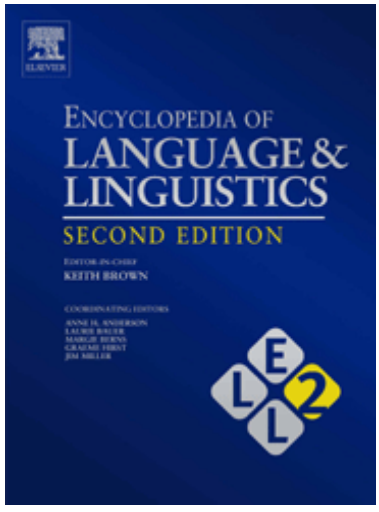
Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use



This article was originally published in the *Encyclopedia of Language & Linguistics, Second Edition*, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

Rooryck J (2006), Generative Grammar. In: Keith Brown, (Editor-in-Chief) *Encyclopedia of Language & Linguistics, Second Edition*, volume 4, pp. 767-769. Oxford: Elsevier.

that can also enrich our understanding of natural language analysis.

Bibliography

Appelt D E (1985). *Planning English sentences*. Cambridge: Cambridge University Press.

Creaney N (2002). Generating descriptions containing quantifiers: aggregation and search. In Kibble R & van Deemter K (eds.) *Information sharing: reference and presupposition in language generation and interpretation*. Stanford, CA: CSLI.

Dale R (1992). *Generating referring expressions*. Cambridge, MA: MIT Press.

Dale R & Haddock N (1991). 'Content determination in the generation of referring expressions.' *Computational Intelligence 7(4)*, 252–265.

Dale R & Reiter E (1995). 'Computational interpretations of the Gricean maxims in the generation of referring expressions.' *Cognitive Science 19*, 233–263.

Davey A C (1979). *Discourse production*. Edinburgh: Edinburgh University Press.

Frege G (1892). *On sense and reference*. [Reprinted in Geach P & Black M (eds.) (1960). *Translations from the philosophical writings of Gottlob Frege*. Oxford: Basil Blackwell.]

Garey W & Johnson D (1979). *Computers and intractability: a guide to the theory of NP-completeness*. San Francisco: W. H. Freeman.

Grice H P (1975). 'Logic and conversation.' In Cole P & Morgan J (eds.) *Syntax and semantics, volume 3: speech acts*. New York: Academic Press. 43–58.

Johnson D (1974). 'Approximation algorithms for combinatorial problems.' *Journal of Computer and Systems Sciences 9*, 256–278.

Krahmer E, van Erk S & Verleg A (2002). 'Graph-based generation of referring expressions.' *Computational Linguistics 29(1)*, 53–72.

Kripke S (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Kronfeld A (1990). *Reference and computation: an essay in applied philosophy of language*. Cambridge: Cambridge University Press.

McDonald D D (1980). *Natural language production as a process of decision making under constraints*. Ph.D. diss., MIT, Cambridge, MA.

McKeown K R (1985). *Text generation: using discourse strategies and focus constraints to generate natural language text*. Cambridge: Cambridge University Press.

Pechmann T (1989). 'Incremental speech production and referential overspecification.' *Linguistics 27*, 89–110.

Searle J R (1969). *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press.

van Deemter K (2002). 'Generating referring expressions: Boolean extensions of the incremental algorithm.' *Computational Linguistics 28(1)*, 37–52.

Author's Personal Copy

Generative Grammar

J Rooryck, Leiden University, Leiden, The Netherlands

© 2006 Elsevier Ltd. All rights reserved.

Like many linguistic terms, the phrase 'generative grammar' means different things to different practitioners of linguistics. The term 'grammar' itself can be loosely defined as the set of rules that accurately describe the combination of elements in a language. A properly 'generative' grammar is, roughly speaking, a grammar whose rules generate (i.e., produce) all and only the correct combinations of elements in a language. This definition corresponds to the use of generative grammar as a common noun. However, the term 'generative grammar' gradually received a much broader meaning as a proper noun, referring to the specific research program that is associated with the mentalist approach to language launched and developed by Noam Chomsky. In this article, the basic characteristics of both meanings of generative grammar are sketched, and the relationship between them is elucidated.

The narrower, common noun meaning of the term 'generative grammar' originates in mathematical recursive function theory and is also used in computer science. A generative grammar, then, is one of two types of formal grammars, generative and analytic. Such a formal grammar comprises a set of rules, which in turn define a possibly infinite set of finite strings composed of a set of fixed elements. These elements need not be words of natural language; they may be any kind of symbol.

A classical generative grammar contains four components: (1) a finite set of nonterminal symbols, (2) a start symbol contained in the set of nonterminal symbols, (3) a finite set of terminal symbols, and (4) a finite set of production rules that rewrite a (string containing at least one) nonterminal symbol into a string of terminal and/or nonterminal symbols. Rules

apply in any order and any number of times until all nonterminal symbols are rewritten as strings of terminal symbols. The set of strings with terminal symbols that can be generated this way is called the 'language' generated by the grammar. Generative grammars can be classified into four types conforming to the so-called Chomsky hierarchy. These types differ by an increasingly stricter format for production rules and concomitantly fewer formal languages. In this formal sense, the term 'grammar' is therefore used only by analogy with the term 'grammar' in natural languages.

Take a generative grammar that consists of the set of nonterminal symbols {X,Y} with X the start symbol, the set of terminal symbols {a,b}, and the rules $X \rightarrow aYb$, $Y \rightarrow Xb$, and $Y \rightarrow ba$. The arrow inside the rules carries the instruction 'is rewritten as.' Applying $X \rightarrow aYb$, followed by an application of $Y \rightarrow Xb$ yields the intermediate string a**Xb**b. This string still contains a nonterminal symbol. Therefore, it requires reapplication of the rule $X \rightarrow aYb$ (yielding a**aYb**bb) and subsequently the rule $Y \rightarrow ba$ to yield a string that consists solely of terminal symbols, in this case aa**ba**bbb. This grammar also generates other strings (infinitely many, in fact), which together constitute the language of this particular generative grammar.

This formal apparatus provides a powerful tool for analyzing natural language. The set of terminal symbols, then, contains the vocabulary of a natural language, say, English. The set of nonterminal symbols contains word class labels (noun, verb, etc.), as well as larger syntactic units, such as noun phrases and sentences. The largest syntactic unit, the sentence, can be taken as the nonterminal start symbol. The production rules of a descriptively adequate grammar generate the set of sentences that an English speaker would understand as uniquely English, and they fail to generate sentences and phrases that are not judged to be proper English by such a speaker. An English speaker will readily accept *the red book* as a grammatical unit, but not **the book red*. A more traditional description would then state that in English, adjectives end up in front of nouns and do not appear after them. In a generative grammar of English, this descriptive fact is a result of the order specified in the rewriting rule. In this rule, a noun phrase (NP) is rewritten as a string in which the Determiner (Det) is followed by the adjectival phrase (AP) and the noun (N) in that order: $NP \rightarrow Det\ AP\ N$. Replacing the nonterminal symbols Det, AP, and N by *the*, *red*, and *book* respectively yields *the red book*. By contrast, the rewriting rule that makes adjectives end up after nouns ($NP \rightarrow Det\ N\ AP$) is not part of the generative grammar of English. The grammar thus provides a fully explicit syntax of English, rather than the informal or implicit characterization often found in traditional grammars. English speakers have rules that allow them to generate an infinite number of new sentences by using a finite vocabulary, and to interpret entirely new combinations of words. A generative grammar can be recursive: the output of one application of a rewrite rule can serve as the input for a later application of the same rule. This mechanism of recursion is also active in natural language; it is responsible for such sequences as *the sister of the brother of the grandfather of the niece of an acquaintance of my aunt*, which can in principle be infinitely extended by recursion, limited only by performance factors.

Applying this line of thinking in a mentalist context, Noam Chomsky in the late 1950s revolutionized the way of looking at natural language. The study of language could reveal "abstract principles that govern its structure and use, principles that are universal by biological necessity and not mere historical accident, that derive from the mental characteristics of the species" (Chomsky, 1975: 3). The computationally 'generative' properties of the grammars of natural languages, and more in particular the property of recursion, is one instance of such abstract, universal, biologically necessary characteristics of the grammars developed by the human species. It should be noted, however, that it is perfectly possible to study properties of formal generative grammars outside of the mentalist context, i.e., generative grammars can be part of nonmentalist approaches to grammar.

Nevertheless, the school of thought on the mentalist nature of language developed by Chomsky and his followers came to be known – rather misleadingly and somewhat erroneously – under the popular name 'generative grammar' or 'generative linguistics,' this time used as a proper noun. In this much broader meaning, the name should be taken as no more than an expedient, historically motivated *pars pro toto*, encompassing not only syntax but a mentalist model of grammar also comprising phonology, morphology, and semantics. Accordingly, this school of thought is not limited to finding the generative grammar generating all and only the grammatical sentences of a language in the formal sense defined earlier in this article. Perhaps even more surprisingly, however, the model of generative grammar (as a proper noun) developed in the Minimalist program does not even contain a generative grammar in the common noun sense. Rewriting rules have been entirely dispensed

Author's Personal Copy

with in favor of more elementary, structure-building operations. Thus, the proper noun meaning of 'generative grammar' has entirely outgrown the common noun meaning. Understandably, then, Chomsky himself opposes the use of 'generative grammar' to describe his approach to language, preferring terms such as the rather laborious 'principles and parameters framework' or the succinctly elegant 'biolinguistics.' Despite such opposition, the term is in widespread use as representing the scientific paradigm that has dominated theoretical linguistics for the past 50 years. The remainder of this article is therefore devoted to characterizing generative grammar as a proper noun in the sense already defined.

The approach to language initiated by Chomsky is radically *internalist*. Language is viewed as a species-specific property, part of the mind/brain. The object of study of linguistics as a part of cognitive science is this mental faculty for language. Chomsky radically equates linguistic theory with the study of *I(nternal)-language*, language as an internal cognitive structure. A theory of grammar is a model of the speaker's linguistic competence, part of the internal structure of his mind. A speaker of a language has internalized a system of rules relating sound to meaning in a particular way. The task of the theoretical linguist is to describe and explain the linguistic competence of a native speaker, the linguistic knowledge of language present in the minds of individuals. A linguist construes hypotheses concerning this internalized system. *E(xternal)-language* is not part of the theory of language and comprises sundry approaches to language as it is used in society, literature, or communication.

The fact that any child can acquire any of the 6000 or so languages in the world in a relatively short period of time and under imperfect input conditions gives rise to the *Innateness Hypothesis*. This hypothesis encapsulates the claim that some set of fundamental characteristics of all human languages must be the same, and part of the *Language Faculty*. The faculty of language is part of the human biological endowment. Children are born with a faculty of language in its Initial State, i.e., the state of the faculty of language before it is exposed to the input of a particular language. This innate knowledge is sometimes called *Universal Grammar*. Language acquisition requires interaction with the environment, in the same way birds need interaction with the environment to 'learn' how to fly or sing. It is assumed that particular properties of Universal Grammar are parametrized, hence the term Principles and Parameters framework. Particular choices for each parameter of UG lead to specific grammars of individual languages. The Initial State is therefore progressively replaced during acquisition with grammars that come to resemble the Steady State, i.e., the internalized grammar of the language of a given adult speaker.

See also: Chomsky, Noam (b. 1928); Cognitive Science: Overview; Communicative Competence; Performance Factors in Spoken Discourse; Principles and Parameters Framework of Generative Grammar; Recursion.

Bibliography

Chomsky N (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky N (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky N (1975). *Reflections on language*. New York: Pantheon.

Chomsky N (1986). *Knowledge of language*. New York: Praeger.

Chomsky N (1988). *Language and problems of knowledge: the Managua lectures*. Cambridge MA: MIT Press.

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

Chomsky N (2002). *On nature and language*. Cambridge: Cambridge University Press.

Smith N V (1999). *Chomsky: ideas and ideals*. Cambridge: Cambridge University Press.

Author's Personal Copy