Overview

Government–binding/principles and parameters theory

Howard Lasnik* and Terje Lohndal

Principles and Parameters Theory is an approach to the study of the human language capacity based on an abstract underlying representation and operations called 'transformations' successively altering that structure. It has gradually evolved from the Government and Binding Theory to the Minimalist Program. © 2009 John Wiley & Sons, Ltd. *WIREs Cogn Sci* 2010 1 40–50

The Principles and Parameters Theory is an approach to the study of the human knowledge of language that developed out of Noam Chomsky's work in the 1970s. Like all of Chomsky's earlier work, it centered on two fundamental questions:

> What is the correct characterization of 'the linguistic capacity' in someone who speaks a language? What kind of capacity is 'knowledge of language'? (1)
>
> How does this capacity arise in the individual? What aspects of it are acquired by exposure to relevant information ('learned'), and what aspects are present in advance of any experience ('wired in')? (2)

Principles and Parameters Theory comes in two incarnations: as Government and Binding Theory (1980s) and as the Minimalist Program (late 1980s until today). In this review, we show how Principles and Parameters Theory developed from its predecessors. We focus, in particular, on characterizing the central traits of Government and Binding Theory and Minimalism, and on how the former developed into the latter. Our view is that Minimalism builds upon and rationalizes the successes of the Government and Binding framework. Not only does Minimalism reach for explaining the properties of the language faculty, it also tries to explain why the specific properties are the way they are. We also focus on the framework underlying the Principles and Parameters Theory, in particular the distinction between universal principles and language-specific parameters. We discuss two major views on how Minimalism should conceptualize parameters and suggest that, in recent years, one of the views has become more prominent than the other. This is in part due to the current focus on reducing Universal Grammar (UG) to its barest essentials.

*Correspondence to: lasnik@umd.edu

Department of Linguistics, University of Maryland, College Park, MD, 20742-7505 USA

DOI: 10.1002/wcs.35

PRINCIPLES AND PARAMETERS THEORY AND ITS ORIGINS

Background and Origins

Chomsky's earliest work, in the 1950s, particularly concentrated on question (1) above, since explicit and comprehensive answers to that question had never been provided before, largely because the question by and large had gone unasked. Chomsky's answer posited a computational system in the human mind that provides statements of the basic phrase structure patterns of languages (phrase structure rules) and more complex operations for manipulating these basic phrase structures (transformations). This framework and its direct descendants fall under the general title Transformational Generative Grammar ('generative' meaning explicit, in the sense of mathematics).

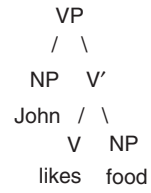
At this point it is useful to introduce a few notions that are important in generative grammar (since Ref 1), and perhaps have become even more so in recent years. The first notion is *descriptive adequacy*. Chomsky argues that a grammar is descriptively adequate to the extent that it correctly describes the intrinsic competence of the idealized native speaker. Correspondingly, a linguistic theory is descriptively adequate if it makes a descriptively adequate grammar available for each natural language. A child, then, has to possess such a linguistic theory, and also has to possess a strategy for selecting a grammar that is compatible with the primary linguistic data. To the extent that a theory succeeds in selecting

such a descriptively adequate grammar, we say that the theory meets the condition of *explanatory adequacy*. The Minimalist Program takes this even further by trying to go *beyond* explanatory adequacy.[2] Not only do we want the theory to succeed in the selection of a descriptively adequate grammar, we also want to know why the theory has the properties it does. We will return to this below.

In the 1960s, the research began to shift more toward question (2). As we have seen, Chomsky coined the term 'explanatory adequacy' for theories that provide a putative answer to that question. A theory of language, regarded as one component of a theory of the human mind, must provide grammars for all possible human languages—since any child can acquire any language under appropriate exposure. To attain a high degree of explanatory adequacy, the theory must in addition show how the learner selects the correct grammar from among all the available ones, based on restricted data, called Primary Linguistic Data. The theories of the 1950s and early 1960s made an infinite number of grammars available, so the explanatory problem was severe.

Through the late 1960s and 1970s, to enhance explanatory adequacy, theorists proposed more and more constraints on the notion 'possible human grammar'. For example, Chomsky's 'standard theory' of the mid- to late-1960s proposed to limit the varieties of transformations.[1,3–5] Chomsky's earliest syntactic theory postulated phrase structure rules (e.g., rules of the type S → NP VP) that create simple structures like e.g., *Mary will solve the problem*, 'singularly transformations' that alter these structures created by phrase structure rules, yielding e.g., *Will Mary solve the problem?*, and lastly 'generalized transformations' that combine separate simple structures into more complex ones, e.g., *John said that Mary will come* from *John said it* and *Mary will come*. Following the observation of Fillmore[5] that the interactions between generalized and singulary transformations were much more limited than predicted, Chomsky[1] proposed eliminating the former in favor of recursion in the base. Recursion is a central trait of human language, and in recent years it has been argued to be the property that separates human languages from other types of animal communication systems.[6] We say that human language is characterized by discrete infinity because even though each sentence itself is finite, there is no limit on the possible number of words in a sentence (e.g., *John said that Mary told him that Peter claimed that*…). Previously, generalized transformations captured this property. Recursion in the base, on the other hand, means that recursion became a property of the phrase structure rules themselves. In addition to rules such as S → NP VP and VP → V, we now also have rules such as VP → V S, thus recursion. This was argued to yield a much more restricted theory, though still not restrictive enough.

The next major move in the direction of explanatory adequacy came in the late 1960s in the form of the 'X-bar theory' of phrase structure, which proposed limitations on phrase structure rules. The basic property is that X-bar theory ensured that phrases are endocentric, i.e., based on a head. In addition to the head of the phrase, the phrase has a complement and a specifier. Derivation (3) shows a typical example of a phrase that conforms to X-bar theory where the noun phrase *John* is a specifier, *likes* is a head and *food* is the complement. For the VP, we see that specifier and verb coincide with the terms subject and object.



(3)

X-bar theory also proposed further limitations on transformations so that they no longer were responsible for derivational morphology of the *destroy-destruction* type (Ref 4, but see also Ref 7). Such morphological relations are quite idiosyncratic, and so better captured in the lexicon, which is, after all, the repository of all the peccadillos characteristic of particular lexical items in various languages. This of course does not mean that the lexicon is a place where 'anything goes'. There are also strict constraints on the structure of human lexicons and on what can constitute a lexical entry (e.g., *pzlip* is not a possible English word, whereas *plip* is).

A human language is a systematic way of relating sound (or gesture, more generally, as in signed languages) to meaning, with syntax mediating between the two. At the point in the development of the theory just summarized (mid-1960s; see Ref 8), the model can be graphically represented as follows, with deep structure, the initial phrase structure representation created in conformity with the requirements of X-bar theory, connected to meaning, and surface structure, the final result of the whole syntactic derivation, connected to sound:

D(eep)-structure – Semantic interpretation
|
Transformations (4)
|
S(urface)-structure – Phonetic interpretation

We can illustrate this architecture by the following example. Take the active sentence *John killed the rabbit* and the corresponding passive sentence *The rabbit was killed by John*. In order to create the passive sentence, the D-structure has to be the same as in the active sentence as it has to convey the input to the semantics such that John is the killer and the rabbit is the killed. Transformations then took care of the fact that *the rabbit* is at the front of the sentence in the passive sentence but not in the active one. This resulting structure after transformations was called S-structure.

While this was the basic architecture, it was known from the earliest work in generative grammar that some aspects of meaning depend on surface structure. In particular, while grammatical relations (subject of, object of, etc.) are most directly related to D-structure, virtually all other aspects of meaning (including scope of quantifiers, anaphora, focus) relate to S-structure. For example, in his earlier work Chomsky[1] already had pointed out that transformations often alter scope possibilities, while leaving understood grammatical relations intact, as in (5) versus (6).

> Everyone in the room knows three languages. (5)
>
> i.e., the sentence can mean that either every person in the room knows three (possibly different) languages or (perhaps slightly less available) that the same three languages are known to ever person in the room.
>
> Three languages are known by everyone in the room. (6)
>
> i.e., the sentence can only mean that the same three languages are known to ever person in the room; the other meaning is not available (or at least not easily so).

This led to a revised model (the 'extended standard theory') in which both D- and S-structures are inputs to semantic interpretation.[9,10]

This model, with some modifications, developed through the 1970s, with more and more restrictions proposed on the phrase structure and transformational options assumed to be available to the child learning a language. These moves were explicitly motivated by considerations of explanatory adequacy, though general considerations of simplicity also played a role, as in all science.

One small simplification in the model from the 1960s was the result of a technical revision concerning how movement transformations operate.[11,12] Trace theory proposed that when an item moves, it leaves behind a 'trace', a silent placeholder marking the position from which movement took place. Under trace theory, the importance of D-structure for semantic interpretation is further reduced, and ultimately eliminated. Once S-structure is enriched with traces, even underlying grammatical relations can be determined at that derived level of representation. Above we looked at an example involving passive, and this can serve well to illustrate this as well. The S-structure of *The rabbit was killed by John* will now also have traces, so it will roughly look something like *The rabbit*$_1$ *was killed* t_1 *by John*, where t_1 indicates the position from which *the rabbit* has moved. Co-indexing shows that *t* and t*he rabbit* are the same entities.

Using the term LF ('Logical Form') for the syntactic representation that relates most directly to the interpretation of meaning and PF ('Phonetic Form') for the one relating most directly to how sentences sound, we have the so-called T-model in (7), which was at the core of Government and Binding (GB) theory.

D-structure
|
Transformations
| (7)
S-structure
/ \
PF LF

The idea is that when the grammar engine has constructed S-structure, this structure needs to get both a phonological/phonetic and a semantic representation. PF is the interface to the articulatory-perceptual systems and LF the interface to the conceptual-intentional system. Since the purpose of syntax is to yield a sound-meaning pair, these two interfaces were assumed to be required in any syntactic theory.

Before we start looking at two recent incarnations of Chomskyan generative grammar, it may be useful to briefly review some overall foundational issues within this approach to the study of the human linguistic capacity. Chomsky has always emphasized the importance of studying language from an internalist perspective. That is, we assume that there is

something like a faculty of language, and our job as linguists is to figure out what the structure of this faculty is. One speaks of a 'language organ', in the sense that the faculty is innate and part of our biology. There is also an obvious way in which Chomsky's approach always has been very 'psychological'. The object we are describing is part of our psychology as it involves how the brain is able to generate language. In the mid-1980s, Chomsky coined the term I-language for this approach. The I stands for intensional, individual, and internal. This stands in contrast to E-language, where E stands for external and extensional. This involves studying the output of production, more specifically language use in various ways. Chomsky himself has never denied the relevance of E-language, but at the present state of inquiry, it is methodologically appropriate to study I-language since all inquiries into E-language ultimately presuppose the existence of I-language, for the simple reason that things that are produced have to be produced somewhere.

Principles and Parameters

The postulated universal ('wired-in') parts of UG are called principles. The (limited) ways in which languages can differ syntactically are called parameters. This model was a sharp break from earlier approaches, under which universal grammar specified an infinite array of possible grammars, and explanatory adequacy required an unfeasible search procedure to find the highest-valued one, given primary linguistic data. The P&P approach eliminated all this. There is no enumeration of the array of possible grammars. There are only finitely many targets for acquisition, and no search procedure apart from valuing parameters. This cuts through an impasse: descriptive adequacy requires rich and varied grammars, hence unfeasible search; explanatory adequacy requires feasible search.

The principles constrain the workings of the computational system underlying the language faculty. These principles are not subject to variation, but were assumed to be identical across all languages. They constrained grammatical operations and ensured for example that the argument structure of verbs was correctly represented (filtering out unacceptable expressions like *John made*, while allowing acceptable ones such as *John made a cake*). This specific principle is called the Theta Criterion.

As for parameters, this notion is used in different ways in the generative literature. One influential idea is that of overspecification,[13–15] namely that there are more parameter values in UG than any human languages have. Put differently, this basically says that UG has a finite number of options which together yield the typologically attested languages. The child then only has to set the correct value (mostly thought to be a choice between two options—like a switchbox as Jim Higginbotham aptly put it) based on the primary linguistic data. One example of such a parameter is the head parameter, which is responsible for a significant word order difference among languages. In head-initial languages, heads invariably precede their complements. For example, in English verbs precede their direct objects, and English has prepositions rather than postpositions. English is head-initial. Languages such as Japanese are head-final and the object precedes the verb and the language contains postpositions.

Another view, originating with Borer[16] and adopted by Chomsky,[17] is that all parameters are lexical, i.e., the variation reduces to differences among grammatical elements (like inflection) between the world's languages. This has in particular been supported by the extensive work on Romance dialects and other closely related languages (see e.g., Refs 18–20), which showed that there are huge variations.

In recent years, the implicit assumption of the two first views, namely that there is a close relationship between language typology and UG, has been questioned.[21,22] Instead it has been argued that parameters should not be conceived of as innate, but rather they are acquired through experience. Exactly how to conceive of parameters will likely be an issue at the forefront of research for some years to come.

Outline

Here we will outline what we take to be core aspects of GB and the Minimalist Program. The reader will notice that the focus is different for each of them and that we do not necessarily discuss the same phenomena. This does not mean that e.g., Logical Form is not important in Minimalism; it just reflects a fact about where the research focus has been directed so far.

GOVERNMENT AND BINDING THEORY

Modularity

On first examination, human languages appear to be almost overwhelmingly complex systems, and the problems, for the linguist, of successfully analyzing them, and for the learner, of correctly acquiring them, seem virtually intractable. But if the system is broken down into smaller parts, the problem might likewise be decomposed into manageable components. In fact,

under this divide and conquer ('modular') approach, languages began to look much simpler in the 1980s. Apparently complex phenomena were seen as the result of the interaction of simple modules. The phrase structure module was virtually reduced to the X-bar schema, with specific instantiations following from properties of particular lexical items. For example, the verb *solve* must be specified in the lexicon as taking a direct object. Given this specification, a specific phrase structure rule saying that a verb phrase (VP) can consist of a V and a noun phrase (NP) would be completely redundant. Further, the X-bar schema itself was extended from just 'lexical' categories (noun, verb, adjective, etc.) to grammatical categories like tense and inflection. It was an irony of the original formulation of X-bar theory that it excluded the most fundamental unit of syntactic analysis—the sentence. All other phrasal units were analyzed as projections of a head, but the sentence was *sui generis*. GB theorizing brought sentence into the fold, by analyzing it as the projection of an inflectional head, Infl, containing tense and agreement information.

GB also simplified the transformational module. In the 1950s and 1960s, the transformational component of the grammar of a particular language was thought to be a long (partially) ordered list of very detailed transformations, some marked optional and others marked obligatory, specific to the language in question. In such a framework, explanatory adequacy is a very distant goal. The GB framework replaced these transformations with very general optional operations, Move α (displace any item anywhere), or even affect α (do anything to anything, cf. Ref 23). There is thus very little transformational syntax that the child has to learn. A grammar this simple and general would seem to massively overgenerate, producing countless numbers of unacceptable sentences. To deal with this overgeneration problem, GB theorists, further developing a line of research begun in the 1960s, posited general constraints on the operation of transformations (locality constraints in particular), and also conditions on the output of the transformational component, 'filters'.

Theta-Theory and the Lexicon

The X-bar schema for phrase structure is one module and the lexicon is another. These modules determine D-structure configurations via the regulation of a third module 'θ-theory'. In a sentence with the verb *solve*, there is a semantic function for a direct object to fulfill, while there is no such function in the case of *sleep*. These semantic functions that arguments (direct objects, subjects, indirect objects, etc.) fulfill are called 'thematic (θ) roles'. Typical examples of thematic roles are agents (*John* in *John killed the cat*), themes (*a cake* in *Mary bought a cake*), and experiences (*Bill* in *Bill heard a shot*). The verb *solve* demands a direct object since the object would fulfill a necessary θ-role determined by the meaning of the verb. Conversely, an intransitive verb like *sleep* does not take a direct object since there would be no θ-role for it to fulfill. These paired requirements on assigners and recipients of theta roles are called the 'θ-Criterion' in the literature.

Case Theory

There are characteristic structural positions that 'license' particular cases, as follows:

Position	Case	Example	
Subject of finite sentence	Nominative	*He* left	
Direct object of transitive verb	Accusative	I saw *him*	(8)
'Subject' of NP	Genitive	*John's* belief	
Object of preposition	Oblique	near *him*	

In many languages (such as Latin, Russian, German), these case distinctions are invariably overtly manifested. In English, only pronouns show an overt distinction between nominative and accusative, but Case Theory posits that all NPs have abstract case (henceforth, Case), even when it is not phonologically visible. In an example like (9), the subject *John* bears nominative Case and the object *Mary* bears accusative Case. If we instead use pronouns, we can see the case distinction overtly, as in (10).

John loves Mary. (9)

He/*him loves her/*she. (10)

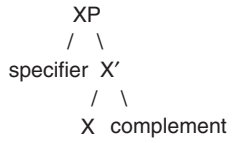
This shows us that there are certain positions that are appropriate for certain Cases. The requirement that all NPs occur in appropriate Case positions is the 'Case Filter', a well-formedness condition on the S-structure level of representation. This Filter rules out an example like **Him loves*.

Government

The notion 'government' is a generalization of the X-bar theoretic head-complement relation. The basic definition is as follows:

> A head H governs Y if and only if every maximal projection dominating H also dominates Y and conversely. (11)

By (11), a head governs its complement and also its specifier. We saw an example of these notions in (3) above, and a more general structure is given in (12).



(12)

Here, X is H and both the specifier and the complement may be Y. A maximal projection is a phrase (like VP, NP); XP in (12).

Case licensing is one property that happens under government, with the governor licensing the governee. A transitive verb governs its direct object NP; a preposition governs its complement NP; Infl governs its specifier (the surface subject of the clause); and N governs its specifier, the 'subject' of the nominal expression. Thus, a Case-licensing head (transitive verb; preposition; finite Infl; N) licenses Case on a nominal expression that it governs.

Types of Movement

The transformational module of the theory recognizes three major subtypes of movement. 'A-movement' is movement to an argument-type position (especially subject position), as exemplified in the passive case discussed above where the understood object surfaces in subject position. Movement here is to the canonical subject position, usually called the specifier of IP or TP, and located (right) above VP in the tree.

'A-bar movement' is movement of an XP to a non-argument position. The movement of an interrogative expression as in (13) (WH-movement) is a central exemplar:

Who will they hire *t*? (13)

In this case, the WH-expression moves to the left edge of the clause.

Lastly, head movement is a case where a head moves. A typical example is that sentence-pairs like those in (14) and (15) are related via movement of the verb *is*, which is the head of the VP.

Mary is a teacher. (14)

Is Mary a teacher? (15)

A major topic at the end of the GB period was the parametric differences among languages with respect to the kind of head movement just illustrated. While English only allows 'auxiliary' verbs (*be*, *have*, modals), other languages allow 'main' verbs to raise as well. In the latter languages, (16) is grammatical.

*Saw Mary a teacher? (16)

This difference shows up in other sentences as well, e.g., in negative sentences.

All three types of movement are regarded as instantiations of one general operation: Move α. The differences follow from independent properties of the items moved and the positions moved to.

Binding

The 'Binding' part of GB theory has as its core anaphoric relations, circumstances under which one expression can or cannot take another as its antecedent, that is, pick up its reference from another. Among the imaginable anaphoric relations among NPs, some are possible, some are necessary, and still others are proscribed, depending on the nature of the NPs involved and the syntactic configurations in which they occur. For example, in (17), *him* can take *John* as its antecedent, while in (18), it cannot.

John said Mary criticized him. (17)

John criticized him. (18)

That is, (18) has no reading corresponding to that of (19), with the pronoun *him* replaced by the 'anaphor' *himself*.

John criticized himself. (19)

A pronoun cannot have an antecedent that is 'too close' to it. This is Condition B of the binding theory. Conversely, an anaphor requires an antecedent quite close to it (Condition A). Compare (19) with (20).

*John said Mary criticized himself. (20)

The pertinent locality is, roughly, being in the same clause (though in certain instances a more complicated notion involving government is implicated).

A third condition (Condition C) excludes an anaphoric connection between *She* and *Mary* in (21), as contrasted with (22).

*She thinks Mary will solve the problem [with *She* intended to refer to Mary] (21)

Mary thinks she will solve the problem. (22)

A structural licensing condition that is relevant here is 'c-command'. Basically, Condition C says that a referential expression cannot be c-commanded by a pronoun that bears the same intended reference. C-command is commonly defined in terms of relative tree height, so that if x is higher than y in the tree, x will typically c-command y. C-command is relevant for many grammatical phenomena, as any introductory syntax textbook will show.

Logical Form

In the core GB model schematized above in (4), LF is not necessarily distinct from S-structure. However, more and more arguments were put forward that transformational operations of the sort successively modifying D-structure, ultimately creating S-structure, also apply to S-structure, creating a distinct LF. One such operation is the analog of overt WH-movement. In sentences with multiple interrogatives, such as (23), at the level of LF all have been argued to be in sentence initial operator position, as illustrated in (24).

Where should we put what? (23)

what_1 [where_2 (we should put t_1 t_2)] (24)

One of the most powerful arguments for covert WH-movement involves constraints on movement. For example, it is difficult to move an interrogative expression out of an embedded question[24] (a question inside another sentence):

*Why_1 do you wonder [what_2 (John bought t_2 t_1)] (25)

If (25) were acceptable, it would mean 'What is the reason such that you wonder what John bought for that reason'. In languages where WH-phrases are *in situ* (unmoved) at S-structure, such as Chinese, their interpretation apparently obeys the same constraints.[25] So in Chinese an example like (26) is possible but one like (27) is impossible on the relevant reading (the one where *weisheme* is understood as having scope over the entire sentence).

ni renwei [ta weishenme bu lai]

you think he why not come (26)

'Why do you think he didn't come?'

(*) ni xiang-zhidao [Lisi weisheme mai-le shenme]

you wonder Lisi why bought what (27)

'*What is the reason such that you wonder what Lisi bought, where the purchase was for that reason?'

This argues that even though the 'why' is not phonetically displaced, it really is moving. But this movement is 'covert', occurring in the mapping from S-structure to LF, hence not contributing to pronunciation, which is exactly what the architecture in (4) derives.

MINIMALISM

The Heritage from GB

The diminishing role of D- and S-structures in the theory suggests that neither is actually a significant level of representation. If a language is to relate sound to meaning at all, it evidently requires the 'interface' levels of LF and PF, the former interfacing with the conceptual-intentional system of the mind, and the latter with the articulatory-perceptual system. Neither D-structure nor S-structure is conceptually necessary in this way. This motivates a shift to a model that is reminiscent of Chomsky's original one in the 1950s, with structure building being done by generalized transformations. The derivation begins with a 'numeration', a selection of elements copied from the lexicon. The lexical items are inserted 'on-line' in the course of the syntactic derivation. The derivation proceeds 'bottom-up' with the most deeply embedded structural unit created first, then combined with another lexical item to create a larger phrasal unit, and so on. We will elaborate on this process below.

Minimalism advances the hypothesis that language is a 'perfect' solution for meeting the requirements imposed by the external systems.[17] It seeks principled explanations instead of purely technical accounts. Recently, Chomsky[2] has suggested that we should go beyond explanatory adequacy (recall that explanatory adequacy involves how to account for the fact that the child converges on the right grammar), and try to explain why the computational system of human language has just the properties it does. A significant component in this venture is the search for what Chomsky has called 'third-factors'.[26] Essentially, the goal is to reduce the principles of UG to their 'barest essentials' and seek principles that are

more general in nature, e.g., as part of our general cognition or even of biological systems more generally. Interestingly, it is possible to see a link between Chomsky's recent focus and what he suggested in his seminal first chapter of *Aspects*, namely that many properties of the language faculty may follow from 'principles of neural organization that may be even more deeply grounded in physical law' (Ref 1, p. 59) Current cutting-edge research is in many ways trying to come to grips with this fascinating hypothesis.

Economy and Last Resort

One major minimalist concern involves the driving force for syntactic movement. From its inception in the early 1990s, Minimalism has insisted on the last-resort nature of movement: in line with the leading idea of economy, movement must happen for a reason and, in particular, a formal reason. The Case Filter, which was a central component of the GB system, was thought to provide one such driving force. Notice that if the Case requirement of a nominal phrase provides the driving force for movement, the requirement will not be satisfied immediately upon the introduction of that nominal expression into the structure. Rather, satisfaction must wait until the next cycle, or, in fact, until an unlimited number of cycles later, because raising configurations can iterate, and it is only the ultimate landing site that licenses nominative Case:

Mary seems [*t* to be likely (*t* to win the race)]. (28)

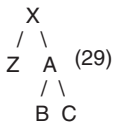
A minimalist perspective favors an alternative in which the driving force for movement can be satisfied immediately rather than indefinitely later in the derivation. In the present instance, suppose the crucial inadequacy lies not in the nominal expression but rather in the item that licenses its Case, e.g., the Tense/Inflection head of the clause. That is, Inflection has a feature, e.g., Number, that must be checked against the NP, which also carries a Number feature. Then, as soon as that head has been introduced into the structure by generalized transformation, it can 'attract' the NP that will check (and consequently delete) its feature. Movement is then seen from the point of view of the target rather than the moving item itself. The Case of the NP apparently does get checked as a result of the movement, but that is simply a beneficial side effect of satisfying the requirement of the attractor. In an elegant metaphor, Uriagereka[27] likens the attractor to a virus. Immediately upon its introduction into the body, it is dealt with (by the production of antibodies in the case of physical viruses, by movement to check the viral feature in the syntactic instance).

The Extension Condition and Merge

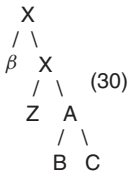
The 'Extension Condition' requires that a transformational operation 'extends' the tree upward. Decades earlier, Chomsky had argued that eliminating generalized transformations yields a simplified theory, with one class of complex operations jettisoned in favor of an expanded role of a component that was independently necessary, the phrase structure rule component. Further, that simplification was a substantial step toward answering the fundamental question of how the child selects the correct grammar from a seemingly bewildering array of choices. Eliminating one large class of transformations, generalized transformations, was a step toward addressing this puzzle. This was a very good argument. But since then, the role of the phrase structure component has virtually vanished. Furthermore, numerous discoveries and analyses have indicated that the transformational component can be dramatically restricted in its descriptive power. In place of the virtually unlimited number of available highly specific transformations of the theories of the 1950s and early 1960s, we can have instead a tiny number of very general operations: Merge (the generalized transformation, expanded in its role so that it creates even simple clausal structures), Move, and maybe a very few others.

Merge combines two things and makes one of them the head of the new structure. For example, if you combine a verb and a nominal phrase, the verb becomes the head of the structure because you then have a verb phrase. In recent years, Chomsky has argued that Merge comes in two flavors: External and Internal Merge.[2] External Merge is when items are first-merged in the syntactic tree, whereas Internal Merge is when a copy of an item is made and remerged elsewhere in the tree (cf. the discussion of passive above). The conceptual advantage of this view is that there is only one basic operation, Merge, and not e.g., two basic operations Merge and Move.

This view of the derivation is also related to 'cyclicity'. In Chomsky's work[1] the requirement that derivations work their way up the tree monotonically was introduced, alongside D-structure. Chomsky used this to explain the absence of certain kinds of derivations. The current name is The Extension Condition. This condition demands that both the movement of material already in the structure (singular transformation) and the merger of a lexical item not yet in the structure (generalized transformation) target the top of the existing tree. Consider in this context the structures in (29)–(31).



(29)



(30)

(31)

where (29) is the original tree, and (30) shows a derivation that obeys the Extension Condition. Here β is merged at the top of the tree. The last derivation, (32), does not obey the Extension Condition because β is merged at the bottom of the tree. Importantly, there is a deep idea behind cyclicity, which again was present in Chomsky's earliest work in the late 1950s. The idea, called the No Tampering Condition in current parlance, seems like a rather natural economy condition. Derivation (30) involves no tampering since the old tree in (29) still exists as a subtree of (30), whereas (31) involves tampering with the original structure.

Linearization

Generative syntax has always been concerned with the hierarchical organization of representations, and the overwhelming majority of syntactically and semantically significant structural relations are hierarchical. Virtually none of these relations involve linear order, although linear order is manifested in phonological representation. Kayne[28] initiated a very influential research line arguing that linear order is actually part of Syntax. This comes through in his system by algorithms that transform hierarchical phrase structure representations into linear order through asymmetric c-command. Chomsky[17] subsequently proposed that linear order is manifested *only* in PF. Interestingly, Chomsky here gives an example of a major goal of Minimalism, namely to reduce all constraints on representation and derivation to 'bare output conditions', determined by the properties of the systems external to the language faculty (but still internal to the mind) that PF and LF must interface with.

Kayne's hypothesis has far-reaching consequences. One is that all structures are of the pattern specifier—head—complement, i.e., SVO languages such as English. These languages are consistent with Kayne's requirement. However, SOV languages such as Japanese are not, since they appear to have the order specifier—complement—head. Kayne's system reanalyzes SOV languages as underlyingly SVO (as all languages must be by this hypothesis) with the SOV order derived by leftward movement. Many phenomena have been productively analyzed in these terms, but one crucial unanswered question at this point is the source of the driving force for all of the required movements.

Interfaces and Multiple Spell-Out

The precise nature of the connection between the syntactic derivation and semantic and phonological interfaces has been a central research question throughout the history of generative grammar. The Minimalist approach to structure building is much more similar to that of the 1950s than to any of the intervening models, suggesting that interpretation in the Minimalist model also could be more like that of in the early model, distributed over many structures. Already in the late 1960s and early 1970s, there were occasional arguments for such a model and for phonological interpretation as well as semantic interpretation. For example, Bresnan[29] argued that the phonological rule responsible for assigning English sentences their intonation contour applies cyclically, following each cycle of transformations, rather than applying at the end of the entire syntactic derivation. There were similar proposals for semantic phenomena involving scope and anaphora put forward by Jackendoff[10] and Lasnik.[30] Chomsky[2,31] argued for a general instantiation of this distributed approach to phonological and semantic interpretation, based on ideas of Epstein et al.[32,33] and Uriagereka,[34] who called the approach 'Multiple Spell-Out'. Simplifying some, at the end of each cycle (or 'phase' as it has been called for the past 10 years) the syntactic structure created thus far is encapsulated and sent off to the interface components for phonological and semantic interpretation. Thus, although there are still what might be called PF and LF components, there are no levels of PF and LF. Epstein argued that such a move represents a conceptual simplification (in the same way that the elimination of D- and S-structures does), and both Uriagereka and Chomsky provided some empirical justification. The role of syntactic derivation, always very important in Chomskian theorizing, becomes even more central on this view because there are no levels of representation at all.

CONCLUSION

Both GB and Minimalism are implementations of the Principles and Parameters approach to the study of human language. Minimalism can best be viewed as a rationalization of the GB model, trying to go beyond explanatory adequacy by focusing on interface constraints as bare output conditions and seeking to rely on third factor conditions as far as possible.

ACKNOWLEDGEMENTS

We are grateful to two anonymous reviewers, associate editor Peter Culicover, and in particular to Noam Chomsky, Norbert Hornstein, and Juan Uriagereka for their helpful comments on a previous draft.

REFERENCES

1. Chomsky N. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press; 1965.
2. Chomsky N. Beyond explanatory adequacy. In: Belletti A, ed. *Structures and Beyond—The Cartography of Syntactic Structures*. Oxford: Oxford University Press; 2004, 104–131.
3. Bresnan J. *Theory of Complementation in English Syntax*, PhD thesis, MIT, 1972.
4. Chomsky N. Remarks on nominalization. In: Roderick J, Rosenbaum P, ed. *Readings in English Transformational Grammar*. Waltham, MA: Ginn; 1970, 184–221.
5. Fillmore CJ. The position of embedding transformations in a grammar. *Word* 1963, 19:208–231.
6. Hauser MD, Chomsky N, Fitch WT. The faculty of language: what is it, who has it, and how did it evolve? *Science* 2002, 298:1569–1579.
7. Postal P. On the surface verb 'Remind'. *Linguistic Inquiry* 1970, 37–120.
8. Katz JJ, Postal P. *An Integrated Theory of Linguistic Descriptions*. Cambridge, MA: The MIT Press.
9. Bach E. *An Introduction to Transformational Grammars*. New York: Holt, Rinehart and Winston; 1964.
10. Jackendoff R. *Some Rules of Semantic Interpretation for English*, PhD thesist, MIT, 1969.
11. Fiengo R. *Semantic Conditions on Surface Structure*, PhD thesis, MIT, 1974.
12. Fiengo R. On trace theory. *Linguistic Inquiry* 1977, 8:35–62.
13. Chomsky N. *Lectures on Government and Binding*. Dordrecht, The Netherlands: Foris; 1981.
14. Baker MC. *Atoms of Language*. New York: Basic Books; 2001.
15. Yang CD. *Knowledge and Learning in Natural Language*. New York: Oxford University Press; 2002.
16. Borer H. *Parametric Syntax*. Dordrecht: Foris; 1984.
17. Chomsky N. *The Minimalist Program*. Cambridge, MA: The MIT Press; 1995.
18. Kayne RS. *Parameters and Universals*. Oxford: Oxford University Press; 2000.
19. Kayne RS. *Movement and Silence*. Oxford: Oxford University Press; 2005.
20. Manzini R, Savoia ML. *L Dialetti Italiani e Romanci. Morfosintassi generativa*. Edizioni dell'Orso, Alessandria, 2005.
21. Newmeyer FJ. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press; 2005.
22. Boeckx C. Approaching parameters from below. In: Di Sciullo A-M, Boeckx C, eds. *Biolinguistics: Language Evolution and Variation*. Oxford: Oxford University Press. In press.
23. Lasnik H, Saito M. *Move α*. Cambridge, MA: The MIT Press; 1992.
24. Ross JR. *Constraints on Variables in Syntax*, PhD thesis, MIT, 1967.
25. Huang CTJ. *Logical Relations in Chinese and the Theory of Grammar*, PhD thesis, MIT, 1982.
26. Chomsky N. Three factors in language design. *Linguistic Inquiry* 2005, 36:1–22.
27. Uriagereka J. *Rhyme and Reason*. Cambridge, MA: The MIT Press; 1998.
28. Kayne RS. *The Antisymmetry of Syntax*. Cambridge, MA: The MIT Press; 1994.
29. Bresnan J. Sentence stress and syntactic transformations. *Language* 1971, 47:257–281.
30. Lasnik H. *Analyses of Negation in English*, PhD thesis, MIT, 1972.
31. Chomsky N. Derivation by phase. In: Michael K, ed. *Ken Hale: A Life in Language*. Cambridge, MA: The MIT Press;.2001, 1–52.
32. Epstein SD, Groat E, Kawashima R, Kitahara H. *A Derviational Approach to Syntactic Relations*. Oxford: Oxford University Press; 1998.
33. Epstein SD. Un-principled syntax: the derivation of syntactic relations. In: Epstein SD, Hornstein N, eds. *Working Minimalism*. Cambridge, MA: The MIT Press; 1999, 317–345.
34. Uriagereka J. Multiple spell-out. In: Epstein SD, Hornstein N, eds. *Working Minimalism*. Cambridge, MA: The MIT Press; 1999, 251–282.

FURTHER READING

Adger D. *Core Syntax*. Oxford: Oxford University Press; 2003.
Boeckx C. *Linguistic Minimalism*. Oxford: Oxford University Press; 2006.

Bošković Ž, Lasnik H (eds). *Minimalist Syntax: The Essential Readings*. Malden, MA: Blackwell Science.
Chomsky N, Lasnik H. The theory of principles and parameters. In: Jacobs J, von Stechow A, Sternefeld W, Venneman T, eds. *Syntax: An International Handbook of Contemporary Research*. New York: 1993, 506–569.
Jackendoff R. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT Press; 1972.
Haegeman L. *Introduction to Government and Binding Theory*. Malden, MA: Blackwell Science; 1994.
Hinzen W. *Mind Design and Minimal Syntax*. Oxford: Oxford University Press; 2006.
Hornstein N, Nunes J, Grohmann KK. *Understanding Minimalism*. Cambridge: Cambridge University Press; 2006.
Lasnik H, Uriagereka J. *A Course in GB Syntax*. Cambridge, MA: The MIT Press; 1988.
Lasnik H, Uriagereka J, Boeckx C. *A Course in Minimalist Syntax*. Malden, MA: Blackwell; 2005.

