都留文科大学研究紀要　第83集（2016年３月）
The Tsuru University Review, No.83（March, 2016）

Transformational and Unification Grammars

Hywel EVANS

Abstract

Noam Chomsky started a revolution in the study of Linguistics as scholars became convinced that the study of syntax revealed something about the structure of the brain itself. Even so, the Chomskyan approach to grammatical analysis is not the only one available and has been aggressively challenged, particularly since the advent of Chomsky's Minimalist Program. This paper offers an overview of the historical development of Transformational Grammar and its main form of competition, Unification Grammar. It is hoped that the reader will understand how and why these different approaches developed in the way that they did. It is also hoped that the reader will begin to question which of these approaches offers the best hope as a simpler, and therefore more truly minimalist, approach to linguistic inquiry.

研究者たちが統語論こそが脳の構造そのものについて何かを明らかにするものであると確信し始めた頃、ノーム・チョムスキーは言語学研究における革命を起こし始めていた。しかし、チョムスキー手法は文法解析への唯一の方法ではなく、特にチョムスキーのミニマリストプログラムが登場して以来、激しい批判にさらされてきた。本稿では、変形文法とその主なる対局である単一化文法の歴史的発達について概説する。この二つのアプローチがなぜ、いかに発達したか読者に理解していただきたい。またどちらのアプローチがよりシンプルに、ゆえに真のミニマリストとして言語学的課題へのアプローチとして最善であるかという問いを投げかけるきっかけとなってほしい。

Introduction

It is no easy matter to teach (at least in an attractive, meaningful, and manageable way) formal syntax to language-major undergraduates. A further challenge arises in case this must be carried out via the medium of the actual foreign language being studied. One option, considered here, is to avoid undesirable complexity, and offer desirable perspective, by presenting a practical, historical overview of the two main approaches to syntactic inquiry, Transformational Grammar and Unification Grammar. It is assumed that students should be encouraged to develop their own skills of descriptive syntactic analysis so that they will naturally begin to notice certain problems that require explanation. The competing frameworks and the various available solutions may then be introduced in a manner that is relevant to the students' needs. Students will then be in a position to make up their own minds about the relative advantages associated with these options while developing a

historical sense of the development of the theories.

Generative grammar

In seeking to account for the unbounded creativity of language, systems employing categories and combinatory rules have a long history: Panini's (1896) original Sanskrit grammar actually goes back two and a half thousand years. Modern European philosophy (Descartes, 1637, p. 47) is grounded in a belief in the infinite power and creativity of reason, itself based on the infinite power and creativity of language. The early nineteenth century philosopher Wilhelm von Humboldt (1836/1999) is regarded as the first European scholar to recognize language as a finite rule-based system allowing infinite creativity. Ferdinand de Saussure (1916) stressed the importance of analyzing the structure of language and the categories involved in linguistic utterances. American structuralists such as Leonard Bloomfield (1914) and Zellig Harris (1951) suggested preliminary versions of hierarchical syntactic structure that would develop into generative grammar. However, the expression *generative grammar* (GG) was not used until the late 1950s, in relation to the syntactic theories developed by Noam Chomsky (1957), who began to look for explanations for the unique power of human language abilities.

Phrase Structure Rules

Chomsky's views on human language ability, arising out of the GG enterprise, were hugely influential and caused a revolution in the way we think about language. For example, it became widely, and often unquestioningly, accepted that much of human syntactic knowledge is innate, alleviating the effort involved for a child learning a language. It was hypothesized that our innate language ability depends in large part on a language acquisition device (LAD), or language module (LM), the part or parts of our brains that in some way houses our innate capacity for language learning. Universal Grammar (UG) is, according to this view, the mental grammar that is hard-wired into our brains from the time we are born. The centrality of syntax in this view of language was entrenched as it became widely accepted that grammar was "autonomous and independent of meaning" (Chomsky, 1957).

From the late 1950s, there were huge changes in how grammatical analysis was conducted and the consequences that were assumed for these analyses. The most famous sentence Chomsky ever wrote was intended as syntactically well-formed nonsense: "Colorless green ideas sleep furiously" (Chomsky, 1957). While the assumption that syntax was autonomous with regard to semantics was to be hotly contested, as will be discussed later, syntax was to be central to linguistic inquiry and an understanding of the brain.

In describing the syntax of any language, phrase structure rules had traditionally been employed to break down a sentence into its constituent parts, which became known as syntactic categories. In Chomsky's 1957 version of GG, the highest-level category was the

sentence (S). S consisted of two central categories, a noun phrase (NP) and a verb phrase (VP). A VP, in turn, was broken down into the verbal complex (Verb or V), including tense or aspect-related structure (Aux), and other related elements such as prepositional phrases (PP). NPs were broken down into determiners (Det), adjectives (Adj), the noun itself (N), and do on.

In the next important development of Chomsky's theory (1965), the VP was given greater complexity as a predicate phrase, and an operator node (Op) was added to create a structure potentially higher than the sentence. This Op node was utilized in order to determine what kind of sentence was being analyzed, a declarative sentence, an interrogative sentence, an imperative sentence, and so on. In this formulation, perceived movement operations — often to the Op node – became central to explaining how language works; in turn, the hypothesis that humans are the only animals that possess a dedicated LAD (or LM) helped account for the rather mysterious, sometimes invisible, movement operations that were being posited by the theorists. These movement operations mediating between different levels of structure created the eponymous transformations in Transformational Grammar.

This general approach to structural analysis was taken a step further with the first version of what became known as X-bar syntax (Chomsky, 1975). In this approach, it was hypothesized that all categories have essentially the same kind of structural properties. The highest level of structure was taken to be X", the intermediate level, X', and the lowest level X. Hence, the highest sentence-level structure was S" with the Complementizer (Comp) node replacing Op and adjoining to S" from the left in English. By 1986, S" had become I" (or IP), on the assumption that a sentence involves a projection of the *inflectional* properties of the verb (parallel to the function of the earlier Aux node). The Comp node also projected to C" with an optional adjoining Specifier (Spec) node. It was assumed that simple phrases lacked the Spec node, and complex phrases possessed the Spec node. In line with this view, NP was reanalyzed as a Determiner Phrase (DP).

Transformational Grammar

It became increasingly obvious that ever more sophisticated phrase structure rules would never be able to explain all linguistic phenomena. For example, how could discontinuous constituents be explained?

1. Taro took out Hanako.
2. Taro took Hanako out.

The prepositional constituent in 2 appears to have moved in a way that is problematic for static phrase structure rules.

3. Taro did something.
4. What did Taro do?

Sentences like 3 and 4 seem to be related in a fairly obvious way. How can 3 be converted into 4?

5. Taro seems to be happy.

In 5, *Taro* is clearly interpreted as the subject of the verb *happy*. However, simple phrase structure rules do not offer any clear explanation for this.

One solution to this would be to expand the lexicon, the information contained in words. For example, perhaps semantic and syntactic information carried in the lexicon by words such as *seems* is crucial to understanding sentences like 5. This is the kind of solution that has been adopted in Unification Grammar frameworks, which will be considered later. The Chomskyan solution, however, was to be satisfied with very limited lexical complexity and structure, and to introduce, instead, an additional rule system that gives rise to transformations. The autonomy of syntax hypothesis meant that any interplay between syntax and semantics driven by the lexicon was either discouraged or explicitly rejected. Instead, explanations for sentences such as 2, 4, and 5 were assumed to require an appeal to transformation operations. These transformations, in turn, gave rise to additional levels of structure: D-Structure (Deep Structure) and S-Structure (Surface Structure). These various levels of structure are interrelated by a set of movement rules, the transformations of Transformational Grammars (TG).

Starting in the 1970s, Robert May (1993) and Chomsky developed a third level of syntactic representation, Logical Form (LF). This was motivated as a level of representation that most fully determines the semantics of a sentence and was extremely influential in offering a way to explain semantic ambiguity via transformational rules. This provided somewhat convincing support for the hypothesis that syntax is an autonomous system with regard to semantics. In the early 1980s, Chomsky (1981) developed Government and binding (GB) theory, offering the hope that human linguistic ability could be explained very largely in terms of syntactic structure. All in all, Chomskyan Linguistics ushered in a scientific revolution in which syntactic analysis (assuming separate syntactic and semantic components in the brain) was seen as central to understanding language (Newmeyer, 1996). In this view, universal syntactic rules are seen as a direct reflection of the capacities of the brain.

Linguistic Wars

Chomskyan linguistics was largely triumphant. However, not everyone was happy with these developments (Harris, 1995). Many were concerned that language was being regarded as too central to psychology rather than being understood simply as an important tool for

communicating meaning. Chomsky's claim that syntax is independent of semantics was often criticized as implausible and unsupported (Hockett, 1968). In the 1960s and 1970s some of Chomsky's colleagues expressed strong and particular resistance to the idea that syntax, rather than meaning, was central to understanding language and the mind. This controversy became known as the Linguistic Wars; many of Chomsky's students and colleagues rebelled, beginning a movement that led to the development of the field of Cognitive Linguistics (Croft & Cruse, 2004). It was not universally accepted that language ability was innate (Anderson, 1995), and Cognitive Linguists explicitly reject that position. In this view, which has become very influential in recent years, it is assumed that the cognitive abilities used in processing language are the same as those used in other non-linguistic tasks.

Unification Grammar

In terms of syntactic theory, the main alternative to TG lies in the area of Unification Grammars (UG). UGs, which developed in large part out of attempts to provide an interpretation of natural language that can be understood by computers, have historically tended to remain neutral with regard to the innateness hypothesis. In the UG approach, language-related facts are explained via complex information written into the lexicon, with no need for movement operations and transformational rules.

As suggested, while movement operations are cut out altogether, lexical entries employed in UG tend to be intimidatingly complicated, convenient for computational analysis but often intimidating for students of grammar. Indeed, such grammars are usually referred to as "unification-based" or "complex-feature-based" (Shieber, 2003). Adding to the complexity of lexical information, UGs encode both syntactic and semantic information in order to explain language-related phenomena.

Attempts to explain linguistic facts by putting more information into the lexicon and thereby rendering transformations unnecessary started to gain steam in the late 1970s and early 1980s (Bresnan, 1982; Gazdar, 1982; Gazdar, Klein, Pullum & Sag, 1985; Kay, 1983). This work led to the development of Generalized Phrase Structure Grammar (GPSG) and the best known of the unification-based theories, Head-Driven Phrase Structure grammar (HPSG).

The best "minimalist" theory

The innateness hypothesis and the autonomy of syntax hypothesis generated a great deal of disagreement and both hypotheses remain unproven. Even so, in 1993, Chomsky raised the stakes even further by instituting the Minimalist Program. The goals of this program were to demonstrate that syntactic movement operations are driven by the need to check morphological features, and that these movement operations in turn are subject to optimality conditions. This assumption was based on a further assumption that language is

perfect in design. Chomsky's work had long been criticized as treating language as an idealized system (Hockett 1968) but Johnson and Lappin (1997) were the first to point out that the Minimalist appeal to "perfection" is both unmotivated and vague. Even so, the vagueness of the program's goals are often defended as stimulating, particularly as the goal of simplification posed by the Minimalist Program can be attained in principle within any theoretical framework (Borsley & Borjars, 2011; Bresnan, 2001; Culicover & Jackendoff, 2005; Evans, 1998).

In the Minimalist conception of things, movement operations were seen to be "costly" in some sense so Minimalists sought to reduce these operations to a minimum. Koster (1986) had previously argued that movement operations were actually unnecessary within a transformational framework. Heny (1970) and Cooper (1975) had shown that quantifier scope ambiguities could be handled without the need for covert syntactic movement at LF.

Furthermore, the so-called *pied-piping* phenomenon (Yoon, 2001) involved in *wh*-movement operations indicates that *wh*- features are involved in a percolation-type operation quite distinct from movement. This raises the possibility that movement can be dispensed with altogether, and reduced instead to something similar to this percolation phenomenon. In particular, it raises the hope that all such operations can be handled most simply and straightforwardly in a UG framework. Even so, Chomskyan Minimalists never seriously consider abandoning movement operations altogether as this would strike at the very heart of the TG enterprise. UG frameworks, however, will perhaps be better able to achieve minimalist goals because movement operations are entirely unnecessary within these formal models of grammar.

Agreement

In HPSG (Pollard & Sag, 1994), for example, structure sharing does the work of movement operations. In handling facts like verb-subject agreement, for example, nouns such as *She* are specified in the lexicon as carrying certain CONTENT semantic features, as follows:

6. PER *3rd*
   NUM *sing*
   GEND *fem*

A verb such as *walks* is specified as subcategorizing for a nominative subject as follows:

7. SUBCAT <NP: [3rd, sing]>

As can be seen in 7, the verb partially specifies semantic information carried by candidate subjects. Now let us consider a simple sentence such as 8:

8. She walks.

In UG approaches such as HPSG, as there is structure sharing between the verb and the subject, the semantic information about the noun is enriched within the verb's specifications. Information comes from two sources, the noun and the verb, which provide partial information about a single linguistic object. This information must be compatible in order to be grammatical. Semantic information carried by the verb may also be enriched in this way. This follows from the lexical specifications carried by verbs in general. Structure sharing is the fundamental explanatory mechanism employed in UGs and offers an effective "minimalist" means of reducing movement operations. Derivational approaches, by contrast, require movement of features from a source (the subject) to a target or (perhaps more naturally) a copying mechanism that complicates the inventory of syntactic operations, unless *all* syntactic operations are explained in terms of such copying operations.

Consider sentences such as 9:

9. John seems to like Mary.

In 9, the verb *seems* subcategorizes for a subject and an infinitival clause complement that has its subject's semantic features unified with that of the matrix subject, *John* in this example. This structure sharing of the subject's semantic features is indicated by [1] in the subcategorization specifications for *seems* below:

10. Subcategorization features for *seems*:
    SUBJ<NP[nom]:[1][3rd, sing]>
    COMPS<VP[inf, SUBJ<NP:[1]]:[2]>

[2] in 10 above refers to the CONTENT, the semantic features, of the infinitival complement. For the example 9, we can imagine this as looking something like 11.

11. CONTENT (semantic features) for *like*:
    RELN *like*
    LIKER [1]
    LIKED [3]

The CONTENT features for *like* can be understood quite simply as indicating a relationship of *liking*. This information unifies in the CONTENT of seems as indicated below:

12. CONTENT (semantic features) for *seems:*
    RELATION appearance

STATE OF AFFAIRS [2]

The verb *seems* specifies for the semantic information in the infinitival clause and this appears in its own CONTENT. The resulting interpretation is that there is an appearance of a state of affairs such that *John likes Mary*, a reasonable characterization of the semantics involving verbs such as *seems*. In this way, it is possible to provide a rather simple and very natural account for the semantics of such sentences while appealing to a single explanatory mechanism. Furthermore, the difference between verbs like *seems* and *tries* can be captured quite adequately.

13. CONTENT (semantic features) for *tries:*
RELATION try
TRYER [1]
STATE OF AFFAIRS [2]

In the CONTENT for *tries*, the CONTENT of the subject [1] appears as the TRYER [1], so that the subject plays a semantic role quite different from that of the subject of *seems*. Consider 14, below:

14. John tried to meet Mary.

Clearly, in 14, John is *trying* to do something in a sense that is quite clearly different from 9, where John is not a *seemer* in any sense. 12 and 13 capture this semantic intuition perfectly well in a way that is profoundly problematic for transformational approaches.

Overt movement

Now consider a sentence such as 15 below:

15. What did you buy?

In order to explain sentences such as 15, transformational approaches assume that there is overt movement of a *wh-* expression to an appropriate target structure. UGs can simply say that the subcategorized object information is subject to inheritance and hence unified on higher phrasal levels. It is then canceled in an appropriately licensed filler-head configuration.

16. To whom did you give the book?

So-called feature-percolation facts, such as indicated in 16, cause huge problems for transformational approaches (Kobele, 2005). In essence, why do *wh-* features seem to be

seeping into larger non-*wh*-expressions that are then for some reason subject to overt movement of some kind? UGs have no problem at all in accounting for such facts. This is because certain features are expected to be subject to inheritance and amalgamated by selecting lexical heads: in this case a preposition.

Conclusion

A critical, comparative presentation of syntactic approaches in a historical perspective offers considerable hope of demystifying linguistics for undergraduate students, who can experiment with operations within the competing frameworks. In particular, given the programmatic (non-theoretical) goals of the Minimalist Program, it is to be hoped that students will think for themselves about which approaches offer the best minimalist solutions. Students should certainly be offered a broad overview of competing theories as too much focus on one form of solution is obviously a rather high-risk enterprise.

References

Anderson, J. R. (1995). *Learning and memory: an integrated approach.* New York: Wiley.

Bloomfield, L. (1914). *An introduction to the study of language.* New York: Henry Holt.

Borsley, R. D., & Borjars, K. (2011). *Non-transformational syntax: Formal and explicit models of grammar.* Blackwell.

Bresnan, J, ed. (1982). *The mental representation of grammatical relations.* Cambridge, Massachusetts: MIT Press.

Bresnan, J. (2001). *Lexical-functional syntax.* Blackwell.

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

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

Chomsky, N. (1975). *Reflections on language.* Pantheon.

Chomsky, N. (1981). *Lectures on government and binding: The Pisa Lectures.* Mouton de Gruyter.

Chomsky, N. (1993). *A minimalist program for linguistic theory.* MIT occasional papers in linguistics no. 1. Cambridge MA: Distributed by MIT Working Papers in Linguistics.

Cooper, R. (1975). *Montague's semantic theory and transformational syntax.* Unpublished Ph. D. dissertation. University of Massachusetts at Amherst.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics.* Cambridge: Cambridge University Press.

Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax.* Oxford: Oxford University Press.

De Saussure, F. (1916). *Course in general linguistics.* Open Court Classics.

Descartes, R. (1637/1960). *Discourse on method and meditations.* Indianapolis, IN: Bobbs-Merrill.

Evans, H. (1998). *A unified theory of inherited features.* Unpublished Ph. D. dissertation. SOAS,

University of London.

Gazdar, G. (1982). Phrase Structure Grammar. In Jacobson, P. & Pullum, G. (eds.) *The nature of syntax representation.* D. Reidel, Dordrecht (pp. 131-186).

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. (1985). *Generalized Phrase Structure Grammar.* Oxford: Blackwell.

Harris, R. A. (1995). *The linguistics wars.* Oxford University Press.

Harris, Z. (1951). *Methods in structural linguistics.* Chicago, University of Chicago Press.

Heny, F. (1970). *Semantic operations on base structures.* Unpublished Ph. D. dissertation, UCLA.

Hockett, C. (1968). *The state of the art.* The Hague: Mouton.

Humboldt, W. von, (1836/1999). *On language: On the diversity of human language construction and its influence on the mental development of the human species.* New York: Cambridge University Press.

Johnson, D. E., & Shalom Lappin (1997). A critique of the Minimalist Program. *Linguistics and Philosophy, 20,* 273-333.

Kay, M. (1983). *Unification grammar.* Tech. Rep., Xerox Palo Alto Research Center, Palo Alto, California.

Kobele, G. M. (2005). Features moving madly: A formal perspective on feature percolation in the minimalist program. *Research on Language and Computation, 3*, 391-410.

Koster, J. (1986). *Domains and dynasties.* Dordrecht: Foris.

May, R. (1993). *Logical Form: its structure and derivatation.* Cambridge, Massachusetts: MIT Press.

Newmeyer, F. J. (1996). *Generative linguistics: A historical perspective.* Routledge.

Panini. *Astadhyayi.* (1896). Book 4. Translated by Chandra Vasu. Benares.

Pollard, C. & Sag, I. (1994). *Head-driven phrase structure grammar.* The University of Chicago Press.

Shieber, S. M. (2003). *An introduction to unification-based approaches to grammar.* Brookline, Massachusetts: Microtome Publishing.

Yoon, J-M. (2001). *Feature percolation, movement and cross-linguistic variation in pied-piping.* Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation.

Received : October, 6, 2015
Accepted : November, 11, 2015