entailment patterns are not simply an invention of logicians or linguists. They must be constitutive, in an unconscious form, of the spontaneous knowledge that endows speakers with their linguistic abilities.

See also: Boole and Algebraic Semantics; Compositionality: Semantic Aspects; Implicature; Monotonicity and Generalized Quantifiers; Presupposition; Quantifiers: Semantics.

Bibliography


Formalism/Formalist Linguistics

P ten Hacken. University of Wales, Swansea, UK © 2006 Elsevier Ltd. All rights reserved.

In formalist approaches to linguistics, the form of language is described independently of other aspects, such as its function. Mentalist approaches, including Chomskyan linguistics, Lexical-Functional Grammar, and Head-Driven Phrase Structure Grammar, aim to describe and explain the linguistic competence of a speaker. Purely formalist approaches, including Generalized Phrase Structure Grammar and Montague Grammar, study language as an abstract object. In Chomskyan linguistics, a grammar describes the individual speaker’s competence, and universal grammar describes the genetically determined language faculty. The language faculty is essential for (first) language acquisition. This model is also used as a basis to explain aspects of language use.

General Characterization

Language can be studied from a variety of perspectives. Formalist approaches to linguistics have in common that they focus on the form of language. This does not mean that other aspects of language – for example, its use – are necessarily ignored. It does mean, however, that the aim is to describe the form of language independently of these other aspects.

An alternative interpretation of the label formalist for an approach to linguistics is that the aim of such an approach is to express generalizations about language in terms of a formalism. Although this property is logically independent of the focus on form, formalist approaches generally satisfy both interpretations.

The assumption that the form of language is determined by an autonomous system is shared by all formalist approaches. Autonomy implies that the system can be described independently of other systems. It does not imply that there is no interaction with other systems. It is this autonomy that makes it worthwhile to study language through its form.

Two main positions can be identified on the question as to how this autonomous system is realized. One is the mentalist view, which assumes that language is a knowledge component in the mind/brain of the speaker. This is the position adopted, for instance, by Noam Chomsky. The other position is formalist in the narrow sense. It assumes that a language is an abstract object. This view is found especially in the logical tradition of formal semantics and implies that what individual speakers know is an imperfect reflection of the actual language.
Formalist approaches to linguistics are opposed to approaches that focus on the function or use of language. The latter include functionalist linguistics and sociolinguistics. The aim of functionalist linguistics is to explain the form of language through its function. An example of a phenomenon often studied in this way is grammaticalization. Grammaticalization is the process by which word concepts become function words or morphemes. Sociolinguistics studies the use of language without drawing conclusions about the nature of the underlying formal system. An example of a phenomenon studied in sociolinguistics is politeness. The relationship of these two approaches to formalist ones is not the same. Functionalist linguistics denies the autonomy of the language system, because it tries to explain it on the basis of external factors. Therefore it is not compatible with formalist approaches. Most work in sociolinguistics, however, can be interpreted as complementary to formalist approaches. Often it is neutral as to the choice of a formalist or a functionalist framework.

Among formalist approaches, Chomsky’s linguistics occupies a prominent position. It is a long-established research programme that attracts significant numbers of scholars. The primary contribution of Chomsky’s theories is the discovery that human language use is constrained by an innate linguistic faculty. Chomsky’s linguistic theory is an empirical science. In an empirical science, a theory explains the form of language through its function. Chomsky’s linguistic theory defines the object of study as the knowledge of language a speaker has. This knowledge is called the linguistic competence or the knowledge of language a speaker has. This model is still the incorporation of more and more data in the explanatory scope of the theory. This model is still adopted, usually implicitly, by many researchers working in empirical science.

Logically, the empirical cycle is not sufficiently explicit to guide scientific research. At every stage, too many alternative possibilities would have to be considered. In the transition from generalizations to theories, for instance, a sufficiently imaginative researcher could come up with more theories than they could test in their lifetime. In empirical science, however, such problems do not arise. In the perception of most researchers, it is difficult enough to come up with a single theory that fits the data. The discrepancy between the logical plethora of theories and their scarcity in scientific practice can be explained by assuming a research program. A research program is the set of assumptions guiding the researcher toward sensible decisions at any step in the empirical cycle. These assumptions need not be conscious to the researcher and are often not formulated explicitly. In the same way as a good chess player does not see all logical possible moves but only the sensible ones, a scientist working in a research program automatically constrains the possible theories to be considered.

Chomsky’s linguistic theory is a research program in linguistics. As such, it should be distinguished from Chomsky’s linguistic theory. While both were conceived by Noam Chomsky in the late 1950s, their aims and later development are strikingly different. Chomsky’s linguistic theory went through a number of stages in its development, including standard theory (ST) in the 1960s, government and binding theory in the 1980s, and the minimalist program in the 1990s. In each of these stages, tree-structure representations of sentences played a role, although the constraints on the production and manipulation of these tree structures varied. Chomskyan linguistics, by contrast, remained stable during this period. It does not refer to tree structures but specifies what a linguistic theory should explain and how such a theory should be evaluated.

Chomskyan linguistics defines the object of study as the knowledge of language a speaker has. This knowledge is called the linguistic competence or internalized language (I-language). It is not open to conscious, direct introspection, but a wide range of its manifestations can be observed and used as data for the study of language.

Competence is opposed to performance. Performance includes various types of language use; for instance, a collection of utterances, a corpus of texts, a set of grammaticality judgments, and a set of results of psycholinguistic experiments. They can be used as data in linguistic research, but in every case...
the relationship they have to competence should be kept in mind in their use. Naturalistic data (utterances, texts) reflect the interaction of competence with a number of other factors. Examples of such factors are knowledge of the situation in which the utterance was made, intention on behalf of the speaker, memory limitations, and degree of fatigue. Because the role of many of these factors is poorly understood, it is often difficult to use naturalistic data as evidence for the nature of competence. Experimental data also reflect competence in interaction with other factors, but in setting up the experiment, an attempt is made to control the non-competence factors. The most frequently used type of experiment is the grammaticality judgment. There is no principled reason for treating it in a special way, but because grammaticality judgments are both highly informative and easy to obtain, they have always been widely used in Chomskyan linguistics.

A grammar is a theory about the competence of an individual speaker. It describes the competence as the system underlying the observed data. As such, the grammar explains the data. The term grammar is sometimes also used to refer to competence, but this use is potentially confusing. In early stages of Chomskyan linguistics, including Standard Theory as outlined by Chomsky (1965), grammars were described in terms of rewrite rules and transformations. At this stage of the theory, rewrite rules of the type \( S \rightarrow NP \ VP \) generated tree structures and transformations that operated on these tree structures to account for phenomena such as subject-verb agreement, passive constructions, and question formation.

The research program as described so far encounters a serious epistemological problem. If we assume a set of grammaticality judgments as our data, there are infinitely many grammars (sets of rewrite rules and transformations) consistent with our data. This is a result obtained in mathematical linguistics, the theory of formal grammars. Because the aim of linguistics is to identify the grammar describing the actual competence, we need a method to distinguish this grammar from the other ones compatible with our data. Adding more data of the same type will exclude some grammars but not reduce the problem in a principled way. At any point in time, our set of data will be finite, and for any finite set of data there are indefinitely many different grammars.

As a solution to this problem, Chomskyan linguistics adopts a cross-linguistic perspective, focusing on language acquisition. In other theoretical frameworks, different languages have been compared to reconstruct a common parent language (historical-comparative linguistics) or to find universals (typology). In Chomskyan linguistics it is assumed that certain genetically determined properties of the human mind are necessary for language acquisition. These properties are therefore reflected in one way or another in any language. By comparing different languages, linguists working in the Chomskyan framework try to find evidence of properties of the language faculty that make the acquisition of these languages possible. These properties have little in common with superficial inductive generalizations. Biologically, they are supposed to be encoded genetically in the same way as, for instance, the property that the human hand has four fingers and an opposed thumb, but they are more difficult to detect because individual languages may reflect them in different ways.

Various indications make the choice of language acquisition as the source of universals plausible. Acquisition of a language is necessary for its existence. Data from child language acquisition show that for many features of syntax, the child does not learn by repetition and resists correction. Moreover, the discrepancy between competence and performance leads to many ungrammatical and incomplete sentences in the input that children receive. The resulting I-language acquired by the child, however, is much more similar to the competence of the people in his or her environment than is a grammar that allows the sentences in the performance the child receives as input. While such considerations make a language faculty as a genetic component of the human mind plausible, conclusive empirical evidence for or against such an assumption is hard to imagine and probably impossible in principle.

In the research program of Chomskyan linguistics, the language faculty is described by a universal grammar (UG). Sometimes the language faculty is also called the language acquisition device or UG, but these names are less appropriate, because the former conflates the knowledge incorporated in the language faculty with its use and the latter conflates it with its description. The interaction of UG and grammars of individual I-languages is an essential factor in the epistemological validity of the Chomskyan research program. Since the purpose is to describe on the one hand all I-languages and on the other the language faculty that makes their acquisition possible, a proper balance has to be struck between what is genetically determined and what is acquired. If UG is too restrictive, it will not be possible to come up with grammars for all I-languages. If UG is not restrictive enough, it will not be possible to explain the learnability of the I-languages.

The interaction of the main components of Chomskyan linguistics can be represented as shown
Language is studied at the universal level (1), at the individual level (2), and at the level of data (3). These levels are represented as the three numbered boxes. On the left we find the real-world entities at each level: the language faculty, the individual competence, and the performance data. On the right we find the constructs of linguistics: UG, individual grammars, and observations about performance. At each level, the linguist’s construct describes the real-world entity. In interpreting the relationships between the different levels, it should be kept in mind that there are many instances of the lower level corresponding to a single instance of the higher level. For example, by describing an I-language, a grammar can explain (aspects of) a large set of performance data. Conversely, these data can be used as a test for the grammar. At the higher level, UG describes the language faculty and explains thereby (aspects of) a large set of I-languages. Again, conversely, I-languages can be used as a test for UG, because UG has to allow for a descriptively adequate grammar that is in addition learnable.

While Chomsky (1965) recognized the epistemological need for UG, he also observed that ST did not incorporate the mechanisms for formulating it. By progressive generalization and cross-linguistic comparison, a model of the language faculty emerged, which was formulated as the Government and Binding Theory (GB theory) by Chomsky (1981). In this model, UG consisted of a number of principles, which were universal. Differences between I-languages were accounted for by including parameters in the principles. A parameter specified a number of options, typically two. In language acquisition, parameters were set by selecting one of these options. The grammar of an I-language was thus determined by the parameter settings selected. The Minimalist Program (MP) outlined by Chomsky (1995) changed many of the theoretical assumptions of GB theory but continued to adhere to the principles and parameters model. An example of a difference between GB theory and the MP is the way movement was constrained. In GB theory, the individual transformation rules of ST were replaced by a single, general rule move \( \alpha \). This meant that anything could move anywhere unless a constraint prohibited it. In the MP, the perspective was reversed: every movement had to be motivated.

A noteworthy consequence of Chomskyan linguistics is that language can be studied at the level of individual I-languages or of the universal language faculty, but not at the level of major languages or dialects. There is no sense of language in Chomskyan linguistics, such that, for instance, David Beckham and Paul McCartney share the same language. Since they have different minds, they cannot share a mental component. The similarity of their I-languages, which by no means amounts to identity, can be explained as a consequence of the limited choice permitted by parameter settings, combined with a large overlap in vocabulary.

**Other Mentalist Approaches**

Adoption of a formalist approach to linguistics and acceptance that language is a knowledge component in the mind of the speaker by no means implies adherence to the research program of Chomskyan linguistics. There are many different ways of elaborating these two basic assumptions into a research program. Although the research program of Chomskyan linguistics has been debated more intensively and developed in more detail, this does not mean that the others are less valid in any sense. In this section, two approaches will be presented that have been developed over the past decades and continue to attract a significant group of researchers: Lexical-Functional Grammar (LFG) and Head-driven Phrase Structure Grammar (HPSG).

Lexical-functional grammar emerged in the late 1970s as a reaction against certain aspects of Chomskyan linguistics. However, the main focus of debate at the time was not the research program.
of Chomskyan linguistics, as described in the previous section, but the interpretation of a number of individual theoretical results. An example of a prominent issue was the so-called psychological reality of grammars for individual languages. In any stage of Chomskyan linguistics, a sentence is represented as a tree structure generated by rewrite rules and subsequently affected by transformations that move elements of the tree to other positions. The principles and parameters of GB theory are constraints on rewrite rules and movement. Some researchers argued that if such a grammar is a psychologically real description of the speaker’s competence, a sentence that requires more movement operations or more complex ones should take longer to process than one with fewer or simpler movement operations. They devised psycholinguistic tests to evaluate this hypothesis and found that processing times were not affected in the expected way. Cognitive complexity could not be related to the complexity of syntactic derivations in a transparent way. This stimulated the development of grammar formalisms in which the role of transformations was reduced or eliminated. Lexical-functional grammar is one of the most successful of them.

As the contributions to Bresnan (1982) indicated, LFG was devised mainly by Joan Bresnan. Instead of tree structures and transformations, LFG has a tree structure and a functional structure for each sentence. The functional structure represents grammatical functions, such as subject and predicate, in a feature structure. The tree structure and the functional structure are linked by a unification-based procedure. In this way, transformations are no longer required, so the problem they pose for the psychological reality of grammars no longer arises.

At first sight, the research program of LFG was not fundamentally different from that of Chomskyan linguistics. Superficially, the main difference seems to be the nature of the grammar formalisms. A grammar in LFG also intends to describe the competence of a speaker. When we consider the universal level, however, a more principled difference can be observed. Lexical-functional grammar assumes that the formal mapping procedure between the tree structure and the functional structure is universal. This assumption constrains the formulation of grammars in a way parallel to the function of UG in Chomskyan linguistics. This means that the language faculty in LFG is interpreted not as the knowledge needed for language acquisition but as a mechanism for language processing. Therefore, the research program of LFG can also be represented as in Figure 1, but rather than UG, a universal mechanism for processing language, i.e., mapping between form and meaning, describes the language faculty.

Another influential framework is HPSG, which emerged in the 1980s. Originally, as represented in Pollard and Sag (1987), it was strongly influenced by logical approaches to syntax (generalized phrase structure grammar; see the following section) and semantics (situation semantics). In more recent representations of the framework, Pollard and Sag (1994) and Sag and Wasow (1999) adopted a mentalist position. However, work in HPSG is more concerned with developing grammars than with discussing their status. In HPSG formalism, tree structures have been replaced entirely by feature structures. Rather than different levels of representation that are related to each other by transformations or other operations, HPSG adopts a single level. All information about what is called a sign (lexical item, phrase, sentence, etc.) is combined into a single structure. This includes phonology, syntax, semantics, and pragmatics. Relationships between components of the structure are indicated by co-indexation. This applies to both, for instance, a pronoun and its syntactic antecedent and to syntactic, semantic, and pragmatic representations of the same item. The grammar formalism in HPSG seems to be understood as a universal mechanism with language-specific parameters for building up feature structures, which suggests a research program similar to the one for LFG. The HPSG literature is in general more interested in describing individual languages than in the nature and status of universals.

**Purely Formalist Approaches**

Formalist approaches to grammar are often labeled as generative grammar. The term ‘generative’ stems from the theory of formal languages. A formal grammar is a set of rewrite rules that generates a set of sentences. The set of sentences is called a language.

In early Chomskyan linguistics, the mechanism of rewrite rules played an important role as a theoretical device. The role of the formulation of individual rewrite rules has declined in the transition to the principles and parameters model in GB theory. The conception of a language as a set of sentences goes counter to the definition of the research topic in any stage of Chomskyan linguistics. Chomsky (1986) called it an E-language (E for externalized), as opposed to the I-language he wanted to study.

Other linguists and philosophers pursued the idea that natural language should be studied as a set of grammatical sentences. Quine (1972) summarized the epistemological implications of this assumption. In his view, a language like English is ultimately a set of grammatical sentences. Although this set is infinite, it can still be determined for each sentence.
whether it is grammatical or not. For any set of sentences, there are infinitely many possible grammars. Rather than finding the correct one among them, this approach considers all of them as equally valid. It is not the knowledge of individual people that is interesting but only the language as an abstract object. English is an abstract object, of which individual speakers have imperfect knowledge, encoded in potentially quite different, more or less imperfect grammars. This position was elaborated by Katz and Postal (1991).

In the domain of semantics, this view of language has a long tradition in the philosophical study of truth and reference. In formal logic, models were developed for calculating whether a particular logical formula is true or false, given a specific assignment of truth values to primitive statements. It was Richard Montague who extended the use of these mechanisms to complex phenomena in natural language, such as quantification. Compared to Chomskyan linguistics, Montague grammar reversed the relative importance of syntax and semantics. Whereas in Chomskyan linguistics, syntax is the central area of linguistics, Montague found syntax not interesting by itself but only as a contribution to calculating the semantics of sentences. He used categorial grammar instead of rewrite rules to represent syntax. In this formalism, constituents are characterized in terms of their relationship to referring expressions and truth values. His model-theoretic semantics was meant to characterize formally the set of possible worlds corresponding to the truth of a sentence. After Montague’s death, a large number of followers continued to pursue this idea.

In the domain of syntax, the progressive formulation of principles generalizing over individual rules in Chomskyan linguistics gave rise to the objection that the grammars were no longer generative because they were not formalized in enough detail to determine whether a particular sentence would be generated or not. Many researchers who had been attracted by Chomsky’s (1965) ST because of its formality did not accept this development. The introduction of these principles was motivated at least in part by the problems caused by the excessive mathematical power of transformations. Therefore, Gerald Gazdar developed a formalism that approached this problem in an alternative way and dispensed with transformations altogether. In Generalized Phrase Structure Grammar (GPSG), even long-distance dependencies are expressed in terms of a purely ‘context-free grammar’ (CFG). A CFG consists only of rewrite rules of the type A → a, where A stands for a single symbol and a for a string of symbols. An example of a long-distance dependency is found in the following: Who did John think Mary loves? Here, who is interpreted as the object of loves. In Chomskyan linguistics, the usual way of expressing this is to say that who originates to the right of loves but is moved to the start of the sentence by a sequence of transformations. Gazdar demonstrated that a CFG can account for this as well as for many other relationships traditionally described with transformations. To achieve this, he decomposed syntactic categories into feature structures and formulated so-called meta-rules. A meta-rule is a rule that takes a CFG rule as input and produces another CFG rule. Meta-rules have access to individual features inside the syntactic categories.

One of the main advantages of CFGs is their easy implementation on a computer. However, Barton et al. (1987) demonstrated that the system of meta-rules adopted in GPSG made grammars computationally intractable. Although GPSG has lost much of its attraction and is no longer pursued as a major grammar formalism, it contributed significantly to the theory of feature structures and was at the basis of the HPSG formalism.

Formalist approaches in the narrow sense do not have a research program parallel to the ones discussed in earlier sections of this article. They do not investigate language as an empirical entity, represented in the speaker’s mind but only as an abstract object. As a consequence, their research is not guided by the empirical cycle. It is more similar to formal logic or mathematics, in which theorems are derived from axioms. The axioms correspond to the grammar and the theorems to the sentences. Contrary to common practice in logic and mathematics, formal linguistics starts with a set of theorems (grammatical sentences) and non-theorems (ungrammatical sentences) and searches for a set of axioms that generates all theorems and none of the non-theorems.

Aspects of Language Use

In all formalist approaches to linguistics, the theory of the language system is intended to be part of a general, empirical account of language-related phenomena. The hypothesis that language can be described as an autonomous system, whether realized as a component of the mind/brain or as an abstract system, implies that the interaction between this system and other factors can be invoked to account for phenomena related to language use.

The use of language in communication is one of the most prominent phenomena that have to be covered in this context. Yet it should be emphasized that language and communication are independent. This contrasts with the view adopted in functionalist
approaches, where communication is invoked to explain language. In formalist approaches, language is not considered as a tool for communication (although it can be used for communication) and communication is not seen as dependent on language (although it often involves language). By recognizing this mutual independence, relevance theory, as developed by Sperber and Wilson (1986), is compatible with formalist approaches to the description of language. In HPSG, the information pertaining to language use can be encoded directly in the feature structures representing signs. In Chomskyan linguistics, it has been suggested that there is a mental component of pragmatic competence interacting with grammatical competence. Kashir (1991) elaborated this idea.

Since the 1980s, a number of other phenomena based on language use have gained a prominent position on the research agenda of Chomskyan linguistics. They include first and second language acquisition and language change. The reason for their prominence is that they can be used as a source of data about the principles and parameters involved in language.

Whereas the logical problem of language acquisition is at the basis of the language faculty hypothesis discussed in the section on Chomskyan linguistics, the practical problem of first language acquisition is analyzed in the principles and parameters model as the process of parameter setting by the child on the basis of performance data provided by people in the child’s environment. The main question here is how the language faculty interacts with general cognitive development. Wexler (1999) advocated the hypothesis that the language faculty matures in the process of language acquisition. Lust (1999) proposed instead that the language faculty remains stable and that the first language competence is a distinct component from the language faculty.

Second language acquisition leads to a state of competence that is usually less than the one achieved for the first language. It is often called interlanguage. In Chomskyan linguistics, an interlanguage is an I-language. The question is what roles are played by such factors as the learner’s cognitive development, the type of access to the language faculty, and the first language competence in the emergence of interlanguage competence. Gregg (1996) gave an overview of some of the main positions.

Language change is the historical development of languages such as English. In Chomskyan linguistics, English is considered an epiphenomenon. Its change means that different generations of speakers have systematic differences in language competence. Lightfoot (1999) presented a model in which it is assumed that parameter settings determining a speaker’s grammar do not change in their lifetime but their use of the grammar does. The use of certain constructions may drop below a threshold necessary for a new generation of speakers to set the parameters in the same way as their parents. When this happens, the language is perceived as having changed.

In this section, various fields have been touched upon that crucially involve language use. The role that these fields play in approaches to linguistics that assume an autonomous system of language demonstrates that this assumption does not prevent the study of language use. Instead, it guides this study in a fruitful and constructive direction. The principles and parameters model adopted in Chomskyan linguistics stimulates this type of research in particular, because it provides external evidence about the nature of principles and parameters.

See also: E-Language versus I-Language; Functionalist Theories of Language; Principles and Parameters Framework of Generative Grammar.

References

Frame Problem

C Viger, University of Western Ontario, London, Ontario, Canada
© 2006 Elsevier Ltd. All rights reserved.

Origins of the Problem

The frame problem, first explicitly identified by John McCarthy and Patrick Hayes (1969), arises in the attempt to design artificially intelligent machines. Intelligent reasoning includes temporal reasoning, such as planning actions or explaining a causal sequence of events. Correct temporal reasoning requires knowing the effects that actions will have in the situations in which they are performed, both the changes and the nonchanges. The frame problem, in its original formulation, is the problem for designers of artificially intelligent systems of how to represent nonchanges, such as when a cup is moved yet its color does not change. For humans such inferences are trivial, just common sense. However, designing systems with common sense has proved to be an extremely difficult problem for artificial intelligence (AI) researchers, making the frame problem of much wider interest than a technical design problem, as it touches on age-old philosophical questions about understanding intelligence – not least as it is manifest in human conversation. In the hands of philosophers, the frame problem has come to refer to a cluster of issues concerning the relevance of the content and speed of human thinking.

McCarthy and Hayes were working within the tradition of using formal logic to design intelligent systems. To handle temporal reasoning, McCarthy had developed the situation calculus, an instance of first-order predicate logic. The formalism indexes time discretely. It includes rules that allow a system to deduce the results of an action, given a description of some situation, i.e., it can produce a description of the situation that results from the action. They applied the formalism to very simple ‘toy worlds,’ such as a block world consisting of only a small number of blocks on a table and a few simple actions, such as moving or stacking, that could be performed on the blocks. What McCarthy and Hayes realized was that the rules specifying the explicit effects of some action were not sufficient for the system to completely describe the situation resulting from that action. For example, if in situation S1, block A is on block B, and blocks C and D are on the table, the formalism does not allow the system to deduce that in S2, the situation resulting from moving block C onto block D, block A remains on block B. An additional rule specifying that block A remains on block B when block C is moved is required for that inference. The rules specifying the nonchanges during an action are called frame axioms, because they specify the stable frame of reference in which the action occurs.

Frame Axioms Result in Computational Overload

The difficulty with using frame axioms is not that the correct inferences cannot be drawn; it is, rather, that a very large number of frame axioms are required even in simple worlds. For example, adding a painting action to the block world requires distinct rules specifying that each block remains the same color when any block is moved or stacked or when any other