Generative Grammar is the label of the most influential research program in linguistics and related fields in the second half of the 20. century. Initiated by a short book, Noam Chomsky's *Syntactic Structures* (1957), it became one of the driving forces among the disciplines jointly called the cognitive sciences. The term *generative grammar* refers to an explicit, formal characterization of the (largely implicit) knowledge determining the formal aspect of all kinds of language behavior. The program had a strong mentalist orientation right from the beginning, documented e.g. in a fundamental critique of Skinner's *Verbal behavior* (1957) by Chomsky (1959), arguing that behaviorist stimulus-response-theories could in no way account for the complexities of ordinary language use. The "Generative Enterprise", as the program was called in 1982, went through a number of stages, each of which was accompanied by discussions of specific problems and consequences within the narrower domain of linguistics as well as the wider range of related fields, such as ontogenetic development, psychology of language use, or biological evolution. Four stages of the Generative Enterprise can be marked off for expository purposes.

1. Transformational Structure and Levels of Description

The foundation of all further developments is to be found in the overwhelming monograph *The Logical Structure of Linguistic Theory* (1975, henceforth LSLT), which was published only 20 years after the manuscript was completed. Starting from the notion of structural analysis as developed by Bloomfield and made precise by Harris (1951), LSLT introduces a radically new perspective on the nature of linguistic structure. On the one hand, LSLT takes up Harris's notion "level of linguistic structure", in terms of which the form of utterances, i.e. their phonetic and syntactic shape, is characterized. On the other hand, however, LSLT develops a notion of linguistic structure that differs fundamentally from the empiricist orientation of American structuralism. More specifically, Chomsky proposes a notion of grammar that provides a full and explicit account of the tacit knowledge a speaker/hearer has of the form of utterances of his or her language. Under this perspective, a grammar G is a theoretical construct, dealing with internal states of knowledge that relate only indirectly to overt behavior of the speaker. In other words, G is viewed as an empirical hypothesis about the mental structure underlying ordinary speech behavior. To make this notion precise, LSLT highlights the importance of intuitive judgments a speaker is able to make about properties of utterances, such as the grammatical wellformedness of the famous nonsense-sentence (1) as opposed to the ungrammaticalness of (2), or the rather different properties displayed by the superficially similar sentences (3) and (4):

(1)   colorless green ideas sleep furiously
(2)   furiously sleep ideas green colorless
(3)   this picture was painted by a real artist
(4)   this picture was painted by a new technique

Hence G of English has to account for the properties on which intuitions like these rely. The study of the general features of G can furthermore be expected to reveal principles of the internal organization of mental states in general and of knowledge of language in particular. This is in fact the basic motivation underlying the research program of Generative Grammar.

More technically, LSLT gives the following account of the concepts needed for this program. First of all, a linguistic level **L** is defined as an algebraic structure of the following sort:

(5) $\mathbf{L} = [\, L, R_1, \ldots, R_m, \mu, \Phi, \varphi_1, \ldots, \varphi_n \,]$, where
             L is the set of primes of **L**;
             R_1 to R_m are sets of relations in **L**, such as linear ordering, inclusion, identity;
             μ is a set of so-called **L**-markers, generated by the primes of **L**;
             Φ is a mapping that maps μ onto the set of grammatical utterances;
             φ_1 to φ_n are operations relating **L** to other levels.

Roughly speaking, an **L**-marker is the representation of the structure of an utterance on the level **L**. The particular levels assumed in LSLT include the phonetic level **Pt**, the phonemic level **Pm**, the morpheme level **M**, the word level **W**, the level of word classes **C**, the phrase structure level **P**, and the transformational level **T**. For **Pm**, the set μ consists of strings of phonemes, for **W**, the elements of μ are sequences of words, and for **P**, the set μ consists of Phrase markers, roughly tree structures indicating the constituency of an utterance. The relation between the different levels defined by the operations φ_i depends on particular conditions of the various levels and is, as Chomsky argues, anything but trivial. The full structural representation of an utterance comprises the **L**-markers assigned to it at the various levels in question. The mapping Φ relating the levels of representation to the actual (or possible) grammatical utterances is correspondingly indirect and complex. The assumption in LSLT is that it may be sufficient to define Φ for phonetic representations, i.e. strings in **Pt**, provided that the relation to **Pt** is defined for all other levels. This accounts for the fact that the structure of an utterance cannot be arrived at by simple operations of segmentation and classification, as assumed in classical structuralism, but depends on highly indirect and complex relations.

Second, a grammar G is a system of operations or rules that generate the sentences of a language with their structural representations. Technically, G determines the set of **L**-markers for each level **L** and the operations by which they are related to the other levels. LSLT deals with the general form of phonological rules, defining the relation between the phonemic and phonetic level, and more extensively with phrase structure rules, generating P-markers, and with grammatical transformations. The role of transformations is a matter of central interest in LSLT (such that for quite a while the term "transformational grammar" was taken as largely equivalent to "generative grammar"), they are assumed to capture structural properties P-markers could not account for. For instance, the difference between (3) and (4), which are not distinct in terms of constituency, is explained by different transformational derivations, informally indicated in (3') and (4'):

(3') a real artist painted this picture ==T$_{Passive}$==> this picture was painted by a real artist
(4') someone painted this picture by a new technique ==T$_{Passive}$==>
      this picture was painted by a new technique by someone ==T$_{Agent-deletion}$==>
      this picture was painted by a new technique

An important point of the theory developed in LSLT is the assumption that grammars must systematically be evaluated. To this effect, a simplicity-metric is introduced by which a grammar is preferred if and only if it expresses more natural generalizations than its competitors. For example, a grammar that systematically accounts for the different properties of sentences like (3) and (4) is to be preferred over one that can only assign to them the same structure. This is the reason why a grammar with a transformational component comes out as simpler than one without it. One of the factors entering the evaluation is the principle of rule ordering, which recognizes the possibility that rules might have a different structural effect if they operate in different ordering. As a simple case in point, in (4') the Agent-deletion can operate only after the Passive Transformation has applied. (The importance of rule ordering was actually observed already in the Sanskrit grammar of Panini.)

The productivity of the proposed framework is demonstrated by an analysis of a large fragment of English with respect to phrase structure and transformations in LSLT and in a stimulating monograph on nominalization by Lees (1960). A seminal study of phonology is given in Halle (1959) for Russian and in Chomsky's unpublished master's thesis on the *Morphophonemics of modern Hebrew.*

A feature that LSLT inherits from the notion of transformational analysis proposed by Harris (1957) is the notion that transformations not only relate individual sentences to their more elementary basis, as shown in (3') and (4'), but are also used to decompose complex

sentences into parts from which they are made up. Thus e.g. (6c) is composed by embedding of (6a) into (6b), replacing the bracketed pronoun:

(6) (a) this picture was painted by a new technique   (b) the director suspected [it]
    (c) ==$T_{Embedding}$==> the director suspected [this picture to be painted by a new technique]

This assumption requires a distinction between simple transformations, like $T_{Passive}$ and so called generalized transformations, like $T_{Embedding}$. Another distinction that must be made is that between optional and obligatory transformations. Passivization, for instance, is optional, while the fronting of relative and wh-pronouns is obligatory, as the ungrammaticalness of (7b) shows:

(7) (a)   I know the man who you met last night
    (b)   *I know the man you met who last night

The clarification of these and a wide range of other complexities concerning the level **T** is one of the driving forces in further developments of Generative Grammar.

Another, more general feature that the Generative Enterprise can be said to inherit from American Structuralism is the assumption that all grammatical structure, from the most concrete level **Pt** of phonetics to the most abstract level **T** of transformational structure, is a matter of linguistic form, which must be embedded in a wider context of a theory of semiotics that accounts for sense, reference, and use of linguistic expressions. Although there have been several proposals to include a level of semantics into the framework of generative grammar, the main emphasis was on the knowledge of linguistic form throughout.

2. The Standard Theory of Generative Grammar

Immediately after the publication of *Syntactic Structures* (actually lecture notes excerpted from LSLT), the theory it proposed attracted interest from various sides, including mathematical linguistics and psychology. An overview of formal properties of grammatical rules, especially of phrase structure grammars, in Chomsky (1963) established what became known as the Chomsky-hierarchy of formal grammars as a branch of sub-recursive function theory. Miller and Chomsky (1963) gave a first outline of the role of Generative Grammar in the study of language comprehension and production. Furthermore, Katz and Fodor (1963) made a first proposal to extend the theory of grammar by a semantic component, using so-called projection rules to compositionally integrate the meaning of lexical items according to their syntactic relations.
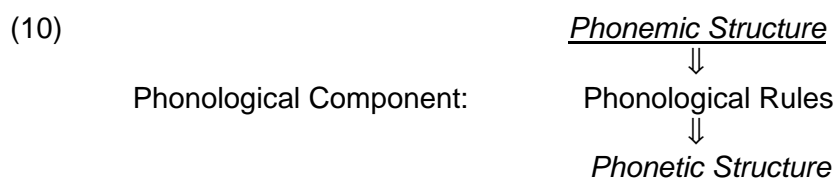
The main development, however, was due to continuous efforts to systematically constrain and strengthen the theory, increasing its explanatory power by reducing the range of possible grammars compatible with a given set of data. Such a reduction dereases at once the task assigned to the evaluation metric. Efforts to this effect led to what was later called the "Standard Theory" of Generative Grammar. Based on preparatory ideas discussed a.o. in Katz and Postal (1964), this is formulated in Chomsky's *Aspects of the Theory of Syntax* (1965). At its core is a re-assessment of the nature of transformational structure. Instead of a separate level **T** of transformations, grammatical transformations are now construed as operations that define the relation between two levels of syntactic representation, both of which are characterized by P-markers. This modification is related to one of the most influential distinctions proposed in the Standard Theory, viz. that between *deep* and *surface structure.* According to this view, each sentence is assigned a deep and a surface P-marker, with Transformations mapping the former onto the latter. Thus except for the different choice of personal or wh-pronouns, (8)(a) – (d) are just different surface realizations of the same deep structure, where the bracketed Agent phrases might furthermore optionally be deleted by $T_{Agent\text{-}deletion}$, known from (4')

(8) (a) John may expect her to meet him
    (b) she may be expected (by John) to meet him
    (c) who does John expect her to meet
    (d) who may be expected (by John) to be met (by her)

On this account, the semantic interpretation of linguistic expressions could apparently be determined on the basis of their deep structure, which would be an interesting step towards an integration of Grammar into an account of meaning and reference. Another important feature of the Standard Theory is a more systematic distinction between (i) the lexical system, i.e. the complex system of rules that introduce lexical items with their (idiosyncratic) phonetic, syntactic, and semantic properties, and (ii) the grammatical rules that integrate the lexical information into the structure of complex linguistic expressions. Finally, the distinction between levels of representation on the one hand and components of grammar generating and interrelating these representations on the other is made more perspicuous than in LSLT. The syntactic levels that appeared to be crucial are deep- and surface-structure, while word- and morpheme structure do not have an independent status. The components of the grammatical rule system are (a) the Base-component, including the lexical system and the phrase structure rules, (b) the transformational component, (c) the phonological component, and, possibly (d) the semantic component. Rules and representations are interrelated according to the schema in (9), with levels of representation underlined:

(9)   Phrase Structure Rules
        Base Component: ⇓
                            Lexical Rules
                                ⇓
                            Deep Structure   ⇒   Semantic Component
                                ⇓
                            Transformations
                                ⇓
                            Surface Structure   ⇒   Phonological Component
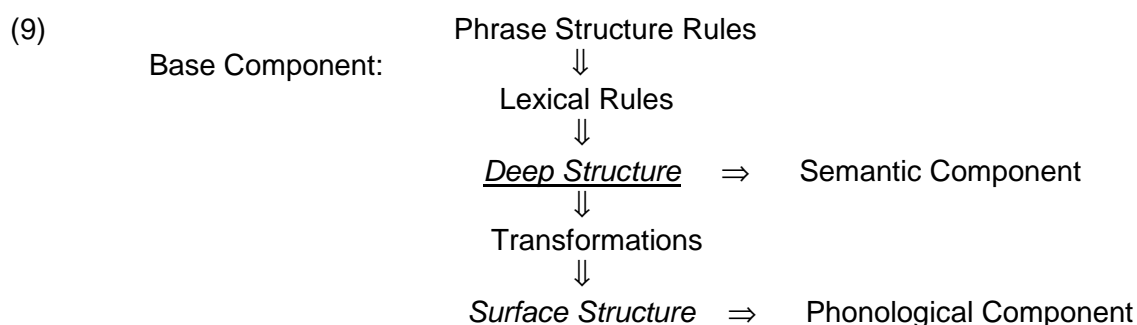
The base component and the transformations together constitute the syntactic component. The semantic component raises problems to be taken up below, while the organization of the phonological component, extensively discussed by Chomsky and Halle (1968), is assumed to have a fairly clear organization, parallel to that of the syntactic component in relevant respects. The underlying phonemic structure is defined by the phonological features of lexical items and their arrangement in surface structure:

(10)   _Phonemic Structure_
                                ⇓
        Phonological Component: Phonological Rules
                                ⇓
                            Phonetic Structure

While there are still optional and obligatory transformations, generalized transformations are dispensed with, since the recursive deep structure of complex sentences is generated already in the base component. A particularly elegant feature that emerges from this reorganization is the principle of cyclic operation of the rules, which is added to the earlier assumption that rules are ordered. Cyclic operation apparently captures a deep property of natural languages: the same sequence of operations applies to increasingly larger parts of an utterance, defining its syntactic properties, its phonetic patterns, and the compositional integration of its semantic structure.

An important general effect of the Standard Theory is the perspicuous account it allows for the traditional notion of linguistic universals, or Universal Grammar UG, considered as the

innate, biologically fixed structure underlying the species-specific language capacity. The range of possible rules and the architecture of their interaction, as schematized in (9) and (10), could now be construed as formal universals, the potential primitive elements entering the rules and representations as substantive universals. Under this perspective, the set of possible grammars G accessible as instantiations of UG determines the diversity of possible natural languages, including the range of variation delimiting processes of linguistic change. By the same token, UG provides the innate, formal basis and predisposition that makes language acquisition possible. These accompanying considerations of the Standard Theory led to a new, principled approach to problems of language acquisition, as well as linguistic change.

3. Towards the Principles and Parameters Theory

Shortly after the Standard Theory was formulated, empirical analyses as well as efforts to improve the structure of the theory led to modifications of the model. Controversial proposals concerned the role of semantics as indicated in (10). Lakoff (1971) and others proposed a model according to which semantic representations are the underlying structure of linguistic utterances, from which surface structures are derived step by step by means of transformations and lexical rules. The insights and difficulties related to these proposals triggered a counterproposal, called the Revised Standard Theory. Based on observations discussed in Jackendoff (1972) it became clear that meaning is not unaffected by surface structure phenomena. (11) and (12) for example are based on the same deep structure, with passivization applying to (12) but not to (11), they do not in general have the same meaning, however:

(11) Three students in the class wrote only one paper
(12) Only one paper was written by three students in the class

(Due to what is called the scope of quantification, in (11) the restriction of having written one paper is ascribed to three students, while in (12) the restriction of being written by three students is ascribed to a single paper.) As phenomena of this sort are by no means marginal, the Revised Standard Theory recognized a more complex relation among syntax and semantics than that indicated in (9): Only certain aspects of meaning are determined by deep structure, while scope and certain other relations depend on surface structure, such that transformations – contrary to the Standard Theory – might now affect the semantic interpretation.

A technical way to handle these problems is the so-called Trace theory proposed in Fiengo (1977). According to this proposal, a constituent that is moved by a transformation leaves behind a trace that is phonetically empty, but may play a role in semantics, as indicated in the following example, where e_i is the trace left in the initial position of the co-indexed pronoun:

(13) who_i do you want to talk to e_i

In many cases, the trace turns out to be similar to what logicians consider a variable bound by the operator contained in the moved constituent. In this sense, invisible elements of the surface structure account for its contribution to semantic interpretation. Other empty elements of the surface structure are "invisible" pronouns, as in (14)(a) and (b), where the actor of the infinitive clause is controlled by (i.e. coreferential with) different constituents of the main clause:

(14)(a) $John_i$ promised Mary [e_i to take the train]
    (b)   John recommended $Mary_i$ [e_i to take the train]

The structural aspect that emerges from transformations, as e.g. in (13), cannot be the Deep Structure of (9), but it also differs from Surface Structure, as it contains elements like traces and empty pronouns. This more abstract surface representation, called S-structure, is assumed to provide the information for phonetic as well as semantic interpretation. On this account, (9) might be replaced by the general schema (15), where LF, viz. the Logical Form, represents all the information G contributes to the meaning of an expression, just as PF, i.e. the Phonetic Form, specifies the conditions G imposes on its pronunciation:

(15)   Deep structure
                                          |
                                   S-structure
                                     /   \
                                  PF   LF

The theory of traces and other "empty" elements is not a purely technical modification, however. Rather it forms part of a general approach to the search for systematic restrictions on the still excessively complex possibilities to construct a grammar according to the Standard Theory. One of the issues motivating this orientation has been called the logical problem of language acquisition. This problem establishes the need to specify the conditions under which an appropriate grammar can be identified on the basis of the restricted and partially defective evidence the learning child is normally exposed to. This is a task which cannot seriously be approached, if UG is as unconstrained as it would be with all the complexities initially allowed for by the transformations of the Standard Theory. On the other hand, it turned out that much of the formal complexities exploited e.g. in operations like the standard passive transformation can be reduced to more elementary operations, once it is recognized that if the application and interaction of these operations is subject to more systematic constraints.

The emergence of Trace theory led to the exploration of a number of systematic conditions which determine the properties and distribution of empty elements of different sorts. These conditions allow transformations to be ultimately reduced to just one simple operation, called *Move* α, where α is any constituent. The systematic and by no means trivial conditions determining the possible choice of α, the domain and bounds of movement, the consequences of movement (such as leaving a trace), are all expressed by general principles needed on independent grounds. One such condition concerns the relations lexical elements impose on their necessary or optional complements. This is called their Argument- or Theta-Structure. Another rather general condition entering these principles is a systematic reorganization of phrase structure rules in terms of the so-called X-bar theory. This theory takes up an observation made already in Harris (1951) about the restricted options according to which constituents X and Y combine in order to form a complex constituent $\overline{X}$. In somewhat simplified terms, properties of $\overline{X}$ are determined by its head X, which furthermore selects Y as its complement or is modified by Y as a free adjunct.

These and a number of other developments led to a new version of the overall framework called the Priciples-and-Parameters-Theory, formulated in Chomsky's *Lectures on Government and Binding* (1981). The basic idea of this theory is to replace the notion of a collection of different types of highly complex rules, from which individual grammars are made up, by a system of universal principles that constrain the effect of basic operations like *Move* α and the options of X-bar theory. The principles in question determine the relation between the positions of a moved constituent, the domain of movement, the possible candidates for movement and the conditions that require the operation. They constitute separate subsystems, called Government, Binding, Bounding, Control, Case, Theta Theory. The properties distinguishing different languages are no longer seen as consequences of intricate systems of different and rather complex rules. Instead, it seems possible to account for the relevant phenomena by the interaction of these general principles. The apparent differences between different languages are assumed to be reducible to different values of a
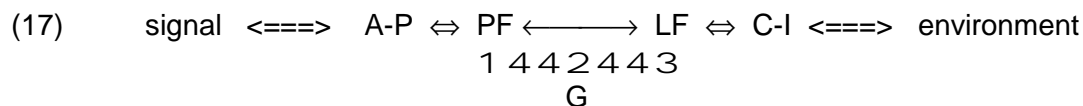
restricted set of parameters contained in these general principles. Thus the head X of a constituent $\overline{\text{X}}$ may be in initial or final position, an embedded clause may need an introductory element or not, etc. Most (or all) of these parameters seem to be related to classes of particular lexical items, so that all language particular properties are basically related to lexical information. This leads to a rather different notion of how UG determines an individual grammar G. Besides lexical items making up the language particular dictionary, grammars of different languages might differ only by their choice of parameter values, while the principles these parameters rely on are given by UG and hence identical for all languages. Under this perspective, language acquisition consists essentially in identifying the lexical system of a language, thereby fixing also the particular values of the parameters contained in UG. Presupposing the schema of levels given in (15), the organization of a grammar G within the Principles-and-Parameters-theory can roughly be indicated as follows:

(16) (a) Lexical System, fixing the idiosyncratic phonetic, semantic, and syntactic properties
        of lexical items
    (b) X-bar syntax,
        Move α
    (c) Principles of Government, Binding, Bounding, Thematic Structure, Control, and
        Case-Assignment
    (d) PF-component
    (e) LF-component

Among the numerous stimulating consequences of this model is a new perspective on historical and typological variation. Like language acquisition, linguistic change and typological differences are not only constrained by UG, but also closely related to choice of parameter values on the basis of universal principles.
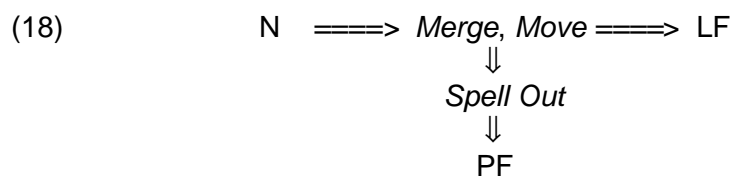
4. The Minimalist Program of Generative Grammar

Research within this framework explored the principles listed in (16c) in more detail, keeping to the overall orientation of the Generative Enterprise, viz. improving the explanatory power by constraining the theory as far as possible. This led in particular to the attempt to derive the principles in question from general conditions of cognitive organization plus the boundaries that are conceptually necessary to identify the domain of language. Minimal assumptions in this sense require linguistic expressions to compositionally relate patterns of the perceptual and articulatory system A-P to representations in the range of conceptually and intentionally organized experience C-I. This requires the internal or I-language detemined by G to provide at least the interfaces PF and LF, relating language to A-P and C-I, respectively, as indicated in (17):

(17)   signal <===> A-P $\Leftrightarrow$ PF $\longleftrightarrow$ LF $\Leftrightarrow$ C-I <===> environment
                         1 4 4 2 4 4 3
                               G

Thus G must determine a pair $< \pi, \lambda >$ for each linguistic expression, where π and λ are representations in PF and LF, respectively. One must assume furthermore, that each pair $< \pi, \lambda >$ must be based on a selection N of elements of the lexical inventory of the language. The research strategy of the Minimalist Program proposed in Chomsky (1995) pursues the aim to derive the properties of UG as far as possible from these minimal assumptions together with the hypothesis that UG is subject to conditions of structural economy. This requires first the lexical items entering the selection N to contain all and only the idiosyncratic specifications by means of which N participates in determining the pair $< \pi, \lambda >$. These include, besides phonetic and semantic features interpreted in P-A and C-I, respectively, certain grammatical or formal features controlling the computation of $< \pi, \lambda >$ from N. The minimal assumption to be made about this computation is an operation *Merge* which

combines two expressions X and Y into a complex expression Z, made up from the features of X and Y, adding the formal features of the head to characterize the formal properties of Z. As a matter of fact, the operation *Merge* is the minimalist version of X-bar syntax, reviving, moreover, a reduced version of the generalized transformations combining two expressions of arbitrary complexity. Against this background, it is an interesting, empirical fact about natural language that the derivation from N to $< \pi, \lambda >$ cannot in general be restricted to the simplest possibility requiring only the operation of *Merge*, but needs the effect of what was called *Move* α, in order to account for the "displacement" of constituents as in *Who$_i$ did he to talk to e$_i$*. As a matter of fact, the operation can be simplified to *Move*, as it is some formal feature of the moving constituent α, which triggers the operation if and only if the feature cannot be eliminated otherwise. More technically, then, the computation of linguistic expressions proceeds as indicated by (18), where is *Spell Out* is an operation that picks up the information relevant for the phonetic interpretation:

(18)   N   ====> *Merge, Move* ====> LF
                                             ⇓
                                         Spell Out
                                             ⇓
                                           PF

A wide range of facts in different languages have been shown to follow from appropriate lexical information together with minimal assumptions about *Merge*, *Move*, plus two sorts of general conditions regulating the economy in representation and derivation of expressions of I-language. What this leads to is indicated in (19):

(19) (a) Conditions on lexical items i.e. sets of phonetic, semantic, and formal features.
     (b) Operations of the computational system of language: *Merge, Move*
     (b) Economy Principles: (i) Representational: Full Interpretation FI
                             (ii) Derivational: Shortest derivation from N to <PF, LF>-pairs

The requirement of Full Interpretation presupposes *Merge* and *Move* to check and eliminate step by step the formal features, leaving only information that can appropriately be interpreted in terms of articulatory and conceptual conditions at the two interface levels. Derivational economy is a fairly abstract notion the appropriate formulation of which is still under exploration. In any event, violation of these principles is now taken to be the source of ungrammaticalness: only optimal derivations allowing for complete interpretation yield well-formed expressions. The control of these conditions is a crucial effect of formal features, coming primarily with special lexical items belonging to so-called functional categories. These include grammatical words like Determiners, Auxiliaries, but also inflectional elements like Tense, Agreement, Number, or Case.

From a more general perspective, the conditions in (19) are rather general, compared to the still rather special principles in the Principles-and Parameters-model (16), just as these principles were a strong generalization compared to the rule systems of the Standard Theory in (9). The shift from (9) to (16) replaced rules of individual grammars by universal, but still language specific principles; the organization of UG indicated in (19) replaces these language specific principles with even more general conditions of cognitive organization. Language specificity is now essentially a matter of formal features and their effect on *Merge* and *Move*, i.e. the operations that compositionally relate P-A to C-I, modules of cognitive organization which are largely independent of the language capacity.

From this point of view, the species-specific language capacity might essentially consist in the availability of discrete infinity and in fact be related, as Chomsky (1982) speculates, to the computational capacity underlying arithmetic. The language capacity clearly recruits and enhances general cognitive and communicative capacities, but its evolution, which Pinker

(1994) persuasively argues to be due to adaptive selection, might well be an independent step accompanying evolutionary changes in the architecture and size of the human brain.

5. General Perspectives

The intellectual force of generative grammar comes from the combination of two equally important factors. First, it provides a suggestive methodological and theoretical perspective for the study of language as part of the biological endowment of the human organism, supporting a central system of the overall cognitive capacity; second, it shows how this perspective can be made precise, turned into a effective research program, and pursued with respect to a large domain of empirical facts not recognized before. It is indeed the wide range of phenomena – from syllable structure and inflection to quantifier scope and syntactic embedding – as well as the diversity of languages – from English, German and Chinese to Hebrew, Malaialam, and Walbiri – that gives the Generative Enterprise its unusual place in the field. The explicit formulation of the theory, its technical means, and its cognitive orientation are the reasons for its strong impact on a wide range of sub-disciplines, from research in language comprehension and production, acquisition and aphasia to typology, historical linguistics, computer science, poetics, theory of music, and philosophy of language.

During the stages sketched in 1. to 4., various offshoots have developed different, though related, ideas. These are, a.o. Categorial Grammar, Dependency Grammar, Case Grammar, Lexical Functional Grammar (LFG), Generalized Phrase Structure Grammar (GPSG), Head-driven Phrase Structure Grammar (HPSG), Relational Grammar, to mention the more influential proposals. Another type of offspring is Optimality Theory (OT), which – starting from a different background – develops certain ideas of the Minimalist Program in a different way, assuming that conditions on linguistic structure may be of different strength and can be violated.

A peculiar relationship must finally be noted between Generative Grammar and the various theories of semantics and pragmatics. Sticking to the initial notion that grammar can and must account for the form, but not properly for the complexities of interpretation and use of language, Chomsky did not extend his theory of language to the domain of semantics and pragmatics. It might noted in this connection, that he always considered Logical Form as a syntactic level, pertaining to the form, rather than the interpretation of language. Therefore, Formal Semantics as developed by Montague (1974), Discourse Representation Theory as proposed by Kamp and Reyle (1993), Speech Act Theory as conceived by Searle (1969) and a fair number of related approaches to aspects of interpretation and use are compatible with and in part stimulated by generative grammar, but they have not been included within its proper domain. Even though this skeptical delimitation of the field was occasionally met with disappointment, it could not really diminish the intellectual attraction exerted by the Generative Enterprise.

References

Chomsky, Noam (1957) *Syntactic Structures*, S'Gravenhague: Mouton

Chomsky, Noam (1959) Review of Skinner's *Verbal Behavior*, in: *Language* 35,26-58

Chomsky, Noam (1963) Formal Properties of Grammars, in: R.D. Luce, R.R. Bush, E. Galanter (eds.) *Handbook of Mathematical Psychology*, Vol. II, 323-418, New York: Wiley and Son

Chomsky, Noam (1965) *Aspects of the Theory of Syntax*, Cambridge, Mass.: MIT Press

Chomsky, Noam (1975) *The Logical Structure of Linguistic Theory,* New York: Plenum Press

Chomsky, Noam (1981) *Lectures on Government and Binding*, Dordrecht: Foris Publications

Chomsky, Noam (1995) *The Minimalist Program*, Cambridge, Mass.: MIT Press

Chomsky, Noam (1982) *The Generative Enterprise: A Discussion with Riny Huybrigts and Henk van Rimsdijk*, Dordrecht: Foris Publications

Chomsky, Noam (1986) *Knowledge of Language: Its Nature, Origin, and Use*, New York: Praeger

Chomsky, Noam (1995) *The Minimalist Program*, Cambridge, Mass.: MIT Press

Chomsky, Noam, and Morris Halle (1968) *The Sound Pattern of English*, New York: Harper and Row

Fiengo, Robert (1977) On Trace theory, *Linguistic Inquiry* 8, 35-62

Halle, Morris (1959) *The Sound Pattern of Russian*, 's-Gravenhage: Mouton

Harris, Zellig S. (1951) *Methods in Structural Linguistics*, Chicago: University of Chicago Press

Harris, Zellig S. (1957) Co-occurance and transformations in linguistic structure, *Language* 33, 293-340

Jackendoff, Ray S. (1972) *Semantic Interpretation in Generative Grammar*, Cambridge, Mass.: MIT Press

Kamp, Hans, Uwe Reyle (1993) *From Discourse to Logic*, Dordrecht: Kluwer

Katz, Jerrold J., Jerry A. Fodor (1963) The structure of a semantic theory, *Language* 39, 170-210

Katz, Jerrold J., Paul Postal (1964) *An Integrated Theory of Linguistic Descriptions*, Cambridge, Mass.: MIT Press

Lakoff, George (1971) On Generative Semantics, in: D. Steinberg, L. Jakobovits (eds.) *Semantics*, 232-296, Cambridge: Cambridge University Press

Lees, Robert B. (1960) *The Grammar of English Nominalzations*, Bloomington: Indiana University

Miller, George A., Noam Chomsky (1963) Finitary models of Language Users, in: R.D. Luce, R.R. Bush, E. Galanter (eds.) *Handbook of Mathematical Psychology*, Vol. II, 419-491, New York: Wiley and Sons

Montague, Richard (1974) *Formal Philosophy*, New Haven: Yale Univrsity Press

Pinker, Steven (1994) *The Language Instinct*, New York: Harper Collins

Searle, John R. (1969) *Speech Acts*, Cambridge: Cambridge University Press

Skinner, Burrhus Frederic (1957) *Verbal behavior,* New York: Appleton-Century-Crofts

Abstract

Generative Grammar is the most influential linguistic theory of the second half of the 20th century. Starting with the publication of Noam Chomsky's *Syntactic Structures*, it went through a number of stages, increasing the systematic insight into the structure of the human language capacity as a species-specific component of the biologically based cognitive endowment.

The first stage proposed a system of levels that systematically characterize the formal properties of linguistic expressions. This system took up assumptions of American structuralism, turning it, however, into a complex theory of mental organization, which strongly contrasted with the behaviorist orientation of previous approaches of Bloomfield, Harris, and others. A characteristic feature of this stage was the central place assigned to the concept *grammatical transformation*, which for some time led to the identification of Generative Grammar with the more specific notion of Transformational Grammar. The second stage of the "Generative Enterprise" was the so-called Standard Theory, according to which grammatical transformations relate two levels of syntactic organization of an expression, viz. its Deep and Surface Structure, determining the semantic and phonetic interpretation, respectively. Further generalizations led to the third stage, called the Principles-and-Parameters-Theory, which showed most of the properties accounted for by idiosyncratic syntactic rules to be derivable from quite general principles. Individual languages seem to differ only with respect to lexical information and a restricted set of parameters to which the otherwise universal principles are sensitive. In a fourth stage, the so-called Minimalist Program, the language specific principles of this theory are further reduced to general conditions of cognitive organization, leaving only a set of formal features that account for the language specific properties of the computational system that correlates invariant patterns of articulation with conceptually organized representations of experience, roughly sound and meaning.

Generative Grammar was right from the beginning one of the driving forces within the field of cognitive sciences, stimulating research in language production and comprehension, language acquisition, language change and typology, computational linguistics, neuro-linguistics, and the biology of language. It has strongly influenced discussions in formal semantics and pragmatics, and the philosophy of language in general. A number of offsprings, such as Case Grammar, Generalized Phrase Structure Grammar, or Lexical Functional Grammar, pursued issues in different, though related ways.