A dynamic network model of grammatical constructions[1]

Holger Diessel
Friedrich-Schiller-Universität Jena

The theoretical framework used by most researchers of child language development is Chomsky's theory of generative grammar. The theory has changed considerably in recent years. The older model, which is still often used in child language research, is called Principles and Parameters, while the newest version is known as Minimalism (cf. Chomsky 1981, 1986, 1995, 2000). Within this framework, it is commonly assumed that children are born with an innate universal grammar consisting of principles and parameters that define the space within which the grammars of individual languages may vary. Grammatical development is seen as a process whereby the parameters of universal grammar are set to a language-specific value by linguistic triggers in the input.

The theoretical framework used in this study is very different; it is based on recent work in functional and cognitive linguistics. The functional-cognitive approach subsumes a variety of related frameworks (cf. Croft 1995; Newmeyer 1998). Two of them are especially important to the current investigation: construction grammar and the usage-based model. Construction grammar subsumes a family of grammatical theories in which constructions are considered the basic units of grammar (cf. Fillmore, Kay, and O'Connor 1988; Lakoff 1987; Langacker 1987a; Fillmore and Kay 1993; Goldberg 1995; Croft 2001), and the usage-based model comprises various network models in which linguistic knowledge is shaped by language use (cf. Bybee 1985, 1995, 2001; Langacker 1987a, 1991; Barlow and Kemmer 2000; Elman, Bates, Johnson, Karmiloff-Smith, Parisi, and Plunkett 1996). Although construction grammar and the usage-based model are in principle independent of each, they are often combined in linguistic analyses (e.g. Langacker 1987a; Croft 2001; Morris, Cottrell, and Elman 2000). This chapter discusses the basic principles of the two frameworks in a unified approach.

1 Construction Grammar

1.1 Constructions

In construction-based theories, constructions are the basic units of grammar. They are commonly defined as grammatical assemblies that are characterized by two features: first constructions combine a specific form with a specific function or meaning (e.g. Lakoff 1987), and second constructions exhibit both general grammatical properties and idiosyncratic features (e.g. Fillmore, Kay, and O'Connor 1988).

The importance of grammatical constructions has been emphasized in various contemporary theories of grammar, notably in Construction Grammar[2] and Cognitive Grammar (cf. Fillmore, Kay, and O'Connor 1988; Fillmore and Kay 1993; Goldberg

[1] This paper has been reprinted in my book The Acquisition of Complex Sentences (Cambridge: CUP, Chap2)

[2] The notion of Construction Grammar, spelled with initial capital letters, refers to a particular construction-based theory developed by Charles Fillmore, Paul Kay, and Adele Goldberg.

1995; Lakoff 1987; Langacker 1987a, 1988, 1991, 2000; Croft 2001), but also in Head Driven Phrase Structure Grammar (Sag 1997), Role and Reference Grammar (Van Valin and LaPolla 1997), and several other related frameworks (cf. Prince 1978; Zwicky 1987, 1994; Pullum and Zwicky 1991; Wierzbicka 1988; Jackendoff 1990, 1997; Culicover 1998).

The traditional notion of construction refers to specific sentence types such as the passive. The main passive construction in English consists of a subject, the auxiliary *be,* and a verb in the past participle, which may be followed by a *by*-phrase. The whole structure is associated with a particular meaning: in contrast to active sentences, passive sentences express the patient (or undergoer) in the subject NP, whereas the actor is optionally expressed in the postverbal *by*-phrase. What is more, although the passive is defined by common grammatical categories and syntactic relations, it is not sufficiently described on the basis of general grammatical rules; rather, an adequate analysis of the passive must take into account that the whole structure constitutes a specific grammatical unit. It combines general grammatical properties with idiosyncratic features that can only be described by construction-particular rules. For instance, the particular expression of the actor in a *by*-phrase is an idiosyncratic property of the passive that cannot be derived from the meaning of *by* and some general grammatical rules.

In construction grammar the notion of construction has been generalized. It does not only apply to structures such as the passive; rather, construction grammar argues that all grammatical assemblies are constructions, i.e. conventionalized symbolic units consisting of a specific form paired with a specific function or meaning (cf. Lakoff 1987; Langacker 1987a; Fillmore et al. 1988; Goldberg 1995). The formal side comprises phonological, morphological and syntactic features, and the functional side subsumes semantic, pragmatic and discourse-pragmatic features.

Grammatical constructions can be seen as complex linguistic signs. In the structuralist tradition of linguistics, the notion of sign is used for words but does not apply to grammatical assemblies (cf. Saussure 1916). However, in construction-based theories the notion of sign has been extended to constructions because constructions are like words in that they represent conventionalized form-function pairings: both can be seen as symbols in which a specific form is paired with a specific function or meaning. Strong evidence for the symbolic nature of grammatical constructions comes from recent experimental studies in which it is shown that speakers associate specific meanings with particular morphosyntactic structures (cf. Bencini and Goldberg 2000; Hare and Goldberg 2000; Kaschak and Glenberg 2000).

The notion of construction is incompatible with central assumptions of generative grammar. According to Chomsky (1965), the system of grammatical rules is divided into three major components: the syntactic component, the phonological component, and the semantic component. Similar divisions hold for more recent versions of generative grammar (cf. Chomsky 1981, 1995, 2000). Each component has its own rules that in principle are independent of each other; that is, grammar comprises syntactic, phonological, and semantic rules that apply in separate compartments or "modules". Since the modules are more or less "autonomous" (i.e. encapsulated compartments of grammar; see Croft 1995 and Newmeyer 1998 for discussion), there is no room for complex linguistic signs in the classical version of generative grammar (but see Jackendoff 1990, 1997). The only conventionalized form-function pairings are words. The meaning and structure of grammatical assemblies (i.e. phrases and clauses) can always be decomposed into semantic and syntactic primitives that constitute the building blocks of complex linguistic elements in this approach. Grammar is thus entirely compositional in Chomsky's version of generative grammar and therefore the notion of construction has been abandoned in this approach:

> The notion of grammatical construction is eliminated, and with it, construction-particular rules. (Chomsky 1995: 4)

Note that the elimination of constructions includes structures such as the passive, which have always been treated as constructions in linguistic theory. All complex linguistic expressions are fully compositional in the current version of generative grammar (i.e. Minimalism). They are derived from a small number of linguistic primitives and some general grammatical rules. The only exceptions are idiomatic expressions, which obviously do not abide by general rules. However, since idioms have the status of words in generative grammar, they do not undermine the general principle that grammar is strictly compositional in this approach.

1.2 The grammar-lexicon continuum

In the standard version of generative grammar, grammar and lexicon are strictly distinguished: grammar consists of principles and rules that account for the systematic or general properties of language, whereas the lexicon contains all idiosyncratic information, i.e. information that cannot be derived from general rules. Construction-based theories have abandoned the categorial division between lexicon and grammar (cf. Langacker 1987a; Goldberg 1995; see also Hudson 1990; Pollard and Sag 1994; Van Valin and La Polla 1997). Since both words and grammatical constructions are considered symbolic units (i.e. form-function pairings) they are uniformly represented in this approach. Specifically, grammar is seen as a continuum ranging from isolated words to complex grammatical assemblies (cf. Langacker 1987a: 25-27; see also Slobin 1997). Idiomatic expressions are part of the grammar-lexicon continuum; in fact, idioms have played a key role in the development of this conception of grammar (cf. Fillmore, Kay, and O'Connor 1988; Nunberg, Sag, and Wasow 1994).

Idioms are obviously conventionalized form-function pairings. Consider for instance the idiom *kick the bucket* (cf. Fillmore et al. 1988). It has a nonliteral meaning that is not predictable from the meaning of its components. Moreover, certain syntactic properties of this expression are idiosyncratic. For instance, the structure cannot be passivized (**the bucket was kicked*) and the object NP is restricted to the singular (**kick the buckets*). However, like most idiomatic expressions *kick the bucket* is not entirely idiosyncratic; rather, it involves grammatical properties that are also found in nonidiomatic expressions. For instance, the verb can occur in different tenses (e.g. *kicked/will kick the bucket*) and is followed by an NP that can be analyzed as the direct object. Thus, the expression *kick the bucket* combines idiosyncratic properties with general grammatical features. This suggests that idiomatic expressions such as *kick the bucket* are not in principle distinguished from regular expressions such as the passive. In fact, idioms can be seen as grammatical constructions that basically carry the same features as nonidiomatic expressions. Idiomatic and nonidiomatic expressions are commonly defined by both regular grammatical patterns and construction-specific features. Compare for instance the previous discussion of *kick the bucket* with Goldberg's analysis of the caused-motion construction.

The form of the caused-motion construction is defined by the following assembly of grammatical categories: '[SUBJ [V OBJ OBL]]' (Goldberg 1995: 152). Semantically, the construction expresses the meaning 'X causes Y to move somewhere'. Examples of the caused-motion construction are given in (1) to (4) (the examples are adopted from Goldberg 1995):

(1) She dragged the child into the car.
(2) He wiped the mud off his shoes.
(3) She forced the ball into the jar.
(4) He pushed the book down the chute.

Note that the verbs of these examples have two semantic features that characterize the meaning of the whole structure: first they are semantically causative (i.e. an agentive subject is acting on a patient), and second they indicate some kind of motion or movement. Based on these examples, one might assume that the caused-motion interpretation is evoked by the verbs that occur in these constructions, but Goldberg (1995: 152-179) shows that the caused-motion reading is also evoked if the construction includes a semantically different verb such as *sneeze*.

(5) She sneezed the napkin off the table.

Sneeze is neither a causative verb nor is it used to indicate motion, but the sentence in (5) has precisely this meaning, which suggests that the caused-motion interpretation is not evoked by the verb. Goldberg argues that the caused-motion reading is a property of the whole structure. In other words, the construction is associated with a specific meaning independent of the lexical expressions it includes. The whole structure constitutes a conventionalized symbolic unit, which cannot be reduced to the properties of its components.

A similar analysis has been proposed for many other constructions such as comparative conditional clauses (Fillmore et al. 1988), presentational and existential *there*-constructions (Lakoff 1987), resultative clauses (Goldberg 1995; Nedjalkov 1983, Boas 2003), verb-initial sentences (Diessel 1997b, 2003b), verb-particle constructions (Gries 2003), and nominal extrapositions (Michaelis and Lambrecht 1996). In fact, construction grammar maintains that all grammatical assemblies are constructions; even the most general structures such as transitive clauses can be seen as conventionalized form-function pairings (cf. Goldberg 1995: 116-119). What distinguishes such general structures from idioms is that they are more abstract and less idiosyncratic. However, that does not constitute a principled difference between idiomatic constructions such as *kick the bucket* and more general constructions such as the caused-motion construction or the passive. Both idiomatic constructions and nonidiomatic constructions are form-function pairings that combine general grammatical properties with idiosyncratic features.

If grammar consists of symbols (i.e. form-function correspondences), there is no principled difference between lexicon and grammar. The only feature that distinguishes grammatical constructions from words is that constructions generally include at least two meaningful components, whereas words may consist of a single meaningful element (i.e. a single morpheme). However, many words are morphologically complex: they consist of multiple morphemes that are combined to complex expressions, which one might analyze as "morphological constructions" (Langacker 1987a: 83-85). Thus, although words do not generally consist of multiple components, there is no principled difference between words and grammatical constructions, and therefore construction-based theories have abandoned the categorial division between lexicon and grammar.

1.3 Schemas and rules

Grammatical constructions vary along two important dimensions (cf. Fillmore et al. 1988; Croft and Cruse forthcoming). First, they vary in terms of syntagmatic complexity. Some grammatical constructions consist of only two elements while others include multiple components. For instance, a prepositional construction such as *in Berlin* contains two structural elements, a preposition and a noun (phrase), whereas the caused-motion construction comprises four elements, namely a subject, a verb, an object, and an oblique (see above).

Second, constructions vary along a scale of schematicity or abstractness. A construction is schematic if it consists of abstract grammatical categories such as NP or subject, and it is concrete if its components are filled by specific lexical items. For instance, idiomatic expressions such as *kick the bucket* are concrete constructions, in which each element is a concrete lexical expression. Abstract structures such as the caused-motion construction on the other hand are highly schematic constructions, which consist of abstract grammatical categories such as NP or subject. Schematic constructions are also called "(constructional) schemas" (Langacker 1987a; Bybee 1995; Ono and Thompson 1995); they account for linguistic generalization, which in other frameworks are described by rules (cf. Rumelhart and McClelland 1986a; Pinker and Prince 1988; Pinker 1999; Bybee 1995; Elman et al. 1996; Marcus 2001; Ramscar 2002).[3]

Constructional schemas are like grammatical rules in that they describe the general properties of linguistic structures. In fact, a constructional schema can be seen as a notional variant of a rule if a grammatical rule is defined as a pattern involving variables (or "placeholders") that can be filled by certain types of elements (cf. Marcus 2001). However, in contrast to traditional grammatical rules, constructional schemas are symbols, i.e. form-function correspondences. They do not only define the way in which elements can be combined but contribute their own (idiosyncratic) properties to grammatical assemblies. In other words, schemas are essentially of the same type as the expressions they combine: both are conventionalized form-function correspondences, whereas traditional grammatical rules (e.g. phrase structure rules such as NP $\rightarrow$ DET (ADJ) N are of a different kind than the elements they combine (e.g. words or phrases).

Since schemas are linguistic expressions, they can be related to other linguistic expressions (i.e. other schemas or concrete constructions). The relationship between schemas (or constructions) is based on similarity: two constructions are closely related if they share a significant number of features. For instance, the ditransitive construction (e.g. *Sally gave Peter the ball*) is closely related to the caused-motion construction (e.g. *Sally gave the ball to Peter*) because the two constructions have similar forms and meanings: both include two arguments and express some kind of transfer (cf. Goldberg 1995: chap 3).

The similarity between constructions constitutes one important aspect of productivity in this approach. Since the similarity between constructions is gradient, rather than absolute, the productivity of constructional schemas varies along a continuum: in the extreme case a constructional schema applies to all instances of a particular type, but very often the application of a constructional schema is much more limited (i.e. restricted to particular types in certain situations).

[3] Although the notions of schema and construction are closely related they must be kept separate. The notion of construction subsumes both abstract grammatical patterns and lexically-specific (or idiomatic) expressions. By contrast, the notion of schema applies only to abstract grammatical patterns, i.e. a schema can be seen as a particular type of construction.

Apart from similarity, the productivity of a constructional schema is determined by the number of expressions that are related to a particular schema: the more types of expressions are linked to a constructional schema, the more productive is its use (cf. MacWhinney 1978; Bybee 1985, 1995). The productivity of rules, on the other hand, is not affected by type frequency. Rules are always fully productive; they automatically apply to all linguistic expressions that carry a certain grammatical feature (cf. Pinker and Prince 1988; Marcus 2001; for a detailed discussion of productivity see Section 2.2.4).

Finally, the symbolic nature of grammatical constructions explains why many grammatical patterns show prototype effects (cf. Givón 1979, 1984; Hopper and Thompson 1980; Bybee 1995). The prototype effects result from the relationships between a constructional schema and its instances. For example, the transitive construction is a constructional schema that is related to a wide variety of subconstructions (i.e. different instances of transitive clauses). In the transitive schema, subject and object are associated with the semantic roles of a prototypical agent and a prototypical patient respectively (cf. Hopper and Thompson 1980). If the verb of a transitive clause has a causative meaning, as in *Peter throws the ball*, subject and object express these roles, but if the verb denotes a psychological state, as in *Peter likes bananas*, the semantic roles are only remotely related to the semantic roles of the transitive schema. In other words, *Peter throws the ball* is a better instance of a transitive clause than *Peter likes bananas.* In a construction-based framework, this can be represented by different types of links relating the various types of transitive clauses to the transitive schema. In a rule-based approach, on the other hand, all transitive clauses are licensed by the same rules, i.e. all transitive clauses have equal status in this approach.

1.4 Prefabricated formulas

In some varieties of construction grammar, low-level constructions are in general underspecified. They include only information that is not provided by more schematic representations (cf. Fillmore and Kay 1993). For instance, an idiomatic construction such as *kick the bucket* would not include general syntactic information about its structure because this information is "inherited" from a constructional schema. In this variety of construction grammar, low-level constructions contain only idiosyncratic information that they do not share with other constructions; all general grammatical features are inherited from constructional schemas. The representations are thus minimal in this approach: every piece of information is only represented in one place in mental grammar.

Other varieties of construction grammar posit that lower-level constructions are fully specified (cf. Goldberg 1995: chap 3; Langacker 1987a: 87). In this view, constructions contain all the information available at a specific level of schematicity, including information that they share with more schematic representations. Lower-level constructions do not inherit information from constructional schemas; rather, they are linked to more schematic representations by instantiation links that indicate the overlap of information. Thus, in this variety of construction grammar, which can be seen as an instance of the "exemplar model" (cf. Nosofsky 1988), the same information is often stored redundantly at different levels of abstraction.

In generative grammar and many other theoretical frameworks, including certain varieties of construction grammar, the storage of information is maximally economical and nonredundant. Economy and nonredundancy are important criteria for the evaluation of scientific models. However, the application of these criteria presupposes that the proposed models provide an adequate account for the phenomena they describe. Adopting an exemplar-based view of categorization, I contend that generative grammar and many other grammatical theories are psychologically inadequate, precisely because these frameworks posit that grammatical representations are maximally economical and nonredundant. Speakers store frequently occurring word collocations and concrete utterances along with constructional schemas; that is, grammar consists of both abstract grammatical representations and prefabricated chunks of concrete expressions that are frequently used in everyday speech (cf. Pawley and Syder 1983; Langacker 1987a, 1991; Bybee and Scheibman 1999; Gregory, Raymond, Bell, Fosler-Lussier, and Jurefsky 1999; Erman and Warren 1999; Wray and Perkins 2000; Thompson and Hopper 2001; Jurefsky, Bell, Gregory, and Raymond 2001). Some of these "prefabs" (Erman and Warren 2000) are fully specified utterances, others consist of concrete expressions that are associated with a specific slot. A few illustrative examples are given in (6).

(6)	*Fully specified utterances*	*Concrete utterances including a slot*
	How are you doing?	Why don't you __ .
	Thank you, I'm fine.	I don't know __ .
	What can I do for you?	Do you mind if __ .
	Get the hell out of here!	I am just about to __ .
	You can't have it both ways.	Would you please __ .
	Either way is fine.	__ is not in the position to __ .
	Say that again.	I can't help Ving __ .
	I don't believe what's happening.	__ never got around to __ .
	You gotta be kidding.	That's just about the __ that __ .
	No, I'm dead serious.	I wonder if __ .

Every native speaker of English knows a very large number of such prefabricated chunks and word collocations. Some of them are entirely familiar expressions that have been used many times before, others are somewhat less familiar and allow for some variation; however, none of the expressions in (6) is newly created.

The frequent use of prefabricated chunks is one of the features that distinguishes the speech of native speakers from the speech of second language learners (cf. Pawley and Syder 1983). The speech of second language learners often sounds unnatural, even if it is grammatical, because second language learners usually do not have enough communicative experience to know the prefabricated chunks that are characteristic of everyday speech.

Although the expressions in (6) are in accordance with general schematic representations, they are not derived on-line by means of constructional schemas; rather, native speakers access these structures directly without activating the corresponding schemas. Thus, from a psychological perspective, the exclusion of redundant information from grammar seems to be inadequate. Grammar (i.e. the grammar-lexicon continuum) includes both prefabricated chunks and constructional schemas. Rather than being "minimal" and "economical", grammar is "maximal" and "nonreductive" (Langacker 2000: 1). It includes a wide variety of constructions that differ in terms of substance and familiarity. Highly abstract constructional schemas, low-level formulas, and prefabricated chunks coexist in the speaker's mental grammar. What is more, the coexistence is motivated because different types of constructions serve different functions.

Constructional schemas allow for the use of novel expressions; they account for the productivity of grammar, making language a flexible tool in novel situations (see below). However, the production of novel expressions is computationally costly. It involves a series of processing decisions that speakers have to make on-line under enormous time pressure (usually within milliseconds). For that reason, speakers tend to draw on prefabricated chunks and low-level formulas when they are available. Unlike novel expressions, these expressions have been computed so often that processing decisions occur with very little effort. In fact, highly routinized expressions are stored as holophrastic units whose internal structure is no longer computed. Thus, the use of prefabricated chunks and low-level formulas has certain advantages over the use of novel expressions: it reduces the amount of utterance planning and sentence processing so that the interlocutors can concentrate on other aspects of the communicative interaction. Spontaneous speech often abounds with formulaic expressions and semi-productive phrases that are organized around concrete lexical expressions. Highly schematic constructions are only activated if prefabricated chunks and low-level formulas are not available.

In general, while the redundant storage of information increases the memory load, it decreases the computational effort in planning and processing. The more information is stored in multiple places (i.e. at different levels of abstraction) the more likely it is that speakers can draw on prefabricated chunks and utterance formulas, minimizing the mental effort for utterance planning and comprehension. Thus, if we measure economy in terms of computational effort, rather than in terms of storage space, the nonreductive model of construction grammar appears to be more economical and efficient than most other grammatical frameworks after all.

2 The usage-based model

2.1 The emergence of linguistic structure

One of the central assumptions of generative grammar is that the basic principles of grammar are innate. Specifically, it is assumed that grammar can be divided into an innate “core” and the “periphery”. The core consists of universal principles and parameters that are part of our genetic endowment, whereas the periphery comprises those aspects of grammar that are not genetically specified.

Challenging the distinction between the core and the periphery, the usage-based approach posits that linguistic structure emerges from language use (cf. Langacker 1988, 2000; Bybee 1995; Elman et al. 1996). In this approach, grammar is seen as a dynamic system that is constantly changing by virtue of the psychological processes that are involved in language use. For instance, one of the central assumptions of the usage-based approach is that the representation of linguistic elements correlates with frequency of occurrence (e.g. Bybee 1985; Langacker 1988). Linguistic expressions and grammatical patterns that occur with high frequency in language use are more deeply entrenched in mental grammar than expressions that are infrequent. Every time a speaker uses a linguistic expression (or grammatical pattern), it reinforces its mental representation, which in turn facilitates the activation of this expression in language use. Thus, the use of linguistic expressions has an immediate effect on the representation and activation of linguistic knowledge.

What is more, language use can change the meaning and structure of linguistic elements and the organization of grammar. This has been amply demonstrated in the literature on grammaticalization (e.g. Heine, Claudi, and Hünnemeyer 1991; Hopper and Traugott 1993; Bybee, Perkins, and Pagliuca 1994; Lehmann 1995). Linguistic expressions are commonly divided into two general types: symbolic expressions and grammatical markers. Symbolic expressions subsume nouns, verbs and adjectives, while grammatical markers comprise elements such as prepositions, conjunctions and auxiliaries. The division between symbolic expressions and grammatical markers is based on two major criteria. First, symbolic expressions and grammatical markers serve different functions. Symbolic expressions denote referents, activities and other concepts, whereas grammatical markers are structural (or topographic) expressions that function to organize constructions. Second, symbolic expressions and grammatical markers differ in terms of class size. Symbolic expressions are open class (except for adjectives, which may be open or closed class; cf. Dixon 1982) while grammatical markers are always closed class items (cf. Talmy 1988).

Grammaticalization theory posits that all grammatical markers are eventually derived from symbolic expressions or from other grammatical markers that previously emerged from a symbolic source (cf. Hopper and Traugott 1993: 104). Frequency of occurrence plays a key role in this process. Linguistic expressions that are frequently used tend to reduce in structure and meaning (cf. Bybee 1985, 2001; Gregory et al. 1999; Juresfky et al. 2001). This may lead to the development of new grammatical markers. To mention two well-known examples, the future marker *gonna* developed from the expression *BE going to INF,* and the conjunction *because* emerged from the adpositional phrase *by cause.* Both *BE going to INF* and *by cause* were frequent collocations before they turned into a future auxiliary and a subordinate conjunction. Other grammatical markers that evolved from frequently used symbolic expressions are prepositions such as *in front of* and *inside,* conjunctions such as *in case* and *while,* modals such as *hafta* (i.e. *have to*) and *gotta* (i.e. *got to*), and bound derivational morphemes such as *–hood* and *–ly* (*–hood* evolved from a noun meaning ‘person’, ‘sex’, ‘quality’, and *–ly* developed from a noun meaning ‘appearance’, ‘form’, ‘body’; OED).

Apart from symbolic expressions (i.e. nouns, verbs and adjectives), demonstratives such as English *this* and *that* and *here* and *there* provide a frequent historical source for the development of grammatical markers (cf. Diessel 1999a: chap 6; 1999b). Since demonstratives are closed class items they are commonly analyzed as grammatical markers that developed from a symbolic source (i.e. from nouns, verbs and adjectives); but as Diessel (19991, 2003a) has pointed out there is no evidence from any language that demonstratives are historically related to symbolic expressions or any other expressions for that matter that do not include a genuine demonstrative. It seems that demonstratives constitute a special class of linguistic expressions that developed very early in the evolution of language.

Demonstratives are commonly used to focus the hearer’s attention on entities in the surrounding situation or elements in the ongoing discourse. In the latter use, they serve an important language internal function; specifically, they are used to track discourse participants and to establish links between chunks of the ongoing discourse. Based on such discourse-related uses, demonstratives frequently develop into grammatical markers. Across languages, demonstratives are commonly reanalysed as definite articles, third person pronouns, relative pronouns, complementizers, sentence connectives, focus markers, copulas, and many other grammatical morphemes (cf. Diessel 1997a, 1999a, 1999b). Like symbolic expressions, demonstratives only grammaticalize if they are routinely used in a specific grammatical construction (cf. Diessel 1999b: chap 6).

Another grammatical phenomenon that is crucially affected by the psychological processes involved in language use is word order. It has been repeatedly argued in the literature that the ordering of linguistic elements is shaped by processing. Specifically, it has been claimed that linguistic elements tend to be arranged in such orders that they are easy to process and easy to produce (cf. Dryer 1992; Hawkins 1994). This concerns both free word order configurations in discourse and fixed grammatical word orders. Linguistic typologists have shown that languages tend to be either consistently left-branching (cf. Figure 1a) or consistently right-branching (cf. Figure 1b) rather than mixed left- and right-branching (cf. Figure 1c.) (cf. Dryer 1992).

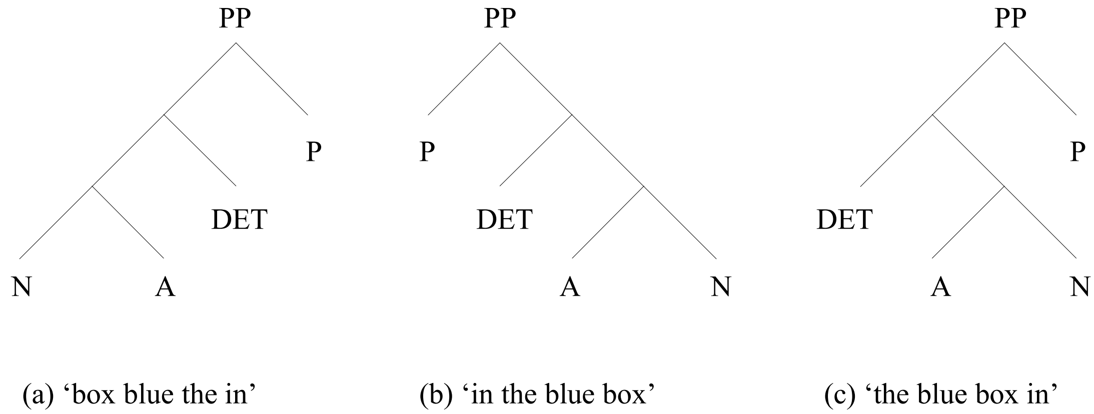

Figure 1. Left, right and mixed branching phrase structures

In generative grammar (notably in the Principles and Parameters framework), the branching directions are assumed to be innate. Languages are either consistently left-branching or consistently right-branching because these are the two options (i.e. parameter values) provided by universal grammar (cf. Frazier 1985; Frazier and Rayner 1988). In the usage-based approach, on the other hand, languages are assumed to employ consistent branching directions because such structures carry a lower processing load than structures with inconsistent left- and right-branching (cf. Dryer 1992; Hawkins 1994). In this view, the consistent branching directions are not innate; rather, they emerge in the historical development of grammar, which is driven by the psychological processes involved in language use. The branching directions can be seen as grammaticalized parsing principles that facilitate the interlocutors' computation of linguistic structures in language use (cf. Hawkins 1990, 1994, 1998).

The psychological mechanism that underlies grammaticalization is "habituation" (Haiman 1994, 1998). Habituation is a general psychological process that does not only affect the use of language but also many other activities such as music and sports. It basically describes the process by which the parts of a complex activity are merged such that the boundaries between the parts are no longer recognized. As a consequence, the complex activity may lose its internal structure and some of its substance, which in turn may lead to the "emancipation" (i.e. separation) of the restructured activity from its historical source (cf. Haiman 1994, 1998). This is exactly what we find in grammaticalization: linguistic expressions that grammaticalize undergo phonological and semantic changes such that they often lose the connection to their historical source, and free word orders that grammaticalize may become so rigid that grammar requires a specific word order regardless of the factors that motivated its development (cf. Hawkins 1994; Wasow 2002).

In sum, while the generative model posits the existence of innate grammatical principles and parameters, the usage-based model assumes that linguistic structure arises from language use. Grammar is shaped by usage—this is the most fundamental principle of the usage-based approach (cf. Bybee 2001; Bybee and Hopper 2001;

Langacker 1988; Hawkins 1994; see also Bresnan and Aissen 2002, who recently expressed a very similar view in the framework of Optimality Theory).

2.2 Network representations

Linguistic knowledge is commonly represented in an activation network in the usage-based model (cf. Bybee 1985, 1995, 2001; Langacker 1987a, 1991; Bates and MacWhinney 1987, 1989; Barlow and Kemmer 2000; Bybee and Hopper 2001). Network representations have a long tradition in cognitive science. In cognitive psychology and computer science connectionist network models are used to simulate cognitive processes (cf. Rumelhart and McClellend 1986a; Elman, Bates, Johnson, Karmiloff-Smith, Parisi, and Plunkett 1996). Two basic types of connectionist models are commonly distinguished: localist networks, which consist of interconnected symbolic units that are similar to symbolic representations in traditional non-connectionist models, and parallel distributed processing networks (i.e. PDP networks), which constitute a more radical departure from traditional models in cognitive science. Both models are self-organizing in that the processing of data can change the representation of conceptual content. However, PDP models are much more radical in this regard than localist models (cf. Elman et al. 1996: 90-97).

In a localist network, each concept is represented by a single node that cannot be decomposed into smaller elements. The specific properties of the nodes are hand-wired by the modeler; that is, the content of each node and its connections to other nodes are designed prior to the simulation (i.e. prior to the processing of data). An illustrative example of a localist network is given in Figure 2.

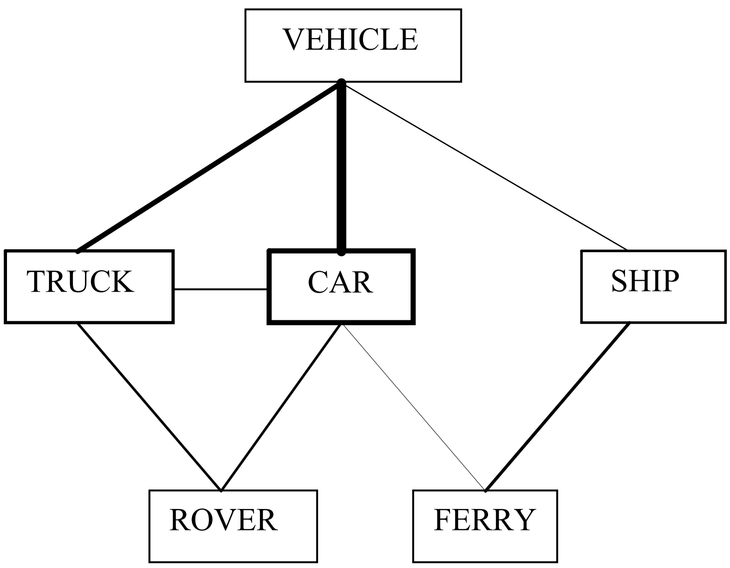

Figure 2 Localist network

In a distributed network, on the other hand, concepts emerge from processing data, i.e. they are not built into particular nodes. A PDP network consists of several layers of nodes and their connections. Both the nodes and their connections have “weights”, or “activation values”, that change in the course of the simulation.

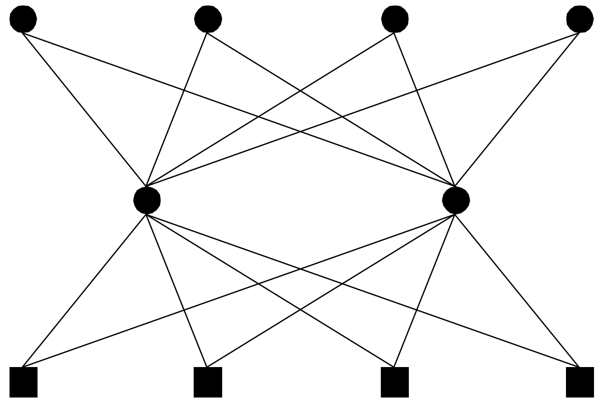

Figure 3. PDP network

The network in Figure 3 has three layers of nodes: the input nodes (represented by the four squares at the bottom), the output nodes (represented by the four circles at the top), and the hidden nodes (represented by the two circles that are connected to both the input nodes and the output nodes). Minimally, a PDP model has two layers of nodes, the input nodes and the output nodes, but most current PDP models have at least one extra layer of hidden nodes, making them more powerful than two-layer networks, which were often used in earlier PDP models (cf. Rumelhart and McClelland 1986a).[4]

Such networks can learn to match a given input pattern (e.g. the root of English verbs such as *walk, hit, sing*) to a particular output pattern (e.g. the past tense forms of these verbs, i.e. *walked, hit, sang*). During training, the activation values of the nodes and their connections change, based on a particular learning algorithm (cf. Rummelhart, Hinton, and McClelland 1986), such that a given input pattern fits the expected output pattern. At the end of training, the network has assumed a global activation pattern that allows the model to process new data in analogy to the input-output patterns that it has processed during training. The global activation pattern that emerges from processing the data can be interpreted as the representation of conceptual content (e.g. the English past tense schema) (cf. Elman et al. 1996: 90-91; see also Rumelhart, Smolensky, McClelland, and Hinton 1986).

PDP models are more flexible than localist representations. In a localist network, the simulation (i.e. the processing of data) can change the activation values of the nodes and their connections, but it cannot alter the content of the concepts. Each concept is represented by a specific node that is designed prior to the simulation. By contrast, the concepts of a PDP model emerge in the course of the simulation—they are immediately grounded in the data that is processed by the network. PDP models are therefore more powerful and more flexible than localist representations.

The network approach has been combined with construction grammar (cf. Langacker 1987a; Lakoff 1987; Goldberg 1995; Croft 2001). Although there are currently no connectionist models of a construction-based grammar (but see Morris,

[4] In addition to the hidden nodes, current PDP models often have a fourth layer of "context nodes", which can simulate the effect of short-term memory in on-line processing. PDP models having context nodes are called "recurrent networks" (Elman 1990).

Cottrell, and Elman 2000), most construction grammarians assume that grammatical knowledge is organized in an activation network (e.g. Langacker 2000; Croft 2001). An illustrative example of a construction-based network is given in Figure 2.4, adopted from Goldberg (1995: 109).

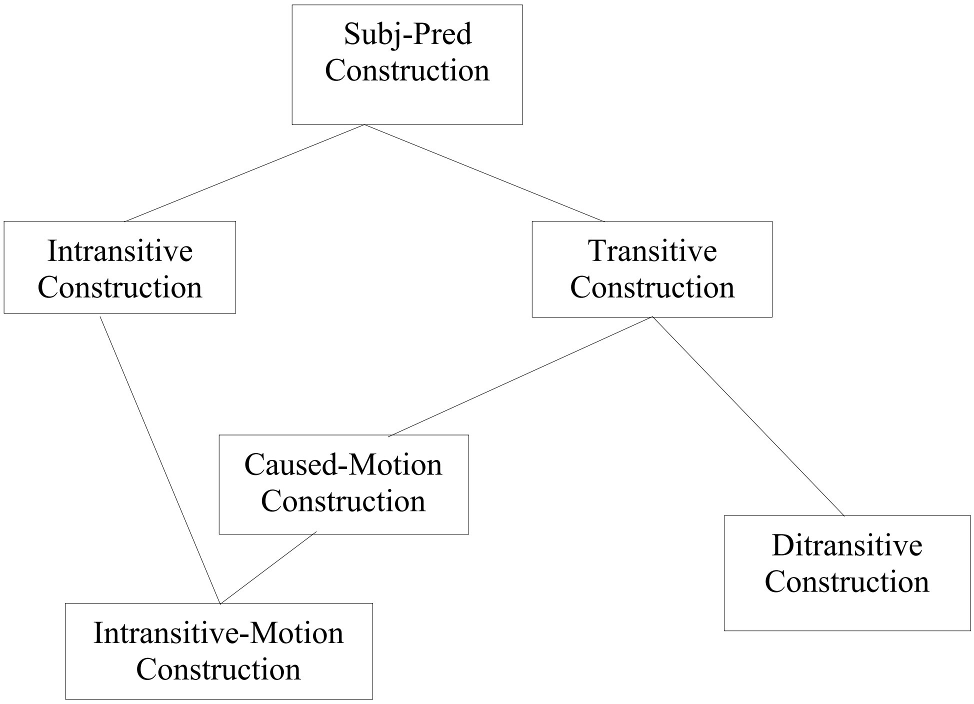

Figure 4. Construction-grammar network

The boxes (or nodes) represent particular constructions that are related to each other. Since each construction is represented by a single node, the network resembles a localist model; however, in principle it could be converted to a distributed representation in which each construction is represented by a global activation pattern. In fact, most functional linguists assume that constructions emerge from the psychological processes that are involved in language use (cf. Hopper 1987; Langacker 1988; Croft 2001; Bybee and Hopper 2001)

2.3 Entrenchment

Constructions have an activation value, which Langacker (1987a) calls the "level of entrenchment". Entrenchment is a psychological notion that corresponds to the "activation value" in a connectionist model. It is directly related to frequency: linguistic expressions that are frequently used are more deeply entrenched (i.e. more highly activated) in the speaker's network of grammatical knowledge than linguistic expressions that are infrequent.

Two types of frequencies must be distinguished: token frequency, which refers to the frequency of concrete expressions in the process of language use, and type frequency, which refers to the number of linguistic expressions that instantiate a constructional schema (cf. MacWhinney 1978; Bybee 1985; Plunkett and Marchman 1991). The two types of frequencies have different effects on the storage, activation, and processing of linguistic expressions.

Constructions that are defined upon the occurrence of concrete words (i.e. prefabricated utterances and lexically specific constructions) are highly entrenched in mental grammar if they occur with high token frequency in language use. Such frequently occurring constructions function as cognitive routines that can be directly accessed without activating a high-level schema. For instance, Bybee and Scheibman (1999) have argued that expressions such as *I don't know* __, *I don't think* __, and *Why don't you* __ have become cognitive routines due to the fact that they are extremely frequent. Although these expressions seem to abide by general grammatical principles, they are processed (both in production and comprehension) without a constructional schema: all three expressions constitute prefabricated chunks or collocations that are stored as holistic units. The independence of these expressions from a constructional schema is reflected in their particular forms and meanings. All three expressions are slightly different from parallel constructions that are less frequent (cf. Bybee and Scheibman 1999): the pronunciation of *don't* is reduced to [\´$], and although the clauses are formally negated they do not really function to negate a proposition: *I don't know* is either used to express the speaker's uncertainty or to indicate polite disagreement, *I don't think* expresses an epistemic stance toward an associated proposition, and *why don't you* marks a suggestion. Both the phonological reduction of *don't* and the particular meaning of these expressions suggest that they have started a life of their own; they have become cognitive routines, which Bybee and Scheibman characterize as "storage and processing units". This is the immediate effect of repetitive language use. In parallel constructions including a less frequent verb and a less common subject (e.g. *we don't eat* __), *don't* is commonly pronounced with an initial stop and a full vowel and serves as a negation marker. In general, what Bybee and Scheibman's analysis shows is that token frequency correlates with the level of entrenchment, which in turn has a significant effect on the storage and processing of lexically specific constructions.

While lexically specific constructions are highly entrenched in mental grammar if they occur with high *token* frequency, constructional schemas are argued to be highly entrenched if they occur with high *type* frequency (cf. Bybee 1985). In a construction-based framework, a type can be defined as a construction that instantiates a particular constructional schema. For instance, *NP pushed NP open* and *NP wiped NP clean* are instances (i.e. types) of the resultative schema (cf. Goldberg 1995: chap 8). The level of entrenchment of a constructional schema correlates with the number of types that are associated with a constructional schema. Other things being equal, schemas that are instantiated by a large number of types are likely to be more deeply entrenched (i.e. more strongly activated at rest levels) than schemas that are only related to a few types. However, since the activation value of individual types is based on their number of tokens, token frequency is also indirectly involved in the entrenchment of a constructional schema.

Interestingly, very high token frequency can weaken the connection of an expression (i.e. a type) to a constructional schema. As Bybee and Thompson (1997) have shown, linguistic expressions that occur with high token frequency are often resistant to diachronic change. For instance, in Middle English questions were constructed by fronting the tensed verb, and negative sentences were formed by placing *not* immediately after the verb. However, in the fourteenth century the patterns began to change: in both questions and negative sentences *do* appeared as a dummy auxiliary. The change affected all verbs except for *have, be, can, may, need, ought, know,* and a few others (cf. Kroch 1989). With the exception of *know,* all of these verbs still occur without *do* in Modern English. Interestingly, all of the verbs that were not affected by the change were extremely frequent at the time when the *do*-

construction emerged. Bybee and Thompson argue that these frequent verbs preserved the old pattern because they were so deeply entrenched in mental grammar that they were not attracted by the new question schema (cf. Tottie 1995). In other words, individual words and constructions may resist analogical change if they are so frequently used that speakers store them independently of a constructional schema.

In sum, the activation value of a constructional schema is determined by both the number of types that instantiate a schema and the number of tokens that determine the activation value of a specific type (for an insightful discussion of the effects of type and token frequencies on the representation of schemas in a PDP model see Plunkett and Marchman 1991).

2.4 Productivity

One of the central characteristics of human language is the productive use of grammatical patterns. In the usage-based approach, productivity can be defined as the likelihood that a constructional schema will be activated for constructing a novel expression (Langacker 2000: 26). Since there are often multiple schemas that are in principle available to construct (or interpret) a novel expression, the activation process usually involves the selection of a specific schema from a set of alternatives (cf. Bock 1977; McClelland and Elman 1986; Bates and MacWhinney 1987, 1989; MacWhinney 1987; Langacker 2000). The selection process is determined by competing factors; two major factors can be distinguished: (i) the level of entrenchment and (ii) the properties of the competing schemas (cf. Bybee 1995; Hare, Elman, and Daugherty 1995; Langacker 2000).

Constructional schemas that are highly entrenched in the speakers' network of grammatical knowledge are more likely to be selected for constructing a novel expression than schemas that are not well entrenched. This has been amply demonstrated in connectionist research. For instance, in their well-known work on the acquisition of the English past tense, Rumelhart and McClelland (1986b) showed that the productivity of the *V-ed* past tense results from the high activation value of this schema (which in turn is based on high type frequency). Specifically, they showed that the *V-ed* schema is the most productive past tense pattern functioning as the default because it licenses the use of a very large number of verb types, which outnumber the verb types of all other past tense schemas, i.e. irregular past tense schemas such as *drink-drank-drunk* (cf. Plunkett and Marchman 1991, 1993; see also Bybee and Slobin 1982 who analyzed the acquisition of the English past tense in a non-connectionist framework from a usage-based perspective).

However, productivity is not only determined by type frequency. If type frequency was the sole determinant of productivity, it would be impossible to account for so-called "low-frequency default patterns" (Hare, Elman, Daugherty 1995). A low-frequency default pattern is a morphological schema acting like the default despite the fact that it is based on low type frequency. A good example is the noun plural in Modern Arabic.[5]

Modern Arabic has several classes of irregular noun plurals, the so-called broken plurals, which outnumber the regular plurals by several times (in terms of types). However, when novel nouns are introduced to the language they tend to form the plural on the basis of the regular pattern unless they are phonologically similar to one of the irregular forms. The regular plural acts thus like a minority default schema

[5] Another frequently discussed minority pattern that is very productive, but probably not the default (cf. Behrens 2002), is the German *s*-plural (see Köpcke 1988, 1993, 1998; Clahsen, Rothweiler, Woest, and Marcus 1992; Bybee 1995; Clahsen 1999).

that is automatically selected if a novel noun does not fit one of the phonological templates that define the irregular forms (cf. McCarthy and Prince 1990).

The Arabic plural suggests that apart form type frequency the productivity of a morphological schema is determined by the phonological properties of the competing schemas. A schema that is defined by specific phonological features can only be selected if the target, i.e. the novel expression, matches these features. If such a narrowly defined schema competes for selection with an "open schema" (i.e. a schema that is defined by very general features; see Bybee 1995) it is very likely that the open schema will win the competition because the probability that the target will match the features of the open schema is much higher than the probability that it will match the specific features of a narrowly defined schema. Thus, in addition to type frequency, the phonological features of the competing schemas determine the productivity of a morphological pattern in the usage-based approach. Other things being equal (i.e. an equal number of types), an open schema is more likely to be selected for constructing a morphological expression than a narrowly defined schema that would be available to construct a parallel expression.

This explains why the regular Arabic noun plural can serve as a low-frequency default pattern. While the broken plurals are defined upon the presence of specific phonological features, the regular plural is basically an open schema, which is selected whenever the target does not match the phonological features of one of the irregular plurals (cf. Plunkett and Marchman 1991, 1993). The reason why the broken plurals are overall more frequent than the default pattern is simply that most existing Arabic nouns match one of the irregular plural schemas. It is thus the specific arrangement of open and narrow schemas that gives rise to a minority default pattern in morphology.

Although low-frequency defaults are relatively rare, they have been discussed extensively in the psycholinguistic literature because they show that productivity is not simply a function of type frequency (cf. Plunkett and Marchman 1991, 1993; Marcus, Ullman, Pinker, Hollander, Rosen, and Xu 1992; Clahsen, Rothweiler, Woest, and Marcus 1992; Prasada and Pinker 1993; Bybee 1995; Hare, Elman, and Daugherty 1995; Elman, Bates, Johnson, Karmiloff-Smith, and Plunkett 1996; Plunkett an Nakisa 1997). While this has never been claimed by the proponents of the usage-based model, some of the early connectionist studies employed two-layer networks in which the productivity of a morphological pattern was in fact a direct function of type frequency (e.g. Rumelhart and McClelland 1986b). However, this has changed in recent connectionist models. Using a three-layer network, Hare, Elman, and Daugherty (1995) showed that a PDP model can in principle account for the existence of a productive morphological pattern that is not backed up by high type frequency. In such a case, it is the overall structure of the similarity space, defined by the features of the competing morphological schemas, that gives rise to the productive use of a minority default pattern (a connectionist model simulating the acquisition of the Arabic noun plural is described in Plunkett and Nakisa 1997).

Like the productivity of morphological schemas, the productivity of syntactic schemas is determined by competing factors. In most situations, there are several syntactic schemas available to realize a specific speech act (or to interpret an utterance). For instance, polar questions can be realized by intonation or by "auxiliary fronting" (e.g. *It is raining? Is it raining?*), and many declarative sentences can be realized with different word orders (e.g. *He picked up the book — He picked the book up; He gave Peter the book — He gave the book to Peter*). In all of these cases, speakers (and hearers) have to select a specific syntactic schema to produce (or interpret) the utterance. The selection is determined by the level of entrenchment (i.e.

by type frequency) and various competing forces. For instance, Diessel (forthcoming) argues that complex sentences including a final adverbial clause constitute two different constructions that speakers select in different situations. The selection process is determined by competing forces from three domains: syntactic parsing, information processing, and semantics. Using Hawkins (1994) parsing theory, Diessel shows that complex sentences including final adverbial clauses are easier to process, and thus more highly preferred, than complex sentences including initial adverbial clauses; but nevertheless the latter are used in certain situations because of semantic and discourse-pragmatic considerations that favor the occurrence of initial adverbial clauses and may override the parsing preference for final occurrence. Similar explanations have been proposed for the productivity of different ordering patterns in verb-particle constructions (cf. Gries 2003a; Wasow 2002), ditransitive constructions (cf. Wasow 2002), genitive constructions (cf. Leech, Francis, and Xu 1994), and complex sentences including infinitival purpose clauses (cf. Thompson 1985).

3 Acquisition

Concluding this chapter, I discuss some of the major differences between the usage-based approach and the generative approach to language acquisition. I first summarize the major arguments of the debate about innateness and then discuss the different views about grammatical development.

3.1 The innateness hypothesis

According to generative grammar children are endowed with innate linguistic knowledge, which crucially determines the process of language acquisition. The initial state of the language faculty is called "universal grammar" or, from a different perspective, the "language acquisition device" (Chomsky 1999: 43). Universal grammar defines the class of possible languages that children are able to acquire. It consists of grammatical principles and parameters that provide a limited set of binary choices. Chomsky (1999: 49) characterizes the parameters as "switches" that are initially unset or set to a default value (see also Hyams 1986). Grammatical development is seen as a process whereby children determine the parameter values of their language based on specific triggers in the input.

The innateness hypothesis of generative grammar is based on arguments from psychology, neurology, and linguistics. One of the most frequently cited arguments supporting the innateness hypothesis comes from brain function studies (i.e. PET and fMRI studies). These studies have shown that different linguistic functions are located in different areas of the brain. The localization of language functions in specific brain areas is often taken as evidence for the innateness hypothesis (cf. Pinker 1994); however, as Elman et al. (1996: 378) have argued convincingly, "localization and innateness are not the same thing". While there seem to be specific brain areas that are involved in different language tasks, the specialization of these areas does not have to be innate; rather, local brain functions might emerge in the process of cognitive development. The brain is a self-organizing organ that develops local specializations as a consequence of processing a specific type of data. Strong support for this view comes from the fact that children with focal brain injuries often develop regional specializations for language in other areas of the brain than normal children (cf. Elman et al. 1996: chap 5).

Other arguments supporting the innateness hypothesis are based on studies examining SLI children. SLI, which stands for Specific Language Impairment, is usually defined as a cognitive deficit that only involves language, in particular grammatical morphology. Since SLI tends to run in families, some researchers suggested that it is based on a genetic defect affecting grammar (cf. Pinker 1994). However, other researchers are not convinced that SLI is really restricted to language, let alone to grammatical morphology. Challenging the definition of SLI as a specific language impairment, they have shown that SLI children have general difficulties in processing information that occurs in rapid temporal sequences and that SLI children also suffer from deficits in symbolic play and spatial imagery (Tallal, Ross, and Curtis 1989). This suggests that SLI is not caused by a genetic defect affecting only grammar or language (for a review of the literature see Elman et al. 1996: chap 7).

In addition to the arguments from brain function studies and SLI children, the innateness hypothesis is commonly supported by linguistic arguments. In particular, it has been claimed that the ambient language is not sufficient to learn grammar from experience alone. According to Chomsky (1972: 78), there is an enormous discrepancy between the grammatical system that constitutes the speaker's competence and the "meager and degenerated data" to which a child is exposed. Based on this assumption, Chomsky maintained that the gap between grammar and experience can only be closed if language acquisition is based on innate linguistic knowledge. This argument is known as "the argument from the poverty of the stimulus" (for a recent discussion of this argument see Crain and Pietroski 2001; see also the articles in the special issue of *The Linguistic Review* 2002).

Challenging this view, Pullum (1996) and Pullum and Scholz (2002) have recently argued that this argument is empirically unfounded. Examining four constructions that according to generative grammarians are so rare that their grammatical properties cannot be learned from linguistic experience (i.e. plurals in compounds, auxiliary sequences, anaphoric *one*, and auxiliary-initial clauses), they show that all four types of constructions are quite frequent in both written and spoken language. While this does not refute the innateness hypothesis, it raises considerable doubt about the validity of the argument from the poverty of the stimulus (see also the corpus-based analysis of child-directed speech by Brent and Cartwright (1996), Cartwright and Brent (1997), Redington, Chater, and Finch (1998), Redington and Chater (1998), and Mintz, Newport, and Bever (2002).

Moreover, a number of recent studies suggested that children's ability to determine linguistic patterns is much better than commonly assumed. For instance, Saffran, Aslin and Newport (1996) found that infants as young as 8 month of age are able to segment a complex string of nonsense syllables into word-like components based on their distribution. Similar results, emphasizing the role of distributional learning in early language acquisition, are reported in studies by Jusczyk (1997), Santelman and Jusczyk (1998), Marcus, Vijayan, Bandi Rao, and Vishton (1999), Höhle and Weissenborn (1999), and Saffran (2001).

Another linguistic argument that generative grammarians have used to buttress the innateness hypothesis might be called "the argument from the universality of grammatical features" (cf. Crain 1991). This argument is based on the assumption that all languages have certain grammatical properties in common. For instance, it has been argued that all languages employ the same grammatical categories such as nouns and verb (cf. Pinker 1984). If this is correct one might ask why these categories are universally attested. Generative grammarians explain the existence of universal categories in terms of innate universal grammar: grammatical categories are universal because they are innate. If they were not innate it would be a complete mystery, according to some generative grammarians, why they are universal (e.g. Crain 1991).

Outside of generative grammar, the existence of universal linguistic categories is highly controversial. Most typologists assume that crosslinguistic generalizations represent tendencies rather than absolute universals (cf. Dryer 1997a). If there are any linguistic categories that exist in all languages, their number is extremely limited. Nouns and verbs are perhaps the only grammatical categories that are truly universal, but even that is controversial (cf. Sasse 1993). However, even if we make the assumption that there are some absolute universals, they would not have to be innate. There are other, cognitive explanations for the existence of linguistic universals. For instance, nouns and verbs might be universal because all languages need these categories to denote two different types of concepts that are essential to human categorization (cf. Langacker 1987b; see also Hopper and Thompson 1984). In general, the usage-based approach assumes that linguistic universals are motivated by functional and cognitive pressures (cf. Givón 1995; Dryer 1997b: Croft 2001, 2003). These pressures increase the frequency of particular linguistic patterns so that they may grammaticalize. Since there are usually several pressures competing with each other, linguistic universals tend to be statistical rather than absolute. For instance, although processing (and/or utterance planning) seems to motivate the use of consistent left- and right-branching (see above), the branching directions of most languages are not entirely consistent. The inconsistency can be explained by the competition between processing and other factors affecting word order. There are, for instance, pragmatic word order principles that can be in conflict with syntactic parsing principles (cf. Diessel 2003b). In addition, it is well known that language contact can have a significant effect on word order. Since individual languages balance the competing pressures in different ways, the branching directions are similar but not identical across languages. Similar analyses have been proposed for many other linguistic universals (cf. Haiman 1983, 1985; DuBois 1985, 1987; Givón 1984, 1990, 1995; Dryer 1997b; Croft 2001, 2003).

In sum, all of the arguments supporting the innateness hypothesis are controversial. There is no compelling evidence that children are endowed with an innate universal grammar. Of course, language acquisition has certain biological prerequisites, but there is no evidence that these prerequisites involve innate linguistic knowledge. Rather, it is conceivable that language acquisition is based on general cognitive mechanisms that are also involved in the development of other cognitive domains.

3.2 Learning vs. growth

In the usage-based approach grammatical development is based on (inductive) learning. It involves general psychological mechanisms such as habituation, entrenchment, and analogy. Habituation involves the routinization or automatization of complex verbal (and nonverbal) activities, entrenchment concerns the strength of mental representations, and analogy acts as a mechanism for the derivation of new knowledge. All three mechanisms are affected by frequency of occurrence: linguistic patterns that are frequently processed become routinized and automatized, their level of entrenchment is strengthened in mental grammar, and they are often involved in analogical reasoning.

Learning is crucially distinct from parameter setting and other mechanisms that in generative grammar explain how children "hook up" their linguistic experience to innate universal grammar. In fact, Chomsky (1999: 43) argues that the notion of learning should be eliminated from the study of language acquisition.

> The term learning is, in fact, a very misleading one, and one that is perhaps best abandoned as a relic of an earlier age, and an earlier misunderstanding. (Chomsky 1999: 43)

Instead of learning, Chomsky (1999) uses the notion of "growth" to characterize the acquisition of grammar. Learning and growth are fundamentally distinguished. The remainder of this chapter discusses the most important differences between learning and Chomsky's notion of growth.[6]

3.2.1. The social-cognitive basis of grammatical development

First, learning and growth make very different assumptions about the social-cognitive foundations of language acquisition. According to Chomsky (1999: 41), grammatical development "is something that happens to the child". In this view, children acquire grammar in a quasi-automatic fashion: if they encounter the appropriate triggers in the input, grammar matures in the same way as the child's body or vision.

In the usage-based approach, grammatical development is considered an active process that crucially involves the use of language. In order to acquire language, including grammar, children have to be involved in social interactions (cf. Tomasello 1999, 2003; Clark 2003). According to Tomasello (1999), human infants are at first exclusively engaged in dyadic situations: they either manipulate objects or focus their intention on other people, who they do not seem to recognize as a person like themselves. At around nine to twelve months of age the situation changes: human infants begin to engage in triadic situations that involve the child, an object, and another person, who is now seen as an "intentional agent" (i.e. a person like the self). Triadic situations require a coordination of interaction with other people; this provides a crucial prerequisite for language acquisition: children are only able to learn the meaning and use of linguistic expressions because they encounter them in pragmatically meaningful situations. Language is essentially an instrument that children acquire in social interactions with other people.

Thus, while Chomsky characterizes grammatical development as a quasi-automatic process that happens to the child, the usage-based model emphasizes the significance of social interactions for the acquisition of grammar.

3.2.2. The role of the ambient language

Second, learning and growth differ with regard to the data that is needed for acquisition. Learning requires robust data: children will only be able to build up representations of grammatical patterns if they are frequently exposed to the relevant data. In other words, frequency of occurrence plays an important role in learning. By contrast, growth is basically independent of frequency: parameters can be fixed based on very little data: "The theory predicts that minimal exposure to data should be sufficient for parameter setting. Ideally, a single example encountered in the input could suffice" (Meisel 1994: 20).

3.2.3. The time course of grammatical development

Third, inductive learning is a gradual process, whereas growth is at least in principle instantaneous (cf. Meisel 1994: 14). As soon as the child is able to identify the

---

[6] It must be emphasized that Chomsky's notion of growth is not generally assumed in generative studies of language acquisition. Thus, the following discussion characterizes only one position in the generative approach.

elements that can act as triggers, a parameter can be set to a specific value. Assuming that most triggers are present in the input data, the theory predicts early and rapid acquisition (see especially Crain 1991 and Crain and Pietroski 2001). Of course, most generative grammarians acknowledge that grammatical development takes a certain amount of time, but this raises the question why triggers do not immediately fix a parameter when children encounter them. Borer and Wexler (1987) call this the "triggering problem". They argue that children initially are not sensitive to all triggers encountered in the data because universal grammar is not fully developed at birth; certain innately determined principles mature only later. Borer and Wexler call this the "maturation hypothesis" (see also Wexler 1999). Based on this hypothesis, they argue that the acquisition of grammar takes time because it follows a "biological program" that evolves only gradually during the early years of life.

Other generative linguists explain the triggering problem with the architecture of universal grammar (cf. Nishigauchi and Roeper 1987; Roeper and Weissenborn 1990; Roeper and de Villiers 1994; Weissenborn 1992). In their view, grammatical development takes time because parameters are interdependent such that a certain parameter can only be set to a specific value after the value of some other parameter has been determined. In this account, it is the arrangement of parameters in universal grammar that explains why parameters are not always immediately set to a specific value once a child encounters a particular trigger in the input.

Since the usage-based model assumes that language acquisition is based on learning, it is expected that grammatical development is gradual. In contrast to growth, learning requires repeated exposure to data over an extended period of time. From this perspective, the triggering problem is a pseudo-problem that arises from specific theoretical assumptions of generative grammar. In fact, the gradual development of grammar is seen as evidence for the usage-based hypothesis that language acquisition is based on learning.

3.2.4. The relationship between child grammar and adult grammar

Finally, learning and growth make different predictions about the nature of children's grammatical categories. In the generative approach, it is assumed that children have the same grammatical categories as adult speakers. Pinker (1984) called this the "continuity hypothesis". It is a logical consequence of the innateness hypothesis: children have adult-like categories because the categories they acquire are predetermined by innate universal grammar.

In the usage-based approach, it is assumed that children's grammatical representations are distinct from the grammatical categories of adult speakers (cf. Tomasello 2000). Children develop representations of grammatical categories by analyzing and systematizing the input data. The development is based on distributional analysis. Based on the distributional patterns that children detect in the ambient language, they construct abstract grammatical representations or schemas. The construction of schemas, which Langacker (2000) calls "schematization", is based on a specific type of analogy that involves the extraction of common features from the ambient language. The extracted features reinforce each other, giving rise to constructional schemas and other abstract representations of linguistic knowledge (see Chapter 8 for a more detailed discussion of this process). Since the extraction of common features from the data is a continuous process, one has to assume that the categories of early child grammar are constantly changing. As children attempt to organize the data, they gradually build up a network of interrelated constructions that successively become more complex and schematic. From this perspective, it is expected that children's grammatical categories are distinct from the categories of adult grammar.

References

Barlow, Michael and Susanne Kemmer (eds.) 2000, *Usage-based Models of Language*, Stanford: CSLI Publications.

Behrens, Heike 2002, Learning multiple regularities: Evidence from overgeneralization errors in the German plural, *Boston University Conference on Language Development* 26: 72-83.

Bencini, Guilia M.L. and Adele E. Goldberg 2000, The contribution of argument structure constructions to sentence meaning, *Journal of Memory and Language* 43: 640-651.

Boas, Hans 2003, *A Constructional Approach to Resultatives*, Stanford: CSLI Publications.

Bock, J. Kathryn 1977, The effect of a pragmatic presupposition on syntactic structure in question answering, *Journal of Verbal Learning and Verbal Behavior* 16: 723-734.

Borer, Hagit and Kenneth Wexler 1987, The maturation of syntax, in Thomas Roeper and Edwin Williams (eds.) *Parameter Setting,* 123-172, Dordrecht: Reidel.

Braine, Martin D.S. 1976, Children's first word combinations*, Monographs of the Society for Research in Child Development* 41.

Braine, Martin D.S. and Barbara Rumain 1981, Development and comprehension of *or*: Evidence for a sequence of competencies, *Journal of Experimental Child Psychology* 31: 46-70.

Brent, Michael R. and Timothy A. Cartwright 1996, Distributional regularity and phonotactic constraints are useful for segmentation, *Cognition* 61: 93-125.

Bresnan, Joan and Judith Aissen 2002, Optimality and functionality: Objections and refutations, *Natural Language and Linguistic Theory* 20: 81-95.

Bybee, Joan 1985, *Morphology*, Amsterdam: John Benjamins.

Bybee, Joan 1995, Regular morphology and the lexicon, *Language and Cognitive Processes* 10: 425-455.

Bybee, Joan 1998, The emergent lexicon, *Chicago Linguistic Society* 42: 1-435.

Bybee, Joan 2000, The phonology of the lexicon: Evidence from lexical diffusion, in Barlow and Kemmer (eds.), 65-85.

Bybee, Joan 2001, *Phonology and Language Use*, Cambridge: Cambridge University Press.

Bybee, Joan and Dan Slobin 1982, Rules and schemas in the development and use of the English past tense, *Language* 58: 265-289.

Bybee, Joan, Revere Perkins, and William Pagliuca 1994, *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World,* Chicago: The University of Chicago Press.

Bybee, Joan and Sandra Thompson 1997, Three frequency effects in syntax, *Berkeley Linguistics Society* 23: 378-388.

Bybee, Joan and Joanne Scheibman 1999, The effect of usage on degrees of constituency: The reduction of *don't* in English, *Linguistics* 37: 575-596.

Bybee, Joan and Paul Hopper (eds.) 2001, *Frequency and the Emergence of Linguistic Structure*, Amsterdam: John Benjamins.

Cartwright, Timothy A. and Michael Brent 1997, Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis, *Cognition* 63: 121-170.

Chomsky, Carol S. 1969, *The Acquisition of Syntax in Children from 5 to 10,* Cambridge, Mass: MIT Press.

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

Chomsky, Noam 1981, *Lectures on Government and Binding,* Dordrecht: Foris.

Chomsky, Noam 1986, *Barriers,* Cambridge: MIT Press.

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

Chomsky, Noam 2000, *New Horizons in the Study of Language and Mind,* Cambridge: Cambridge University Press.

Clahsen, Harald 1999, Lexical entries and rules of language: A multidisciplinary study of German inflection, *Behavioral and Brain Sciences* 22: 991-1060.

Clahsen, Harald, Monika Rothweiler, Andreas Woest, and Gary F. Marcus 1992, Regular and irregular inflection in the acquisition of German noun plurals, *Cognition* 45: 225-255.

Clark, Eve V. 2003, *First Language Acquisition,* Cambridge: Cambridge University Press.

Crain, Stephen 1991, Language acquisition in the absence of experience, *Behavioral and Brain Sciences* 14: 597-650.

Crain, Stephen and Paul Pietroski 2001, Nature, nurture and universal grammar, *Linguistics and Philosophy* 24: 139-186.

Croft, William 1995, Autonomy and functionalist linguistics, *Language* 71: 490-532.

Croft, William 2001, *Radical Construction Grammar,* Oxford: Oxford University Press.

Croft, William 2003, *Typology and Universals,* Cambridge: Cambridge University Press, [second edition].

Croft, William and Alan Cruise, Forthcoming, *Cognitive Linguistics,* Cambridge: Cambridge University Press.

Culicover, Peter W. 1998, *Syntactic Nuts. Hard Cases, Syntactic Theory, and Language Acquisition,* Oxford: Oxford University Press.

Dabrowska, Eva 2000, From formula to schema: The acquisition of English questions, *Cognitive Linguistics* 11: 83-102.

Davidse, Kristin 2000, A constructional approach to clefts, *Linguistics* 38: 1101-1131.

Diessel, Holger 1996, Processing factors of pre- and postposed adverbial clauses, *Berkeley Linguistics Society* 22: 71-82.

Diessel, Holger 1997a, The diachronic reanalysis of demonstratives in crosslinguistic perspective, *Chicago Linguistic Society* 33: 83-98.

Diessel, Holger 1997b, Verb-first constructions in German, in Marjolijn Verspoor, Lee Kee Dong, and Eve Sweetser (eds.), *Lexical and Syntactical Constructions and the Construction of Meaning*, 51-68, Amsterdam: John Benjamins.

Diessel, Holger 1999a, The morphosyntax of demonstratives in synchrony and diachrony, *Linguistic Typology* 3: 1-49.

Diessel, Holger 1999b, *Demonstratives. Form, Function, and Grammaticalization,* Amsterdam: John Benjamins.

Diessel, Holger 2001, The ordering distribution of main and adverbial clauses: A typological study, *Language* 77: 343-365.

Diessel, Holger 2003a, The relationship between demonstratives and interrogatives, *Studies in Language* 27: 581-602.

Diessel, Holger 2003b, Verb-first constructions in English, Leipzig: Max Planck Institute for Evolutionary Anthropology.

Diessel, Holger to appear, Competing motivations for the ordering of main and adverbial clauses, *Linguistics*.

Diessel, Holger and Michael Tomasello 1999, Why complement clauses do not include a *that*-complementizer in early child language, *Berkeley Linguistics Society* 25: 86-97.

Diessel, Holger and Michael Tomasello 2000, The development of relative clauses in English, *Cognitive Linguistics* 11: 131-151.

Diessel, Holger and Michael Tomasello 2001, The acquisition of finite complement clauses in English: A corpus-based analysis, *Cognitive Linguistics* 12: 1-45.

Diessel, Holger and Michael Tomasello 2003, The acquisition of relative clauses in English and German, Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dixon, R.M.W. 1982, *Where Have all the Adjectives Gone,* Berlin: Mouton.

Dryer, Matthew 1980, The positional tendencies of sentential noun phrases in universal grammar, *The Canadian Journal of Linguistics* 25: 123-195.

Dryer, Matthew 1992, The Greenbergian word order correlations, *Language* 68: 81-138.

Dryer, Matthew 1997a, Why statistical universals are better than absolute universals, *Chicago Linguistic Society* 33: 123-145.

Dryer, Matthew 1997b, Are grammatical relations universal?, in Joan Bybee, John Haiman, and Sandra A. Thompson (eds.), *Essays on Language Function and Language Type. Dedicated to Talmy Givón,* 115-143, Amsterdam: John Benjamins.

DuBois, John A. 1985, Competing motivations, in John Haiman (ed.), *Iconicity in Syntax,* 343-366, Amsterdam: John Benjamins.

DuBois, John A. 1987, The discourse basis of ergativity, *Language* 64: 805-855.

Elman, Jeffrey L. 1990, Finding structure in time, *Cognitive Science* 14: 179-211.

Elman, Jeffrey L. 1993, Learning and development in neural networks: The importance to start small, *Cognition* 48: 71-99.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domencio Parisi, and Kim Plunkett 1996, *Rethinking Innateness. A Connectionist Perspective on Development,* Cambridge: MIT Press.

Fillmore, Charles J. 1988, Mechanisms of Construction Grammar, *Berkeley Linguistic Society* 14: 35-55.

Fillmore, Charles J, Paul Kay, and Catherine O'Connor 1988, Regularity and idiomaticity in grammatical constructions: The case of *let alone, Language* 64: 501-538.

Fillmore, Charles J. and Paul Kay 1993, *Construction Grammar,* Berkeley: University of California.

Flickinger, Daniel, Carl Pollard, and Thomas Wasaw 1985, Structure sharing in lexical representations, *Association of Computational Linguistics* 23: 262-267.

Foley, William A. and Robert D. Van Valin 1984, *Functional Syntax and Universal Grammar,* Cambridge: Cambridge University Press.

Frazier, Lyn 1985, Syntactic complexity, in David R. Dowty, Lauri Karttunen, and Arnold Zwicky (eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives,* 129-189, Cambridge: Cambridge University Press.

Frazier, Lyn and Keith Rayner 1988, Parameterizing the language processing system: Left versus right-branching within and across languages, in John Hawkins (ed.), *Explaining Linguistic Universals,* 247-279, Cambridge: Cambridge University Press.

Givón, Talmy 1979, *On Understanding Grammar,* New York: Academic Press.

Givón, Talmy 1984, *Syntax: A Functional-typological Introduction*, Vol. I, Amsterdam: John Benjamins.

Givón, Talmy 1990, *Syntax: A Functional-typological Introduction*, Vol. II, Amsterdam: John Benjamins.

Givón, Talmy 1995, *Functionalism and Grammar*, Amsterdam: John Benjamins.

Goldberg, Adele E. 1995, *A Construction Grammar Approach to Argument Structure*, Chicago: The University of Chicago Press.

Gregory, Michelle, William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurefsky 1999, The effects of collocational strength and contextual predictability in lexical production, *Chicago Linguistic Society* 35: 151-166.

Gries, Stefan T. 2003, *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*, London: Continuum.

Haiman, John 1983, Iconic and economic motivation, *Language* 59: 781-819.

Haiman, John 1985, *Natural Syntax: Iconicity and Erosion,* Cambridge: Cambridge University Press.

Haiman, John 1994, Ritualization and the development of language, in William Pagliuca (ed.), *Perspectives on Grammaticalization*, 3-28, Amsterdam: John Benjamins.

Haiman, John 1998, *Talk is Cheap: Sarcasm, Alienation, and the Evolution of Language*, New York: Oxford University Press.

Halliday, Michael A.K. 1985, *An Introduction to Functional Grammar*, Baltimore: University Park Press.

Hare, Mary, Jeff L. Elman, and Kim G. Daugherty 1995, Default generalization in connectionist networks, *Language and Cognitive Processes* 10: 601-630.

Hare, Mary and Adele E. Goldberg 2000, Structural priming: Purely syntactic?, *Proceedings of the Twenty-first Annual Meeting of the Cognitive Science Society*, 208-211, Mahwah: Lawrence Erlbaum.

Hawkins, John A. 1990, A parsing theory of word order universals, *Linguistic Inquiry* 21: 223-261.

Hawkins, John A. 1994, *A Performance Theory of Order and Constituency*, Cambridge: Cambridge University Press.

Hawkins, John A. 1998, Some issues in performance theory of word order, in Anna Siewierska (eds.), *Constituent Order in the Languages of Europe*, 729-780, Berlin: Mouton de Gruyter.

Hawkins, John A. 1999, Processing complexity and filler-gap dependencies across grammars, *Language* 75: 244-285.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer 1993, *Grammaticalization, a conceptual framework*, Chicago: Chicaogo University Press.

Höhle, Barbara and Jürgen Weissenborn 1999, The origins of syntactic knowledge: Recognition of determiners in one-year-old German children, *Boston University Conference on Language Development* 24: 418-429.

Hopper, Paul J. 1987, Emergent Grammar, *Berkeley Linguistics Society* 13: 139-157.

Hopper, Paul J. and Sandra A. Thompson 1980, Transitivity in grammar and discourse, *Language* 56: 251-299.

Hopper, Paul and Elizabeth Closs Traugott 1993, *Grammaticalization*, Cambridge: Cambridge University Press.

Hyams, Nina 1986, *Language Acquisition and Theory of Parameters,* Dordrecht: Reidel.

Israel, Michael, Christopher Johnson, and Patrica J. Brooks 2000, From states to events: The acquisition of English passive participles, *Cognitive Linguistics* 11: 103-129.

Jackendoff, Ray 1975, Morphological and semantic regularities in the lexicon, *Language* 51: 639-671.

Jackendoff, Ray 1990, *Semantic Structures*, Cambridge: Cambridge University Press.

Jackendoff, Ray 1997, Twistin' the night away, *Language* 73: 534-559.

Johnson, Christopher 1999, Metaphor vs. conflation in the acquisition of polysemy: The case of *see,* in Masako K. Hiraga, Chris Sinha, and Sherman Wilcox (eds.), *Cultural, Typological, and Psychological Perspectives in Cognitive Linguistics,* 155-169, Amsterdam: John Benjamins.

Jurefsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond 2001, Probabilistic relations between words: Evidence from reduction in lexical production, in Bybee and Hopper (eds.), 229-253.

Jusczyk, Peter W. 1997, *The Discovery of Spoken Language*, Cambridge: MIT Press.

Kaschak, Michael P. and Arthur M. Glenberg 2000, Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension, *Journal of Memory and Language* 43: 508-529.

Kay, Paul and Charles J. Fillmore 1999, Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction, *Language* 75: 1-33.

Köpcke, Klaus-Michael 1988, Schemas in German plural formation, *Lingua* 74: 303-335.

Köpcke, Klaus-Michael 1993, *Schemata und Pluralbildung im Deutschen. Versuch einer kognitiven Morphologie*, Tübingen: Narr.

Köpcke, Klaus-Michael 1998, The acquisition of plural marking in English and German revisited: Schemata versus rules, *Journal of Child Language* 25: 293-319.

Kroch, Anthony S. 1989, Reflexes of grammar in patterns of language change, *Language Variation and Change* 1: 199-244.

Lakoff, George 1987, *Women, Fire, and Dangerous Things*, Chicago: Chicago University Press.

Langacker, Ronald W. 1987a. *Foundations of Cognitive Grammar,* Vol. I*, Theoretical Prerequisites*, Stanford: Stanford University Press.

Langacker, Rondald W. 1987b, Nouns and verbs, *Language* 63: 53-94.

Langacker, Ronald W. 1988, A usage-based model, in Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics,* 127-161, Amsterdam: John Benjamins.

Langacker, Ronald W. 1991, *Foundations of Cognitive Grammar*, Vol. II, *Descriptive Application*, Stanford: Stanford University Press.

Langacker, Ronald W. 2000, A dynamic usage-based model, in Barlow and Kemmer (eds.), 24-63.

Leech, Geoffrey, Brain Francis and Xunfang Xu 1994, The use of computer corpora in the textual demonstrability of gradience in linguistic categories, in C. Fuchs and B. Victorri (eds.), *Continuity in Linguistic Semantics*, 57-76, Amsterdam: John Benjamins.

Lehmann, Christian 1995, *Thoughts on Grammaticalization*, Munich: Lincom Europa.

Lieven, Elena V.M., Julian M. Pine, and Gillian Baldwin 1997, Lexically-based learning and early grammatical development, *Journal of Child Language* 24: 187-219.

Marcus, G. F. 2001, *The Algebraic Mind. Integrating Connectionism and Cognitive Science*, Cambridge: Bradford.

Marcus, G. F., M. Ullman, S. Pinker, M. Hollander, T.J. Rosen, and F. Xu 1992, Overgeneralization in language acquisition, *Monographs of the Society for research in Child Development* 57.

Marcus, G. F., S. Vijayan, S. Bandi Rao, and P.M. Vishton 1999, Rule learning in seven month-old infants, *Science* 283: 77-80.

McCarthy, John and Alan Prince 1990, Foot and word in prosodic morphology: The Arabic broken plural, *Natural Language and Linguistic Theory* 8: 209-283.

McClelland, James L. and Jeffrey L. Elman 1986, The TRACE model of speech perception, *Cognitive Psychology* 18: 1-86.

Meisel, Jürgen M. 1994, Parameters in acquisition, in Plaul Fletcher and Brain MacWhinney (eds.), *Handbook of Child Language,* 9-35, Oxford: Blackwell.

Michaelis, Laura and Knud Lambrecht 1996, Towards a construction-based theory of language function: The case of nominal extraposition, *Language* 72: 215-247.

Mintz, Toben H., Elissa L. Newport, and Thomas G. Bever 2002, The distributional structure of grammatical categories in the speech to young children, *Cognitive Science* 26: 393-424.

Morris William C., Garrison W. Cottrell, and Jeffrey Elman 2000, A connectionist simulation of the empirical acquisition of grammatical relations, in Stefan Wermter and Run Sun (eds.) *Hybrid Neural Symbolic Integration*, 175-193, New York: Springer-Verlag.

Nedjalkov, Vladimir (ed.) 1983, *Typology of Resultative Constructions*, Amsterdam: John Benjamins.

Newmeyer, Frederick J. 1998, *Language Form and Language Function,* Cambridge: MIT Press.

Newport, Elissa L. 1990, Maturation constraints on language learning, *Cognitive Science* 14: 11-28.

Nishigauchi, Taisuke and Thomas Roeper 1987, Deductive parameters and the growth of empty categories, in Thomas Roeper and Edwin Williams (eds.), *Parameter Setting*, 91-121, Boston: Reidel.

Nosofsky, R.M. 1988, Similarity, frequency and category representation, *Journal of Experimental Psychology: Learning, Memory and Cognition* 14: 54-65.

Nunberg, Geoffry, Ivan A. Sag, and Thomas Wasow 1994, Idioms, *Language* 70: 491-538.

Ono, Tsuyoshi and Sandra A. Thomposon 1995, What can conversations tell us about syntax, in Philip W. Davies (eds.), *Alternative Linguistics. Descriptive and Theoretical Models,* 213-271, Amsterdam: John Benjamins.

Pawley, Andrew and Hodgetts Syder 1983, Two puzzles for linguistic theory: Native like selection and native like fluency, in Jack C. Richards and Richard W. Schmidt (eds.), *Language and Communication,* 193-226, London: Longman.

Pinker, Steven 1984, *Language Learnability and Language Development*, Cambridge: Harvard University Press.

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

Pinker, Steven 1999, *Words and Rules, The ingredients of language,* New York: Basic Books.

Pinker, Steven and Alan Prince 1988, On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition, *Cognition* 28: 73-193.

Plunkett, Kim and Virginia Marchman 1991, U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition, *Cognition* 48: 43-102.

Plunkett, Kim and Virginia Marchman 1993, From rote learning to system building: Acquiring verb morphology in children and connectionist nets, *Cognition* 48: 21-69.

Plunkett, Kim and Ramin Charles Nakisa 1997, A connectionist model of the Arabic plural system, *Language and Cognitive Processes* 12: 807-836.

Pullum, Geoffrey K. 1996, Learnability, hyperlearning, and the poverty of the stimulus, *Berkeley Linguistic Society (parasession on the role of learnability of grammatical theory)* 22: 498-513.

Pullum, Geoffrey K. and Barbara C. Scholz 2002, Empirical assessment of stimulus poverty arguments, *The Linguistic Review* 19: 9-50.

Pullum, Geoffrey K. and Arnold Zwicky 1991, Condition duplication, paradigm homonymy, and transconstructional constraints, *Berkeley Linguistics Society* 17: 252-266.

Prasada, Sandeep and Steven Pinker 1993, Generalizations of regular and irregular morphological patterns, *Language and Cognitive Processes* 8: 1-56.

Prince, Ellen F. 1978, A comparison of WH-clefts and *it*-clefts in discourse, *Language* 54: 883-906.

Ramscar, Michael 2002, The role of meanng in inflection: Why the past tense does not require a rule, *Cognitive Psychology* 45: 45-94.

Redington, Martin and Nick Chater 1998, Connectionist and statistical approaches to language acquisition: A distributional analysis, *Language and Cognitive Processes* 13: 129-191.

Redington, Martin, Nick Chater, and Steven Finch 1998, Distributional information: A powerful cue for acquiring syntactic categories, *Cognitive Science* 22: 435-469.

Roeper, Thomas and Jürgen Weissenborn 1990, How to make parameters work, in Frazier and de Villiers (eds.), 147-162.

Roeper, Thomas and Jill G. de Villiers 1994, Lexical links in the WH-chain, in Barbara Lust, Gabriella Hermon, and Jaklin Kornfilt (eds.), *Syntactic Theory and First Language Acquisition: Cross-linguistic Perspectives*, Vol. II, *Binding, Dependencies, and Learnability,* 357-390, Hillsdale, NJ: Lawrence Erlbaum.

Rumelhart, David E. and James L. McClelland (eds.) 1986a, *Parallel Distributed Processing. Explanation in Microstructures of Cognition*, Vols I-II, Cambridge: MIT Press.

Rumelhart, David E. and James L. McClelland 1986b, On learning the past tense of English verbs, in Rumelhart and McClelland (eds.) Vol. II, 216-271.

Rumelhart, David E., Geoffrey Hinton, and James L. McClelland 1986, A general framework for parallel distributed processing, in Rumelhart and McClelland (eds.) Vol. I, 45-76.

Rumelhart, David E., Paul Smolensky, James L. McClelland, and Geoffrey Hinton 1986, Schemata and sequential thought processes in PDP models, in Rumelhart and McClelland (eds.), Vol. II, 7-57.

Saffran, Jenny R. 2001, Words in sea of sounds: The output of infant statistical learning, *Cognition* 81: 149-169.

Saffran, Jenney R., Richard N. Aslin, and Elissa L. Newport 1996, Statistical learning by 8-months-old infants, *Science* 274: 1926-1928.

Sag, Ivan A. 1997, English relative clauses, *Journal of Linguistics* 33: 431-483.

Santelmann, Lynn M. and Peter W. Jusczyk 1998, Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space, *Cognition* 69: 105-134.

Sasse, Hans Jürgen 1993, Syntactic categories and subcategories, in Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.), *Syntax*, 646-686, Berlin: Mouton de Gruyter.

Saussure, Ferdinand de 1916, *Cours de Linguistique Générale*, Paris: Payot.

Shopen, Timothy (ed.) 1985, *Language Typology and Syntactic Description,* Vols I-III, Cambridge: Cambridge University Press.

Slobin, Dan I. 1997, The origins of grammaticizable notions: Beyond the individual mind, in Dan I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition,* Vol. V*, Expanding the Context,* 265-323, London: Lawrence Erlbaum.

Tallal, P., R. Ross, and S. Curtiss 1989, Familial aggregation in specific language impairment, *Journal of Speech and Hearing Disorders* 54: 157-173.

Talmy, Leonard 1988, The relation between grammar and cognition, in B. Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, 165-205, Amsterdam: John Benjamins.

Talmy, Lenard 2000, *Toward a Cognitive Semantics,* Vol. I, *Concept Structuring Systems*, Cambridge: Cambridge University Press.

Thompson, Sandra A. and Paul Hopper 2001, Transitivity, clause structure, and argument structure: Evidence from conversation, in Bybee and Hopper (eds.), 27-60.

Tomasello, Michael 1999, *The Cultural Origins of Human Cognition*, Cambridge: Harvard University Press.

Tomasello, Michael 2000, Do young children have adult syntactic competence?, *Cognition* 74: 209-253.

Tomasello, Michael 2003, *Constructing a Language. A Usage-based Theory of Language Acquisition*, Cambridge: Harvard University Press.

Tottie, Gunnel 1995, Lexical diffusion in syntactic change: Frequency as a determinant of linguistic conservatism in the development of negation in English, in Dieter Kastovsky (ed.), *Historical English Syntax,* 439-467, Berlin: Mouton de Gruyter.

Wasow, Thomas 2002, *Postverbal Behavior,* Stanford: CSLI Publication.

Weissenborn, Jürgen 1992, Null subjects in early grammars: Implications for parameter-setting theory, in Jürgen Weissenborn, Helen Goodluck, and Thomas Roeper (eds.), *Theoretical Issues in Language Acquisition: Continuity and Change in Development,* 269-299, Hillsale, NJ: Lawrence Erlbaum.

Wexler, Kenneth 1999, Maturation and growth of grammar, in William C. Ritchie and Tej K. Bhatia (eds.) 1999, *Handbook of Child Language Acquisition,* , 55-109, Oxford: Blackwell.

Wierzbicka, Anna 1988, *The Semantics of Grammar,* Amsterdam: John Benjamins.

Wray, Alison and Michael R. Perkins 2000, The functions of formulaic language: An integral model, *Language and Communication* 20: 1-28.

Zwicky, Arnold 1987, Constructions in monostratal syntax, *Chicago Linguistic Society* 23: 389-401.

Zwicky, Arnold 1994, Dealing out meaning: Fundamentals of syntactic constructions, *Berkeley Linguistics Society* 20: 611-625.