Tutorial: An Introduction to Syntax

Lewis P. Shapiro
Department of Communicative Disorders
San Diego State University
and
Center for Human Information Processing
University of California, San Diego

This paper is intended as an introduction to syntax. Borrowing from Chomsky's Government & Binding and Principles & Parameters frameworks (Chomsky, 1986, 1992, 1995), various aspects of syntactic theory are described. These include lexical, functional, and phrasal categories and how they are put together into clauses and sentences, how words are represented in the mental lexicon, how lexical properties project to the syntax, and how noun phrases are assigned structural and semantic information. Additionally, how sentences that are not canonically ordered are derived and represented, how and to what degree pronouns refer, and the principles that connect all these theoretical notions to form knowledge of language are described. The paper concludes with a summary of work in normal and disordered language, including treatment of language disorders, that has exploited aspects of the syntactic theory described in this paper.

KEY WORDS: syntax, linguistics, language

Within theoretical linguistics, syntax is the study of the architecture of phrases, clauses, and sentences. The modern roots of the study of syntax can be traced to the pioneering work of Noam Chomsky, who in 1957 wrote Syntactic Structures. Chomsky changed the face of linguistics by casting its domain inwardly. That is, the concern shifted from describing external language phenomena to characterizing the mental machinery that purported to explain the native speakers' knowledge of language; that concern remains today. In this summary I borrow rather heavily from Chomsky's Government-Binding Theory and Principles and Parameters (Chomsky, 1986, 1992).

The proper study of (introductory) syntax could easily take up the entire body of a large text (e.g., Haegeman, 1992; Radford, 1988; van Riemsdijk & Williams, 1986). Because this paper is intended as a tutorial, it will emphasize constructs that will help the reader understand the general nature of syntactic theorizing without being too concerned with theory-internal issues; the paper also emphasizes those aspects of syntactic theory that are relevant to research in normal and disordered language performance in children and adults. Leonard and Loeb (1988) previously cut a similar path in their tutorial on Government-Binding theory. The present paper is intended as an update to that effort as well as a more detailed examination of syntactic theory.

So far as the intended audience is concerned, this paper is not

There are several other influential linguistic theories, for example, Generalized Phrase Structure Grammar (GPSG) and its variants (e.g., Gazdar, Klein, Pullum, & Sag, 1985), Lexical Functional Grammar (LFG) (Kaplan & Bresnan, 1982), Relational Grammar (e.g., Perlmuter & Rosen, 1982), and Constructional Grammar (Fillmore, 1988; Goldberg, 1995). The reader is directed to Seils (1985) for a summary of some of these theories.
necessarily directed toward those researchers who are familiar with theories of syntax. Rather, it is intended for the audience of researchers or students of language who may not have a current understanding of syntax or its relevance and would like an introduction. If you are a student of language, a rudimentary knowledge of syntax is essential because it is part of the basic science of language, as are theories of phonology, morphology, semantics, and discourse. Also, syntactic theory forms the backbone of much of the work in normal and disordered language processing. Indeed, understanding syntax and manipulating it in principled ways is becoming more and more important to the treatment of language disorders in both children and adults.

I’d like to begin by asking the reader to keep one thing in mind during this exercise: A theory of syntax must describe all and only all the well-formed, acceptable sentences in any given language. If a syntactic theory not only explains the sentences that are grammatical, but also somehow allows the generation of ungrammatical sentences, then the theory may not be an accurate characterization of our linguistic knowledge. Putting it another way, our knowledge of language allows us to make judgments regarding the sentences that are acceptable and those that are not. One purpose of a theory of syntax is to describe the mental machinery that we must possess in order to make such judgments. The theory, therefore, must explain both how we know that a particular sentence is acceptable and how we know that a particular sentence is not acceptable. As a start to the construction of this theory, consider:

1. Dillon and Joelle went to the beach
2. *Dillon and Joelle put to the beach

Our intuitions about English (driven by our unconscious linguistic knowledge) allow us to effortlessly judge (1) as acceptable (i.e., grammatical, well formed) and (2) as unacceptable (i.e., ungrammatical, ill formed—signified by an *). If you were asked why this is so, you might hypothesize that “sentence (1) is well formed and sentence (2) is not because they have different verbs.” But such a statement doesn’t really tell us anything substantial; it’s just a description of the difference between the two sentences. For example, to explain the grammaticality of (1) and the ungrammaticality of (2), we would have to know something about how we acquire the lexical category of verbs, how phrasal categories are constructed out of lexical categories, how each verb picks its linguistic environment (some verbs go with some phrases and some go with other phrases, to which both [1] and [2] attest), and how this information comes together to yield knowledge of the grammatical sentences of a language. Later, the theoretical machinery that some syntacticians claim is necessary to explain the facts about (1) and (2) will be described.

Our syntactic knowledge also allows us to go beyond making judgments about whole sentences. If (1) were to be divided into two sections, where would the division occur? It is very likely that the sentence would be divided between its subject (Dillon and Joelle) and its predicate (went to the beach). Thus, a first pass at the syntactic structure of the sentence might be:

3. [Dillon and Joelle] [went to the beach]

Note that the outer set of brackets demarcates the entire sentence; the inner sets of brackets divide the sentence up into further parts.

But why wouldn’t (1) be divided into the following?:

4. *[Dillon and] [Joelle went to the beach]] or
5. *[Dillon and Joelle went] [to the beach]]

Sentence (1) would likely be divided as (3) and not as (4) or (5) because you have the mental capacity to effortlessly know that (3) yields well-formed parts (or constituents) and (4) and (5) do not. Now take what we have called the predicate and divide it further, into two parts:

6. [[went] [to the Beach]]

What we have done in (6) is to divide the predicate into a verb (went, the past tense of go) and a prepositional phrase (to the Beach). And, of course, we can further divide the prepositional phrase into a preposition (to) and its noun phrase object (the Beach).

Our grammatical knowledge also allows us to make judgments about reference; we know that some lexical items refer to others in the same sentence or in the discourse (i.e., extra-sentential reference). For example:

7. Dillon hit himself
8. Dillon hit him

We know that himself must refer to Dillon in (7), yet in (8) we know that him cannot refer to Dillon.

Our syntactic knowledge also allows us to make judgments of "sameness of meaning." For example:

9. Dillon hit Joelle
10. Joelle was hit by Dillon

We know that although the focus may be different in (9) and (10), the basic meaning is the same.

Finally, our syntactic knowledge allows us to recognize structural ambiguities, sometimes subtle ones:

11. The mechanic fixed the car in the garage

If some thought is given to what (11) can, in principle, mean, we come to the conclusion that it can mean either
something like "it was in the garage where the mechanic fixed the car" or it can mean "it was the car in the garage (instead of the car outside the garage) that the mechanic fixed." These meanings, as we shall see later, are reflected by where the prepositional phrase (in the car) fits into the phrasal geometry of the sentence.

It is likely the reader cannot verbalize the rules and principles that the judgments about examples (1)–(11) were based on, basically because knowledge of language is tacit or unconscious. The job of the theoretical syntactician is to observe, hypothesize, and test what this knowledge consists of. The data that linguists primarily use to develop and test their theories come from grammatical intuitions of native speakers. From these intuitions, theoretical constructs and hypotheses are developed and placed within a theory and then tested against new observations (new data). Based on these new data, the theory can either be accepted, revised, or rejected. And this process of data gathering, testing, theory development, data gathering, goes on and on. Such is the toil of any science, including the science of language.

In this paper, then, aspects of a particular theory of syntax will be summarized. These include lexical and phrasal categories and how they are put together into clauses and sentences, how words are represented in the mental lexicon (dictionary), how lexical properties project to the syntax, and how noun phrases (NPs) are assigned structural and semantic information. Additionally, I will describe how sentences that are not canonically ordered are derived and represented (in English, Subject-Verb-Object ordering is considered canonical, or basic), how and to what do personal and reflexive pronouns refer, and how the principles that connect all these theoretical notions form knowledge of language. The paper concludes with examples of psycholinguistic and neurolinguistic efforts that show the relevance of the study of syntax to language performance and to the treatment of language disorders.

**Lexical and Functional Categories**

Categories are theoretical constructs that linguists use to explain the fact that some words behave differently than others. Instead of using vague notions like "nouns are persons, places, or things" (the word run can be a noun or a verb), "verbs refer to actions" (destruction is an "action," but is actually a noun), and "prepositions are words referring to locations" (La Jolla is a location, but is a noun), linguists have looked to phonological, morphological, and distributional evidence to determine or rationalize lexical categorization, or parts of speech. For example, phonologically, the primary stress often falls on the first syllable of multisyllabic nouns (e.g., PERMits, RECORDs), yet on verbs the primary stress often falls on the second syllable (e.g., perMITS, reCORDS). Morphologically, nouns can be pluralized (boys, women) and verbs cannot. Nouns and verbs can form complex words made up of more than one morpheme, but prepositions cannot; they are invariant. Distributionally, nouns occur in particular and in different parts of a sentence than do verbs; thus, they cannot be substituted for each other (they are said to be in complementary distribution). For example, nouns can be pre-modified by adjectives (very big boy, pretty woman, etc.) yet verbs cannot (*very big know); nouns can be quantified and specified (e.g., made definite or indefinite) (e.g., the boy, a boy), yet verbs cannot (*a/the know). So, a verb cannot be substituted for a noun, and vice versa (Dillon kissed Joelle, *Kissed Dillon Joelle). Indeed, substitution is one constituency test that linguists use to help determine the category of a lexical item.

Because of these phonological, morphological, and distributional facts, linguists have hypothesized a limited set of lexical categories such as Nouns, Verbs, Adjectives, and Prepositions, as well as a set of functional categories like Determiners (the, this, some, many, etc.), Complementizers (that, whether, for, etc.), and Inflections (modals such as will and should, for example, and tense and agreement morphemes). Claiming that there are distinct categories that behave differently allows the linguist to make general statements, like "nouns can be pluralized, verbs cannot." Such general statements allow lexical items to be represented economically. For example, because only nouns can be pluralized, it is not necessary to represent the plural noun separately from its singular version. All that is needed is a representation of the singular noun and a rule that states that nouns can be pluralized; this will automatically generate the plural form for the noun. Importantly, this productive mechanism simplifies the acquisition process that the child undergoes. That is, the child does not need to "memorize" each plural form for each singular noun counterpart; all the child needs to know is the rule for the plural and the fact that any noun can have a plural form (although the child will indeed have to memorize irregular forms). This emphasis on language acquisition forms the basis for much of linguistic theory.

**Phrasal Categories and Phrase Structures**

Categories such as nouns, verbs, and prepositions are not just arranged in a one-level left-to-right serial order. Instead, evidence suggests that they form phrasal categories, and both lexical, functional, and phrasal categories are arranged in a hierarchical structure to form clauses and sentences, much like a house is built with a foundation, walls, beams, and a roof. Again, consider the sentence The mechanic fixed the car in the garage. The ambiguity that was described earlier is not lexical; none of the crucial words in the sentence has more than one sense. How, then, can we account for the ambiguity?
One way to explain this phenomenon is to assume that phrases are organized into hierarchical structures and that there will be cases where more than one structure can be assigned to a particular phrase. Thus, The mechanic fixed the car in the garage can be assigned two different structures. The two structures can be viewed in different ways. For example, an approximation of the two structures can be viewed by different labeled bracketing:

12. \[ S [NP \text{The mechanic}] [VP \text{fixed} [NP1 [NP2 \text{the car}]] [PP \text{in the garage}]] \]

13. \[ S [NP \text{The mechanic}] [VP \text{fixed} [NP \text{the car}] [PP \text{in the garage}]] \]

In (12), the prepositional phrase (PP) in the garage modifies the noun phrase (NP2) the car; thus, there is one larger NP (NP1), the car in the garage (I have boldfaced the brackets in this case). In (13) the PP in the garage modifies the entire verb phrase (VP) fixed the car (later, I will rid phrasal notation of numbered nodes such as NP1 and NP2 in favor of a more parsimonious representation). Another method of showing phrasal geometry is through the phrase structure tree (phrase marker). (14) and (15) show a first-pass at the two phrase structure trees for the ambiguous sentences, (12) and (13), respectively:

Viewing just the verb phrase (VP) for a moment, in (14) the PP in the garage modifies the noun phrase the car and thus attaches to the “higher” NP1 node, forming an NP the car in the garage; in (15) the PP in the garage modifies the VP and thus attaches to the “higher” VP1 node, forming the “higher” VP fixed the car in the garage. So, the structure of the sentences of a language can be captured by phrase structure representations, where each structure suggests a specific interpretation.

Continuing with examples (14) and (15), consider in more detail the symbols used to label phrase structure nodes. There is a distinction between lexical nodes and phrasal nodes. Lexical nodes are made up of lexical categories such as noun (N), verb (V), and preposition (P). Phrasal nodes include phrasal categories such as noun phrases (NPs), verb phrases (VPs), and prepositional phrases (PPs). For example, the sentence the mechanic ate contains a noun phrase (NP) the mechanic, which in turn contains the noun (N) mechanic. Phrase structure theory contains a principle that captures an important generalization about the structure of these categories:

16. The Head Principle: Every phrasal category contains a head; the head and its phrasal counterparts share the same properties.

An NP must contain an N, which is the head of the NP, a VP must contain a V, and so on; the head and its phrase share properties. For example, if a head noun is plural, so too is the entire NP (e.g., The boys are wild). This principle serves as an important constraint on phrase structure representations; if there were no such constraint, phrase structures would allow the generation of impossible structures (e.g., an NP containing a V).

Continuing, an analysis of English sentences tells us that a sentence consists of both a noun phrase and a predicate (i.e., a verb phrase). (17) is a shorthand way of stating this property:

17. \[ S \to NP\text{ VP}: [S [NP \text{The mechanic}] [VP \text{fixed the car in the garage}]] \]

The arrow, for now, means “is rewritten as” or “consists of,” so (17) says that an S can be rewritten as an NP and a VP. An NP, in turn, consists of a noun (obligatory), which may or may not be preceded by a determiner (optionality is denoted by a set of parentheses), as in:

18. \[ NP \to (DET) N: [NP [DET \text{The}] [N \text{mechanic}]] \]

A verb phrase consists of at least a verb, and potentially many other optional elements, including another NP, a PP, or even another Sentence (clause). (19) contains examples of some of these possibilities:

19. \[ VP \to V: [VP [V \text{slept}]] \]
V NP: [VP [V fixed] [NP the car]]
V NP PP: [VP [V sent] [NP the letter] [PP to his mother]]
V S: [VP [V discovered] [S that the manuscript was stolen]]

A prepositional phrase may include a preposition followed by an NP, as in:

20. PP → P NP: [PP [P in] [NP the garage]]

Finally, there is one important constraint on phrase structures that has been left out. Consider:

21. S → NP VP
   VP → V NP

To generate a sentence given a particular set of phrase structures, lexical items are inserted into the category slots via lexical insertion. A grammatical sentence corresponding to (21) might be:

22. Joelle kicked the door,

where Joelle is lexically inserted into the N slot of the subject NP, kick into the V slot of the VP, and the door into the direct object NP. But what about the following?

23. *Joelle thinks the door

Sentence (23) has the same structure as (22), can fit into the phrase structure representation described in (21), and yet our intuitions tell us that it is not a well-formed sentence in English. What makes (22) different from (23)? The only difference lexically between the two sentences is, of course, the verb. If the verb kick is inserted in the phrase structure of (21), the sentence is well formed; if the verb think is inserted, the sentence is ill formed. So, the theory of grammar as it stands now is simply too powerful; it generates ungrammatical as well as grammatical sentences. The theory, therefore, must have a way to restrict the output of phrase structure representations like those in (21) to generate only the well-formed instances of our language. Before this problem is tackled, some generalizations about phrase structures that have given rise to what is called X-bar theory need to be detailed.

X-Bar Theory

X-bar theory is a formal way of characterizing what is common about phrase structures. Recall that each phrasal category must contain a head (16). For NPs, the head is the N, for VPs, the V, and so on. But what could be the head of the S-node in, for example, (14) and (15)? To answer this question, note that sentences have inflection; they are inflected for tense (TNS) and agreement (AGR). At first it might appear that it is not

the sentence that has tense, but, instead, the VP. For example, in the sentence, The mechanic fixed the car in the garage, it appears that the VP is past tense, since the head of the VP (the V) has past tense morphology (-ed). But consider that the past tense can be separated from the VP, as in What the mechanic did was fix the car, where the past tense is now part of the auxiliary did and is no longer “attached” to the verb itself. Also consider that overt, non-affix tense markers like will, for example (e.g., The mechanic will fix the car), are separated from the VP. For these and other reasons, linguists now consider tense to be represented separately from the verb and verb phrase, forming what is called an Inflection Phrase (IP). The head of IP is the functional category INFL or I (for Inflection).

Consider again (14), repeated in (24), but this time with a more fully specified phrasal architecture:

24. The mechanic fixed the car in the garage

First, note the following terms: branching means that a node splits into other nodes, dominates means that a given node is “higher up in the tree” than other nodes, and immediately dominates means that a given node is directly above another node in the tree, with no other nodes intervening between them. In (24), the S node is now replaced by an IP (Inflection Phrase) that dominates all other nodes of the tree. The IP branches and immediately dominates an NP and an intermediate structure (called I-bar; written as I') whose head is INFL (which, in this case, is past tense). The subject NP branches and immediately dominates DET (determiner) and an intermediate category, N'. N' immediately dominates the head N. The I' also has a VP attached “to the right.” This schema thus retains the generalization that all phrases must have a head while accepting our intuitions that sentences have inflections that are independent from verbs and verb phrases.

Example (24) has several other generalizations. Note that the V has as its complement an NP (the car in the garage). And note that the head V (fix) falls “to the left”
of its complement. If we were to draw out the PP, we would also note that the head P (in) also falls to the left of its NP complement the garage. The same holds true for the INFL node (ed), which falls to the left of its complement, the VP (fix the car in the garage). So it seems that one generalization about (24)—and indeed all phrase structure (PS) rules of English—is that the head occurs to the left of its complements. This particular order of heads and their complements is not a universal property of all languages; languages generally fall into two camps, head-first or head-last.\(^3\)

Another generalization is that the determiner modifies or specifies the NP; the NP can be definite, indefinite, quantified, personalized, and so forth. So, for example, we could have a mechanic, the mechanic, some mechanic, all mechanics, and so forth. We will assume that the subject NP inhabits a functional category position called Spec (for Specifier); I will continue, however, to fill this Spec position with an NP.

Note that instead of using node labels like NP1 and NP2, I am now using XP and X' (where XP stands for NP, VP, PP, etc.). This X-bar notation captures the generalization that all phrase structure representations have the same form.\(^4\)

Consider now the structure of clauses:

25. Joelle wondered whether the boy ate

The lexical item whether is a Complementizer (as are that, if, and for in English) that often signals an embedded clause. Because each phrasal type has a head that shares the properties of the phrase, the Complementizer whether heads a Complementizer Phrase (a CP). Given these assumptions, consider the approximate structure of (25), shown in (26) (see next column).

In (26), the verb wonder appears in the matrix clause (the top-most or main clause) and the verb eat appears in the embedded clause. The IP dominates all nodes and branches to immediately dominate an NP and an I'. The subject NP immediately dominates an N', which immediately dominates an N (Joelle). The I' branches to immediately dominate its head, the INFL node, and the VP. The VP immediately dominates a V', which branches and immediately dominates its head V (wonder) and an embedded CP (whether the boy ate), which is the complement (argument) of the V. The CP, in turn, branches to a Specifier position (which is empty; more on this later) and to a C'. The C' branches and immediately dominates its head Complementizer (COMP) (whether) and an IP (the boy eat). The IP branches and immediately dominates the DET position (the) and an NP. The NP branches and immediately dominates an N' (which dominates the N) and an I', which branches and immediately dominates its head INFL node (+TNS) and a VP. The VP immediately dominates a V', which dominates its head V (eat).

To simplify matters a bit, consider the following generalizations shown in (27) and (28):

26. [Image of a diagram showing the structure of clauses]

27. All phrase structures have the same form: An XP (maximal projection), X's (intermediate projections), a head X (a lexical category), a complement (ZP) of the head that is on the same phrasal level as the head, a Specifier position, and, perhaps, an adjunct phrase (modifier; VP) that can attach above the head.

28. X-bar schema:

---

\(^3\)If the head occurs to the left of its complements in English, then what about the structure of NPs (e.g., the mechanic)? The article the seems to modify the N (mechanic), yet the "head" Noun appears to its right. One possibility that has been suggested by Abney (1987) and that is favored by many linguists is to consider NPs as Determiner Phrases (DPs), where the category DET is now the head of the DP and therefore falls to the left of its complement: DP → DET N.

\(^4\)More recent analyses (e.g., Pollock, 1989) suggest that the INFL projection can be divided into separated phrases for Tense (TNS) and Agreement (AGR), and AGR subdivided into AGR-S (subject) and AGR-O (object). Each of these has a functional category as head. This modification has become known as the "Split Infl Hypothesis."

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
According to (28) the head (X, sometimes referred to as X*), is an atomic element (a category) drawn from the lexicon. Phrase markers are projected from the head to intermediate levels (e.g., N*, V*, etc.) and to a maximal projection (e.g., NP, VP, etc.). The complement of the head is often called its argument, which is syntactically on the same phrasal level as the head (the head and its arguments are therefore said to be “sisters”). The specifier is immediately dominated by the maximal projection and is a sister to the X* level, and the adjunct is often immediately dominated by an intermediate projection. Importantly, the terms specifier, argument, and adjunct are not formal category terms (like NP, N*, etc.) but are, instead, relational terms so that we speak of the argument of X, the specifier of XP, and so on. For example, a subject NP is the specifier of IP, a direct object NP is an argument of the verb, and so forth.

This schema and the trees depicted in (14), (15), (24), and (26) assume binary branching (Kayne, 1984), whereby an X-bar category can at most immediately dominate two nodes. One reason behind constraining trees to binary branching has to do with learnability. It could be argued that if the binary branching constraint is part of a child’s tacit knowledge of language, then fewer decisions will need to be made regarding the structure of the language. Hence, acquiring a language will be “faster” and “less effortful” than if the grammar allowed unconstrained branching.

Because all phrase structures conform to the X-bar schema in (28), the acquisition of the phrasal geometry of sentences becomes a matter of acquiring the order in which the specifier, head, and adjuncts fall. According to the Principles and Parameters approach, children have, as part of their innate capacity to acquire language, knowledge of this X-bar schema. There is also a limited set of parameters that must be set via experience in their native language, parameters, for example, that allow children to order these constituents. As will be evident shortly, acquiring the lexicon (i.e., mental dictionary) of the language makes this task easier.

**Structural Relations**

Consider now two important formal relations among the nodes of a tree; the first of these is known as c-command:

29. A node X c-commands a node Y if and only if the first branching node dominating X also dominates Y.

Using (26) as an example, the subject NP c-commands I’ (and vice versa) because the first branching node dominating the NP is IP, which also dominates I’. The V c-commands CP (the verb’s complement or argument) since the first branching node dominating V is V’, which also dominates CP.

The fact that heads c-command their arguments sets the stage for another local relation, known as government:

30. A node X governs a node Y if and only if:
   (a) X c-commands Y,
   (b) X and Y are in the same (maximal) projection, and
   (c) no maximal projection intervenes between X and Y.

On most versions of government the governor must be a head (either a functional or lexical category). Using (26) once again as an example, a head (e.g., V) governs its internal argument(s) (e.g., CP in the matrix clause); that is, V c-commands CP, V and CP are in the same maximal projection (VP), there are no intervening maximal projections between V and CP, and V—the governor of CP—is a lexical category.

It turns out that the principle of Government yields several other principles, including subcategorization, thematic role assignment, Case assignment, and trace-antecedent relations. These will be discussed in the following sections.

**The Lexicon and Theta-Theory**

Before describing X-bar syntax, the point was made that phrase structure rules, by themselves, are too powerful; that is, their output can generate both grammatical and ungrammatical sentences. Again, since—at this point in our theorizing—lexical insertion can pull out any verb to fill the V-node, some sentences will turn out to be grammatical (e.g., “Joelle kicked the door”) and some will turn out ungrammatical (“Joelle thinks the door”). One way for the theory to constrain or to restrict the output of phrase structures to allow only well-formed sentences is through information represented in the Mental Lexicon or mental dictionary. The Lexicon contains representations of the list of words of a language that a speaker has acquired and stored in memory. Included in each word’s lexical entry is information about its phonological form (how it is spoken or read), its lexical category (verb, noun, preposition, etc.), semantic information (what it refers to in the real world), and importantly, the legal sentence environments the word is allowed to enter. As was shown above, for example, not
all words, in this case, verbs, can fit into all types of sentence structures.

It will turn out that verbs are crucial to accounts of syntactic representation as well as to accounts of sentence processing. Each verb carries two types of syntactic well-formeded information: strict subcategorization and argument structure; argument structure, in turn, interacts with the more semantic notion of thematic roles.

**Subcategorization**

By subcategorization we mean the type of syntactic environment into which a verb can enter. In particular, it is a characterization of the type of phrasal category that follows the verb (in head-first languages). Consider the following two examples:

31. [NP Dillon] [VP [V kissed] [NP Joelle]]
32. [NP Dillon] [VP [V put] [NP the toy] [PP on the shelf]]

The verb *kiss* requires a direct object NP (Joelle) and the verb *put* requires both an NP (the toy) and a PP (on the shelf). Thus, *kiss* is said to subcategorize for an NP, and *put* for an NP and PP. The fact that our knowledge of English tells us that *kiss* must be followed by an NP renders the sentence *Joelle kissed incompletely, or ungramatically; the subcategorization for *kiss* is not satisfied in this sentence. Verbs in principle can subcategorize for any phrase type, or combinations of phrases, but each verb chooses its allowable phrasal environment.

Such information is represented as part of the verb’s entry in the lexicon. Partial entries for *hit* and *put*, for example, would be:

33. *kiss* V (lexical category)

/ks/ (phonology)

[__.NP] (subcategorization)

34. *put* V (lexical category)

/put/ (phonology)

[__.NP PP] (subcategorization)

As shown in (33) and (34), information in the lexicon (i.e., subcategorization information) restricts the output of the phrase structure rules of a grammar to only well-formed instances of a language. That is, if it is assumed that the subcategorization frame for each verb must be satisfied in the sentence in which the verb is contained, then we ensure that the correct verb will be inserted in its proper structural configuration. Again, if the phrase structure allows the generation of a sentence like “Joelle ___ the door,” only a verb that allows a direct object NP can be inserted in the (blank) verb slot.

Indeed, there is a principle that formally describes this constraint on well-formedness:

35. The Projection Principle: Lexical representations (e.g., subcategorization possibilities) are projected to the syntax.

Given the lexical representations in (33) & (34), *Dillon kiss* and *Dillon put on the shelf* are ruled out by the Projection Principle because the subcategorization frame for each verb does not project to the syntax. That is, the verb *kiss* requires a direct object NP, and it is missing in the sentence *Dillon kiss*; the verb *put* requires both an NP and a PP, and the NP, again, is missing from *Dillon put on the shelf*. This principle is quite powerful and rationalizes syntax in the following sense: *When we acquire a lexical item, we acquire facts about the item’s syntactic privileges of occurrence.*

As shown in (26) with *wonder*, verbs can also subcategorize for sentential clauses:

36. [NP Joelle] knew [NP the answer]
37. [NP Joelle] knew [CP that the answer was correct]

(36) & (37) show that the verb *know* selects either a direct object NP or a sentential clause (CP), and thus would have the following (abbreviated) lexical entry:

38. *know* V

/now/

[__.NP CP]

It turns out that for a verb to subcategorize for a particular complement, it must govern that complement. For example, in (26) the matrix *V wonder* governs its complement, the CP *whether the boy ate*.

**Argument Structure**

Most sentences can be considered representations of relations between a predicate and its arguments, hence the term argument structure. An NP or a CP can be an argument of a verb if it occupies what is called an argument position (i.e., subject and complement positions). Unlike subcategorization information, argument structure is not concerned with the syntactic form of the phrasal categories a verb allows, but instead is concerned with the number of participants described by the verb. The transitive verb *kiss* requires two participants, a “kisser” and a “kissee”; thus, it selects for two arguments and hence a two-place argument structure. The verb *put* requires three participants and thus entails a three-place argument structure. Arguments and argument structures are often represented as variables (x, y, and z):

39. [Dillon x] [VP kissed [Joelle y]]
40. [Dillon x] [VP put [the ball y] on [the shelf z]]

Note that in (39) one of the arguments (y) falls within
the bracketed domain of the VP, and in (40) two of the arguments (y & z) also fall within the VP. These arguments are often called internal because they fall within the internal configuration of the VP; the subject argument (x) is often called external because it falls outside the maximal projection of the verb phrase.

The verb knowing also allows a two-place argument structure, although the arguments can have different syntactic realizations:

41. [Joelle x] [VP knew [NP the answer y]]
42. [Joelle x] [VP knew [CP that the answer was correct y]]

In (41) the internal, y-argument has a different semantic status than does the internal, y-argument in (42), although the verb is the same. Roughly, in (41) the argument is the Theme of “knowing” whereas in (42) the argument is an Event or Proposition. Thus, arguments have different syntactic as well as semantic roles to play in the sentence; these “semantic” roles are discussed in the next section.

**Thematic and Semantic Roles**

Consider the following sentences:

43. Joelle melted the wax.
44. The wax melted.

First, notice that the NP (the wax) appears to play the object role of the verb melted in (43), yet plays a subject role of melted in (44). Second, notice that although the NPs in the two sentences each serve a different role to the verb, they nevertheless seem to have something in common: The NP in both cases reflects an entity undergoing some sort of transformation (“melting”). In a sense, then, the NP seems to be playing a similar role in both sentences. Consider also:

45. The wax was melted by Joelle
46. It was Joelle who melted the wax
47. It was the wax that was melted by Joelle

Although the arguments of the verb melted (Joelle, the wax) occur in different positions in (45)–(47), they also seem to play a similar role. In each case Joelle appears to be the Agent of the melting, and in each case the wax appears to be the Theme of the melting. The generalization that arguments can play similar roles while appearing in different syntactic positions rationalizes, in part, the notion of thematic roles, which are thematic-semantic notions of the sort that answers the question “who did what to whom.” Each argument, then, takes on a certain thematic role (e.g., Agent, Experiencer, Theme, Patient, Goal, Benefactive, etc.). Each verb selects sets of thematic roles assigned to its arguments; each set of roles is called a thematic grid. Like both subcategorization and argument structure, thematic properties are written into the representation of the head (e.g., the Verb) as part of the lexical entry.

Thematic roles are assigned to arguments in the sentence, usually by the head, which has the property of being a theta-assigner. Theta-assigners are typically the lexical categories of verb, preposition, noun, and adjective (in contrast to functional categories, which are not theta-assigners). For example, the verb kiss requires a two-place argument structure (x y) and assigns to each argument a thematic role taken from its thematic grid; the preposition in is also a theta-assigner:

48. kiss V
   (Agent Patient) in P
   (Location)

[Dillon AGENT] kissed [Joelle PATIENT] in [the park LOCATION]

As (48) shows, the verb kiss assigns its thematic roles of Agent and Patient to the subject and direct object positions, respectively (technically, the Agent role is assigned by the VP, that is, the entire VP. Of course, the properties of the VP depend on the properties of the head V, so for present purposes we will assume that the V assigns a role to the external as well as internal arguments). *Theta Theory* states that the predicate assigns its associated thematic roles to particular grammatical positions. A transitive verb like kiss will have two thematic roles to assign, one associated with the subject position and one associated with the object position (in the Principles and Parameters approach, object position is internal to the VP, whereas subject position falls outside the VP (but see footnote 4). Within the PP (in the park), the preposition in is the head of the PP and assigns the role of Location to its argument (the park).

Two important principles constrain assignment of thematic roles in the syntax: the Projection Principle (16) and the Theta Criterion:

49. Theta Criterion: Every argument (e.g., NP) in a sentence must receive one and only one thematic role; each thematic role associated with a theta-assigner must be assigned to one and only one argument.

The Theta-Criterion, much like the Projection Principle, ensures that a verb's (and a preposition's) thematic properties specified in a lexical entry will be assigned one-to-one to the arguments represented in the syntax. The syntax of a sentence, then, is determined to a large extent by the lexical properties of the head of each phrase. For example, if the verb requires one argument, only
that argument should be observed in the sentence. If a verb requires two arguments, both arguments must be observed; and the same holds for three argument verbs.\(^5\) Importantly, then, the Projection Principle can now be revised to not only include subcategorization information, but, crucially, argument structure and thematic information as well. Because the verb kiss, for example, requires a two-place argument structure, the thematic roles written into the verb's thematic grid must be "projected" to the syntax.

**Lexical Contrasts**

There are numerous lexical alternations that interact with the syntax in interesting ways, showing, again, how lexical properties project to the syntax. One of the most well-known of these contrasts is spray/load:

50a. Dillon sprayed Jello on the wall
b. Dillon sprayed the wall with Jello

51a. Dillon loaded the blocks on the truck
b. Dillon loaded the truck with blocks

The verbs in (50) \& (51) are said to alternate. In (50a) the direct object (the Theme) is Jello whereas the object of the preposition (serving the role of Location) is the walls. Yet, in (50b) the order is reversed; the location alternates syntactic positions with the object of the verb. There is also a subtle semantic difference between the (a) and (b) versions. When the object of the preposition does not express the location (as in [b]), the sentence has the sense that the entire wall was sprayed; this is often called the holistic interpretation. When the location argument is signaled by the object of the preposition (as in [a]), it has a partitive interpretation. This

\(^*\)Verbs with two internal arguments (e.g., double-object verbs) present interesting problems for the binary branching hypothesis discussed in the section on X-bar theory. Consider:

(i) Joanna sent the letter to Mitzi
(ii) Joanna sent Mitzi the letter

(i) and (ii) appear to be related since they contain similar arguments and thematic roles, though the arguments in each appear in different syntactic positions. How does the theory account for these facts? Not so easily, it turns out. Recall that a head and its arguments are said to be "sisters" (that is, they are on the same phrasal level in the tree), and that the head is said to co-command its arguments. Skipping the details, if binary branching is retained, only one argument can be sister to the head V; the other argument is attached to a V that is higher in the tree. Thus, in (i) V co-commands its argument NP the letter, but it doesn't co-command the second internal argument, the NP Mitzi in the PP to Mitzi. Relatedly, in (ii) a different argument is sister to the head (the NP Mitzi) whereas the second internal argument (the letter) attaches higher up in the tree and is not co-commanded by V. If both (i) and (ii) are related, as they indeed seem to be, then why should the head co-command a different argument in each case?

Various solutions have been proposed to account for double-object verbs. A full discussion of these issues is far beyond the scope of this paper. The reader is referred to Barsa and Lassik (1986), who suggest a "flat" rather than a binary branching structure, and to Larson (1988) and Jackendoff (1990b) where the relations among various principles of the grammar and double-object verbs are explored.

same semantic distinction holds true for the examples in (51). The semantic properties of the verb, then, translate to distinct syntactic forms.

Another well-studied example is the causative. Consider the verb melted again:

52a. Joelle melted the wax.
b. The wax melted.

In (52a) the verb is used as a transitive (that is, with a direct object) and in (52b), as an intransitive. In the transitive use the subject causes the action described by the verb. In the intransitive use the verb describes a state. Note that these semantic properties of the verb translate to different syntactic forms.

There are many of these alternations found in language, and their existence suggests that a verb's argument structure is determined, to a large extent, by its meaning. Indeed, without getting into the details here, work suggests that theoretical notions like thematic roles might be better explained by reference to a theory of lexical-conceptual structure (LCS); that is, to a theory of semantics (see, for example, Jackendoff, 1990a; for more on lexical contrasts, see Levin, 1993; for more on argument structure, see Grimshaw, 1990). A theory of LCS would claim that thematic roles like Agent, Patient, and so on, are descriptive simplifications for another level of representation. We now move out of the Lexicon and examine further how lexical properties interact with the syntax.

**Trace-Theory and Move-Alpha**

Consider now the following sentences, and how they fit into the theory thus far:

53. Dillon kissed Joelle (active)
54. Joelle was kissed ______ by Dillon (passive)
55. It was Joelle who Dillon kissed ______ (object cleft)
56. Who did Dillon kiss ______? (wh-question)

Despite appearing in different grammatical positions, Dillon seems to be the Agent and Joelle seems to be the Patient in (53)–(56) (in [56], Joelle is replaced by the NP Who). The Projection Principle and the lexical entry for kiss require that kiss have a direct object argument position to which the Patient role can be assigned. And the Theta-Criterion requires that each argument position be assigned a thematic role represented in the verb's thematic grid. But it appears that (54)–(56) should be ruled out (ruled as ungrammatical) by both principles since there doesn't seem to be a direct object position, as shown by the "gap" (____). Our intuitions, however, tell us that these sentences are grammatical. But, how can they be grammatical if there
doesn’t seem to be an argument position to which to assign the Patient role?

One possibility is that there is, indeed, a direct object position to which the role of Patient is assigned, and this position is just in the place where we expect it to be, immediately after the verb. According to the theory under consideration, it turns out that there is such a position in (54)–(56), just as there is in (53). However, in the former the positions are said to be lexically unfilled (or “empty”).

Assume that the direct object (Joelle in the examples above) originates in the canonical post-verb direct object position and moves to a pre-verb position. The Projection Principle enforces proper thematic role assignment by requiring insertion of an empty category or trace into the position from which the category has moved. A trace is like a “ghost” that is left behind when the NP moves; it is a lexically unfilled position acting as a “place holder.” The trace is then linked or co-indexed with the moved category, forming a co-reference relation between the two positions. The thematic role—in this case, the Patient—is assigned to the (original, direct object) position occupied by the trace, and the moved category, called the antecedent to the trace, inherits the thematic role. Specifically, the trace of the movement and the NP that moved form a chain. Briefly, a chain may consist of two or more members that are co-indexed; each chain is considered a single argument (there are one-member chains as well, those NPs that do not co-refer with anything, like the NP Dillon in [53]–[56]). The theta-criterion can now be revised to include chains:

57. Theta-criterion: Thematic roles in a lexical entry are assigned to chains, and each chain receives one and only one thematic role.

All this brings us to, arguably, the most important insight of the Principles and Parameters theory (born from Chomsky’s original work in the 1950s and 60s; see Chomsky, 1957, 1965): There are, at this point in the discussion, (at least) two levels of representation for the sentence. The syntactic structure that exists before the NP is moved is known as the underlying or d-structure and that which exists after the NP is moved is the s-structure; the syntactic operation that moves the NP is known technically as the transformation move-alpha. An approximation of the underlying structure for the wh-question (56) “Who did Dillon kiss?” is shown in (58a), and the s-structure for the sentence is shown in (58b):

58a. Dillon did kiss [who]?
   — Underlying form (D-structure)
   [Did Dillon kiss [who]?
   — Subject-aux inversion
   (head-to-head movement)

58b. [Who]; did Dillon kiss [trace]?
   — S-structure

In the d-structure shown in (58a), the direct object position is filled by the wh-word who. After who moves to the front of the sentence via move-alpha, it leaves behind a trace with which it is co-indexed. Note now that in the s-structure, there is a direct object position to receive the role of Patient, although it is now occupied by the trace (i.e., it is lexically unfilled). Also, note that the thematic role of Patient is assigned to the trace position at S-structure, and the role is inherited by the moved wh-word who because the trace and its moved wh-phrase co-refer and form a theta-chain. Thus, the lexical properties of the verb kiss, the theta-criterion, and the Projection Principle are now satisfied.

Sentences (54) & (55) receive a similar analysis, as shown by the abbreviated s-structures in (59) & (60):

59. [Joelle], was kissed [trace], by Dillon
60. It was [Joelle], [who], Dillon kissed [trace],

Again, the lexical properties of the verb kiss require two arguments that are assigned the roles of Agent and Patient from its thematic grid, and there are now two positions to which the thematic roles can be assigned: the subject position that is lexically filled by the NP Dillon, and the direct object position that is filled by an empty category, the trace. In (59), the direct object NP (Joelle) has been moved out of its canonical (i.e., “basic”) post-verb position to a position at the “front” of the sentence, leaving behind a trace of that movement. The trace and its antecedent co-refer, and thus form a theta-chain that is assigned the Patient role by the verb. In (60) who has moved from its post-verb position to a pre-verb position. The trace, who, and Joelle are co-indexed (i.e., they all co-refer) and thus form a theta-chain; the chain is assigned the Patient role.

Another theoretical distinction is observed when the passive construction is compared to wh-constructions (wh-questions and relative clause structures). These two constructions are distinguished by what has been called NP-movement and Wh-movement, respectively (both movement types are subsumed under the general rule Move-alpha). Such a distinction has to do with, among
other things, what has been called the landing site of the moved constituent, or the position to which a constituent moves. NP-movement involves moving an NP to an argument position (A-position). An A-position, again, is a position that can receive a thematic role by a theta-assigning lexical category (a head). Typical A-positions in English include, again, the subject, direct object, and indirect object (object-of-preposition). Wh-movement involves moving a constituent to a non-argument position (called an A-bar position; where “bar” has the meaning of “null”), a position that does not receive a thematic role and is typically a functional category such as the Specifier position. To see the difference, first consider the following partial d-structure (i.e., underlying form) for the passive *Joelle was kissed by Dillon*:

In (61), the direct object NP *Joelle* originates in the post-verb argument position and moves to the subject position of the sentence, occupied by an empty NP position (e) immediately dominated by the IP (we will rationalize this movement in a later section). The NP is in an argument position (more specifically, it is the external argument of the verb and thus serves the role of Subject). When the direct object moves, it leaves behind a trace that is then co-indexed with the NP (the position to which it moves). The verb *kiss*, as a theta-assigner and head, assigns the Patient role to the chain consisting of the trace and the moved NP (*Joelle*). Omitting some details here, the preposition by, also a lexical category and theta-assigner, assigns the role of Agent to the NP *Dillon*.

Now consider once again the underlying form or d-structure for the wh-construction *Who did Dillon kiss?*, this time in terms of its phrase structure representation (see next column).

In (62), the post-verb NP (wh-morpheme) who moves to the empty (e) Specifier position, which is an A-bar position that does not receive a thematic role. Additionally, there is a second movement in (62), whereby INFL “raises” to COMP (i.e., *did* moves to COMP). This is called head-to-head movement because the head of one phrasal type (INFL) moves to the head of another phrasal type (COMP); this used to be called Subject-to-Aux inversion, explaining how, for example, *Will John go?* is derived from *John will go*. To review, then, NP-movement has an argument position landing site, yet wh-movement has a non-argument position landing site. Both types of movement leave behind a trace that is co-indexed with the position to which the constituent moves, and both form a chain whereby the moved constituent inherits the thematic role assigned to the internal argument position by the verb.

The Empty Category Principle (ECP)

Consider the following constraint on NP- and wh-traces:

63. The Empty Category Principle (ECP): A trace must be properly governed.

X properly governs Y if and only if:
(a) X governs Y (see (31), and
(b) X is a lexical category (so that X theta-assigns Y), or
(c) X and Y are co-indexed.

Using (62) as an example, the trace (in the direct object position, after *who* has moved to Spec of CP) is properly governed by V because V both governs the trace (it commands it and is in the same maximal projection as the V) and assigns a theta-role to the position occupied by the trace.

There are, then, two kinds of empty categories that have been described thus far: wh-traces and NP-traces. There is another empty category that shares some of the same properties of traces. Consider that clauses can be either finite, as in “Joelle wondered whether the boy ate,” wherein both the matrix and embedded clauses INFL is marked [+TNS] (refer back to [26]), or
non-finite. The subject of a non-finite clause is an empty category called \( \text{PRO} \) (or "big PRO"):  

64. Joelle wondered [whether \( \text{PRO} \) to go home]  

Although the matrix clause is finite and INFL is marked as \([+\text{TNS}]\) (i.e., "Joelle wondered"), the embedded clause containing the infinitive to go is \([-\text{TNS}]\). The subjects of the matrix clause (Joelle) and embedded clause (PRO) are said to co-refer much like what was shown with traces and their antecedents.

### Case Theory

Case theory, another component of our grammar, accounts for the distribution of NPs. Consequently, it also has important implications for movement of NPs in non-canonical constructions (e.g., the passive). Case theory was borne out of the traditional notion of case, whereby NPs and pronouns can be distinguished by, for example, accusative, nominative, and genitive morphological case. Modern English has an impoverished morphological case system compared to many other languages; now only pronouns retain case in the traditional sense, and not full NPs. For example, in the sentence *He gave his car to him*, he is nominative, his is accusative, and him is genitive. However, what the theory is concerned with here is not this traditional notion of case, but, instead, a more technical notion whereby each NP must be assigned structural Case. Consider, then, the following principle:

65. Case Filter: *NP if NP has phonetic content and no Case.

The Case Filter means that a sentence is ungrammatical if it contains a phonetically realized NP (an NP that is "sounded out" as opposed to an empty category) that is not assigned Case.

Here is a very brief summary of Case theory:

- Verbs and prepositions assign Accusative Case to their complements (arguments).
- If a clause is finite \(([+\text{TNS}])\), then Nominative Case is assigned to the subject NP by I (INFL). Note that the Case Filter (65) says nothing about the subject of non-finite clauses (e.g., "Joelle wondered [whether PRO to go home]") because the subject of the clause is PRO, which does not have phonetic content and therefore is invisible to the Case Filter.
- Case is typically assigned under Government

Examining Case theory further, by example, consider:

66. Joelle thought Dillon kissed Chester on the mouth  
The matrix verb *think* is tensed (i.e., \([+\text{TNS}]\)), thus INFL Case-marks the subject (Joelle), assigning to it Nominative Case. The embedded verb *kiss* is also tensed; thus, the embedded INFL assigns Nominative Case to the embedded subject (Dillon). The embedded verb *kiss* assigns Accusative Case to its object (Chester). The preposition *on* assigns Accusative Case to its object (*the mouth*). In each of these, the Case-assigner is a governor to its assignee.

Consider now how the Case Filter interacts with movement and many of the other principles that have been described thus far to yield the passive.

67. The toy was given ___ to Dillon (by Joelle)  
give (Agent Theme Goal)  

In (67), the direct object NP *(the toy)* has moved from its post-verb position to the front of the sentence (via NP-movement), leaving behind a trace. Note that the verb *give* has a set of thematic roles to assign, that the Projection Principle (35) states that lexical properties must be observed in the syntax, and that the Theta-Criterion (49) requires that all arguments must receive a thematic role and all thematic roles in a theta-grid must be discharged. Given these principles, why is it that (67) is not ill-formed by our theory? That is, if there is no external argument in the truncated passive of (67) that can receive the role of Agent, how is the Theta-Criterion satisfied? It turns out that the passive morphology *(en)* associated with the verb has several consequences, one of which is that the Agent role is absorbed. The verb assigns the Theme role to the chain consisting of the moved NP *(the toy)* and its trace. The Goal is assigned to the indirect object NP *(Dillon)*, and the Agent role is absorbed by the passive *-en*, allowing all three thematic roles to be discharged and thus satisfying the Theta-criterion. In the full passive version of (67), Agent cannot be assigned to the external argument *(Joelle)* since the passive morphology has absorbed this role (Agent is then assigned by the by-phrase to the external argument).

Like theta-role absorption, another property of the passive morphology is to absorb Case. *Give* is the Case-assigner to its complement *(the toy)* yet the passive morphology takes away the ability of *give* to assign structural Case. This leaves the NP without Case, suggesting that the Case Filter will intervene and rule (67) as ill formed. To circumvent a violation of the Case Filter, the direct object NP must move. Movement is to subject position, where Nominative Case can then be assigned by finite INFL. In an important sense, then, Case Theory requires movement of the NP in the passive construction.

There are several commonalities between Case and Theta theories. Both Case and Thematic roles are assigned under Government. Many Case-assigners are also Theta-assigners; many Case-assignees are also Theta-assignees. Passive morphology absorbs both Case and
Theta-roles. Without getting into the details here, it turns out that many of the properties of Case theory can indeed be derived from the properties of Theta theory, thus simplifying the task that children have in acquiring the facts that both theories entail.

**Binding Theory**

Recall from the beginning of this paper that our grammatical intuitions allow us to recognize that some lexical items co-refer and some do not:

68. Dillon hit himself

69. Dillon hit him

In (68), the lexical item in the direct object argument position (himself) refers back to the subject NP argument Dillon. In (69), however, the direct object argument him does not refer back to the subject NP. Thus, English allows two types of pronouns, a reflexive pronoun (or anaphor) shown in (68), and a personal pronoun (or pronominal) shown in (69). What generalization can be discovered about these two pronouns? First, consider that each of these constructions contains only a single clause. Next, notice that the reflexive pronoun appears to co-refer with an NP within that clause; such a relation between the two positions is called co-reference. The personal pronoun, on the other hand, does not co-refer with an NP within the clause. Given such a generalization about the two types of pronouns, consider constructions with two clauses (note that we are using co-indexation to describe the relation between anaphors/pronominals and their antecedents):

70. [Rico knows] [CP that Scott, hit himself]
71. *[Rico knows] [CP that Scott hit himself]
72. [Rico knows] [CP that Scott hit him]
73. *[Rico knows] [CP that Scott, hit him]
74. [Rico knows] [CP that Scott hit Lola]

Does the generalization about pronouns and their antecedents (those NP positions to which they co-refer) still hold in multi-clause constructions? It appears that it does, since in (70) the anaphor himself co-refers with an NP (Scott) within its own clause, and seems well-formed. However, the anaphor in (71) co-refers with the subject of the matrix clause (Rico) and seems ungrammatical (himself does not refer to Rico). Thus an anaphor (i.e., reflexive pronoun) requires reference within its local clause. In (72) the pronominal element him co-refers with the subject of the matrix clause (Rico) and seems well-formed, yet in (73) him co-refers (or is co-indexed) with the subject of the embedded clause (Scott) and indeed seems ungrammatical. Thus, pronominals seem to not require reference within a local clause. It turns out also that the personal pronoun does not have to co-refer with the subject of the matrix clause, as it does in (72). If we "strip" off the co-index from the subject NP (Rico), the interpretation of the construction would be that the personal pronoun him would refer to someone other than Rico, someone mentioned in the discourse, perhaps. Importantly, the generalization about personal pronouns still holds, however, because the pronoun would not get its reference in its own clause. Finally, (74) shows that the referential expression (R-expression) Lola does not co-refer with any other position, and thus cannot be co-indexed with any position in the sentence.

Given these data and the notions of C-command & Government described in the section on Structural Relations, the following describes more formally the principles of the Binding Theory:

75. Binding Theory:

X binds Y if and only if X commands Y and X and Y are co-indexed.

Principle A: An anaphor must be bound in its governing category (it must be c-commanded by an antecedent with which it is co-indexed)

Principle B: A pronominal must be free in its governing category (it must not be c-commanded within the same clause)

Principle C: An R-expression (lexical NP) must be free everywhere (it must not be c-commanded by anything)

A governing category is the smallest domain (NP or CP) that contains the anaphor/pronominal and its governor (see 30 for a definition of government). In (70) the anaphor himself is c-commanded by the subject of the embedded clause (Scott) and co-indexed with it. Therefore, it is bound in its governing category, which is the bracketed CP. Yet in (71) the anaphor is not c-commanded by the subject of the matrix clause (Rico). Because an anaphor must be bound in its governing category, (71) is ruled out by Principle A of the Binding Theory. In (72) the pronominal element him is not c-commanded by its antecedent Rico, and thus is free in its governing category. Yet in (73) the pronom is bound in its governing category, but because it is a pronoun (which must be free in its governing category) and not a reflexive, (73) is ruled out by Principle B. Finally, in (74) the R-expression Lola is not c-commanded nor co-indexed with any position.

**The ECP and its Relation to Binding**

The ECP (63) also suggests that there are similarities between the Theories of Binding and Traces, since both theories refer to the notion of Government. Indeed,
it turns out that what holds for overt NPs also holds for "empty" NPs or traces. We should expect this because both empty and phonologically realized categories are similar distributionally (i.e., they occur in the same positions) and share similar properties. Therefore, the principles of Binding also appear to hold for traces. Starting with Principle A (75), an anaphor (e.g., herself) must be bound in its governing category (roughly, within its own clause); so too must an NP-trace be bound in its governing category. Principle B states that a pronominal (e.g., her) must be free in its governing category. Although there are no empty categories in English that have this property, there is an empty category pro (or "little pro") in other languages (e.g., Spanish, Italian) that indeed does. These languages allow the pronoun to be "dropped" (hence the term pro-drop languages), yielding an empty category having all the properties of the lexically filled pronoun. Finally, Principle C states that an R-expression (e.g., Lola) must be free everywhere; so too must a wh-trace be free.

Features and Their Role in Binding

We have been assuming that lexical and functional categories are primitives; that is, they cannot be further decomposed. However, the theory suggests otherwise, and that, indeed, categories can be decomposed into a set of features and that these features specify what the categories have in common. For example, the lexical categories noun (N), verb (V), adjective (A), and proposition (P) each can be described in terms of a unique feature complex:

\[76. \quad N = [+N, -V] \]
\[V = [-N, +V] \]
\[A = [+N, +V] \]
\[P = [-N, -V] \]

Note that the categories V and P each are designated by the [-N] feature. This suggests that they behave similarly, and, indeed, they both assign Case and Theta-roles to their complements. Similarly, INFL can have the feature [+TNS]. INF marked as [+TNS] assigns Case to its subject; INF marked as [-TNS] does not. C (Comp) is also considered by some to be assigned the feature [+WH], thus either heading an embedded interrogative [+WH] or an embedded clause [-WH].

The Binding principles just described also interact with feature complexes. Reflexive pronouns, subject to Principle A, are said to have the feature complex [+Anaphor, -Pronominal]. Personal pronouns, subject to Principle B, have the feature complex [-Anaphor, +Pronominal]. Referential expressions, subject to Principle C, have the features [-Anaphor, -Pronominal]. Thus, the Binding theory has often been reconstituted into the following:

- An element that has the feature [+Anaphor] must be bound in its governing category.
- An element that has the feature [+Pronominal] must be free in its governing category.
- An element that has the features [-Anaphor, -Pronominal] must be free.

Finally, non-overt categories (e.g., wh-trace and NP-trace) can also be decomposed into sets of features. The idea here is that if features of Binding can characterize overt NPs, then they should also characterize non-overt NPs. Without delving into the details, recall that NP-movement (61) has an argument position landing site and yields an NP-trace. It turns out that the NP-trace is bound in its governing category, much like a reflexive pronoun with the feature [+Anaphor] is. Thus, both an anaphor and an NP-trace are subject to Principle A. Wh-movement has an A-bar (non-argument) position landing site and yields a wh-trace. Like an NP-trace, a wh-trace is c-commanded by its antecedent and is coindexed with it; thus, it is also bound in its governing category. However, the Binding theory is only concerned with binding from an A-position, and because the antecedent to the trace is in an A-bar position (Spec), it is not A-bound by anything in its governing category, just like an R-expression. Both an R-expression and a wh-trace, then, are subject to Principle C.

Principles and Parameters

Some of the principles of the theory of syntax under consideration here are suggested to be part of Universal Grammar (UG), the part of the grammar that is innately given by virtue of our human genetic code and that which constrains language acquisition. For example, the Head Principle, the Projection Principle, the Empty Category Principle, aspects of the Binding, X-bar, and Trace theories, the notions of d- and s-structure, are all assumed to be part of UG. Of course, languages are different, so it can't be that all languages treat these principles in exactly the same way. The aspect of the theory that captures the variation among languages is the notion of parameters. For example, the Head Principle states that each phrase must have a head that shares the same properties of the phrasal node projections. But it doesn't say anything about the order of the head to its complements. Evidence from English suggests that the head appears to the left of its complements or arguments, hence, English is a head-initial language. However, there are many languages where the head appears after its complements; these languages are known as head-final (e.g., Japanese). So, there is a two-choice parameter that the child is assumed to set given appropriate linguistic triggers (input). The same choice is made...
for the position of the Specifier and Adjunct Phrases. There are several other proposed parameters, including whether or not wh-movement is allowed by the grammar, whether or not pronouns can be omitted (the pro-drop parameter), the structural barriers to movement, and other features.7

Summary and Implications for Normal and Disordered Language

In this review of syntax, it was claimed that a native speaker’s knowledge of language involves lexical and functional category representations, phrasal representations that can be ordered hierarchically to describe sentence structure (captured by the X-bar schema), the principles of C-command and Government that describe relations among nodes in the tree, the lexicon and properties of lexical items (Theta theory and Case theory) and how they contribute to, and interact with, syntax, d- and s-structure realizations of sentences along with transformations or movement rules that relate the two, the three principles of Binding theory and how they related to Trace theory, and feature complexes. Many of these theoretical constructs have had a profound impact on accounts of language acquisition, sentence production, and sentence comprehension in both normal and disordered language users. A full discussion of this research is beyond the scope of the present paper; however, the following examples taken from the literature are offered:

The Garden Path Theory of Sentence Processing (e.g., Frazier, 1978; Frazier & Clifton, 1995)

For almost 2 decades this work has shown how linguistic theory is unavoidably linked to accounts of human sentence processing. In brief, this account claims that the human sentence processor constructs tree structures using syntactic representations from X-bar theory. The construction of these syntactic representations is constrained by various processing principles (e.g., minimal attachment: do not postulate any unnecessary nodes;

late closure: attach new items into the clause or phrase currently being processed). This account is closely associated with the theory of modularity, whereby the language processing system is claimed to consist of a number of independent, autonomous processors that do not interact until each has run its course of operation.

The Constraint Satisfaction Account of Sentence Processing (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Garnsey, 1994)

This work, as well, has shown the inescapable link between linguistic theory and sentence processing. However, instead of focusing on syntactic representations, this account claims that lexical representations (e.g., thematic roles, semantic properties) are the primary source of establishing the structure and meaning of a sentence. Unlike the garden path account, this account is highly interactive because all information (e.g., semantic, pragmatic, etc.) interacts initially during sentence processing.

Gap-Filling

This is psycholinguistic terminology for the trace-antecedent relation. Swinney and colleagues (e.g., Swinney, Ford, & Bresnan, 1989; Swinney & Osterhout, 1991) have found that when a lexically unfilled position (i.e., a trace) is encountered during the process of attempting to understand a sentence, that position is filled automatically and immediately by the legal antecedent, that is, by the co-indexed NP to which the unfilled position co-refers. Relatedly, Nicol (1988) has found that at the point where a reflexive pronoun is encountered in a sentence, only the grammatically appropriate, locally bound antecedent is accessed, and when a personal pronoun is encountered, only the non-locally bound antecedent is accessed.

Lexical Representations

Shapiro and colleagues (Shapiro, Brooks, Gordon, & Nagel, 1991; Shapiro, Zurif, & Grimshaw, 1987, 1989; see also Rubin, Newhoff, peach, & Shapiro, 1996) have found that when a verb is encountered in a sentence, all of the verb’s argument structure possibilities are activated. Once argument structure information is made available to the system, most sentence processing accounts hypothesize an operation that then links thematic roles to arguments in the sentence.

Sentence Comprehension in Aphasia

Here there is a host of findings that are intimately tied to linguistic theory. One of the most important sets of works that has shown the importance of linguistic theory to the description and explanation of aphasia is
from Grodzinsky (1990; 1995). Briefly, he has shown that the sentence comprehension patterns evinced by agrammatic Broca’s aphasic patients can be described in terms of the deletion of traces (see also Hickok, Zurif, & Canseco-Gonzales, 1993; Mauner, Fromkin, & Cornelli, 1993). Relatedly, Zurif and colleagues (Swinney & Zurif, 1995; Zurif, Swinney, Prather, Solomon, & Bushell, 1993) have shown that their agrammatic Broca’s aphasic patients do not normally re-access the antecedent to a trace during sentence comprehension, yet their Wernicke’s aphasic patients do. On the other hand, Shapiro, Gordon, Hack, and Killackey (1993) have found that Wernicke’s aphasic patients do not normally access the argument structure possibilities for verbs, yet Broca’s aphasic patients do. This “double-dissociation” suggests that the process of connecting a trace to its antecedent is independent from the process of accessing the argument-taking properties of the verb, and that, perhaps, different brain regions are responsible for the normal operation of each.

Sentence Production in Aphasia

Friedmann and Grodzinsky (in press), and Hagiwara (1995) have recently shown how the sentence production patterns of agrammatic aphasic patients can be described in terms of the phrase structure tree. For example, adopting the Split-Infl hypothesis, Friedmann and Grodzinsky propose that the different levels of severity observed in agrammatic production might be explained by where in the phrase structure tree a defective node (e.g., Tense, Agreement, COMP) occurs; the lower the node in the tree, the more severe the impairment since the tree cannot project higher than the impaired node.

Treatment of Language Deficits

Thompson and colleagues (Thompson & Shapiro, 1994; 1995; Thompson et al., 1997) have shown in a series of studies that the lexical and syntactic properties of sentences need to be considered in treatment programs for agrammatic aphasia. Training production of complex structures that contain traces (e.g., wh-questions, relative clauses, passives) results not only in learning of trained structures, but also results in generalization to untrained sentences that share linguistic properties. For example, training wh-movement structures (e.g., wh-questions; see [60] above) generalizes to other structures relying on wh-movement (e.g., relative clauses) but not to structures relying on NP-movement (e.g., passives; see [59] above).

Language Acquisition and Language Disorders in Children

Several recent accounts of normal language acquisition have referred to X-bar theory to explain the relatively impoverished use of functional categories in early grammar. For example, Lebeaux (1988) and Radford (1990) have claimed that early grammar lacks the functional projections DET, INFL, and COMP. Thus, the difference between early and adult grammars is structural. Hyams (1996), however, has claimed that early grammar has intact functional projections, but that the projections are underspecified in terms of their link to the discourse. Relatedly, there have been several attempts to describe specific language impairment (SLI) in children by referring to the development of use of functional projections (e.g., Leonard, 1995; Rice, Wexler, & Cleave, 1995).

There are, of course, many more examples of work that relate linguistic theory to language performance, far too numerous to do justice to here. The reader is referred to the recent special issues on Agrammatism in Brain and Language (1995) and to the Journal of Psycholinguistic Research (specifically, the special issues from the CUNY Human Sentence Processing Conference), Language Acquisition, as well as to articles in the Journal of Speech and Hearing Research.

To conclude, understanding much of the research in language acquisition, comprehension, and production is difficult without at least an introductory knowledge of syntax. I hope that this tutorial will serve as an entry into this research.

Acknowledgments

I would like to acknowledge the support from NIH (NIDCD) grants DC00494 and DC01848. I would also like to thank Dr. Cynthia Thompson and two anonymous reviewers for their very helpful comments on an earlier draft of this paper.

References


Berkeley Linguistics Society, University of California, Berkeley.


Received September 26, 1995
Accepted September 30, 1996

Contact author: Lewis P. Shapiro, PhD, San Diego State University, Department of Communicative Disorders, 5500 Campanile Drive, San Diego, CA 92182-1518. E-mail: Shapiro@mail.sdsu.edu