

語と統語の起源論

—第1部 統語論と形態論の本質—

Theories on the Origin of Words and Syntax

Part 1: The Nature of Syntax and Morphology

ジョン・D・オーエン

John D. Owen

要旨

本稿は2部構成の論文「語と統語の起源論」の第1部に当たる。第1部「統語論と形態論の本質」では、語形成を支配する法則と、言語学的表現で発生する語彙の構成要素と統語的構成要素の接点を検証する。語と統語の起源に関しては、次の2つの相関的問題を提示しなければならないだろう。1. 何が言語発達を左右する法則を決定するのか。2. どのような法則が統語的言語処理の発生において語の配列を決定するのか。第2部の「形態論に関する考察と仮説」では、この2つの問題が実際には別個の問題ではないかもしれないことを検証する。言い換えるなら、おそらくこの議論のさらに適切な提言となるだろうが、語を発生させる文法要素は、語句や文章を発生させる要素とは別個でかけ離れたものなのか、という問題である。さらに進めた議論として主張したいのは、この二分法の認識は、発話の構成要素を作る「言語レベル」を話者が決定し特徴づけることに大きく影響を受けると思われることである。

Abstract

This paper contains the first of a two part discussion of Theories on the Origin of Words and Syntax. In this first discussion, *The Nature of Syntax and Morphology*, we will examine the rules that govern word production and the interface between lexical and syntactical elements, which arises in the occurrence of linguistic expression. We shall address two relative issues in regards to the origin of words and syntax: 1. What defines the rules that govern the development of words? 2. What rules determine the sequence of words in the production of syntactic computations? In the forthcoming second discussion, *Debates and Suppositions of Morphology*, we will examine arguments that these two questions may not actually be distinct issues. In other words, perhaps a more relative line of questioning for our discussion might be to ask, are the components of grammar that produce words distinct and strictly separate from the components that produce phrases and sentences? Our perception of this dichotomy, as arguments put forth shall attempt to assert, seems largely to depend on where we determine and distinguish the *linguistic levels* we assign to the elements of utterances.

●キーワード: Linguistic Level, Transformations, Deep Structures, Surface Structures

The central notion in linguistic theory is that of
“linguistic level.”

— Chomsky

Introduction

Words are by common recognition the building blocks in the expression of meaning, or, more specifically, they are the observable elements of meaning that propagate in strings of phrases and sentences. The words of any particular language, taken in their entirety, constitute that language's lexicon or total inventory of *morphemes*, which are the smallest lexical elements of

meaning within a particular language. Management of this inventory system is enabled by a separate filing scheme of *phonemes*, labeling each morpheme with discrete and distinguishable sounds. In this sense linguistic elements can be seen as reducible within a hierarchy of meaning. We can view *phrases* and *sentences* as the overarching expressions of syntactic meaning, which are constructed from sequences of *words*. In turn, each word is itself a complex unit of lexical meaning, composed of combinations of *morphemes*, which circumscribes lexical meaning at simplex. Though meaning is irreducible beyond the

morphemic level, individual morphemes are subject to demarcation by a contrastive coding of distinguishable sounds, which we call *phonemes*. As such, *phonemes* represent the concise inventory of distinct vowel and consonant sounds unique to each language.¹⁾ These sounds, while themselves vacuous of meaning, in effect function as inventory managers of meaning by mapping sounds onto other lexical elements. In summary of this description, 'words' are units of meaning that operate in a syntactic environment of phrases and sentences; their meanings, however, are the composition of smaller more reducible lexical elements. That is to say, in appearance, words have both a syntactical and lexical nature.

Words are "atomic" at the level of phrasal syntax and phrasal semantics. The words have "features," or properties, but these features have no structure, and the relation of these features to the internal composition of the word cannot be relevant in syntax. – this is the thesis of the atomicity of words, or the lexical integrity hypothesis, or the strong lexicalist hypothesis (as in Lapointe 1980), or a version of the lexicalist hypothesis of Chomsky (1970), Williams (1978;1978a), and numerous others.

Di Sciullo & Williams, 1987, p.49

In the disciplines of physics and chemistry, atoms are observed to be made up of smaller subatomic particles of matter. However, for practical purposes of identity, we classify the observable elements of matter, as found in periodic tables, by their atoms not by the composition of particles within atoms. By analogy to syntactic theory, for practical purposes of identifying meaning at the phrasal and sentential level, words are our lowest level of concern. Likewise, development of meaning at the word level is thought to ascend from the arrangement of sub-particles that constitute words, rather than descend from expressions of syntax that are composed of words.

Levels of Linguistic Development:

phonemes→morphemes→words→phrases→sentences

In the traditional development of Structuralist Morphology, at least since Saussure, morphemes have long been taken for granted to be the atoms of words. Moreover, word development was widely believed to be of the same nature as syntactic development. That is, words are discrete structures of meaning developed by the same rules of grammar responsible for the production of phrases and sentences. This would have seemed to be an undisputable understanding of the nature of both the development of syntax and morphology. However, this view came under palpable scrutiny in the second half of the twentieth century, when the linguistic level at which meaning is determined was brought into question.

If we understand words to be a distinct inventory of readily accessible meaning, we can then also hypothesize that development of such meaning may be influenced by distinct rules. In other words, if we view words as atoms of meaning, we can thus view everything to the left of words as smaller sub-particles of such meaning and everything to right of words as larger syntactic expressions of meaning. In this sense we define words as atomic elements working in a syntactic environment, and we are defining their morphemic constituents as the subatomic particles of syntactic atomicity. The linguistic level at which we place atomicity, or so it would seem, directly reflects how much access we determine the syntax of sentences to have on the morphology of words.

Discussion

1. The Nature of Syntax and Morphology

One of the central issues in relation to morphology-syntax interaction is the determination of what, if any, morphological operations take place at the syntactic level and what morphological operations take place at the lexical level. This problem, of where to mark the parameters for linguistic levels of morphology, is perpetually deliberated with vigor throughout the literature, both implicitly and explicitly. However, its emergence was neither precipitous nor dramatic. Rather, it culminated through a series of observations over a period of roughly two decades, spawned from three seminal publications by Chomsky, *Syntactic*

Structures (1957), *Aspects of the Theory of Syntax* (1965) and *Remarks on Nominalization* (1970).

The strongest possible proof of the inadequacy of a linguistic theory is to show that it literally cannot apply to some natural language. A weaker, but perfectly sufficient demonstration of inadequacy would be to show that the theory can apply only clumsily; that is, to show that any grammar that can be constructed in terms of this theory will be extremely complex, ad hoc, and 'unrevealing', that certain very simple ways of describing grammatical sentences cannot be accommodated within the associated forms of grammar, and that certain fundamental formal properties of natural language cannot be utilized to simplify grammars. We can gather a good deal of evidence of this sort in favor of the thesis that the form of grammar described above, and the conception of linguistic theory that underlies it, are fundamentally inadequate.

Chomsky, 1957, p.34

Syntactic Structures (1957) was essentially a single long argument, based on available descriptive methods, for making determinations about linguistic levels and predictions of linguistic outcomes. An early observation Chomsky makes is that the interdependence between words is not reliant on their individual adjacency to one another, but rather the grouping of their phrasal constituency. Thus, a descriptive grammar that stipulates sentences to be merely strings of consecutive words, without recourse to the wider context of its phrasal structure and constituents, fails to adequately give an explanation of a language. What makes words observably consequential to sentences is not their connection to each other, but rather their relationship to the syntactical structures to which they are grouped, in larger phrasal units. Chomsky's contribution was not simply to add further accounts to the prescriptive or descriptive observation of phrasal rules, but rather he proposed that the phrase structure rules (as well as the lexicon) must be generated in what he termed the *Deep Structure* of sentences. Moreover, his investigation of

these rules revealed patterns of movement among the components of the said structures, what Chomsky would eventually term the *Transformational* component. Chomsky argued that a descriptively adequate grammar must, among other things, define a concept of grammatical identity, capable of syntactically mapping these deep structure movements onto the *Surface Structures* of sentences. Such a notion entails a capacity to recognize each and every sentence by a structural description, which correctly reflects the native speaker's knowledge of a given syntactic structure, as demonstrated by their performance. He delineated a number of such structures, among which his description and analysis of English base and auxiliary verbs not only brought him to prominence, but launched a revolution in the way syntactic structures are analysed.

Chomsky was the first to properly map the movement or structural transformation of the auxiliary and base verb interface, commonly known today as *affix-hopping*. Chomsky's basic insight was that no matter how complex a particular English verbal expression, every auxiliary verb determines the form of the verb that succeeds it. Thus, if we compare *John has read the book* with *John has been reading the book*, the auxiliary verb *have* in both sentences is followed by a past participle form; simply put, the auxiliary verb *have* must be followed by a past participle. Chomsky accounted for this correlation, between an auxiliary and whatever follows it in two steps: it begins with a phrase structure rule that pairs each auxiliary with an abstract affix (*have-en, be-ing*, etc.), and then undergoes deep structure movement that switches each abstract affix with the succeeding verb, resulting in a transformed verb-suffix combination.

Specifically, for utterances utilizing the auxiliaries *BE* and *HAVE*, Chomsky revealed this to be only their surface appearance. At a deeper linguistic level of syntax, below their surface appearance, they actually exist in the more complex composite form of *BE -ING* and *HAVE -EN*. However, the *-ING / -EN* endings, after incurring separation from their auxiliary stem,

are subject to compulsory movement (transformation) to a position behind the base verb, at which point they reaffix themselves to the base verb as the suffix -ING or -EN. Thus, AUX -> (BE -ING) (HAVE -EN) VERB, as illustrated in figure 1.

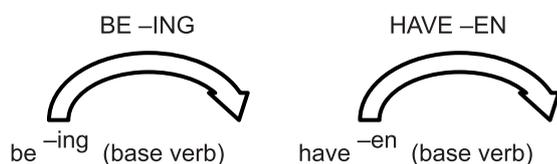


figure 1

Chomsky's hypothesis, of the transformations the auxiliary verb's abstract affixes with the succeeding base verb, proved able to predict precisely why the base verb always receives the correct form of suffix. The account explained that the form of suffix a base verb requires is determined by its preceding auxiliary. Additionally, in order to secure this deep structure sequence, the first element (i.e. the auxiliary) of transformation will always identify the tense marker. For example: *John has eaten* (PRES marking on *have=has*). *John had eaten* (PAST marking). Along these same lines of analysis, modal verbs (*can, may, must, etc.*) can be seen to always be followed by a base verb form: *John may win.* vs. *John may have won.* vs. *John may be winning;* and the progressive auxiliary *be* must be followed by a present participle. Again Chomsky was able to diagram this correlation in two steps, first with a phrase structure rule that pairs each auxiliary with an abstract affix (*have-en, be-ing, etc.*), and then with a transformation that switches each abstract affix with the following verb, resulting in a verb-suffix combination, thus, providing a structurally adequate description of grammar at an identifiable linguistic level. Chomsky's observation is in summary: AUX -> TNS MKR (MODAL) (HAVE -EN) (BE -ING) (BE -EN) VERB.

Such insights, into the protocol of auxiliaries, work well to orchestrate Chomsky's overall achievements brought to bear in *Syntactic Structures*. Among which, he advanced a demonstrative understanding that a word, by its nature, is not necessarily restricted to a

standalone lexical identity existing at the surface or base level of a sentence. Rather, it is simultaneously a dynamic entity, with a particular function, operating at a deeper linguistic level. As evidenced above by the auxiliary *be*, which while discernably present at the surface of an utterance, it is in reality of a complex two-part nature subjected to transformations that also functions at a (previously uncharted) deep linguistic level (i.e. the transformative: BE -ING). Likewise, the auxiliary *have* is, in reality, a complex word of a similar quality (as in the transformative: HAVE -EN). Furthermore, the deep structure transformation of auxiliaries causes them to constrain the words that they precede on the base level. Though some of the theorizing of *Syntactic Structures* would inevitably not stand the test of time, the notion of "deep structures" as we shall see has had lasting implications.

Prior to the publication of *Syntactic Structures* linguists regarded the purpose of their discipline as merely being the classification of the elements of human languages. Linguistics was seen as nothing more than a classificatory science, more or less analogous the taxonomy of plants. The traditional approach, of what fell under the rubric *structural linguistics*, had been dictated to the collection of the systematic facts encountered in a language, i.e. a corpus comprising phonemes (basic sounds), morphemes (basic units of meaning), varying types of phrases and sentences. Chomsky more ambitiously sought to construct a theory with a testable hypothesis for predicting and explaining the composition and purpose of every meaningful expression of possible construction. In *Aspects of the Theory of Syntax* (1965), henceforth referred to as *Aspects*, Chomsky sought to apply the premises empowered in *Syntactic Structures*, to define and describe a complete grammar of language. As he delineates below, this feat was attempted by subdividing the grammar into three parts, a *syntactical* component that generates and describes the internal structure (of the infinite number of possible sentences the language is capable of producing), a *phonological* component that describes the sound structure) of the sentences generated by the syntactical component)

and a *semantic* component that describes the meaning structure of the sentences.

A grammar contains a syntactic component, a semantic component, and a phonological component. The latter two are purely interpretive; they play no part in the recursive generation of sentence structures. The syntactic component consists of a base and a transformational component. The base, in turn, consists of a categorical subcomponent and a lexicon. The base generates deep structures. A deep structure enters the semantic component and receives a semantic interpretation; it is mapped by the transformational rules into a surface structure, which is then given a phonetic interpretation by the rules of the phonological component. Thus the grammar assigns semantic interpretations to signals, this association being mediated by the recursive rules of the syntactic component. The categorical subcomponent of the base consists of a sequence of context-free rewriting rules. The function of these rules is, in essence, to define a certain system of grammatical relations that determine semantic interpretation, and to specify an abstract underlying order of elements that makes possible the functioning of the transformational rules.

Chomsky, 1965, p.141

For our purposes, what is momentous about Chomsky's grammar model is that it requires the partitioning of syntax into two corresponding linguistic levels, a *base* and *transformational component*. Further still, the base is segmented into two subcomponents, one of which provides for a lexicon. The net effect is that the lexicon became expressed as, with some (undefinable) degree of distinction, segmented from the *transformational component* of syntax. This determination of distinct linguistic levels between syntax and lexicon, as we shall observe, has had lasting consequences. Such categorizations, as outlined in the model above, were brought about in part as response to a challenge by Fodor and Katz (1964). "As a rule, the meaning of a word is a compositional function of the meanings of

its parts, and we would like to be able to capture this compositionality" (p. 501). By compositionality, Fodor and Katz asserted what needs to be understood was what they termed the *Projection Problem*, which translates as the extraction of the meaning of a sentence from the meanings of its parts. Among Chomsky's innovations in *Aspects*, two that are central to the development of lexical semantics are the syntactic innovations that led to "deep structure" and the incorporation of the Katz-Postal hypothesis, in which transformations function to preserve meaning. The inevitable necessity of this will become self-evident in the examples below.

Chomsky's response to the Projection Problem was to devise a grammar that, in addition to syntax, phonemes and semantics, also entails a transformation component. In summary, the syntactical component generates and describes the internal structure of all the sentences that a language can generate, the phonological component describes the sound structures of the sentences that are generated and the *semantic* component designates the meaning of the sentences. However, Chomsky further conceived that the *syntactical* component must comprise two subcomponents, the *base rules* (i.e. the surface structure) and *transformational rules* (i.e. deep structure). The base component of Chomsky's grammar contains the phrase structure rules, which enable the *deep structure* of each sentence. In turn, the transformational component transfigures the deep structure of a sentence back into the corresponding *surface structure*.

The locus of a traditional structural grammar is the syntax; the phonology and the semantics are purely interpretative modules, in the sense that they describe the sound and the meaning of the sentences produced by the syntax but do not generate any sentences themselves. Chomsky identified several problems with such a grammar model.

Firstly, there is no natural way to describe the ambiguities of a sentence such as "I like her cooking." Phrase structure rules alone provide only one extraction

for this sentence resulting in a *projection problem*. The sentence, being compositionally ambiguous, entreats the grammar to reflect that fact by providing contrasting syntactical derivations and hence applicably alternative syntactical descriptions.

Secondly, phrase structure grammars have no way to depict and account for any of the possible differences in meaning of such an ambiguous sentence. To illustrate this point Chomsky offered two (now well-renowned) sentences: (1) *John is easy to please*. (2) *John is eager to please*. Both sentences have precisely the same phrase structure sequence of a noun · copula · adjective · infinitive verb, however, the grammar of the two is decidedly dissimilar. In the first sentence, though it is not apparent from the surface word order, “*John*” functions as the direct object of the infinitive verb “*to please*”; meaning: “*it is easy for someone to please John*”. Whereas in the second “*John*” functions as the subject of the verb *to please*; the sentence means: *John is eager that he please someone*.

Thirdly, just as in the above example, the surface structure of a sentence may not reflect its deep structure and the deep structure may not have a direct manifestation in the speech signal of the surface. There is no available way to mark between such phrases and explain these facts with merely phrase structure analyses, moreover, any assumptions of language would be incomplete without providing an account of these underlying structural ambiguities.

Let’s consider in more detail the example, from above: “*I like her cooking*”. We observe that what we have is not just one marked phrase but several different underlying sentences each capable of being marked with a different meaning. Hence, underlying this one sentence “*I like her cooking*” are the analogous phrases “*I like what she cooks*,” “*I like the way she cooks*,” “*I like the fact that she cooks*”. Further still, the underlying meanings of the latter two could provide the phrase marker: “*I like it that she is being cooked*”. The ambiguity in this sentence requires the grammar to account for some number of

varying phrase markers. Thus, Chomsky compels us to conclude, without evidence to the contrary forthcoming, that different phrase markers (i.e. meaning) produced by the phrase structure rules are transformed into identical surface structure phrases by the application of the transformational rules at some deep linguistic level.

Chomsky’s examinations, of these sorts of ambiguous morpho-syntactical interactions, brought into question the disparities between the morphological operations that take place in the *syntax* and the morphological operations that take place in the *Lexicon*. Moreover, it was not until a *Lexicon*, as a separate component of the grammar, was introduced in the theory that such inquiries could even arise in this form. Before *Aspects* derivational and inflectional morphology was done necessarily in the *syntax*. However, the incorporation of an independent *Lexicon*, in what would eventually be called “The Standard Theory”, did not immediately bring about a change concerning the way in which morphological operations were understood.

The *Lexicon* that Chomsky expressed in *Aspects* was basically an inventory of lexical items, with their (idiosyncratic) phonetic, semantic and syntactic properties. Moreover, lexical items were identified for inherent and non-inherent features by means of syntactic transformations. While Chomsky suggested that verbs like *refuse* and *destroy* should appear as lexical entries in the *Lexicon* independent of syntactical effects, he made no such assertion about the derivational morphology of corresponding nouns like *refusal* and *destruction*. Instead they were seen as derived through nominalization rules of the *syntax*, in which as the lexical elements (verbs) *destroy* and *refuse* they were converted to (nouns) *destruction* and *refusal*.

This introduction of a *Lexicon*, separate and distinct from rewriting rules of *syntax*, however, paved the way for reanalyzing how word formation processes, and morphological operations in general, were to be approached in subsequent works in early *generative grammar*. It was not until *Remarks on Nominalization*

(Chomsky, 1970), hence forth referred to as *Remarks*, that lexicalist approaches to the morphology-syntax interface started to emerge. In *Remarks*, Chomsky argued that some derived nouns like *destruction* were derived lexically, rather than transformationally developed. That is, in order for a process to be considered transformational, it must be readily identifiable and invariably productive. However, if we suppose the operations between linguistic forms are irregular and idiosyncratically unproductive, the elements involved should be considered a product of the Lexicon.

Chomsky cited the relatively unproductive affiliation between many verbs and their derived nouns, as well as the semantic idiosyncrasies between gerundive nominalizations (GNs) and derived nominalizations (DMs). GNs, as illustrated below in sentence (1), are far more productive, regular and predictable than DMs, as in sentence (2). Whereas, the former is derived by the application of highly unvarying syntactic transformations (i.e. the affixing of *-ing*), the latter is derived without uniformity and by inexplicable means. Such lexical items must therefore be considered a product of the Lexicon rather than transformationally derived through the syntax.

- (1) Mary's giving a book to Ann
- (2) Mary's gift of a book to Ann

This notion that some nominalizations of derivational morphology are lexical rather than syntactical came to be known as the *lexicalist hypothesis*. From this line of inquiry two conceptually different theoretical positions of derivational processes developed: The Weak Lexical Hypothesis (Allen 1978, Siegel 1974, Aronoff 1989, Anderson 2000, Spencer 2005) contends most DM's are lexically derived, but maintain that inflection is a morphological component contained in syntax. The Strong Lexical Hypothesis (Jackendoff 1975, Lapointe 1985, Selkirk 1982, Kiparsky 1983, Scalise 1988, Lieber 1992) defines inflection and derivation, as well as many phonological rules, to be lexically determined and wholly separate from syntax.

Conclusion

In summary, interface between lexical and syntax elements is based in part on two arguments. First, derived nominals do not exhibit a behavior that differs in any sense from that of gerundive nominals, and therefore they would not need to appeal to transformational mechanisms to explain anything about their syntactic importance. The second is that nouns possessing a superficial morphology, functionally identical to nominals derived from verbs, need in principle a comparable account development. For example, nouns such as *eagerness* in standard nominalization would entail the existence of an abstract verb, such as to *eager*. In the absence of such a verb, how do we account for such nouns? Chomsky provided an example of how a lexicalist approach to nominalizations in which syntactic elements play a role can be used to predict the behavior of nominalized nouns, without resorting to the transformational component by means of computational rules. Otherwise, word development must be seen as proviso of verb normalization in a closed system. The insights gained from these three publications are not stipulations of analysis or even theory, but rather the nature of what can be described. These principals and the methods Chomsky set down have turned out to provide deeper understanding into the internal structures of words, especially derivational morphology, which in many respects is quite distinct from the descriptions that apply to syntactic rules. In comparison to syntax, morphology appears to be as much about exceptions as about rules. This does not mean that it is not amenable to rigorous inquiry, only that it cannot be described in the same sorts of patterns that can be mapped in syntax. As such, morphology is concerned with the composition of words, while syntax is concerned with the combinations of words.

Notes

- 1) The number of phonemes can vary greatly among languages, while some Slavic language operate with more than 50 explicit sounds, Standard Japanese functions efficiently with only 22 phonemes.

References

- Allen, Margaret R. 1978. Morphological investigations. Ph.D. thesis, University of Connecticut.
- Anderson, S. R. (2000), "Some Lexicalist Remarks on Incorporation Phenomena", in B.Stiebels & D.Wunderlich, *Lexicon in Focus* (Studia Grammatica 45), 123-142
- Aronoff, M. (1989), "Two Senses of Lexical", in Proceedings of ESCOL '88, Ohio State University, CLC Publications
- Chomsky N (1957) *Syntactic Structures*. Mouton, The Hague
- Chomsky N (1965) *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA
- Chomsky, Noam. 1970. Remarks on nominalization. In R. A. Jacobs and P. S. Rosenbaum, eds., *Readings in English transformational grammar*, pages 232-286. Waltham: Ginn and Company
- DiSciullo, Anna Maria, and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.
- Jackendoff, R. (1975) "A system of semantic primitives" in Schank & Nash-Webber (eds.) *Theoretical Issues in Natural Language Processing*~ (Cambridge, Mass.: BBN
- Kiparsky, P. (1983), Word formation and the lexicon, in: Ingemann, F. (ed.), *Proceedings of the 1982 Mid-America Linguistics Conference*, University of Kansas, pp. 3-32.
- Lieber, R. (1992), *Deconstructing Morphology*, University of Chicago Press, Chicago
- Scalise, S. (1988), Inflection and Derivation, in "Linguistics", 26, 561-581
- Selkirk, E. O. (1982), *The syntax of words*, Cambridge (Mass.), MIT Press
- Siegel, Doris. 1974. *Topics in English morphology*. Ph.D. thesis, MIT. Ph. D. thesis, MIT.
- Spencer, A. (2005), *Word-Formation and Syntax*, in P. Štekauer and R. Lieber (eds.), *Handbook of Word-Formation*, Amsterdam, Springer.