

The Generative Word

Hagit Borer, QMUL

1 Querying Words

1.1 Introduction

Consider what speakers of English know about the word *head*. Beyond its pronunciation, or phonological form (PF) /héd/, they know its various meanings. They also know how to combine it with other words, that is, they know how this word relates to its larger syntactic context. One way to encode this knowledge is with categorial labels: the PF /héd/ in a *noun* (N) syntactic structure means *BRAIN-CONTAINING BODY PART*. Of course, this is not the only possible meaning of /héd/-N. Additional established meanings, possibly of a metaphoric origin, would be *TOP*, *BRAIN*, *LEADER*, etc.

The form /héd/ may also occur with the syntactic structure *verb* (V). Thus consider (1a-c):

1. a. Kim headed the team
- b. Kim headed toward the team
- c. *Kim headed

While /héd/ is V in each of (1a-c), any English speaker would know that *head* may denote *LEAD* in (1a) but not *ADVANCE*. The converse applies in (1b) where *head* may denote *ADVANCE*, but not *LEAD*. Thus in the syntactic context of a direct object, *the team*, /héd/-V may only mean *LEAD*, while in the context of a directional expression (*toward*, *away*, *home*), /héd/-V may only mean *ADVANCE*. Either way, the verb must have *some* complement, or ungrammaticality results, as in (1c). Any attempt to represent knowledge of the word *head*, then, must include some relatively rich syntactic information about the contexts in which it may occur with particular meanings.

Consider now the occurrence of *head* within (N-N) compounds and derivatives. In English, both have quite well-defined syntactic properties. Compounds such as *headgear* or *towel rack* combine two nouns to yield another noun with a meaning composed of its parts (*GEAR* for *HEAD*(s); *RACK* of *TOWEL*(s)). Within derivatives, a fairly productive structure combines nouns with the suffix *-y* to give rise to an adjective denoting a property related to that noun (*dirt-y*, *brain-y*, *silk-y* etc.) These preliminary generalizations are structurally represented in (2):

2. a.

N	
/	\
N1	N2
<i>head</i>	<i>gear</i>
<i>towel</i>	<i>rack</i>
- b.

A	
/	\
V	-y
<i>dirt</i>	
<i>brain</i>	

A closer look, however, reveals that matters are considerably trickier. While the meaning of *headgear* emerges from its discrete parts, that is not the case for *headway* with the meaning *PROGRESS* (and compare with the meaning of *head away*). Similarly, the meaning of *bulkhead* cannot be computed from composing any of the synchronic meanings of *bulk* (N) or *head* (N). A similar case holds for derivatives. Combining *head*-N with *-y* gives rise to *heady*, meaning *GIDDY* or *INTOXICATING*, which is not related to any of the established meanings of *head*-N (and compare with compositional *brainy*.)

Clearly, context remains crucial for the interpretation of non-compositional compounds and derivatives. That context, however, is not semantic: while *bulk* is crucial for the meaning *BULKHEAD* to emerge, the synchronic meaning *BULK* plays no role in it. Nor is the context syntactic. Like *towel rack*, *bulkhead* consists of N+N and could fit into (2a). Similarly, *heady* as N+y would fit into (2b), with little *syntactic* difference from *dirty* or *brainy*. Since the meaning of *heady*, *bulkhead* or *headway* cannot be predictable from their respective parts, that meaning must be independently *listed* in speakers' mental word list, their *lexicon*. Insofar as the meaning of *towel rack* or *brainy* is predictable from the already independently listed *towel*, *rack* or *brain*, separate listing may not be needed. Not so, however, for *heady* or *headway*.

But if *heady*, *bulkhead* or *headway* alongside many other compounds and derivatives constitute independent lexical entries, effectively independent words, what is the relationship between such words and *head*, itself already (at least one) independent word? Are we justified in claiming that the English word *head* (N) is an actual grammatical part of either *heady* or *bulkhead*?

One conclusion is inevitable from this brief exposition. If the term *word* is to be used within a rigorous, explanatory grammatical theory, it is necessary to have a theoretically grounded understanding of its phonological, syntactic and semantic combinatorial properties, while leaving room for at least some degree of listedness for simple as well as for complex formations such as compounds and derivatives. Theoretical approaches to words and to the lexicon, and the scientific debates surrounding them have shaped major developments as well as major controversies within the generative tradition, some of which will be reviewed below.

2 The birth of the syntactic lexicon: Chomsky 1965, 1970

From the onset, an important aim within generative grammar has been to formulate syntactic operations that allow for maximal formal simplicity and generality. The lexicon, in turn, became the repository of all information not otherwise predictable from formal properties of the system. A prime example is PF, clearly a *listed*, *arbitrary* and *learned* property of utterances. In his 1965 *Aspects of the Theory of Syntax* (henceforth *Aspects*) Chomsky also undertook to list the unpredictable syntactic contexts in which a word could be placed.

Recall that *head-V* can occur in the two distinct syntactic contexts in (1a-b). Alongside *head*, consider the verb *think* which may have a sentential complement (impossible for *head*), and the verb *read*, which unlike *head* allows the omission of the object. Before the augmented lexicon of *Aspects*, these differences were represented by means of a cumbersome rewrite system that broke the class of Vs into subclasses, as in (3), where the formal commonality of V_T and V_P is obscured, redundant categories abound, and descriptive adequacy is compromised:

3. $VP \rightarrow V_T NP$
- $VP \rightarrow V_I$
- $VP \rightarrow V_D PP$
- $VP \rightarrow V_P S$
- $V_T \rightarrow head, think, read$
- $V_I \rightarrow read, think$
- $V_P \rightarrow think, read$
- $V_D \rightarrow head$ (T=transitive; I-intransitive; D=directional; P=propositional)

Following *Aspects*, general rewrite rules as in (4) became possible:

4. a. $VP \rightarrow V (NP) (S)$
- b. $V \rightarrow think, head, read$

These rewrite rules are a major improvement: they capture the fact that *head*, *think* and *read* are all instances of V and dominated by a VP with an identical syntactic distribution. There is a cost, however. Eliminating lexical terms from the rewrite component requires a distinct formal mechanism, *Lexical Insertion*, which matches a lexical item already listed with its syntactic context, with the appropriate syntactic structure generated by the general rewrite rules in (4). Within such a system, the syntactic *insertion frame* (or *subcategorization*) of *read* would look like (5), where the underlined space indicates where *read* is inserted relative to its obligatory syntactic context:¹

5. *read*: [+V, +[] {NP,∅}], +READ, +/ ríd/]
- |_____↑

From the perspective of this approach, the multiple instantiations of the PF /héd/ are best conceived as separate lexical entries. While their PF is shared, nothing much else is, for even the verbal instantiations share little beyond the V label. A disjunctive set of brackets could be put around the relevant insertion frames, but that would only obscure what is endemic in lexical representations. Redundant rewrite rules in (3) are eliminated, but at the cost of strengthening a linguistic component, the *Lexicon*, which systematically values the particular over the general:

6. a. *head1*: [+V, +[] NP], +LEAD, +/héd/]
- b. *head2*: [+V, +[] DIR], +ADVANCE, +/héd/]
- c. *head3*: [+N, +count, +BODY PART, +/héd/]
-

Chomsky explicitly endorses the point, noting that while the set of possible insertion frames is limited by universal principles, the association of syntactic properties with any particular PF-meaning pair (e.g. /héd/-LEAD; /héd/-BODY PART) cannot be expected to display regularities. Nothing about either /héd/ or the meaning LEAD can predict an obligatory direct object or absence of a sentential complement. Nor are *head*'s meanings LEAD or ADVANCE predictable from the insertion frames in (6a-b). These meanings and their syntactic insertion frames must be listed as independent, unrelated properties:

The lexicon consists of an unordered set of lexical entries and certain redundancy rules. Each lexical entry is a set of features...Some of these are phonological features, drawn from a particular universal set of phonological features.... Some of the set are semantic features. These, too, are presumably drawn from a universal "alphabet", but little is known about this today, and nothing has been said about it here. *We call a feature 'semantic' if it is not mentioned in any syntactic rule, thus begging the question of whether semantics is involved in syntax.*^[15] The redundancy rules of the lexicon add and specify features wherever this can be predicted by general rule. *Thus the lexical entries constitute the full set of irregularities in the language.* (Aspects: 142, emphasis added)

Chomsky's 1970 "Remarks on Nominalization" (henceforth Remarks) renders Aspects's lexicalist agenda considerably more powerful. Remarks postulates as 'lexical' not only properties of listed items, but certain aspects of relatedness, thereby substituting listing not only for phrase structure rewriting rules, but also for what, at the time, was commonly assumed to result from transformational rules. Remarks was a detailed study of the syntactically predictable (by

¹ (5) combines proposals in Aspects with some subsequent modifications.

The set of featural properties built into the representation in (5) and similar can be translated, almost without modification, into the system of features proposed within the Minimalist Program (Chomsky 2000, 2001 and much subsequent), where a verb such as transitive *head* would be listed with a feature that would require the syntactic presence of a nominal direct object.

assumption transformational) when contrasted with the syntactically unpredictable (by assumption lexical). Its ultimate aim was to explicitly exclude from syntax all 'relations' that require access to lexically specified, item-specific information. As such, the agenda questions the *grammatical* relatedness of items that can only be related to each other by making reference to item-specific unpredictable information.

Consider briefly the specific contrasts Chomsky studies. In (7) we have sentences encoding verbs with their direct objects or direct *arguments*. In (8), we have the correlating *verbal gerunds*:

7. the scientist knew the solution
the enemy has destroyed the city
the builder enhanced the foundations
8. (the scientist) knowing the solution (*verbal gerunds*)
(the enemy) having destroyed the city
(the builder) enhancing the foundations

To characterize the relationship between the sentences in (7) and the gerunds in (8), we could postulate a transformational rule that maps the former to the latter. Such a rule must accomplish the following:

9. a. Eliminate Tense
b. Make the subject optional
c. Add *-ing* to the verb (or highest auxiliary)

There are no other changes. (9) makes no reference to the insertion frame of the target verb, to its meaning or to its PF, correctly predicting that the transformation generalizes to all sentences containing a verb or an auxiliary.

Contrastively, consider the relationship between the sentences in (7) and the nominals in (10), all plausibly derived from verbs (deverbal nominals, DN), some with an *-ing* nominal ending:

10. (the enemy's) destruction (of the city)
(the student's) perception (of her school)
(the scientist's) knowledge (of the solution)
(the builder's) enhancement (of the foundations)
(the building's) transformation (of the landscape)
(the politician's) reading (of his defeat)

While arguably *destruction*, *enhancement*, *read* etc. are related to the verbs *destroy*, *enhance*, and *read*, respectively, correctly characterizing an operation that would transform these verbs to DN requires reference to *listed* properties of the verb. Thus at least the following item-specific properties need to be considered when we form the DNs in (10):

11. Phonological unpredictability:
 - a. Item-specific choice of nominalizer (note that *-ing* ending is always possible):
destroy → *destruction*
prove → *proof*
enhance → *enhancement*
know → *knowledge*
 - b. *item-specific stem allomorphy*:
destroy ⇔ *destruction*; *(to) *destruct* **destruction*
perceive ⇔ *perception*; *(to) *percept*; **percivation*

12. Semantic unpredictability:

- a. DNs may have meaning unrelated to the source verbs:
transformation (technical grammatical meaning); *proofs*; *reading* (=INTERPRETATION) (as well *recital*; *transmission*, and many others);
- b. Even 'predictable' meanings (e.g. *destruction*) are ambiguous between action and result readings, with the latter unavailable for sentences or gerunds.
- c. Source verbs are not independently attested altogether (*vision*, *fiction*)

In addition, and unlike gerunds, DNs have properties that are difficult to reconcile with a syntactic VP source, but are expected if they are (underived) nouns:

- 13. a. Objects are optional in DNs where obligatory for sentential VP and gerunds
- b. (Logical) objects marked with *of*, on a par with nouns (and unlike gerunds)
- c. Adverbs, possible for gerunds, are barred in DNs

As Chomsky notes, syntactic transformations that affect VPs within both sentences and gerunds may be barred in DNs. In and of itself, this does not require consulting item-specific information, but it supports the absence of a sentential-type VP within DNs when contrasted with verbal gerunds. One such case is illustrated in (14):

14. Datives		Double Object Constructions:	
Verena gave a book to Laurie	↔	Verena gave Laurie a book	
Verena giving a book to Laurie	↔	Verena giving Laurie a book	<i>verbal gerund</i>
Verena's gift of a book to Laurie	!!	*Verena's gift (of) Laurie of a book	<i>DN</i>

Chomsky concludes that while gerunds are transformationally derived from sentences, DNs constitute an independent, underived lexical entry. By extension, as derivatives and compounds may allow unpredictable meaning and item-specific phonological properties, they must all be lexically listed independently.

The Lexicon with this augmented role soon led to a wide range of investigations focusing on the syntactic, semantic, phonological, and morphological properties of words, and the results of these investigations continue to have significant consequences for generative theories and for our understanding of the human linguistic faculty. Some of these results are summarized in the next four sections. Section 7 turns to more recent theoretical developments designed to shift syntactic power away from the lexicon and back into syntax.

3 Lexicalism

Aspects and Remarks introduced a notion of relatedness that was based on information stored with individual lexical items, and conceived of as non-generative. Setting aside Generative Semantics (which summarily rejected the move) it is interesting to note that some scholars that adopted it, in some form, eventually abandoned the generative tradition altogether as a consequence.² Once a system is in place that allows large, cognitive storage of syntactic representations as in (5)-(6), even if such representations are universally constrained, an almost inevitable consequence is the emergence of suggestions that such listed storage alone accounts for the human syntactic capability. The temptation is even greater if the syntactic properties of these representations are explicitly divorced from the emergence of their meaning. The most influential (and lasting) of these approaches is *Construction Grammar* which replaces the lexicon with a *Constructicon* - a list of constituent-structure strings and frames, potentially as large as a sentence

² For an insightful review of the debate with Generative Semantics see Newmeyer (1980).

but at times populated by shortish idiomatic expressions (e.g. *let alone*), and associated with holistic non-compositional meaning beyond which there is little need, if any, for derivational or compositional operations (see in particular Fillmore et al, 1988; Goldberg 1994 and subsequent).

Within approaches that maintained generative, hierarchy-building combinatorial operations, much of the research from the mid-70's onwards focused on the division of labour between the lexicon and the syntax. The emerging consensus in the early 80's is probably best characterized by Wasow's (1977) seminal work on adjectival and verbal passives, where he proposed the following typology for lexical vs. syntactic operations:

15.	<i>Lexical Rules</i>	<i>Transformations</i>
A	do not affect phrase structure	may alter the output of phrase structure rules
B	may change categorial labels	do not change category labels
C	are 'local' – involve only material specified in the insertion frame (e.g. arguments such as subject, object etc.)	need not be 'local'; formulated in terms of structural properties of phrase markers
D	apply before any transformations	may follow (other) transformations
E	may have idiosyncratic (listed, item-specific) exceptions	have few or no true exceptions

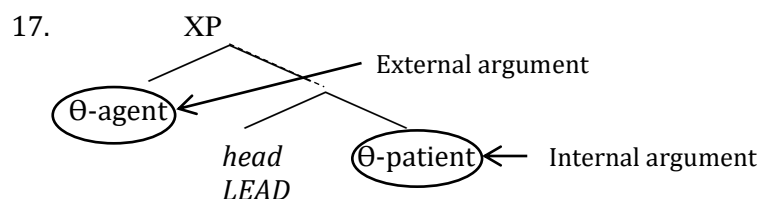
Wasow's typology, with surprisingly minor modifications, survived the numerous adjustments that carried the Standard Theory into the Extended Standard Theory, then to Government and Binding (GB) and finally Minimalism (including the overwhelming majority of Chomsky's Minimalist writings with relatively minor potential departures in Chomsky 2013 and later). Approaches to what lexical entries are have changed, and perspectives on the operations that may impact lexical entries have changed as well. Within syntactic approaches, the term 'transformation' with its implication of construction-specific structural change (e.g. Passive Transformation, Topicalization Transformation) has, by and large, fallen out of use, replaced by more general 'Move' or 'Merge' that apply uniformly to construction-neutral structures. Yet, the partition of labour between the lexical and the syntactic has survived, obeying largely the edicts postulated by Wasow: lexical operations are structure preserving and local, may result in the change of categorial labels and may incorporate item-specific exceptions, all properties that, by prevailing assumptions, cannot involve syntactic operations.

4 The Government-Binding Model and Lexical Semantics

With operations contingent on properties of listed items removed from syntax, syntactic research in the 70's turned its attention to long distance dependencies. Alongside these investigations, however, a growing number of generativists turned their attention to facets of lexical items that were no longer deemed 'syntactic' in and of themselves, but which clearly interact with the syntax through syntactic features and lexically-specified syntactic insertion frames. We thus find increasing interest in properties of verbal complements, argument structure and linguistically-expressed events, assuming these are all informed by the listed properties of words in general and verbs in particular. By the early 80's these research foci play a major role not only within purely lexicalist approaches, such as Lexical Functional Grammar, but also within GB, taking

on board the results of extensive research on argument and event structure within Relational Grammar (Perlmutter 1983 i.a.) Of particular significance were attempts to derive the categorial, syntactic properties of insertion frames (c-selection, Grimshaw's 1979 terminology) from semantic properties. At their crudest, semantic properties involve the association of particular words, mostly verbs, with a set of argumental roles, or θ -roles, alongside linking conventions which regulate the mapping of particular roles to syntactic positions. A verb such as *head-LEAD* could be said to have a meaning that requires two θ -roles, an *agent* and a *patient*. In turn, general linking principles would decree that *agent* participants must be structurally higher than *patient* participants, or in Williams' (1981) terminology, structurally *external* to the domain containing the verb and its complement. (16) now emerges (underlining for the external argument), with the listed entry for *head*, by virtue of its meaning, constituting a set of instructions to the syntax, leading to the (schematic) structural configuration in (17):

16. *head*: θ -agent, θ -patient



Starting with Grimshaw (1979) and strongly advanced in Pesetsky (1982), this research agenda becomes an explicit effort to derive syntactic insertion frames from an articulated theory of lexical semantics. To quote from Pesetsky (1982):

The primitives of θ -theory – notions like "agent", "patient", "goal" etc. probably meet the criterion of epistemological priority [...]. On the other hand, the primitives of c-selection – syntactic categories like NP, S', Small Clause etc. – do not meet the conditions of epistemological priority. They are not, in Chomsky's words, "concepts that can ... provide the primary linguistic data that are mapped by the language faculty to a grammar."*If this discussion is correct, it follows that we want to derive the theory of c-selection from some other theory, whose primitives are epistemologically prior. Such a theory would be a semantic theory – specifically a theory of lexical semantics.* (180-181, emphasis added)

This agenda led to considerable cross-linguistic results, many under the umbrella of the Lexicon Project, run at the MIT Cognitive Science Centre 1983-1989. That project set itself the aim of "discovering elements of meaning which recur systematically in the definitions of words and the principles which determine the mapping from lexical semantics to morphosyntax."³ Among many constructions investigated in a broad range of languages are transitivity alternations, resultatives, applicatives, double object constructions, locative inversion, and many others (English illustrations in (18)). Many now would argue that the lexical semantics of individual listed words is not the appropriate source of syntactic generalizations, yet it remains the case that knowledge of correlations between meaning and structure acquired within this agenda is at the core of all current modelling of argument structure, event structure and complementation.

18. a. load the hay on the wagon/load the wagon with hay (transitivity alternation)
 b. the garden swarmed with bees/bees swarmed in the garden (locative alternation)
 c. water the tulips flat (transitive resultative)

³Levin (2011).

- d. the river froze solid (*intransitive resultative*)
- e. in the forest lies a hidden treasure (*locative inversion*)

Departing in this respect from *Aspects and Remarks*, Chomsky (1986) broadly endorses the reduction of c-selection to semantic selection. He proposes that lexical-semantic considerations determine 'Canonical Structure Realization', thus leaving the door open for residual listed formal properties which cannot be semantically derived. For instance, following Pesetsky (1982), Chomsky assumes an obligatory object (as with the case of *head-LEAD*) emerges from the *obligatoriness* of accusative Case, listed for *head-LEAD*, but not for *read*. Unreducible to lexical semantics, such obligatoriness must be listed.

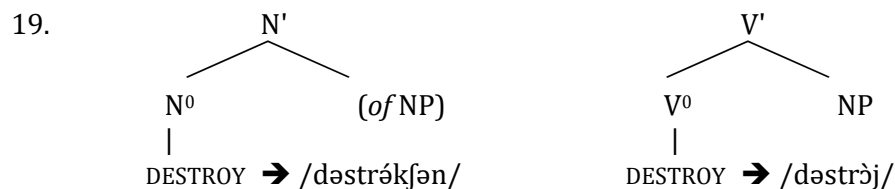
Anticipating section 7.2, note that the results of the lexical semantics agenda establish dependencies between some syntactic structure and some semantic effects. However, the claim that these connections are mediated through the lexical semantics of listed items (and are not direct correspondences between structure and interpretation) may, and has been, challenged.

5 Word Formation (WF)

Remarks, as noted, is devoted primarily to studying the exceptional. Yet Chomsky himself notes that some DNs *are* systematically related to their verbal source. Thus alongside *GEARBOX transmission* also means *the act of TRANSMITting* and *proof(s)* may certainly refer to *the act of PROVing*.

To capture these regularities, Chomsky proposes that a pair such as *destroy/destruction* constitutes a single lexical entry with a fixed meaning, and with a fixed syntactic insertion frame which is expected to hold across all its instantiations. What is missing from that entry, however, is a categorial label. The labelling as V or N, in turn, is acquired by inserting this a-categorial item under the syntactic nodes N or V respectively. The execution crucially presupposes that hierarchical structure is uniform across categories. Technically, Chomsky introduces an overarching constraint on phrase structure – the X'-scheme. Within that scheme, every phrasal head (X⁰) projects two additional structural levels (X' and X'', the latter the maximal phrase, XP), and relations between a head and its syntactic insertion context hold constant across category types (e.g., objects would always be sisters of X⁰, adjuncts would always be sisters of X' etc.). An additional assumption provides nouns, across the board, with optional rather than obligatory arguments.

An important auxiliary claim now involves the mapping from structure to PF. Specifically, if the relevant entry, say DESTROY or CIVILIZE, is inserted under V⁰ it would be pronounced *destroy* or *civilize*, respectively, but if inserted under N⁰, it would be pronounced *destruction* or *civilization*. The (schematic) *Remarks* structures are in (19):



The representations in (19) do not involve a *change* of categorial label. Rather, the label emerges solely from the syntactic structure, with the distinction between *destroy* and *destruction* reducible to a distinct PF in distinct syntactic contexts. With Wasow's (15B) in mind, one must ask whether category-changing rules, as such, *ever* exist, lexically or syntactically, as the mechanism which relates /dəstrəj/, the verbal instantiation of DESTROY, and /dəstrəkʃən/, its nominal instantiation, is neither syntactic nor lexical. Furthermore, within that approach, the relationship

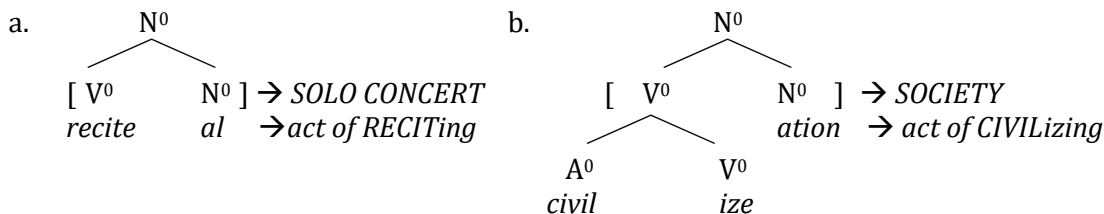
between *destroy* and *destruction* is altogether not derivational or even directional in nature. In (19) *destroy*, occupying a verbal head, and *destruction*, a nominal head, are equally complex – both are X^0 terminals dominating the very same entry. That the noun is morpho-phonologically complex and includes within it a stem that is largely phonologically identical to the verbal realization is certainly not a syntactic fact, and it is not easy to see how it can be modelled lexically, given the assumption that DESTROY is a single a-categorial entry.

This perspective on complex words was challenged by Morris Halle in his seminal 1973 *Prolegomena to Word Formation*. Halle enhances the case for an articulate lexicon by highlighting the item-specific nature of inflectional marking, but rejects a non-derivational approach to complex words. In its stead he argues for the development of a rigorous Word-Formation rule system (WF), which derives the set of all *possible* words in some language L. As such, this rule system parallels syntactic rule systems, whose output consists of all *possible* phrases in a given language. Unlike the output of syntactic rules, however, the output of the WF system is subject to a filtering procedure. The filter, in this case, is the *Lexicon* of L, which consists of all *actual* words in L, together with their unpredictable properties, where present, and which acts to exclude *possible* but not *actual* words, and provide some output items with properties otherwise not predictable from the derivation (e.g., it would provide the interpretation *SOLO CONCERT* for *recital*).

Halle's *Prolegomena* was immensely influential, with a burgeoning community of WF scholars developing systems that derive *possible* words that are nonetheless sensitive to item-specific listed restrictions.⁴ Departing from Chomsky's Remarks, lexical entries within these approaches are typically categorial, and may be combined by WF rules to yield increased structural complexity. As for the combinatorial systems, they are not only generative, but also suspiciously syntax-like, including (but not limited to) rewrite rules for WF, heads of words, on a par with heads of phrases, insertion frames for word subparts and affix categorial projections. Within these approaches rules that alter categorial labels as in (20), are typically absent, and the relationship between the verb *recite* and the noun *recital* is derived from combinatorial processes which increase structural complexity, and where categorial labels remain unchanged, as they would be in a syntactic phrase. The resulting WF structures, as in (21), only differ from canonical *syntactic* constituent structures in one respect – they do not involve an increase in projection level. I return to this matter in section 7.1:

20. $N \rightarrow A$ (in some context); $V \rightarrow N$ (in some context); etc.

21. WF:



⁴ A notable exception is Jackendoff (1975). The Jackendovian lexicon consists of an economy-driven network of non-derivational cross-references between listed items, thereby relating the relevant aspects of [_vdestroy] and [_ndestruction], otherwise listed separately, complete with their category, PF and interpretation.

While the methodological distinction between possible and actual words continues to enable substantial progress within WF, its ultimate success is contingent on understanding why such distinction should exist for (complex) words but not for (non-trivial) syntactic phrases.

6. The Complete Word

The end of the 20th century sees most generative approaches postulating an extremely powerful lexicon, in which words control at least the following:

22. a. lexical semantic information, resulting in argument array and event properties
- b. syntactic features and syntactic insertion frames (to the extent not derived from (a))
- c. combinatorial operations of words and word parts (WF)
- d. word (and sub-word) phonological properties

Within this augmented lexicon, an individual entry constitutes a set of instructions to syntax, phonology, morphology, and semantics. Into the present century, D-structure (GB) or licit (external) Merge (Minimalism) are by consensus trivial outputs of lexically-encoded information combined with general principles of constituent-structure building. Constraints that bar syntactic operations from applying in the lexicon and prevent syntactic derivations from altering lexically-specified features guard this central role for the word. The former constraint is called The Lexical Integrity Hypothesis (Lapointe, 1979) or Atomicity:

23. "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 relations of these features to the internal composition of the word cannot be relevant in syntax." (Di Sciullo and Williams 1987: 49)

In GB, the Projection Principle (Chomsky 1981) preserved lexical properties in syntactic derivations; in Minimalism the Inclusiveness Condition has that role:⁵

24. "Given the numeration N ...any structure formed by the computation ... is constituted of elements already present in the lexical items selected for N; no new objects are added in the course of computation apart from rearrangements of lexical properties." (Chomsky 1995: 228)

Words, then, affect syntactic and semantic computations but not formed by either syntactic or semantic operations. They are syntactically and semantically opaque - neither the syntax nor the semantics can modify or even refer to their internal composition. Thus, these principles clearly delineate the boundary between the lexical and the syntactic. Consider the verb *recite*, selecting an *agent* and a *patient* when compared with the noun (*piano*) *recital*. Clearly, deriving the latter from the former would be in violation of the Projection Principle/Inclusiveness. First, both *agent* and *patient* are eliminated, and second, the label V is eliminated and the label N is added. Perforce, the relationship between *recite* and *recital* cannot be syntactic, and must be lexical.

The *Word* thus conceived is both atomic and *complete* with properties which instruct the syntax, the semantics, the phonology and the morphology. As such, it is a unique formal object. Syntactic terminals such as V or T(ense) are not complete in the same sense, nor are their combinatorial outputs (e.g. TP, VP). They have syntactic properties, but not even T has formal semantic properties; it must be converted to a semantically appropriate representation to be interpreted. Phonological and semantic objects are not complete. Nor are most morphological affixes (e.g. *-ation* is conceivably N and has PF, but no fixed semantics). Syntax creates *syntactic* objects from *syntactic* terminals that are translated into *semantic* objects, formulas, which utilize their own terminals and modes of composition. Semantic objects do not correlate with unique phonological objects, and even the claim that they correlate with unique syntactic objects is largely

⁵ Proponents of Atomicity and the Projection Principle/Inclusiveness disagree on lexical operations, potentially with significant syntactic consequences. For example, while Williams (1981) and others allow modifications that affect syntactic linking, these are rejected in Baker (1985) and much subsequent work.

not accepted. Finally, while phonological units (of varying complexity) frequently correspond to syntactic constituents, including terminals, that is not always the case. Words, however, are by assumption simultaneously phonological, morphological, semantic, and syntactic units, with none of these distinct sets of properties derived from the other.

Yet the only widely accepted diagnostic for what a (substantive) word is remains phonological, which in a language such as English is tied to the occurrence of a single main stress. The centrality of PF for diagnosing words should be evident from the logic of preceding discussion. I compared the properties of *head* across its occurrences because they all share the PF /héd/. I did not suggest that the verb /héd/-ADVANCE, and the verb /ədvráens/-ADVANCE constitute a single lexical entry, although they share more syntax and semantics than they do with /héd/-LEAD. Similarly, when Chomsky proposes that DESTROY is an a-categorial item giving rise in different contexts to both /dəstrákʃən/ and /dəstròj/, he presupposes the significance of phonological relatedness. It is unlikely that he would have embarked upon dissuading us from believing in the transformational relatedness of *write* and *essay*. This, of course, does not exclude the possibility that they are. It serves to emphasize, rather, that most of our knowledge about linguistic relatedness emerges from theoretical perspectives predisposed to relating similar PFs and excluding unrelated PFs as instantiation of a single item.⁶

The over-riding question, then, is why a unit recognized and defined primarily through its pronunciation should have properties that affect syntactic and semantic computations. The issue is particularly crucial because phonological words may convey the very same syntactic and semantic content otherwise conveyed by multiple words, both inter- and intra- language. The English pair in (25) provides an example:

25. a. [VP *become* [A *red*]] → /bɪkám/+ /réd/
 b. [VP[V [A *red*]-en]] → /rédən/

7 Challenges: Neo-Constructional and Root-based Approaches

The past two decades have seen gained influence for research agendas that redraw the line between listedness and syntax. These agendas adopt advances in WF and results within semantics, syntax, and phonology that offer alternative solutions to the problems outlined in Aspects and Remarks.

There is little doubt of the need to *somewhere* list for words their morpho-phonological properties in isolation and in context (*walk-walked* but *give-gave*; *govern-ment* vs. *recit-al*; *destroy* vs. *destruct-ion* etc.) and their meanings where they are not compositional (*cat*→*CAT*; *liquidate*→*ASSASSINATE*; *headway*→*PROGRESS* etc.). The challenge is directed at associating listed sound-meaning pairs with labelled syntactic terminals and syntactic insertion frames.

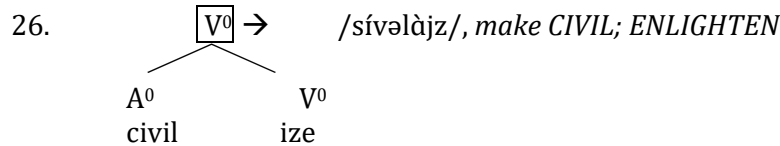
7.1 Insertion Frames, the WF perspective

Recall that objects obligatory for verbs may nonetheless be optional in DNs (13a). This is a general property. Nonetheless, in Remarks it is argued to support a non-transformational approach, as if DNs were derived from verbal structures, the object would be obligatory. However, this conclusion does not follow from the formal properties of an operation that would derive nouns

⁶As a cursory glance at English *be* reveals, some distinct PFs *are* grammatically related. The overwhelming majority of linguistic traditions, however, severely circumscribe such non-phonological relatedness.

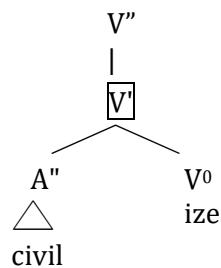
from verbs (e.g. categorial change) or from unpredictable irregularities. Rather, it follows from the obligatory insertion frame for verbs, which forces e.g. *destroy* to have a direct object.

Consider now the word structure in (26), where PF and meaning are associated with the higher instantiation of V^0 (boxed):

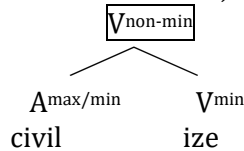


I noted previously the absence of level increase in WF representations. With level increase, the structures would be (27a) or (27b), depending on execution, where it would be V' or $V^{\text{non-min}}$ that would be associated with PF and meaning (boxed):

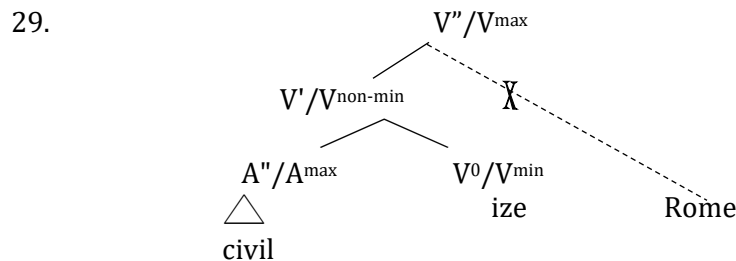
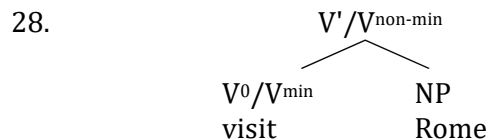
27. a. X' -scheme



b. Bare Phrase Structure, Chomsky (1995)



However, WF can adopt neither structure in (27a-b). The insertion frame of *civilize* necessitates a direct object, an interpretation that requires the object NP to be a sister of V^0/V^{min} . This would be straightforward for the underived verb *visit* in (28), but complex *civilize* with the boxed structure in (27a-b) is already $V'/V^{\text{non-min}}$, and thus *Rome* could not merge with V^0/V^{min} , to be interpreted as its direct object:⁷



Assuming distinct direct object positions for *visit* and *civilize* is clearly undesirable. If, however, the output of WF rules, regardless of their complexity, is always *syntactically* X^0/X^{min} , as in (26), adding the direct object as a syntactic sister of X^0/X^{min} would be possible. It therefore follows that the formal nature of syntactic combinatorial rules must differ from that of WF combinatorial rules.

⁷ The discussion presupposes the formal impossibility of ternary branching.

Upon closer scrutiny, this conclusion as well follows from the syntactic insertion frame of *civilize*, together with the requirement for direct objects to be sisters of terminals, i.e. V^0/V^{\min} .

What, however, if we could sever the verb from its (apparent) arguments or its (putative) syntactic insertion frame? If that were done, there would be little reason to postulate distinct combinatorial operations for WF and syntax. Nouns could be derived from verbs and a complex derivational history could be embedded under a verbal projection, both without making reference to arguments.

7.2. Insertion frames, the Neo-Constructional Perspective

Independently of the architectural considerations above, research on argument structure and verb complements from the 1990s onwards converges on the conclusion that the syntax and the interpretation of argument structure are not contingent on properties of selecting words, but emerge from larger syntactic and semantic configurations. I noted in reference to (22a-b) that correlations between syntactic structure and semantic event interpretations need not be mediated by the lexical semantics of individual entries, but could involve a direct mapping between some (non-trivial) phrasal structure and interpretative rules. Under such a scenario the *patient* occurring with *head-LEAD* is not its argument, and neither is the directional expression occurring with *head-ADVANCE*. Rather, both are interpreted through their independent syntactic positions. Interpretation of this sort contributes to the emergence of a larger syntactico-semantic domain within which the terminal *head-/héd/* is interpreted as *head-LEAD* or *head-ADVANCE* respectively. For such an approach, there is little reason to assume that the pair in (1a-b) results from the structural realization of two individually listed words with their distinct insertion frame. It could do, instead, with a single *head-V* which is assigned a finer-grain meaning in accordance with its syntactic context.

There is considerable support from formal semantic approaches inspired by Davidson (1967) for severing the interpretation of arguments from phrasal heads, primarily verbs. Such approaches view arguments in terms of (grammatical) events. In Parsons' 1990 Neo-Davidsonian approach, the semantic representation of (30) would be essentially as in (31), with *agent* or *patient* (or equivalents) naming a relationship between participants and events, not mediated through properties of the verb. Within this approach, the verb functions as an event modifier (e=event):

30. Mary headed the team

31. $\exists e$ [*head* (e) & *Agent* (Mary, e) & *Patient* (the team, e)]

Support for severing the (logical) subject from the verb, advocated at least as early as Marantz (1984), has gained momentum as a result of integrating a Davidsonian syntactico-semantic approach (e.g. Kratzer 1996). Alongside these developments, and beginning with Baker (1985), we see a drive to associate all argumental roles with a fixed hierarchical configuration, thereby placing more and more weight on configurational interpretations. The model developed in Hale and Keyser (1993 and subsequent) has been pivotal to this enterprise. However, they viewed their investigations as fundamentally lexical, providing a hierarchical explanation for the emergence of word properties. Fully syntacticalized approaches followed shortly, all based on the assumption that argumental interpretation is configurational and independent of selecting words.

The thrust of such proposals, at times called Neo-Constructional, is illustrated in (32a-e), originally from Clark and Clark (1979):

32. a. The fire stations sired throughout the raid
- b. The factory sired midday and everyone stopped for lunch
- c. The police sired the Porsche to a stop
- d. The police car sired up to the accident
- e. The police car sired the daylights out of me

If the syntax of (32a-e) were determined by listed insertion frames, we would need five different insertion frames for *siren*, of which at least four would convey interpretational information that cannot be deduced from sounding sirens alone. The interpretations of (32a-e) clearly pattern with those of the syntactic configurations in (33a-e):

33. a. The bells rang throughout the raid
- b. The factory signaled midday and everyone stopped for lunch (*e.g. by sirening*)
- c. The police forced the Porsche to a stop (*e.g. through sirening*)
- d. The police car rushed up to the accident (*e.g. while sirening*)
- e. The police car scared the daylights out of me (*e.g. with its sirening*)

A compelling account would attribute interpretations of the events in (32a-e) to their distinct syntactic properties, mirrored in their interpretational correlates in (33a-e), by the syntactic position of the participants and grammatical roles played by prepositions and particles. The verb *siren* is best viewed as a modifier of the emerging event, playing a role roughly equivalent to that of an adverb (i.e. 'sireningly').

The emerging agenda postulates fragments of phrasal structure that are mapped onto particular interpretations, with the meaning of individual words resulting from a combination between listed core conceptual content and whatever interpretational constraints emerge from the embedding structure.⁸ To yield (32a-e), *siren* could be listed as /sájrən/-*SIREN* with the latter indicating particular noise emissions, the only meaning component common to all verbal occurrences, and possibly nominal ones too. Remaining aspects of event interpretation come from syntactic architecture. From this perspective, the infelicity of (34) emerges not from the *grammatical* properties of *fall* but from a clash between the basic meaning of FALL and the event interpretation that emerges from the syntax of (34). This infelicity parallels that of juxtaposing *sleep* and *furiously* or *colourless* and *green*, as in Chomsky's (1957) famous example in (35):

34. The police car fell up to the accident

35. *Colorless green ideas sleep furiously*

If, however, *the Porsche* in (32c) is not part of the listed insertion frame of *siren*, and by extension nor is any other complement, the problems for unifying WF and syntactic combinatorial processes noted in section 7.1 disappear.

7.3 An A-Categorial Lexicon and PF

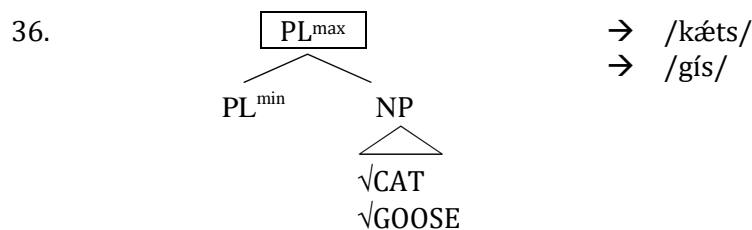
While I rejected some arguments against unifying WF and syntax, neo-constructional analyses neither *force* such unification nor *exclude* lexically-listed syntactic information. The converse, however, does not hold. Any account that dispenses with lexically listed *syntactic* properties such as category label and insertion frames is perforce committed to the neo-constructional agenda,

⁸ In contrast with Construction Grammar, Neo-Constructional approaches assume that the correlations between syntax and interpretation follow from universal principles of the syntax-to-semantics mapping, and are neither language specific nor analogical in nature.

because in the absence of syntactico-semantic information listed with terminals, events and arguments can only be interpreted through syntactic structure. By the same logic, such accounts are committed to unifying WF with syntax, for categorial labels are the playing chips of WF systems. For these reasons, the neo-constructional agenda has come to be linked with models that eliminate syntactic information from lexical entries.

Recall that in Remarks, category label and PF depend on the syntactic insertion environment. Effectively, objects such as [_v*destroy*] and [_N*destruction*] do not exist independently and can have PF only once embedded in some syntactic structure. One obvious way to achieve this result would be to insert PF *following* rather than *preceding* the syntactic derivation. Going beyond Remarks, suppose a-categorial entries, henceforth *roots*, that come with neither category nor insertion frame, and consist of little more than indices tracking their derivational history. By the end of the syntactic derivation, and depending on structure, [_v...√DESTROY...] or [_N...√DESTROY...] emerge, and in these contexts, the proper PF becomes available. For such a view, largely already implicit in Remarks, words such as *destroy* or *destruction* are neither atomic nor complete, but assembled piecemeal throughout the syntactic and the phonological computation, on a par with phrases.

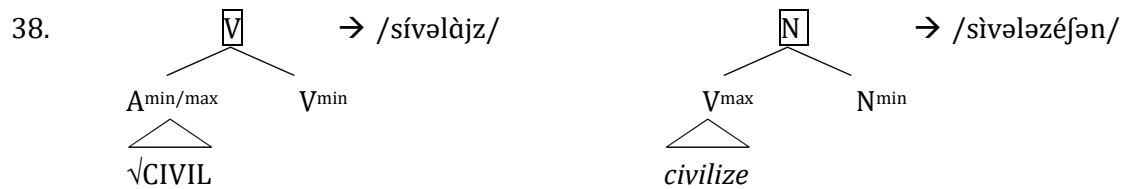
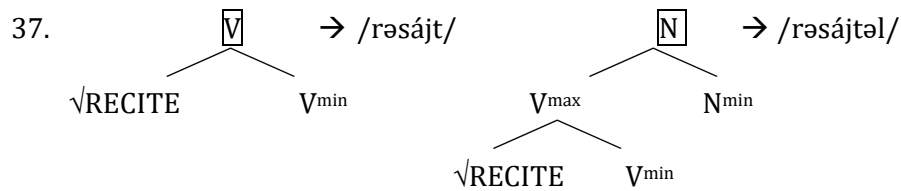
Now consider syntax ⇔ PF mapping in greater detail. For lexicalism, the *syntactic* domain of the phonological word (the domain of a single main stress in English, recall) cannot exceed X⁰/X^{min}, because words with their PFs are inserted as atomic syntactic terminals. If, however, PFs are available following the syntactic derivation, we expect correlations between larger constituents and single phonological words. Consider pairs such as *cat-cats* and *goose-geese*. For lexicalism, *cats* and *geese* would be inserted under N and provided with the feature [+plural]. This feature restricts emerging syntactic configurations (e.g. by allowing the plural auxiliary *are* and excluding *is*). This is the checking system proposed in Chomsky (1995, i.a.). The alternative is in (36):



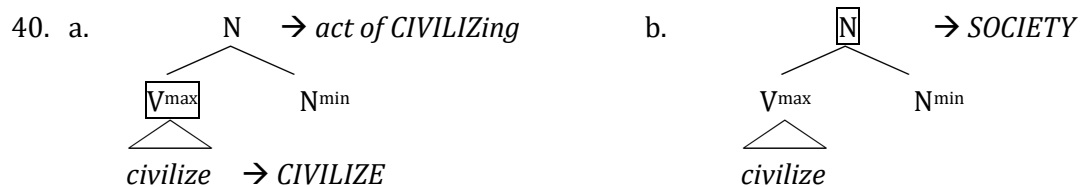
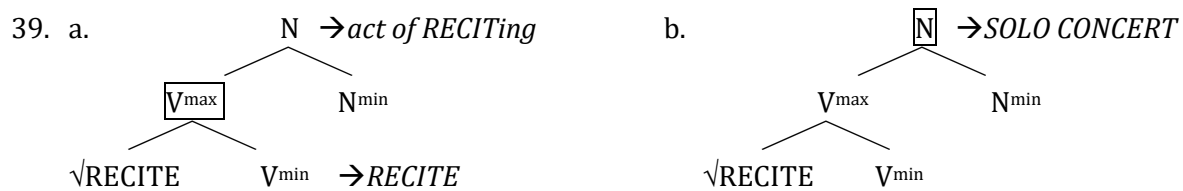
Here, roots consist of little beyond indices tracking their occurrences. They are inserted under N and combined with the phonologically abstract syntactic node PLURAL. PF, in turn, is assigned to the entire PL^{max} constituent in (36), yielding a single *phonological* word corresponding to two syntactic terminals. At this stage the listed properties of √GOOSE are consulted, yielding *geese* rather than the default *-s* marking of *cats*. Representations such as (36) thus permit integrating into syntactic structures discrete node features such as PL, allowing an exploration of limitations on PFs that emerge from syntactic structures, and conversely, limitations on syntactic structure that emerge from PF.

The rationale applies in a similar manner to derivatives, yielding the *syntactic* structures in (37)-(38) with PF determined by syntactic structure in conjunction with the phonological properties of the root:⁹

⁹To illustrate, English N^{min} corresponds to *-ness, -ity, -ment, -ation, -al* and more. Of these, the first two are syntactically conditioned by A (*kindness, ability*) and the others by V. The choice between syntactically identical suffixes, however, (e.g. *-ment* and *-al*) is root dependent and hence [_v√REFER]-*al* vs. *[_v√REFER]-



One notable piece missing from this sketch concerns the availability of meanings not predictable from parts, so requiring listing. This is the case for *recital* (*SOLO CONCERT*) and *civilization* (*SOCIETY*), as well as *headway* (*PROGRESS*). How can such listing be captured within a syntactic approach to their formation? The difficulty, however, is only apparent, emerging from associating listedness with syntactic Atomicity. Once Atomicity is dispensed with, there is little reason to exclude a correspondence between listed meaning and more complex syntactic constituents. Since both *recital* and *civilization* are (at least) ambiguous, listed meanings may be associated with the embedded boxed constituents in (39a), (40a) to yield compositional *act of RECITing* or *act of CIVILIZing*. Alternatively, meaning may be associated with the larger structure, in (39b), (40b), yielding an unanalysed meaning unpredictable from its parts:



There is no challenge here to the need to list unpredictable meanings, whether that of *cat* or of *recital*. The challenge, rather, is to the claim that listedness entails the absence of syntactic complexity. The task facing root-based approaches is to successfully delimit the syntactic domains within which listed meaning could emerge.¹⁰

ment, but *[$\sqrt{\text{GOVERN}}$]-*al* vs. [$\sqrt{\text{GOVERN}}$]-*ment*.

Prefixes (*de-* in *destroy* or *trans-* in *transform*), whose theoretical status is altogether poorly understood, are set aside.

¹⁰ See Doron (2014) and Alexiadou et al. (2015) on debates within the Neo-constructional community concerning the mapping of constituent structures to PF and interpretation and properties of roots in general.

For a fuller syntactic account of DNs (see (13)-(14) and related discussion) see, i.a. Roeper and van Hout (2009) as well as Borer (2013).

Chomsky himself remains committed to the existence of lexical entries consisting of bundles of features, many of which impact the syntactic derivation. This said, in his most recent work (2013, 2015) on the emergence of categorial labels (for complex constituents), he once again endorses the view originally suggested in Remarks, of the lexicon as potentially consisting of a-categorial roots labelled in the context of their syntactic environment.

8 A Brief Note on Grammatical Formatives

Aspects and Remarks are concerned with the properties of substantive vocabulary that corresponds, by and large, to conceptual knowledge. Because of that, I have not touched on properties of non-substantive vocabulary – grammatical formatives. Grammatical formatives mark specifically *grammatical* functions and are commonly articulated through inflectional markings (plural, tense, voice), categorial affixation (-ing; -ation; -er), and discrete function words, such as articles and demonstratives, modals, quantifiers, and others. Any cursory perusal reveals that the properties of grammatical formatives are clearly distinct from those of substantive items. Unlike substantive items, they do not correspond to concepts, nor do they have a ‘lexical semantics’ in any meaningful sense. To be sure, *many* or *the* do have a semantic function, but that function is found in formal semantic notions such as quantification, not in conceptual knowledge. Further, while the substantive vocabulary constantly expands, grammatical vocabulary consists of a small, virtually non-expanding set. Finally, while much of (English) substantive vocabulary oscillates quite freely between categorial types ([_{N/V}email], [_{N/V}internet], [_Vtext] ...) the category and function of grammatical vocabulary is fixed and unchanging.

While the grammatical role of substantive vocabulary has declined within theoretical modelling, that of formal vocabulary has become central, with explicit proposals attributing to it not only syntactic and semantic properties, but crucial roles in language acquisition, language variation, and language evolution. Given their distinct formal properties and their distinct theoretical significance, they require and deserve discussion not attempted here.

References

- Alexiadou, Artemis, Hagit Borer and Florian Schäfer (2015) (Eds.). *The Roots of Syntax, the Syntax of Roots*, Oxford: Oxford University Press
- Baker, Mark. (1985). ‘The Mirror Principle and morphosyntactic explanation’, *Linguistic Inquiry* 16:373–416.
- Borer, Hagit. (2013). *Taking Form, Structuring Sense Vol. 3*. Oxford: Oxford (
- Chomsky, Noam. (1957) *Syntactic Structures*. The Hague/Paris: Mouton
- Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. (1970). ‘Remarks on nominalization’, in Jacobs, R. A. and P.S. Rosenbaum (eds.) *Readings in English Transformational Grammar*. Waltham, MA: Ginn, 184–221.
- Chomsky, Noam. (1981). *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Chomsky, Noam. (1986). *Knowledge of language: Its Nature, Origin and Use*. New York: Praeger.
- Chomsky, Noam. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. (2000). ‘Minimalist inquiries: the framework’, in Martin, R., D. Michaels, and J. Uriagereka (eds.) *Step by Step: Essays on Minimalist Syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press.
- Chomsky, Noam. (2001). ‘Derivation by Phase’, in Kenstowicz, M. (ed.) *Ken Hale: A life in Language*. Cambridge, MA: MIT Press, 1–52.
- Chomsky, Noam. (2013). ‘Problems of projection’, *Lingua* Volume 130, June 2013, Pages 33–49
- Chomsky, Noam. (2015) ‘Problems of projection: extensions,’ in Di Domenico, E., C. Hamann and S. Matteini (eds.) *Structures, Strategies and Beyond*, Studies in honour of Adriana Belletti. John Benjamins
- Clark, Eve and Herbert Clark. (1979). ‘When nouns surface as verbs’, *Language* 55:767–811.
- Davidson, Donald. 1967. “The logical form of action sentences.” In *The Logic of Decision and Action*. Ed. N. Rescher. 81-95. Pittsburgh: University of Pittsburgh Press.

- DiSciullo, Anna Maria and Edwin Williams. (1987). *On the Definition of Word*. Cambridge, MA: MIT Press.
- Doron, Edit (2014) ed. *Theoretical Linguistics: On the Identity of Roots* Volume 40, Issue 3-4
- Fillmore, Charles J., Paul Kay and Mary Catherine O'Connor. (1988). 'Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*'. *Language* 64:501-538.
- Goldberg, Adele. (1994) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.
- Grimshaw, Jane. (1979). 'Complement selection and the lexicon', *Linguistic Inquiry* 10:279-326.
- Hale, Kenneth and S. Jay Keyser. (1993). 'On argument structure and the lexical expression of syntactic relations', in Hale, Kenneth and S. Jay Keyser. (Eds.).
- Halle, Morris. (1973). 'Prolegomena to a theory of word formation', *Linguistic Inquiry* 4: 3-16.
- Jackendoff, Ray. (1975). 'Morphological and semantic regularities in the lexicon', *Language* 51: 639-71.
- Kratzer, Angelika. (1996). 'Severing the external argument from the verb', *Phrase Structure and the Lexicon*, Johan Rooryck and Laurie Zaring. Dordrecht: Kluwer, 109-137.
- Lapointe, Steven. (1980). 'A Theory of Grammatical Agreement', Ph.D. dissertation, Amherst, University of Massachusetts.
- Levin, Beth (2011). 'The Lexicon Project and its legacy', PowerPoint Presentation, Stanford University
- Marantz, Alec. (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
- Newmeyer, Frederick. (1980). *Linguistic Theory in America*. New York: Academic Press.
- Parsons, Terence. 1990. *Events in the Semantics of English. A Study in Subatomic Semantics*. Cambridge, Mass: MIT Press.
- Perlmutter, David (1983) *Studies in Relational Grammar 1*, Chicago: University of Chicago Press
- Pesetsky, David. (1982) *Paths and Categories*. Ph.D. dissertation, MIT.
- Roeper, Thomas and Angeliek van Hout. (2009). 'The representation of movement in *-ability* nominalizations: Evidence for covert category movement, edge phenomena, and local LF', in Giannakidou, A. and M. Rathert (eds.) *Quantification, Definiteness, and Nominalization*. Oxford: Oxford University Press, 344-364.
- Wasow, Tom. (1977). 'Transformations and the lexicon', in P. W. Culicover, T. Wasow and A Akmajian (eds.) *Formal Syntax*. New York: Academic Press
- Williams, Edwin. (1981). 'Argument structure and morphology', *The Linguistic Review* 1, 81-114