Jurafsky D and Martin JH (2000) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Upper Saddle River, NJ: Prentice Hall.

Mitchell DC, Cuetos F, Corley MMB and Brysbaert M (1995) Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research* **24**: 469–488.

Pickering MJ, Traxler MJ and Crocker MW (2000) Ambiguity resolution in sentence processing: evidence against frequency-based accounts. *Journal of Memory and Language* **43**: 447–475.

Tabor W, Juliano C and Tanenhaus MK (1997) Parsing in a dynamical system: an attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes* **12**: 211–271.

Thomas J and Short M (eds) (1996) *Using Corpora for language research*. London: Longman.

Syntax

Introductory article

Colin Phillips, University of Maryland, College Park, Maryland, USA

CONTENTS

Syntactic theory aims to explain how people combine words to form sentences, and how children attain knowledge of sentence structure.

GOALS OF SYNTACTIC THEORY

Syntactic theory aims to provide an account of how people combine words to form sentences. A common feature of all human languages is that speakers draw upon a finite set of memorized words and morphemes (i.e. minimal meaning-bearing elements) to create a potentially infinite set of sentences. This property of *discrete infinity* allows speakers to express and understand countless novel sentences that have never been uttered before, and hence forms the basis of the creativity of human language. Syntactic theory is concerned with what speakers know about how to form sentences, and how speakers acquire that knowledge.

For example, speakers of English know that 'dogs chase cats' and 'cats chase dogs' are possible sentences of English, but have different meanings. Speakers know that 'chase dogs cats' is not a possible sentence of the language, and that 'cats dogs chase' is possible in specific discourse contexts, as in 'cats, dogs chase, but mice, they flee'. Speakers' knowledge of possible word combinations is often referred to as the *(mental) grammar*.

An accurate model of a speaker's knowledge of his or her language should minimally be able to generate all and only the possible sentences of the language. For this reason, syntactic theory is often known as *generative grammar*. In the 1950s, early attempts by Noam Chomsky and others to create explicit generative grammars quickly revealed that speakers' knowledge of syntax is a good deal more complex than had been anticipated. Research on syntactic theory has relied primarily upon speakers' intuitive judgments about the well-formedness ('grammaticality') of sentences of their language. Since grammaticality judgments can be gathered relatively easily, syntactic theory has amassed a large database of findings about an ever more diverse set of languages.

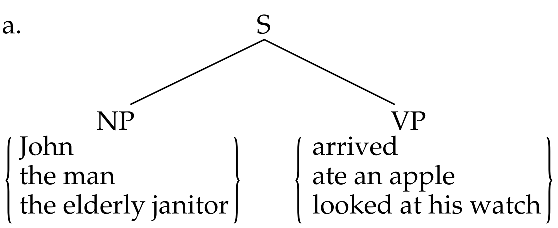
The complexity of syntactic knowledge sharpens the problem of how language is learned. Research on language acquisition has demonstrated that children know much of the grammar of their language before they are old enough to understand explicit instruction about grammar. Therefore, a primary challenge for syntactic theory has been to understand how a child can learn any language, relatively effortlessly, and without explicit instruction. Research on comparative syntax has met this challenge by seeking to characterize human languages in terms of universal syntactic properties, which may reflect the child's innate knowledge, and non-universal clusters of syntactic properties that pattern together across languages, and hence may be learned as a group. Thus, the study of comparative syntax and the study of language learning are closely related. (*See* **Innateness and Universal Grammar**)

FUNDAMENTALS OF SYNTACTIC THEORY

Discrete Infinity

Almost all accounts of the discrete infinity property of natural language syntax start from the notion that sentences consist of more than just sequences of words. In the minds of speakers and listeners, sentences are hierarchically structured representations, in which words are grouped together to form phrases, which in turn combine to form larger phrases. For example, a minimal sentence of English, such as 'John arrived', contains a subject and a predicate, but the roles of subject and predicate may be replaced by phrases of arbitrary complexity. By representing possible subjects and predicates as *noun phrases* (NPs) and *verb phrases* (VPs) respectively, the structure of many possible sentences (S) can be captured. This basic 'template' for sentences of English can be expressed as a tree structure, as in (1a), or as a phrase structure rule, as in (1b).

a.



$$\text{b. } S \rightarrow NP\ VP \qquad (1)$$

Just as rules like S → NP VP provide templates for sentences, templates can also be specified for the internal structure of noun phrases, verb phrases, and many other phrase-types. Even a small number of phrase structure rules and a small lexicon can generate large numbers of sentences. With only the five phrase structure rules in (2) and a 30-word lexicon (consisting of 10 nouns, 10 determiners, and 10 verbs) 122,100 different sentences can be generated.

$$\begin{aligned} S &\rightarrow NP\ VP \\ VP &\rightarrow V\ NP \\ VP &\rightarrow V \\ NP &\rightarrow Det\ NP \\ NP &\rightarrow N \end{aligned} \qquad (2)$$

Rules that allow a phrase to be embedded inside another phrase of the same type are known as *recursive* rules. Coordination (3), modification (4), and sentential complementation (5) all involve recursion. They can thus be invoked arbitrarily many times in a single sentence. Such rules increase the expressive power of the grammar from merely vast to clearly infinite. There are obvious practical limitations on the length and complexity of naturally occurring sentences, but such limitations are typically attributed to independent limitations on attention and memory.

$$\begin{aligned} NP &\rightarrow NP\ Conj\ NP \\ VP &\rightarrow VP\ Conj\ VP \\ Conj &\rightarrow \textit{and} \end{aligned} \qquad (3)$$

$$\begin{aligned} VP &\rightarrow VP\ PP \\ NP &\rightarrow NP\ PP \end{aligned} \qquad (4)$$

$$\begin{aligned} VP &\rightarrow V\ S' \\ S' &\rightarrow Comp\ S \\ Comp &\rightarrow \textit{that} \end{aligned} \qquad (5)$$

Although the rules listed in (1)–(5) fall far short of the expressive power of English, even this small fragment shows how natural language syntax uses finite means to generate infinitely many sentences. (*See* **Phrase Structure and X-bar Theory**)

Motivating Structures: Constituency

The syntactician's toolbox includes a number of structural tests that can be used as aids in diagnosing sentence structures; for example, *constituents* of sentences can generally be conjuncts in coordinate structures, as is shown for NPs and VPs in (6a, b). Other tests that show the constituency of VPs include substitution of the expression 'do so' for a VP (7a), and fronting of the VP to a clause-initial position (7b).

a. Wallace fetched[$_{NP}$ the cheese] and [$_{NP}$ the crackers]
b. Wallace[$_{VP}$ sliced the cheese] and [$_{VP}$ opened the crackers] (6)

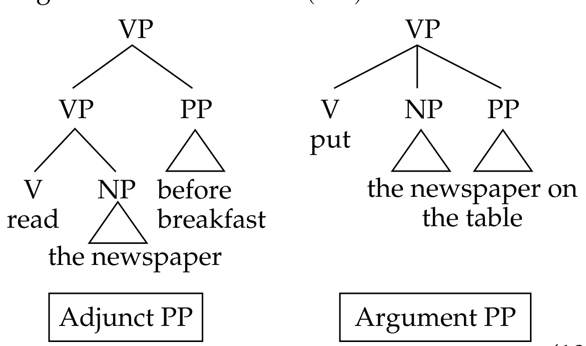
a. Wallace [$_{VP}$ read the newspaper] and Gromit [$_{VP}$ did so] too.
b. Wallace wanted to [$_{VP}$ impress Wendolene], and [$_{VP}$ impress Wendolene] he did. (7)

Constituency tests like those shown in (7) can be used to demonstrate that prepositional phrases (PPs) that are adjuncts (i.e. optional phrases) of VP recursively expand the VP, whereas PPs that are arguments of the verb (roughly, required phrases) do not. (8) shows that when *do so* substitution applies to a VP containing an adjunct-PP, the PP may be targeted or ignored by *do so* substitution. This indicates that there is a smaller VP-constituent that excludes the adjunct-PP. In contrast, (9) shows that an argument-PP cannot be ignored by *do so* substitution. If the argument-PP is 'stranded' by substitution (9b), the result is ungrammatical (indicated by the asterisk). This indicates that argument-PPs are contained within the smallest VP constituent. These contrasts motivate the VP-structures shown in (10).

a. Wallace [$_{VP}$[$_{VP}$ read the newspaper] [$_{PP}$ before breakfast]], and Gromit [$_{VP}$[$_{VP}$ did so]] too.
b. Wallace [$_{VP}$[$_{VP}$ read the newspaper] [$_{PP}$ before breakfast]], and Gromit [$_{VP}$[$_{VP}$ did so]] [$_{PP}$ at lunchtime] (8)

a. Wallace [$_{VP}$ put the newspaper [$_{PP}$ on the table]], and Gromit [$_{VP}$ did so] too.
b. *Wallace [$_{VP}$ put the newspaper [$_{PP}$ on the table]], and Gromit did so [$_{PP}$ on the floor]. (9)

Adjunct PP : VP → VP PP
Argument PP : VP → V (NP) PP

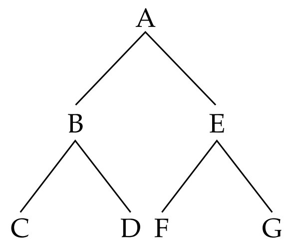


(10)

Motivating Structures: Hierarchy

In addition to tools that show which groups of words form constituents, other tests diagnose the *hierarchical* relations among positions in a structure. A striking finding of syntactic research is that many grammatical phenomena are sensitive to a structural relationship known as *c-command*, which is similar to the logical notion of *scope*. A node c-commands its sister and any nodes contained inside its sister. Thus, in the structure in (12), node B c-commands its sister, node E, and nodes F and G contained inside its sister. On the other hand, node C c-commands its sister, node D, and no others.

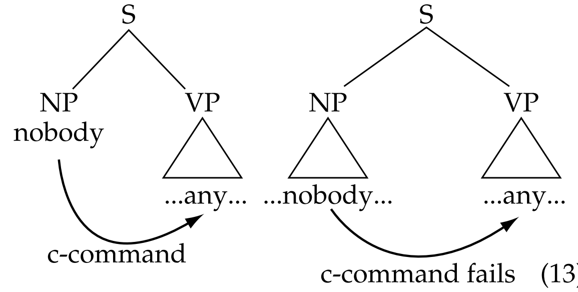
C-Command
A node c-commands its sister and all nodes dominated by its sister. (11)



(12)

One syntactic phenomenon that is sensitive to the c-command relation involves *Negative Polarity Items* (NPIs), such as 'anybody' or 'ever', which are only possible when they are c-commanded by an appropriate licenser, typically a negative expression such as 'not' or 'nobody'. NPIs are possible when the negative expression c-commands the NPI (13a–d) but are impossible when the negative expression fails to c-command the NPI (13e–h), because the negative expression is embedded inside a subject NP.

a. Wallace didn't find any cheese.
b. Nobody found any cheese.
c. Wallace didn't think that he would ever refuse cheese.
d. Nobody thought that Wallace would ever refuse cheese.
e. *[$_{NP}$ The fact that Wallace didn't like the cheese] amazed anybody.
f. *[$_{NP}$The fact that nobody liked the cheese] amazed anybody.
g. *[$_{NP}$ The person that Wallace didn't notice] thought that Gromit would ever return.
h. *[$_{NP}$The person that nobody noticed] thought that Gromit would ever return.



(13)

Multiple Roles: Transformations

In any sentence, each word or phrase occurs in a unique position in the linear sequence of words. However, many words and phrases appear to participate in *multiple* structural relations in the sentence. A central concern of syntactic research since the 1950s has been to understand how individual phrases can assume multiple roles in a sentence.

The multifunctionality of phrases has been most fully explored in the case of NPs. First, speakers represent the *thematic role* of each NP in a sentence. *Agent* thematic roles are canonically realized on subject NPs, *theme* thematic roles are canonically realized on direct object NPs, and thematic roles such as *goal*, *beneficiary*, or *location* are canonically realized inside PPs (14).

Wallace sent the flowers to Wendolene
agent *theme* *goal* (14)

However, it is important to distinguish thematic roles from *grammatical relations* such as *subject* and *object*, since thematic roles can be realized in different grammatical relations. The theme argument of the underlined verb 'steal' is realized as a direct object in the active sentence in (15a), as a subject in the passive sentence in (15b), and as the subject of a higher clause in the raising construction in (15c).

a. The penguin stole the diamond.
theme *active*
b. The diamond was stolen by the penguin.
theme *passive*
c. The diamond seemed to have been stolen.
theme *raising + passive*
(15)

It is also important to distinguish thematic roles and grammatical relations from the structural positions in which NPs receive their *scopal* interpretation. In English, the scope of *wh*-phrases such as *who*, *what*, or *which boy* is marked by their surface position: they are positioned sentence-initially in direct *wh*-questions (16a), and in clause-initial position of an embedded clause in indirect *wh*-questions (16b).

a. Which story did the teacher know that the children like?
theme
direct object
matrix scope
b. The teacher knew which story the children always like.
theme
direct object
embedded scope (16)

In English the scope of quantified NPs such as 'every boy' is not encoded in surface word order, with the consequence that sentences like (17) are scopally ambiguous. However, in other languages, such as Hungarian and Japanese, the scope of quantificational NPs is often fixed by surface word order.

Some student answered every question correctly.
some > every: *a single student got all the questions correct*
every > some: *for each question, there was at least one student who got it right* (17)

Since speakers must distinguish for any NP a representation of its thematic role, its grammatical relation, and its scope, it is important to ask how these different structural positions are related to one another. One well-known answer to this question is that speakers represent multiple structures for each sentence, one structure each for thematic role assignments, grammatical relations, and scope relations. The thematic representation is typically known as the *deep structure* (or *d-structure*) of the sentence, and it is related to the other levels of representation – *surface structure* (or *s-structure*) and *logical form* (LF) – by transformational processes which move NPs from thematic positions to positions associated with grammatical relations, and scope positions. This *Transformational Grammar* approach was proposed by Chomsky in the 1950s and has been elaborated and modified by many linguists since that time.

(18) shows a sample transformational derivation of a *wh*-question. The NP 'who' receives the theme thematic role as the sister of the verb in the d-structure representation in (18a). In (18b) it moves to subject position of the embedded passive clause. Finally, in (18c) it moves from subject position to its clause-initial scope position. Following a now standard elaboration of the transformational approach, (18) represents the residue of movement as phonologically null elements known as *traces* (*t*).

a. the announcer believed had been elected who
b. the announcer believed who_i had been elected t_i
c. who_i did the announcer believe t_i had been elected t_i
(18)

Some alternative approaches share with Transformational Grammar the assumption that sentence structure representations consist of multiple levels, but deny that there is a derivational relationship among the levels of representation (e.g. Lexical Functional Grammar, Categorial Grammar). Other approaches assume that the thematic, grammatical relation, and scope properties of an NP are all represented in a single enriched phrase structure, but that only one of these positions is normally encoded phonologically (e.g. Head-Driven Phrase Structure Grammar: see further below).

Types of Dependencies

Speakers' syntactic representations encode a variety of structural relations among words and phrases. The most basic relations are the groupings of words and phrases that form hierarchical constituent structures. These relations can be expressed entirely in terms of *sisterhood* and *dominance* among phrase structure nodes. The discussion above has introduced a number of other syntactic relations, or *dependencies*.

The syntactic relation between a pronoun or reflexive and the NP from which it takes its reference is known as a *binding* dependency. Such dependencies have been studied extensively under the rubric of *Binding Theory*.

a. Wallace$_i$ likes himself$_i$ — *local*
b. Wallace$_i$ thinks that Wendolene likes him$_i$ — *nonlocal* (19)

(*See* **Binding Theory**)

In *VP-ellipsis* constructions the VP in the second conjunct is dependent on the VP in the first conjunct for its interpretation (20). Transformational analyses of *wh*-questions (21a) and relative clauses (21b) treat the relationship between the *wh*-phrase and the trace as a binding relation between a *wh*-operator and a variable.

Wallace [$_{VP}$ likes cheese]$_i$ and Gromit does [$_{VP}$]$_i$ too. (20)

a. Who$_i$ did the voters elect t_i
b. The man who$_i$ the voters elected t_i (21)

(*See* **Ellipsis**)

A leading question in research on referential dependencies involves how closely related the different types of referential dependencies are: does each type of dependency follow independent principles, or do they follow the same principles?

CONSTRAINTS ON DEPENDENCIES

Wh-movement and related phenomena have been among the most extensively investigated topics in syntactic research, giving rise to a wealth of findings. By virtue of the length of *wh*-dependencies, syntacticians can manipulate which structural positions participate in the *wh*-dependency, and which structural positions the dependency crosses. *Wh*-dependencies have thus served as a kind of 'magnifying glass' for the investigation of syntactic dependencies.

Wh-dependencies can span many clauses, in fact arbitrarily many clauses, and thus they are often referred to as *unbounded dependencies*. In (22), the *wh*-phrase has been *extracted* from a number of embedded clauses, each of which is the *complement* (direct object) of the next higher verb.

Which candidate$_i$ did the court fear [that the public might conclude [that the voters had elected t_i]] (22)

However, in a tradition of research beginning with influential work in the late 1960s by John Robert Ross, it has been found that there are many syntactic environments which *wh*-extraction cannot cross. Following Ross's terminology, the environments that block extraction are known as *islands*, and restrictions on extraction are known as *island constraints*.

Relative clauses create islands for extraction (23a), as do indirect questions (23b), complements of NPs (23c), subjects (23d), and adjunct clauses (23e). Extraction from definite NPs or NPs with a possessor is highly marked, although indefinite NPs create no such difficulties (24). If a phrase is extracted from one conjunct of a coordinate structure, it must also be extracted from the other conjuncts (25a, b).

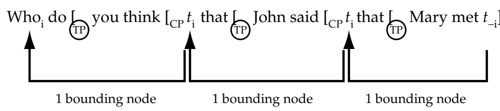
a. *Who$_i$ did the court upset the voters [who favored t_i]
b. *Who$_i$ did Bill wonder [whether his new outfit would shock t_i]
c. *What$_i$ did Sarah believe [the rumor that Ed was willing to spend t_i]
d. *Who$_i$ did [the fact that the president nominated t_i] upset the opposition party?
e. *What$_i$ did Wallace eat the cheese [while he was reading t_i] (23)

Who$_i$ did Sally hear a? the *Helen's story about t_i? (24)

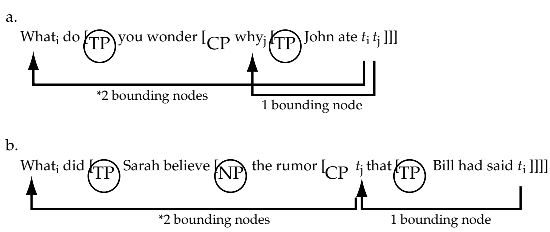
a. *What$_i$ did [Gromit read the newspaper] and [Wallace eat t_i]
b. What$_i$ did [Gromit read t_i] and [Wallace eat t_i] (25)

The examples in (22)–(25) raise the question of why *wh*-dependencies can be arbitrarily long as in (22), but not as in (23)–(25). Work on this question led to the proposal that, contrary to appearances, all *wh*-dependencies are *local*. If all *wh*-movement is local, proceeding one clause at a time, then apparent long-distance *wh*-movement turns out to be a series of local movements, each one targeting a landing site in the next higher CP (*Complementizer Phrase*), as in (26). If all *wh*-movement must be local, then it follows that relative clauses and embedded *wh*-questions create islands, because these are cases in which an intermediate landing site of movement is already filled (27a).

The best-known implementation of this proposal is the *Subjacency Constraint*, proposed by Chomsky in the early 1970s. In its original formulation the constraint blocks any *wh*-dependency that spans more than two *bounding nodes*, where the bounding nodes are defined as NP and TP (*Tense Phrase*: a more recent term for the 'S' node, recognizing *Tense* as the head of a clause). This formulation also explains the complex NP constraint violation in (23c) and (27b).



(26)



(27)

The proposal that a long-distance *wh*-dependency involves a sequence of local dependencies receives interesting support from a number of languages that show a syntactic residue of local movement. In certain varieties of Spanish, for example, subject–auxiliary inversion occurs in every clause in the path of *wh*-movement ((28): compare this to the English translation, in which inversion occurs only in the highest clause).

a. Juan pensaba que Pedro le había dicho que la revista había publicado ya el articolo.
Juan thought that Pedro him had told that the journal had published already the article
'Juan thought that Pedro had told him that the journal had published the article already.'

b. Qué pensaba Juan que le había dicho Pedro que había publicado la revista?
What thought Juan that him had told Pedro that had published the journal
'What did Juan think that Pedro had told him that the journal had published?'

c. *Qué pensaba Juan que Pedro le había dicho que la revista había publicado?
What thought Juan that Pedro him had told that the journal had published (28)

The island constraints restrict the nodes that a *wh*-dependency may cross. All of the examples presented so far involve extraction of a direct object *wh*-phrase. In addition, subject and adjunct *wh*-phrases in English are subject to tighter restrictions than object *wh*-phrases. For example, extraction of an embedded direct object *wh*-phrase is possible, irrespective of whether the embedded clause contains an overt complementizer 'that' (29a, b). However, extraction of an embedded subject *wh*-phrase is impossible if the complementizer is overt (29c). This constraint is known as the *that-trace constraint*, and it has been observed in many languages, as discussed further below.

a. Who$_i$ do you think t_i that John met t_i?
b. Who$_i$ do you think t_i John met t_i?
c. *Who$_i$ do you think t_i that t_i met John?
d. Who$_i$ do you think t_i t_i met John?

(29)

There are also differences in extraction possibilities between argument *wh*-phrases such as 'what' and 'which books', and adjunct *wh*-phrases such as 'why' and 'how' (see Further Reading).

A long-standing goal of syntactic research on unbounded dependencies has been to uncover a set of general principles that can explain the full variety of constraints on *wh*-dependencies. Although there have been many different attempts to unify the constraints on movement, two observations have been pervasive, and have featured in many different theories. First, if movement is required to be *local*, then it is subject to *intervention effects*, when a required landing site of movement is occupied by another element. Second, movement paths that include noncomplement nodes (subjects or adjuncts) are consistently more restricted than paths that include only complement nodes (sisters of heads). (*See* **Constraints on Movement**)

CROSS-LANGUAGE SIMILARITIES AND DIFFERENCES

A fully general theory of the mental representation of syntax clearly must handle the facts of all human languages. In addition, cross-linguistic investigations are important to accounts of how natural language syntax is learnable. *Universals* of syntax, or *principles*, may be part of the child's innate endowment, and thus not need to be learned. *Non-universal* syntactic properties must also be learnable within the constraints imposed by the time and evidence available to the child. When a set of syntactic properties *covaries* across languages, it is possible that the learner only needs to learn one member of the set of properties in order to draw appropriate conclusions about the entire set. Thus, an important goal of cross-linguistic syntax research is to find clusters of covarying syntactic properties, or *parameters*. This *Principles and Parameters* (P&P) approach to syntax has been most intensively investigated in transformational approaches to syntax but it can be applied equally well to other syntactic approaches. (*See* **Government–Binding Theory**)

Research on comparative syntax has discovered a number of striking cross-linguistic parallels between languages that appear very different on the surface. An example from Mohawk serves as an illustration.

One constraint on pronouns in English is that a pronoun cannot co-refer with an NP that it c-commands. This constraint ('Binding Condition C') accounts for the contrast between (30a) and (30b). The pronoun inside the subject NP fails to c-command the direct object in (30a), thereby allowing co-reference. On the other hand, the subject pronoun c-commands the name inside the object NP in (30b), thereby preventing co-reference.

a. [$_{NP}$ The book that he$_i$ bought] offended John$_i$
b. *He$_i$ bought [$_{NP}$ the book that offended John$_i$] (30)

Unlike English, which exhibits strict subject–verb–object (SVO) word order, Mohawk, an Iroquoian language spoken in Quebec and upstate New York, exhibits free word order, allowing all six possible permutations of subject, verb, and object, and also allows 'discontinuous constituents' in the form of split noun phrases (31). Based on such properties, languages like Mohawk have sometimes been described as 'nonconfigurational'. However, Mark Baker has demonstrated that Mohawk exhibits similar configurational asymmetries to English, as the contrast in (32) shows. This contrast can be explained by Binding Condition C, just as in English, provided that we attribute some degree of underlying configurational structure to Mohawk sentences.

Ne kíke wa-hi-yéna-' ne kwéskwes
NE this FACT-1sS/MsO-catch-PUNC NE pig
'I caught this pig.' (31)

a. Wa-ho-nakuni-' tsi Sak wa-hi-hrewaht-e'
FACT-NsS/MsO-anger-PUNC that Sak FACT-1sS/MsO-punish-PUNC
'That I punished Sak$_i$ made him$_i$ mad.' (co-reference possible)
b. Wa-shako-hrori-' tsi Sak wa-hi-hrewaht-e'
FACT-MsS/FsO-tell-PUNC that Sak FACT-1sS/MsO-punish-PUNC
'He$_i$ told her that I punished Sak$_i$.' (co-reference impossible) (32)

The similarity between English and Mohawk is striking, given how different the languages appear on the surface. Furthermore, evidence from many other languages suggests that Binding Condition C is a universal of natural language syntax, which may be part of innate linguistic knowledge. Consistent with this suggestion, studies by Stephen Crain and his colleagues have shown that children exhibit knowledge of Binding Condition C by their third birthday, which is as early as it has been possible to test this knowledge. This finding has been replicated in children learning other languages (e.g. Italian, Dutch, Russian). It is particularly encouraging news for a child learner of Mohawk that Binding Condition C need not be learned, since it is unlikely that the presence of the constraint could be guessed from the input to the Mohawk child.

Although syntactic research has uncovered many universals of language, there are clearly many properties that vary across languages and hence must be learned. The search for parametric clusters of syntactic properties has turned up a number of cross-language correlations. The most useful correlations are those that link abstract (and hence difficult-to-observe) syntactic properties with more easily observable syntactic properties. For example, the *that-trace* constraint introduced above does not apply in all languages: it applies in English (33a), but not in Italian (33b). This is not easily inferred from the language input to children, since it is not easy to observe the *absence* of a particular construction. However, the availability of *that-trace* sequences correlates cross-linguistically with the availability of postverbal subjects, which are readily observable in the input to the learner. Italian allows postverbal subjects (34b), but English does not (34a). This connection has been reinforced based on the study of many other languages. Therefore, the child learner may be able to learn whether the *that-trace* constraint applies in his language, by observing whether postverbal subjects are available.

a. *Who$_i$ did you say that t_i has written this book?
b. Chi$_i$ hai detto che t_i ha scritto questo libro?
who have-you said that has written this book
'Who did you say has written this book?' (33)

a. *Have arrived many students
b. Hanno arrivato molti studenti.
have-3pl arrived many students
'Many students have arrived.' (34)

VARIANTS OF SYNTACTIC THEORY

Since the 1960s syntactic theory has undergone a number of changes, and has spawned a variety of different grammatical theories, each with a different title, such as Relational Grammar (RG), Head-Driven Phrase Structure Grammar (HPSG), Lexical-Functional Grammar (LFG), Categorial Grammar (CG), Government-Binding Theory (GB), Tree Adjoining Grammar (TAG), etc. While it is tempting to view these as monolithic alternatives, to do so would be misleading.

First, all approaches provide only fragments of a full theory of grammatical knowledge; sometimes these fragments only partially overlap between approaches. Second, there are many fundamental points of agreement between the different approaches. Third, the differences among practitioners of the same general framework can be as large as or even larger than the differences between frameworks. The differences of opinion that engender different 'named' grammatical theories draw greater attention, but they have no special status. Therefore, rather than reviewing different named grammatical theories, this section focuses on a selection of fundamental issues on which syntactic theories diverge.

Syntactic Atoms and How They Combine

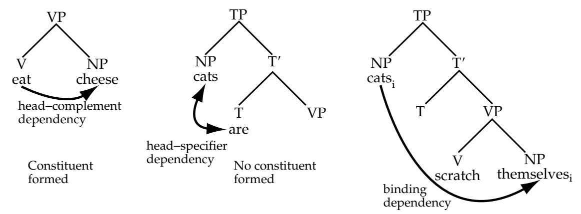
First, syntactic theories differ on the issue of what are the 'atoms' of syntax, that is, the pieces of sentences that are stored in a speaker's long-term memory. At one extreme are certain versions of Transformational Grammar (including the recent *Minimalist Program*), which claim that the atoms of syntax are smaller than words – either morphemes or individual syntactic features. Under this approach, underlying syntactic structures are formed by selecting a set of these atomic units, and combining them based on highly general principles of structure-building. Under this approach, syntax is responsible even for the formation of word-sized units. For example, an inflected verb such as 'runs' may be formed by independently selecting the verbal head 'run' and the inflectional head *[3rd person singular, present]*, and applying a transformation which combines them to form a complex syntactic head, which is spelled out as the word 'runs'.

At the other end are approaches that assume much larger atoms, in the form of templates for phrases or even clauses. Construction Grammar and some versions of Tree Adjoining Grammar are examples of such approaches. Under these approaches, the representation of idiomatic expressions is little different from the representation of other types of phrases. Construction Grammar has provided some insightful analyses of constructions that have been largely overlooked in mainstream transformational syntax. (*See* **Construction Grammar**)

Despite disagreements about the size of the atoms of syntax, there has been a quiet convergence of opinion on the role of the atoms of syntax. In early generative theories it was standard to distinguish the terminal elements of syntax (i.e. lexical items) from the phrase structure rules that determine how the terminals combine. In most current theories this distinction has been eliminated, and the work once done by phrase structure rules is replaced by a set of highly general conditions on how syntactic atoms combine. In these *lexicalized* grammars, information about the combinatorial possibilities of syntactic atoms is built into the lexical entries of the atoms themselves. Lexicalism is a common feature both of theories that assume very small syntactic atoms and of theories that assume much larger syntactic atoms.

Types of Structural Dependencies

A second issue involves the question of how syntactic elements enter into structural dependencies. As a starting point, in a typical phrase structure grammar syntactic elements enter into two basic types of dependencies, illustrated in (35). First, when syntactic elements enter into a *sisterhood* relation, this both forms a dependency between the two elements and creates a new syntactic *constituent*. Constituents may participate in a variety of different syntactic processes, such as coordination, movement, and ellipsis. On the other hand, many syntactic dependencies do not involve the formation of new constituents, for example subject–verb agreement and reflexive binding.



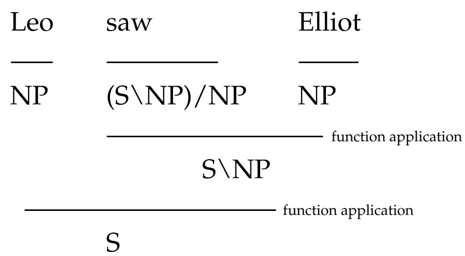
(35)

The phrase structure notions that create this division among structural dependencies continue to dominate thinking about syntax, but there are a number of interesting alternative proposals that reduce or eliminate this distinction. First, *Dependency Grammars* treat all syntactic dependencies in a parallel fashion, and do not single out constituent-forming dependencies as special (36).

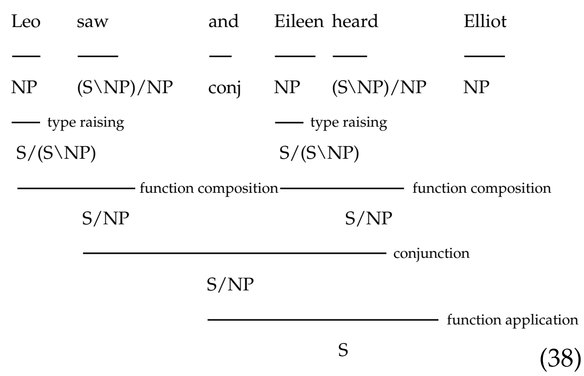
Dependency Grammar representation of argument and agreement relations (36)

Combinatory Categorial Grammars (CCG) also reduce the division of dependency types, but in the opposite manner from Dependency Grammars. In CCG grammars, information about the elements that a syntactic atom may combine with is encoded in enriched category labels. For example, an intransitive verb such as 'run' might have the category label S\NP, read as 'a category that combines with an NP to its left to form an S'. Dependencies between syntactic sisters are formed by the rule of *function application*. Dependencies between non-sisters are also formed by function application, thanks to the mediating effects of *function composition* rules, which allow the combinatorial requirements of a syntactic atom to be passed up through a series of larger units.

By virtue of their more uniform treatment of structural dependencies, both Dependency Grammar and Combinatorial Categorial Grammars have led to innovative proposals about the treatment of constituency phenomena. Whereas most phrase structure grammars impose a clear boundary on which syntactic relations form constituents, DG and CCG do not. Therefore, these approaches have been used to analyze syntactic phenomena that do not fall straightforwardly under standard notions of constituency, such as 'nonconstituent coordination' involving subject–verb sequences. In CCG, flexible constituency relations are made possible by the introduction of *type raising* rules, which allow the combinatorial requirements of a category to be satisfied in a different order (38).



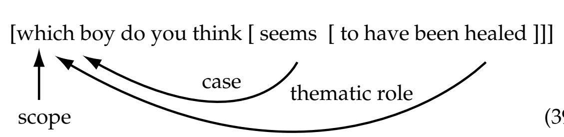
(37)



(38)

Multiple Roles: Alternatives to Transformations

All syntactic theories must address the fact that individual syntactic elements can enter into a variety of different structural relations. In a multi-clause sentence, a single phrase may take scope in one clause, be case-marked by the predicate of a second clause, and receive a thematic role in a third clause (39).



(39)

Since the 1950s and 1960s, transformational approaches to syntax have famously argued that phrases can bear multiple syntactic roles because there are multiple syntactic levels of representation, which are related to one another via movement through a series of different structural positions. As a result, a leading area of research on transformational syntax has been the investigation of constraints on possible movement operations, as outlined above.

The syntactic frameworks *Lexical-Functional Grammar* (LFG) and certain versions of *Combinatory Categorial Grammar* (CCG) share with Transformational Grammar the assumption that words and phrases bear multiple roles because they appear in multiple different levels of representation. These approaches diverge from Transformational Grammar in the respect that they do not assume that each different level is a hierarchical constituent structure, or that the levels are related by movement transformations. LFG assumes independent levels of *a-structure* (argument structure: representation of argument/thematic roles), *f-structure* (function structure: representation of subject, object, etc. roles), and *c-structure* (constituent structure: surface syntax). There are rules for mapping between these levels, but the mappings are not assumed to be transformational. In some versions of CCG separate representations of *argument structure* and *surface structure* are posited. (*See* **Lexical-Functional Grammar**)

In *Head-Driven Phrase Structure Grammar* (HPSG) only one syntactic level of representation is assumed. This single level of representation combines words and phrases into the surface constituent structure that is familiar from many other syntactic theories. However, the terminal elements of these structures contain highly articulated feature structures, which encode a great deal of information about argument structure, phonology, 'moved' arguments, and so on. Whereas transformational grammars use movement operations to handle the multiple roles problem, the same work is done largely internal to individual syntactic heads in HPSG. For example, a verbal head may encode the information that a *wh*-phrase, which is represented as the *focus* argument of the clause, is to be treated as the filler of one of the slots in the verb's argument structure list. Constraints on movement operations in transformational approaches must be replaced in nontransformational theories by related constraints on the relations between scope and argument slots. (*See* **Phrase Structure Grammar, Head-driven**)

Causes of Ungrammaticality

Standard approaches to syntactic theory assume a set of syntactic atoms and a relatively small number of formal principles or constraints that determine how these atoms may be combined to form sentences. A sentence is assumed to be grammatical if it violates no constraints. An ungrammatical sentence is a sentence that violates one or more constraints. Some variants of syntactic theory have explored broader notions of what causes a sentence to be (un)grammatical.

Functional grammars typically emphasize the role of meaning or of communicative efficiency in determining the well-formedness of a sentence. Such approaches typically do not appeal directly to semantics or processing efficiency to explain ungrammaticality. Rather, semantics or processing efficiency are used to provide a functional motivation for a set of formal grammatical constraints.

Both *Optimality Theory* and certain versions of the *Minimalist Program* question the standard assumption that a grammatical sentence is a sentence that violates no constraints. This characterization is replaced in these approaches with the requirement that a well-formed sentence is the *optimal* candidate from a set of possible structures/derivations for that sentence. In other words, a sentence may be deemed ungrammatical for the simple reason that there exists a better way of expressing the same thing. (*See* **Optimality Theory**)

CHALLENGES AND FUTURE PROSPECTS

Universal Grammar

The Principles and Parameters approach to syntax, which is compatible with any of the syntactic frameworks discussed above, aims to explain how a child can attain rich knowledge of any language. It does so by seeking universal syntactic principles and clusters of syntactic properties that covary across languages. Syntactic research since the 1970s has uncovered a wealth of cross-linguistic findings, and a number of good candidates for universals of syntax have been found. However, the search for parameters has met with mixed success. The prospect that each parametric cluster may be linked to an easily observable surface property of the language appears to be viable, but parametric clusters of properties appear to be both narrower and more numerous than originally expected. It remains to be seen whether all of natural language syntax, including the idiosyncrasies of each language, can be handled in terms of a Principles and Parameters approach.

The Unification Problem

Syntactic theories are theories that aim to characterize the mental representations underlying knowledge of language, and of how people acquire that knowledge. However, most syntactic theories characterize knowledge of which sentences are grammatical and which sentences are ungrammatical, with few suggestions about how speakers successfully access this knowledge in real-time speaking or understanding, or about how children acquire this knowledge. Even less is known about how to encode this knowledge in brain structures. The overall goals of syntactic theory may be significantly affected by findings in these areas. In addition, a complete syntactic theory will have to provide answers to questions about how sentence structures are learned, used in real time, and encoded in the brain.

Further Reading

Baker M (2001) *The Atoms of Language*. New York: Basic Books.

Baltin M and Collins C (2000) *Handbook of Contemporary Syntactic Theory*. Malden, MA: Blackwell.

Bresnan J (2000) *Lexical Functional Grammar*. Malden, MA: Blackwell.

Chomsky N (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.

Crain S and Lillo-Martin D (2000) *An Introducion to Linguistic Theory and Language Acquisition*. Malden, MA: Blackwell.

Culicover P (1997) *Principles and Parameters*. New York: Oxford University Press.

Radford A (1988) *Transformational Grammar*. Cambridge, UK: Cambridge University Press.

Roberts I (1997) *Comparative Syntax*. London: Edward Arnold.

Sag I and Wasow T (2000) *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications.

Steedman M (2000) *The Syntactic Process*. Cambridge, MA: MIT Press.

Syntax, Acquisition of

Introductory article

Stephen Crain, University of Maryland, College Park, Maryland, USA
Rosalind Thornton, University of Maryland, College Park, Maryland, USA

CONTENTS

Research shows that young children adhere to core grammatical principles early in the course of language development. Experimental studies of child language confirm the theory of universal grammar, and support the continuity assumption for language development.

INTRODUCTION

From the vantage point of linguistic theory, all normal children are expected to have full command of a rich and intricate system of linguistic principles as soon as these principles can be assessed, roughly around age three. Experimental investigations of child language, however, have often led to a different picture of language development. Several studies are interpreted as showing that language learning takes several years, and that children make numerous missteps along the way. The article begins by reviewing the reasons, based on current linguistic theory, for anticipating the rapid growth of linguistic knowledge. Then it turns to the laboratory, to consider a sample of findings that do not sit well with the expectations of linguistic theory, as well as some findings that do comport with theory.

PRINCIPLES AND PARAMETERS

Despite the complexity of human languages, children rapidly converge on a grammatical system