Ninio, A. (2014). **Syntactic development: Dependency grammar perspective**. In P. Brooks, V. Kempe & J.G. Golson (Eds.), Encyclopedia of Language Development. Thousand Oaks, CA: SAGE Publications.

**Syntactic development: Dependency grammar perspective**

Dependency Grammars (DGs) are a family of linguistic theories built on the dependency relation between two words as the building block for syntactic structure. DG is an attractive framework for developmental theory due to its generative use of a single mechanism of binary combination between two words, which is iterated to produce infinitely long sentences. A review of the developmental evidence indicates that children learn a dependency-type internalized grammar to produce their word combinations.

DG originated in the so-called Prague School of linguistics of which the French linguist Lucien Tesnière was a member. Grammars belonging to this family include I. Mel'cuk's Meaning-Text Theory, R. Hudson's Word Grammar, and S. Starosta's Lexicase, among others. These types of grammars are extremely prominent in European linguistics, but they are not very well known in the United States. In such syntactic theories, grammatical relations such as Subject-Verb and Verb-Direct Object are seen as subtypes of a general, asymmetrical dependency relation: one of the words (the head) exhibits a host of local control phenomena towards another word (the Dependent). First, heads determine the syntactic and semantic features of the head-dependent combination, so that in most cases the combination inherits the features of the head. For example, a modified noun such as *white cat* is still a nominal just like *cat* as far as its semantics and its syntactic combinatory behavior are concerned. Second, heads control the characteristics and placement of their dependents: for example, the transitive verb *hit* in *John hit Fred* requires a pre-verbal subject nominal complement (*John*) and a post-verbal direct object (*Fred*). According to theory, the syntactic structure of a sentence as a whole is built up from dependency relations between individual pairs of words; it has the structure of a completely connected, directed tree. N. Chomsky’s Minimalist Program employs a very similar binary operation named Merge as its major structural-building process; hence the two theories are formally identical.

DG is a favorite among automatic parser-builders and is used in many applications of Natural Language Processing. Among other successes, IBM’s computer Watson that won the game *Jeopardy* against two human champions, uses a parser built on the principles of Slot Grammar which is a DG-type syntax developed by M. C. McCord.

DG provides an elegant framework for developmental theory due to its generative use of a single mechanism of binary combination between two words, which can be iterated to produce sentences of any length. The two-word atomic unit employed by the theory is the simplest possible word-combination; hence it provides an elegant description of young children’s word-combinations at the start of syntactic development.

A dependency-theory oriented acquisition theory can, potentially, view all of syntactic development as a continuous process, beginning with the earliest, two-word combinations. As has been pointed out by researchers such as L. Bloom, M. Bowerman, M. Braine, R. Brown, D. Ingram, L. McCune, A. Ninio, and M. Tomasello, early multiword utterances are built around a random collection of "predicate words" which children combine with words representing their arguments in the relevant propositional function. The structure of such early utterances is usually described in semantic terms, but a review of the developmental evidence indicates that the earliest word combinations are in most probability syntactic head-dependent combinations. The earliest combinations demonstrate children's having acquired the dependency and linear ordering relations between pairs of words, one of whom is a headword, the other, one of its semantic and syntactic dependents. Indeed, dependency might play a decisive role in determining word order in children’s early combinations. In an analysis of two-word long child utterances of a child acquiring Dutch, W. van Langendonck found that dependents systematically appear before heads, regardless of the specifics of grammatical relation involved, with the exception of the finite verb and its non-nominal dependents, which appear in this speech sample post-verbally. This pattern of ordering replicates the one found in Dutch child-directed parental speech.

M. Vihman in an important study of the development of early word-combinations in her son Raivo who grew up bilingual in Estonian and English, found that in his first four months of word combinations, Raivo produced many mixed-language utterances, joining words from English and Estonian in the same utterance. Such mixed language constructions cannot derive directly from the input, as the adults in this child's environment did not mix languages. In all cases the combinations represent dependency-couples, although the heads and the dependents are drawn from different lexicons. Other bilingual children have also been observed to produce a high level of mixed-language examples in their first word combinations, such as children mixing Spanish and English or French and German. Such findings demonstrate that children generate their productive combinations on the basis of words’ predicate-argument semantics, expressing the relations as syntactic combinations, and not by the rote-learning of positional formulae from the input.

English sentences are built on three core grammatical relations of which are Subject-Verb (SV), Verb-Object (VO) and Verb-Indirect object (VI). These core syntactic relations are important for discriminating between the applicability of DG and that of Construction Grammar to models of first language acquisition. Constructions are meaningful linguistic signals of any size representing learned form-meaning pairings. It is, however, well known that the core syntactic relations (also known as core grammatical relations or case roles) are not meaningful constructions, but merely purely formal roles defined by their coding properties. Among others, R. S. Jackendoff argues that basic phrase structure, structural case marking, and agreement are syntactically autonomous, and the relevant phenomena are better not included among the meaningful constructions covered by the theory of Construction Grammar. The reason, as pointed out for instance by T. Givon, is that the basic transitive and intransitive constructions of English do not possess prototypical semantics but, rather, grammatical subjects and objects can fill a great variety of semantic roles such as agent of action, patient of state, dative and benefactive. I. Schlesinger reviewed some of the linguistic literature on semantics of direct objects and concluded that objects possess a practically infinite variety of thematic roles. Jackendoff pointed out the same regarding the double-object construction of the ditransitive. It follows that the acquisition of the simplest and most important syntactic relations of nominative-accusative languages can be accounted for only if we consider them purely formal syntactic relations of the head-dependent type as in DG, and not meaningful constructions as in Construction Grammar.

In a study exploring the acquisition of core syntax in English speaking children, A. Ninio used a large corpus of 421 children based on the English language observations transcribed and stored in the CHILDES (Child Language Data Exchange System) archive, covering the early period of word-combinations known as Brown's Stage I grammar. The mean age of the children in the sample was 2;3.17 (SD 0;4.1). The children’s speech was compared to that of a group of 506 parents. Parents produced 338,970 tokens of core grammar. The relative frequency of the three core grammatical relations was SV 57.6%,VO 40.6% and VI 1.8%. Children produced 25,795 tokens of core grammar, and the relative frequency of the three grammatical relations was SV 55.7%, VO 43.1%and VI 1.2%. That is, in Stage I grammar the relative proportion of children's use of the three syntactic complements of verbs is already very similar to the distribution of the same grammatical relations in parental utterances.

The mastery of core syntactic relations is apparently facilitated by the use of pronouns and other indexical referential signs. A. Ninio found that English-speaking parents’ two-word long sentences of the SV, VO and VI relations have indexical signs, mostly pronouns, as complements in 94% of the time. L. McCune investigating the SV, VO and SVO patterns in four children acquiring English found that pronoun use spikes before the children begin to produce grammatical relations.

At this first stage of syntax, children need only to learn the combinatory behavior of individual predicate words, on a lexically specific basis. These lexically specific word-combinations are not, however, isolated from each other. For instance, when a child learns to produce verb-object combinations with a new verb, her learning is facilitated by previously having learned the same coding rules for other verbs. For a given set of coding rules, learning exhibits a typical accelerating *learning curve* over the sequence of lexemes learned into that pattern, so that, for example, learning the tenth verb with a direct object takes much less time and effort than learning the first, second or third verb in this pattern. For example, Tomasello’s daughter Travis began verb-object combinations with the verb *get,* placing the direct-object pronoun or noun in a post-verbal position. Although she generated quite a few sentences of this type with the verb get, it took heralmost a full month to proceed from the first verb to the second, *find,* in this pattern. After this slow start, however, other verbs such as *open*, *catch* and *hold*, appeared one after the other with a post-verbal direct object; the learning of their syntactic behavior apparently facilitated by the previously learned verbs with the same syntactic rule. Transfer of learning within a given syntactic pattern is made easier by the fact that children have a tendency to learn early on general lexical items such as all-purpose or generic verbs (e.g., the transitive verbs *do, make* and *want* and the intransitive verbs *come* and *go*); such “pathbreaking verbs” may provide easy-to-apply templates for the syntactic pattern for other, more specific ones, as claimed by E. Clark. Transfer of learning and facilitation are based on similarity of form and not of meaning, hence a subject which is an agent of action can still facilitate the acquisition of a subject which is an undergoer or a patient. In addition, syntactic constructions facilitate each other’s acquisition if they are merely similar but not identical in their form, as was found in a study by K. Abbot-Smith and H. Behrens comparing two German passives, one of which (the *sein*-passive) was apparently easier to learn than the other (the *werden*-passive) due to the former’s similarity to previously learned constructions using the same auxiliary verb such as the copula-construction.

After learning the basic dependency combinatory mechanism and the generation of two-word syntactic atoms, the next developmental task is extending the syntactic structure to include more than one dependency relation. That is, children need to learn to apply the dependency operation iteratively. This skill is needed in order to build three-word sentences, and, with repeated iterations, sentences of any length. Evidence for the acquisition of the principle of iteration of the head-dependent relation as a separate developmental stage comes from researchers such as L. Elbers, G. Ewing, J. C. Hill, and S. Powers who have observed that moving from a single two-word combination to a three-word long sentence in which a second dependency relation is built on one of the original words, poses a special difficulty for some children. In particular, sometimes children repeat the shared word participating in the two combinations, creating a sequence of two separate dependencies instead of a combined one, such as in the sentence *Take this, this ball*. With further development, the shared word occurs only once, generating the correct *Take this ball*. The processing load posed by the dependency structure of different three- and four-word constructions predicts the order of their acquisition. Analysis by A. Ninio of an English-speaking child's first 102 sentence types of 3- and 4-words (produced when the child was between 1;6.8 and 1;7.18) showed that, with the exception of just 10 sentences, in the overwhelming majority (90.2%) were the all-adjacent type, namely, all dependency couples were immediately adjacent to each other. For instance, in the sentence *Mommy made book*, both *Mommy* and *book* are dependents of the verb *made* and both are adjacent to it. In the minority of sentences such as *Draw me man,* both *me* and *man* are dependents of the verb *draw* but the first dependent (*me*) separates the second dependent (*man*) from its head *draw*. In usual syntactic terminology, co-dependents of the same head are sisters, thus this sentence is an exemplar of the pattern where a dependent is separated by a sister-dependent from their shared head. There were practically no sentences among the first 102 where a dependent is separated from its head by its own dependent, as would be in the sentence *Draw yellow ball*, where *yellow* is a dependent of *ball*. For comparison, the next 272 sentence types of more than two words were also analyzed; these were produced when the child was between 1;7.19 and 1;8.8. Of these sentences, 74.3% were the all-adjacent type as all dependents were immediately next to their syntactic heads, as in the sentence *Maria hit me* where both *Maria* and *me* are adjacent to their head *hit*; the rest, 25.7%, had at least one dependent separated from its head. This result shows a considerable and significant increase in the more complex constructions at this second period of the development of three-plus word combinations. The next pattern to develop after the all-adjacent one is where a dependent is separated by a sister-dependent from their shared head, as in the sentence *Maria told me draw* where both *me* and *draw* are dependents of *told* and *me* separates its sister *draw* from the head *told*. The third pattern involving a dependent which is separated from its head by its own dependent (as would be in the sentence *Bring big apples* where the attributive adjective *big* separates its head *apples* from that word’s head *bring*) is a later acquisition and there were practically no sentences of this kind in the relevant age range.

The results of this study were replicated on a sample of children acquiring Hebrew. Analyzed were 10-month-long longitudinal home observations of 16 children, half from 18-28 months and half from 22-32 months old. There were 2,510 analyzable three-word utterances in the corpora. Sentences in which all dependents are immediately adjacent to their heads constitute about 80% of all three-word utterances at this age group; moreover, their proportion continues to be very high throughout the 22-32 months period, with a very slight tendency to decrease. Sentences in which a dependent is separated in the string from its head by a co-dependent constituted about 17.5% of all sentences; this pattern gains in frequency with increased age. The third type -- sentences in which a dependent is separated from its head by its own dependent -- account only for less than 5% of the overall production, and its frequency does not show a tendency to increase with age in the investigated age range. Again, all-adjacent dependencies were the first to develop, followed by sister-separated dependents.

These findings provide a window into the cognitive processes through which the means of sentence generation and interpretation are carried out. When two syntagmatically related words are in each other’s immediate neighborhood in the vocal string, it means that these two words may be accessed in immediate succession. In psycholinguistic terms, a dependency relation is a computational command to the effect that during sentence generation and comprehension, the information carried by the two separate words comprising the dependency couple is to be combined or synthesized; if these words are immediately consecutive, this procedure can be carried out without recourse to storage in, and retrieval from, short-term memory. During sentence generation, if the two members of a dependency couple are to be separated by some intervening material in the sentence, the processing of one dependency relation is interrupted by the processing of another, thus creating an “open dependency” for the duration. Until the second member of the couple is generated, the speaker has to keep in short-term memory the fact that such a closure is pending. For example, the generation of *Give him apples* requires that the speaker store the demand for the second dependent of *give* until after the first dependent (*him*) is produced.

Open dependencies also demand extra processing during the interpretation of sentences. Listeners have to store in memory all words whose requirement of a head have not yet been satisfied until the head word is encountered. For instance, in the sentence *Children in school and day care often catch common infections from their peers,* the word *children* is separated by six other words from its syntactic head the inflected verb *catch.*, requiring that the listeners keep the word in storage while processing the six words in between. In addition, listeners have to backtrack their way in the sentence when they encounter a word whose head is a word that has already been processed (e.g., *apples* in *Give me apples*, or *quickly* in *Come here quickly*). There is evidence from adults that the storage and retrieval processes involved take up considerable short-term memory; there seems to be a very strict absolute limit on the number of open dependencies operative at the same time in a sentence, so that if their number passes three, the sentence is incomprehensible (e.g., *The rat the cat the dog chased ate died*.)

Except for the basic principles of dependency and its iterative application, the further growth of a DG-type grammar in children is driven by the development of vocabulary. Not only is each predicate to be learned on an individual basis, together with the syntactic relations it needs to or can be involved in, but the level of organization of children’s syntactic system appears to be considerably influenced by the acquisition of so-called functional items such as determiners, prepositions and auxiliary verbs. B. Corominas-Murtra, S. Valverde, and R. V. Solé, working in the Complexity Science tradition, analyzed the longitudinal corpus of a child obtained from the CHLDES database and built a syntactic network where words are connected through syntactic links to all other words they were in a head-dependent relation with in the transcribed child sentences. They found that around two years of age a sharp transition in the structure of the syntactic network, from a less-organized structure to a much more adult-like and complex organization. They connect this leap to a sudden lexical explosion of functional words in the child’s vocabulary just prior to this time. That is, both the specific components and the global organization of the child’s syntactic system is driven by advances in the lexical items the child is mastering.

Recently, an automatic syntactic parser (MEGRASP) was added to the programs available to CHILDES users, updating a previous version called GRASP. This program, developed by K. Sagae and colleagues, uses information on the morphological tier of the CHILDES files to generate labeled syntactic dependencies. This is a parser in the tradition of Dependency Grammar. The parser has been used to tag the grammatical relations in the English, Spanish, Mandarin, and Hebrew segments of the CHILDES database. Hopefully, the availability of an automatic dependency parser will open the door to many future studies of syntactic development from the DG perspective.

Anat Ninio

The Hebrew University of Jerusalem

See Also: Argument structure (acquisition of); Item-based learning; Relational terms (acquisition of); Syntactic development: Construction grammar perspective; Syntactic development: Generative grammar perspective; Child Language Data Exchange System.

Further Readings

Corominas-Murtra, Bernat, Valverde, Sergi, and Solé, Ricard. V. “The Ontogeny of Scale-free Syntax Networks: Phase Transitions in Early Language Acquisition.” *Advances in Complex Systems,* v.12 (2009).

Hill, Jane C. “Combining Two-term Relations: Evidence in Support of Flat Structure.” *Journal of Child Language,* v.11 (1984).

Hudson, Richard L. *English Word Grammar*. Oxford: Blackwell, 1990.

Mel'cuk, Igor A. *Dependency syntax: Theory and practice.* Albany, NY: State University of New York Press, 1988.

Ninio, Anat. *Language and the learning curve: A new theory of syntactic development.* Oxford: Oxford University Press, 2006.

Ninio, Anat. *Syntactic development, its input and output*. Oxford: Oxford University Press, 2011.

Robinson, Peter. “Constituency or Dependency in the Units of Language Acquisition? An Approach to Describing the Learner's Analysis of Formulae.” *Lingvisticae Investigationes, International Journal of French Linguistics and General Linguistics, v.*10 (1986).

Sagae, Kenji, Davis, Eric, Lavie, Alon, MacWhinney, Brian, and Wintner, Shuly.“Morphosyntactic Annotation of CHILDES Transcripts*.*” *Journal of Child Language,* v.37/3 (2010).

van Langendonck, Willy. “Word Grammar and Child Grammar.” *Belgian Journal of Linguistics*, v.2(1987).