linguistics, especially in the morphological theory of Chomsky's mentor, Zellig Harris. Chomsky's analysis of English verbs and their morphology might have been possible within another framework, but a deeper understanding of the assumptions that undergird it makes the analysis even more beautiful.

In his preface, Chomsky described SS as a comparison of "three models for linguistic structure" (p. 6) and a demonstration "that a certain very simple communication theoretic model of language and a more powerful model ... 'immediate constituent analysis' cannot properly serve the purposes of grammatical description" (ibid.). He showed that the even more powerful transformational model that he developed "provides a good deal of insight into a wide range of phenomena" (ibid.).

I had always accepted Chomsky's presentation of SS as a single argument in favor of the need for the mathematical power of transformations in linguistics,

and I have taught that to my students. I now understand that the book also resembles Darwin's *Origin of Species* (1859), which Darwin called one long argument for the theory of evolution by natural selection. Darwin was no mathematician. His arguments were all grounded in the insights that his deceptively simple theory made possible. I now understand that much the same is true of SS. The value of transformational method lies especially in the beautiful truths that it allows us to express in a simple fashion. In terms that Chomsky would use later, while it might be possible to describe English verbs without resorting to a transformation, the transformation allows us to gain understanding, and there are no formal methods for determining whether we have understood anything.

I will not relate *Syntactic Structures* to later developments in morphology. My goal is to understand the morphology of *Syntactic Structures* on its own terms, in its own time, and in relation to what came before it. I am especially interested in how Chomsky's abandonment of discovery procedures led him to construct an analysis that could be judged in terms of beauty and truth.

The morpheme of SS lies within the post-Bloomfieldian tradition of what Peter Matthews (1993) and others have termed *distributionalism*. This primarily methodological movement arose in the wake of Bloomfield's *Language*; Bloomfield may have been an inspiration but he was not an adherent. The movement formed the core of American structuralist linguistics from around 1940. Its adherents advocated doing linguistic analysis based solely on distribution and without resorting to lexical semantics. This was no empty exercise, but rather stemmed from the deep belief that one's practice must have a firm foundation. The scientific study of meaning had none. Even today, although much progress has been made in understanding how the meanings of complex utterances are built up from the meanings of their parts, the meanings of the most basic parts, individual lexical words, remain largely a mystery.

2 What the thunder said

Chomsky is famous for his footnotes. This article was sparked by my attempt to understand two enigmatically startling footnotes in SS. If we think of Chomsky's footnotes as a kind of gemara to the text, then the purpose of the article is commentary on the specific gemara that these two notes make up.

The first footnote is found on page 29. A couple of pages earlier, Chomsky had introduced a set of rewrite rules for very simple sentences, one of which, (13ii),

rewrites NP as T + N. Two pages later, in observing that some rewrite rules must be context sensitive, he writes the following (pp. 28-29):[1]

> One generalization of (13) is clearly necessary. We must be able to limit application of a rule to a certain context. Thus T [article] can be rewritten *a* if the following noun is singular, but not if it is plural; similarly, Verb can be rewritten "hits" if the preceding noun is *man*, but not if it is *men*. In general, if we wish to limit the rewriting of *X* as *Y* to the context *Z – W*, we can state in the grammar the rule
>
> (16) $Z + X + W \rightarrow Z + Y + W$
>
> For example, in the case of singular and plural verbs, instead of having *Verb → hits* as an additional rule of (13). we should have
>
> (17) $NP_{sing} + Verb \rightarrow NP_{sing} + hits$
>
> indicating that *Verb* is rewritten *hits* only in the context NP_{sing}—.
> Correspondingly, (13 ii) will have to be restated to include NP_{sing} and NP_{pl}. This is a straight-forward generalization of (13).

The first of our footnotes (fn. 3) is appended to this last sentence:

> Thus in a more complete grammar, (13ii) might be replaced by a set of rules that includes the following:
>
> $$NP \rightarrow \left\{ \begin{array}{c} NP_{sing} \\ NP_{pl} \end{array} \right\}$$
>
> $NP_{sing} \rightarrow T + N + \varnothing \ (+ Prepositional\ Phrase\ )$
>
> $NP_{pl} \rightarrow T + N + S \ (+ Prepositional\ Phrase\ )$
>
> where *S* is the morpheme which is singular for verbs and plural for nouns ("comes," "boys"), and $\varnothing$ is the morpheme which is singular for nouns and plural for verbs ("boy," "come"). We shall omit all mention of first and second person throughout this discussion. Identification of the nominal and verbal number affix is actually of questionable validity.

These two morphemes each have one form and share two contextually distributed meanings, singular and plural. They share these meanings in a crossing manner, as in Table 1:

Misgivings about such a crossing pattern may be what led to the last line of the note, which ascribes 'questionable validity' to the identification of the nominal

---

1 I preserve the numbering of SS throughout, since subsequent cited passages will refer to that numbering. All otherwise unattributed quoted passages are from SS.

Table 1: the two morphemes *S* and ∅

morpheme	∅	S
Context	**meaning**	
Noun	SINGULAR	PLURAL
Verb	PLURAL	SINGULAR

and verbal affix. But the analysis was tempting enough for Chomsky to return to it a few pages later, in a footnote to the centerpiece of the book, the analysis of English verbs (pp. 38-40). I present the analysis here in full:

> Consider first the auxiliaries that appear unstressed; for example, "has" in "John has read the book" but not "*does*" in "John *does* read books." We can state the occurrence of these auxiliaries in declarative sentences by adding to the grammar (13) the following rules:

(28)　　(i) $Verb \rightarrow Aux + V$
　　　　(ii) $V \rightarrow hit, take, walk, read$, etc.
　　　　(iii) $Aux \rightarrow C(M)(have + en)(be + ing)(be + en)$
　　　　(iv) $M \rightarrow will, can, may, shall, must$

(29)　　(i)　$C \rightarrow$ $\left\{ \begin{array}{l} S \text{ in the context } NP_{sing}- \\ \varnothing \text{ in the context } NP_{pl}- \\ past \end{array} \right\}$

　　　　(ii) Let *Af* stand for any of the affixes *past*, ∅, *en*, *ing*. Let *v* stand for any *M* or *V* or *have* or *be* (i.e., for any non-affix in the phrase Verb). Then:
　　　　　　　　$Af + v \rightarrow v + Af\#$,
　　　　　　Where # is interpreted as a word boundary
　　　　(iii) Replace + by # except in the context $v - Af$. Insert # initially and finally.

> The interpretation of the notations in (28iii) is as follows: we must choose the element *C*, and we may choose zero or more of the parenthesized elements in the given order. In (29i) we may develop *C* into any of three morphemes, observing the contextual restrictions given. As an example of the application of these rules, we construct a derivation in the style of (14), omitting the initial steps.

(30)   *the + man + Verb + the + book*   from (13i-v)

   the + man + Aux + V + the + book   (28i)

   the + man + Aux + read + the + book   (28ii)

   the + man + C + have + en + be + ing +   (28iii)—we select the
      read + the + book   elements *C*, *have + en*, and *be + ing*.

   the + man + S + have + en + be + ing + read + the + book   (29i)

   the + man + have + S# be + en#read + ing #the + book   (29ii)—three times

   #the# man# have + S #be + en# + read + ing #the + #book#   (29iii)

The morphophonemic rules (19), etc., will convert the last line of this derivation into:

(31)   the man has been reading the book

(29i) sets the stage for (29ii), a transformation, which we now call affix hopping. (29i) also bears our second footnote of interest, which is short:

> We assume here that (13ii) has been extended in the manner of fn. 3, above, p. 29, or something similar.

We don't know what "something similar" might be, but morphemes like *S* and $\varnothing$ of fn. 3 are astonishing to almost any morphologist because they contradict two fundamental tenets. First, every morpheme must have a single meaning (though we may not always know what that meaning is), though it may have many forms. Second, no two morphemes can have the same meaning.[2] As shown in Table 1, these two morphemes each have only one form but two meanings, depending on context, and they share these meanings. If we are to understand how Chomsky could have entertained morphemes with such strangely-crossed meaning patterns, we must review the history of the use of the term *morpheme* since its invention, with special attention to the place of meaning in its definition.

---

2 The claim that there are no synonymous words dates to Girard (1718). Bloomfield accepted it for morphemes (see below) and the notion of allomorphy depends on it, as does morphological blocking. Not everyone understands its value or ubiquity. Fought, for example, noted: "I have not found any effort in his [Bloomfield's] published work to justify his unusual position on synonymy...Indeed, the outright rejection of synonymy is a position not often taken" (1999, p. 323). The opposite is true.

3 The morpheme through time³

The term *morpheme* was coined by Baudouin de Courtenay before 1880, on the model of the term *phoneme*.⁴ He defined the morpheme as follows (1895/1972, p. 153):

> ...that part of a word which is endowed with psychological autonomy and is for the very same reason not further divisible. It consequently subsumes such concepts as the root (radix), all possible affixes, (suffixes, prefixes), endings which are exponents of syntactic relationships, and the like.

3.1 Morphemes and meaning: Bloomfield's quandary

We cannot know exactly what Baudouin meant by 'psychological autonomy'. Fifty years later, in his (1933) the bible of American structuralist linguistics, Leonard Bloomfield, who by then had given up on the utility of psychological interpretation of language for linguists, defined morphemes in terms of meaning or semantics instead of psychological autonomy. He declared that "each linguistic form has a constant and specific meaning" (p. 145) and a few pages later that "[a] linguistic form which bears no partial phonetic-semantic resemblance to any other form is a *simple* form or *morpheme*" (p. 161). Although Bloomfield defined the morpheme in terms of meaning, he cautioned in the next paragraph that "[a] morpheme can be described phonetically, since it consists of one or more phonemes, **but its meaning cannot be analyzed within the scope of our science** [emphasis mine]." And on the next page:

> The meaning of a morpheme is a *sememe*. The linguist **assumes** [emphasis MA] that each sememe is a constant and definite unit of meaning, different from all other meanings, including all other sememes, in the language, but **he cannot go beyond this** [emphasis MA]. There is nothing in the structure of morphemes like *wolf*, *fox*, and *dog* to tell us the relation between their meanings. This is a problem for the zoölogist. The zoölogist's definition of these meanings is welcome to us as a practical help, but it cannot be confirmed or rejected on the basis of our science. (ibid., p. 162)

---

3 Anderson (2015) provides an insightful comprehensive survey of the morpheme since the term and notion were first created. Matthews (1993) is broader in scope but his treatment of the history of the American structuralist conception of the morpheme and its place in both structuralist theory and early generative grammar is masterful.

4 Jakobson (1971) traces Baudouin's first use of the term to his lectures of 1877-78. A program of these lectures was published in 1880.

An entire chapter of *Language* bears the title *Meaning*, but this chapter consists mainly of an apologia for not discussing meaning more precisely. Bloomfield believed: "In order to give a scientifically accurate definition of meaning for every form of a language, we should have to have a scientifically accurate knowledge of everything in the speaker's world" (p. 139). He concludes: "The statement of meaning is therefore the weak point in language-study, and will remain so until human knowledge advances very far beyond the present state" (p. 140). Nonetheless, although he cannot define meanings, he makes the following axiomatic declaration:

> Since we have no way of defining most meanings [of morphemes MA] and of demonstrating their constancy, we have to take the specific and stable character of language as a presupposition of linguistic study, just as we presuppose it in our everyday dealings with people. We may state this presupposition as the *fundamental assumption of linguistics* (§5.3), namely:

> *In certain communities (speech-communities) some speech-utterances are alike as to form and meaning.* (p. 144).

Bloomfield's attitude towards meaning is a classic case of taboo: he assumes that each morpheme is a distinct pairing of a meaning and a form, but he insists that meaning is so powerful a notion that he can't talk about or even think about what it is.[5] It is true that Bloomfield favored what he called a materialist or mechanistic theory of psychology in the analysis of meaning over what he memorably termed mentalistic psychology, but that had no effect on his attitude towards meaning itself or towards his practice. As Bernard Bloch so cogently remarked of Bloomfield in his obituary, "He had convinced himself ... that it does not matter what particular brand of psychology a linguist finds attractive, so long as he keeps it out of his linguistic writing" (Bloch 1949/1970, p. 526).

---

5 This taboo surrounding meaning most closely resembles the taboo that has long surrounded the name of the god of the Hebrews: we know that this god has a name, and we even know how to spell the name in letters (יהוה the *tetragrammaton* 'four-letter word'), but we have never been permitted to speak it, because the name is so sacred as to be taboo. According to tradition, if anyone ever were to successfully pronounce the complete true name of the deity, the world would end immediately.

3.2 The triumph of distributionalism [6]

The small band of Sapir and Bloomfield's immediate successors in American Structural linguistics came to be called the post-Bloomfieldians. They treated his *Language* as their bible. The boldest among them dealt with the taboo on meaning in a more radical way than their mentor: by banishing meaning from linguistics entirely and putting their faith in distribution alone.[7] The result was what Matthews (1993) calls distributionalism. Most prominent among the distributionalists was Chomsky's mentor, Zellig Harris, and most radically in his *Methods in Structural Linguistics* (Harris, 1951). The distributionalists themselves regarded Harris as their leading theoretician. For example, in his review of *Methods in Language*, Norman McQuown (1952) calls it "epoch-making" and declares: "Not since Bloomfield's Language has there been such an ambitious attempt to cover a whole field." And we are still in the first paragraph of a laudatory ten-page review! In the next sentence, McQuown cautions: "Unlike Bloomfield's, however, this book is limited to the presentation of one principle and one method of linguistic analysis and description. The principle is that of relative distribution, the method that of controlled substitution." The entertainment of morphemes like Chomsky's S and $\varnothing$ makes sense only in the context of Harris's version of distributionalism. The entire treatment of the English verb in SS is profoundly distributionalist. I will now turn to the roots of distributionalism. Its connection to the phoneme is well known but worth rehearsing. The tie to Whorf, especially Whorf (1945) is new.

---

6 My discussion of the post-Bloomfieldians and distributionalism relies heavily on the sections devoted to those topics in Matthews (1993). Matthews does not discuss Whorf's abstractionism, which I propose as an important factor in the rise of distributionalism.

7 The idea of exploring lexical semantics in terms of use rather than definition was not confined to linguistics. Ludwig Wittgenstein, for instance, who had no known contact with linguists, came to similar conclusions in the second half of his career, with his famous dictum: "For a large class of cases of the employment of the word 'meaning'—though not for all—this word can be explained in this way: the meaning of a word is its use in the language" (Wittgenstein 1953, §43). Wittgenstein had worked on the ideas in Philosophical Investigations since returning to philosophy and to Cambridge in 1929. He submitted a complete draft to the publisher in 1945, but withdrew the final version in 1946, authorizing posthumous publication. Among American linguists, only Harris, who was steeped in formal logic, might possibly have known of Wittgenstein, but none of Wittgenstein's work on ordinary language was published until after Harris had completed his book in 1946, making even this implausible. In short, there was likely no direct contact between the two. In any case, Wittgenstein never formulated any specific methods for exploring his use theory of meaning in the way that Harris did. The ordinary language philosophers who followed his lead suffered from a similar absence of explicit methodology and the effort sadly dissipated.

3.3 The abstract phoneme, Whorf's covert categories, and distributional methods

The greatest achievement in the field of linguistics in the second quarter of the 20th century was the discovery of the phoneme as a distinctive sound unit of individual languages, vindicating Saussure's structuralist approach to language.[8] The phoneme, no matter how one attempts to define its essence, is defined in practice by the two distributional methods by which the phonemes of a language are found or discovered. These are substitution/contrast and complementary distribution. If two sounds contrast (are judged by native speakers of the language to result in distinct words) when substituted for one another (as with /p/ and /b/ in English *pat* and *bat*), they belong to separate phonemes. If linguists discover that they are in complementary distribution (as with $[p^h]$ and $[p]$ in *pit* and *spit*), they are variants or allophones of the same phoneme. The allophones of a single phoneme are gathered together into a set. The phoneme is the name of the set. So, we may say that the set called the English phoneme /p/ has the members $\{ p, p^h, ... \}$. In the 1930's and later, there was much discussion over the 'reality' of the phoneme and over whether one variant of each phoneme had a privileged status in the mind of the language user or otherwise. But the methods for discovering phonemes and allophones were not touched by this discussion: the methods were entirely distributional. It was crucial to know whether two words meant different things, but their actual meanings were irrelevant. The phoneme was a true breakthrough, a property of all languages, and though the debate over whether a phoneme was a single 'real' sound or simply the name of a set of allophones persisted well after the structuralist phoneme succumbed to the generative underlying representation, the debate had no effect on the distributional method and this remarkable result.

Benjamin Lee Whorf was a central member of the group of linguists who had gathered around Edward Sapir at Yale in the 1930's. Sapir's coterie constituted the first critical mass of modern linguists in North America. It continued to thrive after Sapir's death in 1939 under Bloomfield (who arrived the next year and was active until his stroke in 1949) and Bloch (from 1943). This group formed the core of American structuralist linguistics. Whorf's influence during his short lifetime (he died at age 44 in 1941) is evinced most by terms that he coined that were later to become standard, notably *allophone* and *lexeme*.

---

8 Readers interested in details of the history of the concept of the phoneme should consult Anderson (1985) and Dresher (2011)

Whorf's article, "grammatical categories," was published in *Language* in 1945. John Carroll writes the following about its provenance:

> According to a note supplied by the editor of Language [Bernard Bloch], "This paper was written late in 1937 at the request of Franz Boaz, then editor of the *Int. J. Amer. Linguistics.* The manuscript was found in the Boaz collection by C. F. Voegelin and Z. S. Harris." (Whorf 1956, p. 87).

The fact that Bloch, the editor of *Language* for over a quarter century and a major gatekeeper of American structural linguistics, chose to publish this, the most theoretical of Whorf's articles, posthumously almost ten years after it had been written and over four years after Whorf's death, tells us that Bloch believed it was important to the structuralist program. The central innovation of the article is the notion of a *covert category*, for which, as opposed to an *overt category*, there is no overt identifiable mark of category membership. Instead, membership can be determined by what Whorf (ever a chemist) called the *reactance* of the category, its distribution. His first example of a covert category is the English intransitive verb, and his definition is a masterpiece of distributional analysis:

> In English, intransitive verbs form a covert category marked by the lack of the passive participle and the passive and causative voices; we cannot substitute a verb of this class (e.g., 'go, lie, sit, rise, gleam, sleep, arrive, appear, rejoice') into such sentences as 'It was cooked, It was being cooked, I had it cooked to order.' An intransitive thus configurationally defined is quite a different thing from the "dummy" intransitive used in traditional English grammar. It is a true grammatical class marked by these and other constant grammatical features, such as nonoccurrence of nouns or pronouns after the verb; one does not say 'I gleamed it, I appeared the table.' Of course compound formations involving these same lexemes may be transitive, e.g., 'sleep (it) off, go (him) one better.' In the American colloquial forms, 'go haywire, go South Sea Islander,' etc., the word or phrase after the verb is a covert adjective, cf. 'go completely haywire.' (Whorf 1956, pp. 89-90)

This passage, in which a category of elements is defined solely in terms of its privileges of occurrence, is perfectly understandable to any modern syntactician. It is modern syntax. As always with Whorf, it is beautifully written. This is distributionalism at its finest.

And it does not depend in any way on meaning, only on reactance. Whorf provides two examples showing that covert categories are "not reflections in speech of natural and noncultural differences." (p. 91). The first is English gender, revealed only in the choice of coreferent pronoun, which must be feminine, masculine, or neuter, and which does not depend on "knowledge of any 'natural properties" (p. 90). The second is Navajo nouns "based actually or ostensibly on shape":

> Some terms belong to the round (or roundish) class, others to the long object class, others fall into classes not dependent on shape. No overt mark designates the class in every sentence. The class mark as in English gender is a reactance; not a pronoun, however, but a choice between certain verb stems that go definitely with one class and no other, although there are very many verb stems indifferent to this distinction. I doubt that such distinctions, at least in Navaho, are simply linguistic recognitions of nonlinguistic objective differences that would be the same for all observers, any more than the English genders are; they seem rather to be covert grammatical categories. Thus, one must learn as part of learning Navaho that 'sorrow' belongs in the "round" class. (ibid. p. 91).

Whorf wrote this article at the request of Franz Boaz and his covert categories are Boazian. They are not given in advance (by nature, Whorf would say) but rather emerge from the structure of each language. Sorrow belongs in the round class because it shares a distribution with the other members of that category. So too with English gender, which is a covert distributional category of English and no other language. Every language has its own categories, many of them covert, which lie hidden in the linguist's data, waiting to be discovered. Chomsky's analysis of English verbs lies squarely within this Boazian tradition.

Both phonemes and Whorf's covert categories demonstrate the success of an objective distributional method that, as Bloomfield had proposed, depend only on words and categories having distinct meaning, not on having to know what these meanings are. The method is also objective and reliant only on the linguist's ability to collect reliable data from a fluent speaker/hearer. It made synchronic linguistics a successful science. The obvious next step was to extend the method to morphology.

3.4 Harris' Method

The structuralist phoneme and Whorf's covert categories demonstrated the power of structure over substance. Zellig Harris struck the final blow. Harris was an outsider, who spent his academic life from the age of 18 at the University of Pennsylvania. For a decade or so, though, he was the standard bearer of the distributional method. He published numerous articles on distribution beginning in 1942 and his major book (Harris, 1951) bore the title *Methods in Structural Linguistics*. It was devoted entirely to this method.[9]

Harris started out as a Semitist and published two well-received philological monographs early on, one on Phoenician (1936) and the other on Canaanite di-

---

9 The title of later editions was shortened to *Structural Linguistics*. One can only speculate on the reasons for the change.

alects (1939). He was famously private all his life and had little personal contact with other American descriptive linguists. Beginning in 1942, though, and continuing for more than a decade, he published a series of articles in *Language* that were tremendously influential in the theoretical mainstream. The first of these set the tone. Harris tells the reader in the first sentence: "The purpose of this paper is to suggest a technique for determining the morphemes of a language, as rigorous as the method used now for finding its phonemes." (Harris 1942, p.169). He begins with what he calls "the present treatment of morphemes" (ibid.), for which he cites Bloomfield: "Every sequence of phonemes which has meaning, and which is not composed of smaller sequences having meaning, is a morpheme." (ibid.). He calls Bloomfield's morphemes morpheme alternants but extends the possible signifiers "by taking sequence to mean not only additive sequence (the addition of phonemes), but also zero (the addition of no phonemes), negative sequence (the dropping of a phoneme), and phonemic component sequence (the addition of a physiological feature of phonemes." (ibid. 170). He encapsulates his tripartite method for finding morphemes in the summary section of the book:

> The method of arranging the phonemes of a language consists of three steps: 1. dividing each phonemically written linguistic expression into the smallest parts which recur with the same meaning in different expressions, each such part to be called a morpheme alternant; 2. grouping into a distinct morpheme unit all alternants which satisfy the following conditions: (a) have the same meaning, (b) never occur in identical environments, and (c) have combined environments no greater than the environments of some single alternant in the language; 3. making general statements for all units which have identical difference between their alternants. (ibid. 179-180)

The morpheme that results from this method does not have a single signifier in the form of concrete sequence of phonemes, à la Bloomfield. Instead, analogous to the post-Bloomfieldian phoneme and allophone, each morpheme names a set of signifiers (allomorphs) in complementary distribution. And some of these signifiers (such as ablaut) are sequences in name only, as noted above. The resulting morpheme is even more abstract than the structuralist phoneme and close to Whorf's covert category. The members of the distributionalist phoneme are at least tethered to one another by the concrete criterion of phonetic similarity, which famously prevents uniting English [h] and [ŋ] under a single phoneme, even though [h] occurs only syllable initially and [ŋ] in the coda. But morphological variants are free from substantive relations. As Eugene Nida (1949) declared in his seminal textbook on morphological analysis in American structuralist linguistics, "There are absolutely no limits to the degree of phonological difference between allomorphs" (44). All that matters are complementary distribution and synonymy.

4 Meaning and method in *Syntactic Structures*

Harris (2002) recognized Noam Chomsky as his most successful student. In the Preface to Harris (1951), dated January 1947, he noted that "N. Chomsky has given much-needed assistance with the manuscript" (p.v). Chomsky similarly wrote: "My introduction to the field of linguistics was in 1947, when Zellig Harris gave me the proofs of his Methods in Structural Linguistics to read" (Chomsky (1975a), p. 25). In the same article, he noted: "While working on LSLT [Chomsky 1955/1975b] I discussed all aspects of this material frequently and in great detail with Zellig Harris, whose influence is obvious throughout" (1975a, p. 4). SS is, in his own words, "a sketchy and informal outline of some of the material in LSLT" (ibid., p. 3). We find homage to Harris in the symbol C in rule (28iii), which became T (for tense) in later work. Harris (1948) named the set of person, number, gender, and tense affixes in Modern Hebrew C. The name is as opaque in Harris as it is in Chomsky but it is the same name.

4.1 Meaning

In SS, Chomsky followed Harris in insisting on the independence of linguistic analysis from semantics. He explicitly refused to rely on meaning in defining morphemes or even in doing linguistics. Chapter 9 of SS is entitled "Syntax and Semantics" and is devoted to the question "whether or not semantic information is required for discovering or selecting a grammar" (p. 93). For morphology specifically, he lists the claim that "morphemes are the smallest elements that have meaning" as "[a]mong the more common assertions put forth as supporting the dependence of grammar on meaning" (p. 94). He concludes that not all morphemes can be usefully defined as minimal meaning-bearing units:

> Such morphemes as "to" in "I want to go" or the dummy carrier "do" in "did he come?" (cf. §7.1) can hardly be said to have a meaning in any independent sense, and it seems reasonable to assume that an independent notion of meaning, if clearly given, may assign meaning of some sort to such non-morphemes as gl- is "gleam," "glimmer," "glow." Thus we have counterexamples to the suggestion (117ii) that morphemes be defined as minimal meaning-bearing elements. (p. 100)

In asserting his freedom from semantics, Chomsky emphasizes in this summary chapter that the entire framework of SS is purely formal:

> In §§3-7 we outlined the development of some fundamental linguistic concepts in purely formal terms. We considered the problem of syntactic research to be that of constructing a

device for producing a given set of grammatical sentences and of studying the properties of grammars that do this effectively. Such semantic notions as reference, significance, and synonymity played no role in the discussion. (pp. 102-103)

In §§3-7, then, we were studying language as an instrument or a tool, attempting to describe its structure with no explicit reference to the way in which this instrument is put to use. (p. 103)

4.2 Method

Chomsky's most lasting innovation in SS was to cast off the shackles by which Harris had so tightly bound distribution to discovery. Harris and his distributionalist colleagues truly believed that they could formulate (and perhaps had formulated) a purely inductive distributional method free of semantics, which, when carefully applied to a large enough set of data, would supply the linguist with the correct analysis of linguistic phenomena, or at least decide which analysis was correct. Harris set all this out carefully at the start of his 1951 book, where, on page 6, he declared: "The whole schedule of procedures outlined in the following chapters...is designed to begin with the raw data of speech and end with a statement of grammatical structure."

In the 1975 preface to Chomsky (1955/1975b), Chomsky describes how, by 1953, he had "abandoned any hope of formulation taxonomic discovery procedures" (p. 33), under the influence of Nelson Goodman and W.V.O. Quine. In Chapter 6 of SS, entitled *On the Goals of Linguistic Theory*, Chomsky rejected the quest for a discovery procedure or even a *decision procedure* that would determine whether a proposed grammar "is, in fact, the best grammar of the language from which this corpus is drawn" (SS p. 51), in favor of what he called a *practical evaluation procedure*, which would choose which of two grammars was better.[10] Characteristically, he makes the argument most cogently in a footnote: "Our main point is that a linguistic theory should not be identified with a manual of useful procedures, nor should it be expected to provide mechanical procedures for the discovery of grammars" (p. 55).

In LSLT, in SS, and for the next decade, Chomsky advocated using a simplicity measure to compare between grammars, but any such measure was more honored in the breach than in the observance. In practice, the most important consequence of abandoning discovery procedures was to free the analyst from any need for self-justification:

---

10 The use of the term *procedure* throughout this passage shows that the theory of SS was constructed against the backdrop of Harrisian linguistic method.

> In short, we shall never consider the question of how one might have arrived at the gram-
> mar whose simplicity is being determined; e.g., how one might have discovered the analysis
> of the verb phrase presented in §5.3. Questions of this sort are not relevant to the program
> of research that we have outlined above...We are thus interested in describing the forms of
> grammars...rather than in showing how, in principle, one might have arrived at the grammar
> of a language. (p. 56)

This stance, more than any technical or analytical innovations, was revolutionary. Combined with the absence of either a practical simplicity measure or any ties to semantics, it freed the analyst to posit whatever structures lay hidden within the language. The primary justification left for defending these structures was beauty. The analysis of the verb phrase in §5.3 is beautiful.

5 Beauty and Truth

The task that Chomsky had assigned to himself beginning in the mid 1950s was to "provide simple and revealing grammars that generate all of the sentences of English and only these" (Chomsky 1956, p. 113). The key phrase is *simple and revealing*. There is a myth that Chomsky proved in SS that transformations are mathematically necessary for the description of natural languages. The roots of this myth lie in such direct claims in the book as that "discontinuities cannot be handled with [Σ, F] grammars" (p. 41) but Chomsky immediately tempered this claim in the very long footnote that follows it:

> We might be tempted to extend the notions of phrase structure to account for discontinuities.
> It has been pointed out several times that fairly serious difficulties arise in any systematic
> attempt to pursue this course...If we were to attempt to extend phrase structure grammar to
> cover the entire language directly, we would lose the simplicity of the limited phrase struc-
> ture grammar and of the transformational development." (pp. 41-42)

Here, Chomsky is acknowledging that his argument for the truth of his approach is rooted in beauty and not in mathematical necessity. His term is *simplicity*, which he says we would lose without transformations. Indeed, a quarter century later, Gazdar et al. (1982) famously showed that a phrase structure grammar could account for the English verb. No one who has read that account, however, would argue that it is anywhere close to beautiful. Like all subsequent accounts that dispense with affix hopping (e.g., Chomsky 1995; Lasnik 2005), its complexity cannot pin down the central observations that the SS analysis puts front and center: the form of each verb in the English verb complex is dictated by the preceding verb word; and, since no verb word precedes the first verb in the complex, the form of

the first verb cannot be determined by its predecessor. Within the SS analysis, a single operation is at play in both. The operation is *affix-hopping*: the affix immediately preceding a verb 'hops over' that verb and attaches to it. The form of the first verb in the complex is made to conform to the hopping generalization by having its affix start off in front of the verb, as the tense marker for the entire complex. Do-support provides further support for this placement of the tense marker.

The affix-hopping analysis is no simpler than a PS analysis in any measurable technical or mathematical sense. The whole of SS is constructed around the fact that transformations constitute a third dimension that considerable extends the analytical power of the theory. The affix-hopping analysis is therefore less simple than the PS analysis. Its virtues are beauty and truth. Chomsky remarks on "the simplicity of the limited phrase structure grammar and of the transformational development" (ibid.). The PS rule in (28) and the transformation in (29ii) split the structure of the complex verb into two parts, each of which is succinct and straightforward. Their beauty lies in their combination. I repeat (28) and (29) here:

(28)   (i) *Verb* → *Aux + V*
       (ii) *V* → *hit*, *take*, *walk*, *read*, etc.
       (iii) *Aux* → *C*(*M*)(*have + en*)(*be + ing*)(*be + en*)
       (iv) *M* → *will*, *can*, *may*, *shall*, *must*

(29)   (i) $C \rightarrow \left\{ \begin{array}{l} S \text{ in the context } NP_{sing}- \\ \varnothing \text{ in the context } NP_{pl}- \\ past \end{array} \right\}$

       (ii) Let *Af* stand for any of the affixes *past*, $\varnothing$, *en*, *ing*. Let *v* stand for any *M* or *V* or *have* or *be* (i.e., for any non-affix in the phrase Verb). Then:
       $$Af + v \rightarrow v + Af\#,$$
       Where # is interpreted as a word boundary
       (iii) Replace + by # except in the context *v − Af*. Insert # initially and finally.

Chomsky shows at some length that (29) "violates the requirements" of PS grammars. He defends the entire analysis by arguing that a PS grammar with the same coverage would be less revealing, though he never uses that word:

> The reader can easily determine that to duplicate the effect of (28 iii) and (29) without going beyond the bounds of a system [Σ, F] of phrase structure, it would be necessary to give a fairly complex statement. ...[S]ignificant simplification of the grammar is possible if we are permitted to formulate rules of a more complex type than those that correspond to a system of immediate constituent analysis. By allowing ourselves the freedom of (29ii) we have been able to state the constituency of the auxiliary phrase in (28iii) without regard to the inter-

dependence of its elements, and it is always easier to describe a sequence of independent elements than a sequence of mutually dependent ones. To put the same thing differently, in the auxiliary verb phrase we really have discontinuous elements - e g., in (30), the elements *have...en* and *be...ing*. ... In (28iii) we treated these elements as continuous, and we introduced the discontinuity by the very simple additional rule (29ii). (pp. 41-42).

The key phrase here is "in the auxiliary verb phrase we really have discontinuous elements." This is a claim about truth, not about simplicity or complexity. The SS analysis won the day because it appeared to reveal a truth, discontinuous dependency, and it represented that revealed truth in a simple and elegant way by splitting it into two parts: (28iii) expressed the dependency and (29ii) expressed the linear discontinuity.

5.1 Chomsky's morphemes and the end of certainty in the search for truth

SS was the work of a post-Bloomfieldian distributionalist freed from discovery procedures. Its morphemes were analytical elements unified only by their distribution and their utility in an analysis. Without discovery procedures, though, there could be no way to know whether the analysis was correct, no certainty. The hope of finding an evaluation metric might provide some solace, but, truth be told, no useful concrete evaluation metric was ever found, and certainly none played a role in SS. The value of an analysis was, we now see, determined largely by the criteria of beauty and truth.

Chomsky uses the word *morpheme* 58 times and *affix* 6 times in SS. What did Chomsky mean by these terms and how did he use what he called morphemes and affixes? Chomsky does not discuss in SS what he means by affix. He has a bit to say about morphemes. For one, he follows Harris in explicitly rejecting the traditional Bloomfield definition of the morpheme as "having actual phonemic 'content' in an almost literal sense" (p. 58) in favor of a more abstract entity:

> This leads to trouble in such well-known cases as English "took" /tuk/, where it is difficult without artificiality to associate any part of this word with the past tense morpheme which appears as /t/ in "walked" /wɔkt/, as /d/ in "framed" /freymd/, etc. We can avoid all such problems by regarding morphology and phonology as two distinct but interdependent levels of representation, related in the grammar by morphophonemic rules such as (19). thus "took" is represented on the morphological level as *take + past* just as "walked" is represented as *walk + past*. The morphophonemic rules (I9ii), (19v), respectively, carry these strings of morphemes into /tuk/, /wɔkt/. The only difference between the two cases is that (19v) is a much more general rule than (19ii). (p. 58)

The footnote to this passage is even more revealing:

> Hockett gives a very clear presentation of this approach to levels in *A manual of phonology* (1955), p. 15. In "Two models of grammatical description," *Linguistics Today, Word* 10.21033 (1954), Hockett rejected a solution very much like the one we have just proposed on the grounds that "*took* and *take* are partly similar in phonemic shape just as are *baked* and *bake*, and similar in meaning also in the same way: this fact should not be obscured" (p. 224). But the similarity in meaning is not obscured in our formulation, since the morpheme *past* appears in the morphemic representation of both "took" and "baked." And the similarity in phonemic shape can be brought out in the actual formulation of the morphophonemic rule that carries *take* + *past* into /tuk/. We will no doubt formulate this rules [sic] as
> ey → u in the context t–k + *past*[11] in the actual morphophonemic statement. This will allow us to simplify the grammar by a generalization that will bring out the parallel between "take"-"took," "shake"-"shook," "forsake"-"forsook," and more generally, "stand"-"stood," etc.

Chomsky lists "the affixes *past*, *S*, $\varnothing$, *en*, *ing*" in the affix hopping transformation. Of these affixes, only *S* and *ing* can be called phonologically concrete in any sense. We must conclude that affixes are morphemes (perhaps bound morphemes) and that morphemes are abstract entities made up of synonymous allomorphs in complementary distribution, as Harris had shown in 1942. Since Chomsky rejects semantics as a criterion, these abstract morphemes can only be determined through distributional analysis, again as Harris had shown.

The abstract 'affixes' that Chomsky names *past* and *en* have no reality outside his analysis of English verbs. What makes the analysis even possible is that each 'affix' is an abstract entity defined entirely in distributional terms. Each one comprises several phonologically distinct variant forms, including non-affixal morphological operations like the ablaut relating *take* and *took*. Also, for the overwhelming majority of verbs the two morphemes are homophonous, realized as <-ed>. We can unify the realizations of each under one morpheme and gather the morphemes under the category affix only within this specific very abstract analysis.[12] What, then, made this analysis so compelling? It's beauty and the belief on the part of readers that this beautiful analysis helped them to understand the phenomenon in a new way.

Which brings us back to the analysis of *S* and $\varnothing$ in footnote 3 and its "questionable validity." A decade earlier, in his warning against the use of purely distributional methods without regard for semantics, Nida had presented this same analysis as a *reductio ad absurdum* argument:

---

11 This is Bloch's (1947) analysis of the relation between *take* and *took*, and it has been repeated many times, even recently, often, as here, without attribution.
12 Affix-hopping is the evidence that these elements are all affixes, but affix-hopping has no validity outside the analysis in SS.

> Hockett considers that complementary distribution is all that fundamentally counts. But the implications of this method seem to be greater than he may have anticipated. For example, consider what could be done on that basis with number distinctions in English. Not only would it be possible to combine all the plural affixes of nouns in one morpheme (a step which we should all agree to), but one could say that these are in complementary distribution with the partly homophonous third-singular suffix of verbs. A single morpheme could then be set up with the meaning 'number distinctiveness' and with the additional distributional characteristic that if an alternant occurs after the noun it does not occur after the verb, and vice versa, e.g. /ðə boyz rən/ the boys run vs. /ðə boyz rənz/ the boy runs. By slight extensions it might be possible to construct a descriptive system by which practically all the features of concord, government, and cross-reference could be treated on a submorphemic level. If this were done, we should only have succeeded in changing the meaning of the word 'morpheme' to apply it to certain distributionally related forms. (Nida 1949 p. 418).

Nida only hinted at the fatal flaw of this analysis, the wish to have one's cake and eat it too: if we combine all the plural affixes of nouns in one abstract noun-plural morpheme on the grounds of complementary distribution, which is indeed "a step which we should all agree to," then *S* cannot be a morpheme on its own, but must rather be one of the allomorphic realizations of the abstract noun-plural morpheme, along with the *−en* of *oxen*, the ablaut of *geese*, the ∅ of *deer*, and many others. Identifying this allomorphic *S* with the *S* of the third person singular is either comparing apples (if the verbal *S* is an abstract morpheme) and oranges (the plural *S* is an allomorphic realization, the output of morphophonemic rules in the framework of SS), or both instances of *S* (noun plural and verb singular) are allomorphs, in which case the likeness of the two affixes is no more than a curiosity, driven, as Nida so discreetly hints, by distributionalist zealotry and deliberate disregard for meaning. It is as accidental as the overwhelming homophonic realization of *past* and *−en* as <-ed>. This raises an entirely different question, about syncretism of morphs rather than morphemes, which has been entirely ignored.

Besides overzealousness, though, this curio of an analysis reveals the risk inherent in Chomsky's daring discard of the safety net of discovery procedures within a distributionalist ethos. Those who truly believed in discovery procedures had the comfort of faith: they knew that their distributionalist methods would always lead to the correct analysis. Without this net, there is no certainty. Some analyses will always fail because they are wrongly constructed, as is this one. Other analyses, including the analysis of English complex verbs, must be judged only in the court of academic opinion, for their beauty and their truth. Few will succeed there and even fewer masterpieces will endure for the ages, as this one has.

6 One: Number 31, 1950

Why do we sit transfixed in front of Jackson Pollock's *One: Number 31, 1950* at the Museum of Modern Art in New York? Its crushing beauty moves us to tears. Giotto's Scrovegni Chapel in Padua has the same effect. Both reveal a truth that the viewer has never experienced before. The same holds for the analysis of English verbs in (28) and (29) of *Syntactic Structures*. Chomsky writes of its simplicity, but simplicity alone is not enough to explain it. Rather it is the simplicity with which the analysis expresses what we quickly grasp as a new truth. The simplicity of the analysis is beautiful because, like the Pollock and the Giotto, it reveals a truth that we have never experienced before. All three share one common feature: their creators discovered an entirely novel technique and, more rarely, they used this technique to express a beautiful truth that no one had revealed before. For Pollock, it was the drip method that came to be called action painting. For Giotto it was naturalism augmented by perspective. For Chomsky it was the formal generative transformational technique exemplified in affix-hopping. But technique alone does not make a masterpiece and it does not make us weep.

Acknowledgements

Bibliography

Anderson, Stephen R. 1985. *Phonology in the twentieth century: Theories of rules and theories of representations.*. Chicago, IL: University of Chicago Press.

Anderson, Stephen R. 2015. The Morpheme: Its Nature and Use. In *The Oxford Handbook of Inflection*, ed. Matthew Baerman, 11–33. Oxford, UK: Oxford University Press.

Bloch, Bernard. 1947. English verb inflection. *Language* 23:399–418.

Bloch, Bernard. 1949. Leonard Bloomfield obituary. *Language* 25:87–94. Reprinted in Charles F. Hockett, ed., A Leonard Bloomfield Anthology. 525-532. Bloomington: Indiana University Press.

Bloomfield, Leonard. 1933. *Language*. New York, NY: Harcourt Brace.

Chomsky, Noam. 1955/1975b. The Logical Structure of Linguistic Theory. MS, MIT.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague, The Netherlands: Mouton de Gruyter.

Chomsky, Noam. 1975a. Preface to The Logical Structure of Linguistic Theory. In Chomsky 1975b.

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

Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection*. London, UK: John Murray.

Dresher, B. Elan. 2011. The Phoneme. In *The Blackwell Companion to Phonology, Volume 1*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 241–266. Malden, MA: Wiley Blackwell.

Fought, John G. 1999. Leonard Bloomfield's Linguistic Legacy: Later uses of some technical features. *Historiographia Linguistica* 26:313–332.

Gazdar, Gerald, Geoffrey K Pullum, and Ivan A Sag. 1982. Auxiliaries and related phenomena in a restrictive theory of grammar. *Language* 591–638.

Girard, Gabriel. 1718. *la justesse de la langue francoise ou les différentes significations des mots qui passent pour synonimes*. Paris, France: Laurent d'Houry.

Harris, Zellig S. 1936. A Grammar of the Phoenician Language. Doctoral Dissertation, University of Pennsylvania.

Harris, Zellig S. 1939. *Development of the Canaanite Dialects: An Investigation in Linguistic History*. American Oriental Series. New Haven, CT: American Oriental Society.

Harris, Zellig S. 1942. Morpheme alternants in linguistic analysis. *Language* 18:169–180.

Harris, Zellig S. 1948. Componential analysis of a Hebrew paradigm. *Language* 24:87–91.

Harris, Zellig S. 1951. *Methods in Structural Linguistics*. Chicago, IL: University of Chicago Press.

Harris, Zellig S. 2002. The background of transformational and metalanguage analysis. In *The Legacy of Zellig Harris: Language and Information into the 21st Century. Volume 1: Philosophy of Science, Syntax, and Semantics*, ed. Bruce E. Nevin, 1–15. Amsterdam, Netherlands: John Benjamins.

Jakobson, Roman. 1971. The Kazan' school of Polish linguistics and its place in the international development of phonology. In *Word and Language: Selected Writings, vol 2*, ed. Roman Jakobson, 394–428. The Hague, Netherlands: Mouton.

Lasnik, Howard. 2005. *Minimalist investigations in linguistic theory*. London, UK: Routledge.

Matthews, Peter Hugo. 1993. *Grammatical theory in the United States: from Bloomfield to Chomsky*. Cambridge, UK: Cambridge University Press.

McQuown, Norman A. 1952. Review of *Methods in Structural Linguistics* by Zellig S. Harris. *Language* 28:495–504.

Nida, Eugene A. 1949. *Morphology: The Descriptive Analysis of Words (2nd edition*. Ann Arbor, MI: University of Michigan Press.

Whorf, Benjamin Lee. 1945. Grammatical categories. *Language* 21:1–11.

Whorf, Benjamin Lee. 1956. *Language, Thought and Reality, edited by John B. Carroll*. London, UK: John Wiley & Sons.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. London, UK: John Wiley & Sons.