Frequency and lexical specificity in grammar: A critical review

To appear in Heike Behrens and Stefan Pfänder (eds.),
Experience Counts: Frequency effects in language. Berlin: Mouton de Gruyter.

Holger Diessel[1]

Abstract

It is a central assumption of usage-based linguistics that syntactic structures, i.e. constructions, are stored and processed at different levels of abstractness and that lower-level constructions are associated with particular lexemes. There is evidence from a wide range of studies that the language user's knowledge of constructions includes detailed information about specific words, which influences sentence comprehension and production and diachronic language change. The current paper provides an overview of this research and considers the cognitive mechanisms that underlie the emergence of lexically particular constructions. Researchers agree that frequency affects the interaction between lexemes and constructions; but the cognitive effects of frequency on lexical specificity are not yet fully understood. Some usage-based linguists have argued that the exemplar mechanism of category learning accounts for the redundant storage of lexically particular constructions alongside more general, i.e. schematic syntactic representations; but the current paper argues that the exemplar model needs to be augmented by a linear processing mechanism, i.e. automatization, in order to explain how lexemes and constructions are related.

1 I would like to thank Heike Behrens and Karin Madlener for helpful comments and suggestions.

1. Introduction

Frequency has become a central topic of linguistic research (see Ellis 2002 and Diessel 2007 for reviews). There is a wealth of new results indicating that frequency influences language use and language development, both in history and acquisition, which in turn affects the representation of linguistic elements in memory (see Bybee and Hopper 2001 and Bybee 2007 for examples). This research has given rise to a new theoretical paradigm known as the "usage-based model" (Langacker 1988). In this new approach, language is seen as a probabilistic system of emergent structures and fluid constraints that are grounded in the language user's experience with concrete words and utterances (cf. Bybee 2006, 2010).

The usage-based model has challenged some central assumptions of grammatical research in structuralist and generative linguistics: It has questioned the common division between the language system and language use, or grammar and performance; it has abandoned the structuralist view that the study of synchronic linguistic states should be separated from the study of language change; and it has rejected the generative assumption that syntactic representations are independent of information about particular lexemes (see Diessel 2011a for discussion).

In structuralist and generative linguistics, syntactic structures are usually analyzed without any reference to specific words; but in the usage-based approach it is commonly assumed that syntactic structures, i.e. constructions, are stored and processed across a continuum of abstractness ranging from patterns that are tied to specific words to fully schematic structural representations (e.g. Langacker 1987; Goldberg 2006; Bybee 2010). There is abundant evidence from linguistic and psycholinguistic research that constructions are often associated with specific words or chunks of lexical expressions.

The (statistical) relationships between words and constructions (or lexemes and structure) are at the center of current research in corpus and computational linguistics (e.g. Manning and Schütze 2000; McEnery and Hardie 2011). In addition, this relationship has been investigated in research on first and second language acquisition (e.g. Ellis 1996; Tomasello 2000; see also Diessel 2004, 2013), sentence processing (e.g. MacDonald et al. 1994; Trueswell 1996), and language change (e.g. Tottie 1991; Bybee

2010). What all of these studies suggest is that the language user's knowledge about constructions includes a great deal of lexical information.

There is general consensus among usage-based linguists that frequency is an important determinant of the relationship between lexemes and constructions; but the cognitive processes that underlie the emergence of lexically particular constructions are not yet sufficiently understood. In the usage-based literature it is often assumed that exemplar learning accounts for the redundant storage of lexically particular constructions alongside more general, i.e. schematic syntactic representations (e.g. Abbot-Smith and Tomasello 2006; Bybee 2006, 2010; Goldberg 2006). The central idea behind exemplar theory is that categories are emergent from individual tokens of experience that are grouped together as exemplars, i.e. clusters of tokens with similar or identical features, which are then used to license the classification of novel tokens. Since emerging token clusters can be more or less complex, exemplar representations vary across a continuum of generality and abstractness (cf. Nosofsky 1988).

The exemplar model provides a theoretical foundation for usage-based analyses of grammar and language acquisition. However, in this paper I argue that an important aspect of language, notably grammar, is not adequately modeled in this approach: Language unfolds in time, i.e. all linguistic tokens occur in sequence, and sequentiality, or more specifically the processing of sequential information, is an important determinant for the emergence of lexically particular constructions. Since the exemplar model has not been developed for the analysis of sequential information, it is not fully adequate to explain how lexemes and constructions are related. It is the central hypothesis of the current paper that the "bare" exemplar model needs to be augmented by a processing mechanism, i.e. automatization, in order to account for lexical phenomena in grammar.

The paper is divided into two main sections. The first section provides an overview of research on lexical phenomena in grammar from different quarters of the language sciences, and the second section considers the empirical findings in light of the exemplar mechanism of category learning and the psychological notion of automatization.

2. Lexical specificity in grammar: A brief overview

There is evidence from various subfields of linguistics that syntactic structures are commonly associated with specific words that influence the language user's behavior in sentence comprehension and production, and the development of linguistic structure in acquisition and change. This section provides an overview of research in corpus linguistics, sentence processing, first language acquisition, and language change that has emphasized the importance of lexical expressions for the analysis of grammatical (or syntactic) phenomena.

2.1. Corpus linguistics

The most compelling evidence for lexical specificity in grammar comes perhaps from research on idioms and prefabs. Natural discourse abounds with idiomatic phrases and prefabricated strings of lexical expressions, which have long been ignored in grammatical research (cf. Bybee and Scheibmann 1999; Erman and Warren 2000; Nunberg et al. 1994; Pawley and Syder 1983; Sinclair 1991; Wray 2002). Some of these prefabs have idiosyncratic properties that are not predictable from their components. The meaning of *bite the bullet*, for instance, cannot be inferred from the meaning of *bite* and the meaning of *bullet*, but is associated with the entire expression. Other lexical prefabs have idiosyncratic structural properties. The prefab *all of a sudden*, for instance, includes a sequence of words that does not conform to general word order patterns (or word order "rules") of Present Day English (cf. Fillmore et al. 1988). Some prefabs are fully specified by lexical expressions (e.g. *How are you? Thank you, I'm fine. Get out of here*); others include open "slots" that can be filled by various expressions (e.g. *I wonder if __, a great deal of __, from __ onwards*) (see Erman and Warren 2000 for a typology of prefabs). Not all lexical prefabs, however, are semantically idiosyncratic or deviant from general grammatical patterns. A sentence such as *I am happy,* for instance, can also be seen as a prefab. The sentence is so frequently used in everyday language that it is likely to be stored and processed as a prefabricated string of words even though it does not have any obvious idiosyncratic properties. In general, every native speaker (of English) knows a very large number of lexical prefabs and collocations, which may or may not have idiosyncratic properties, and these prefabricated strings of lexical expressions constitute an important aspect of their linguistic knowledge.

Idioms and lexical prefabs provide strong evidence for the hypothesis that grammar includes lexically particular constructions. However, most research on grammar is concerned with more abstract and variable syntactic patterns. Of particular importance for grammatical analysis are verb-argument constructions, which can occur with a wide range of lexical expressions. However, verb-argument constructions are often statistically biased to include particular words in certain structural positions. For instance, the English ditransitive construction can occur with a large number of verbs, e.g. *give, show, bring, send,* and *teach,* to mention just a few. Most of these verbs can also occur in other, semantically related grammatical patterns, notably in the *to*-dative construction, in which the recipient is expressed by a prepositional phrase after the theme (cf. examples 1-2).

(1) Peter sent John a letter.
(2) Peter sent a letter to John.

However, there are strong statistical biases to use particular words in particular syntactic frames. The verbs *give* and *tell*, for instance, are significantly more frequent in the ditransitive construction than in the *to*-dative; whilst for the verbs *bring* and *take* it is the other way around (cf. Gries and Stefanowitsch 2004). What is more, the postverbal NPs of the ditransitive tend to be expressed by different nominals. It is well-known that the recipient role of the ditransitive is often encoded by a personal pronoun or proper name (cf. example 3) and that the theme role is usually expressed by a full nominal and only rarely by a personal pronoun (cf. example 4).

(3) Sue sent me/Peter a letter.
(4) Sue sent Peter it.

The lexical biases of the ditransitive construction are motivated by general semantic and pragmatic principles. There is a large body of research arguing that the occurrence of particular verbs (and nominals) in the ditransitive construction (and other grammatical patterns) can be explained by general semantic and/or pragmatic properties (e.g. Goldberg 1995; Levin 1993; Haspelmath 2012). However, irrespective of the fact that the speaker's choice of particular words is semantically and/or pragmatically motivated, there is evidence that the lexical biases of verb-argument constructions are also represented in memory. Speakers "know" that the ditransitive construction typically occurs with particular verbs (and particular nominals) because they have experienced this construction so frequently with certain words that the lexical patterns have become an integral part of their knowledge about the ditransitive.

One piece of evidence for this hypothesis comes from the fact that the co-occurrence patterns of lexemes and constructions are not entirely predictable from general semantic and/or pragmatic criteria. In the usage-based literature it is commonly assumed that verb-argument constructions "fuse" with nouns and verbs that are semantically compatible with their (constructional) meanings (cf. Goldberg 1995: 50). The ditransitive construction, for instance, denotes an act of transfer and is thus expected to fuse with transfer verbs. However, this assumption is not entirely borne out by the data: There are verbs such as *donate* (or *carry* and *push*) that are readily conceptualized as transfer verbs, but do not occur in the ditransitive construction (e.g. Wonnacott et al. 2008). In corpora, *donate* is exclusively found in the semantically related *to*-dative construction (cf. example 5) (cf. Gries and Stefanowitsch 2004), and most speakers find *donate* unacceptable in the ditransitive (cf. example 6), which is difficult to explain by semantic criteria.

(5) Peter donated money to the Red Cross.

(6) *Peter donated the Red Cross money.

The most plausible explanation why *donate* is excluded from the ditransitive construction seems to be that people have never experienced *donate* in this grammatical pattern. If this is correct it is consequent to assume that other lexical co-occurrence patterns of the ditransitive, which one could in principle derive from general semantic and/or pragmatic criteria, are also weakly conventionalized; i.e. they are part of what speakers know about verb-argument constructions (cf. Rosemeyer, this volume, for a related account).

Additional support for this hypothesis comes from corpus research on English varieties. In the syntactic literature, it is commonly assumed that the ditransitive construction cannot occur with a noun as recipient and a personal pronoun as theme (e.g. **She gave the woman it*). However, in some northern varieties of British English, notably in the dialect of Lancashire, this pattern is attested in corpora and accepted by native speakers of this variety (cf. example 7; adopted from Siewierska and Hollmann 2007).

(7) Show your father them.

What is more, it is not uncommon in the Lancashire dialect to change the order of recipient and theme. In standard English, the recipient precedes the theme in the ditransitive; but in this variety the reverse order is also quite common, especially when the theme is a pronoun and the recipient a noun (e.g. *She gave it the man*), but also when recipient and theme are both expressed by pronouns (e.g. *I'll give it you*). Thus far, however, this pattern has only been found with three verbs, namely with *give, send*, and *show*, indicating that the order theme-recipient is lexically particular, i.e. tied to particular verbs (see Siewierska and Hollmann 2007 for details).

Taken together, these data suggest that the lexical biases of the ditransitive (and other verb-argument constructions) are not entirely predictable from semantic and pragmatic criteria; rather they involve a certain degree of conventionalization. People "know" the lexical patterns of the ditransitive and other verb-argument constructions from their linguistic experience and this knowledge affects their behavior in sentence comprehension and production.

2.2. Sentence comprehension

A wealth of recent results indicate that sentence comprehension is crucially influenced by individual lexemes. In the generative approach, it is commonly assumed that online sentence processing is guided by structure-based heuristics such as "minimal attachment" and “late closure”, and that lexical and semantic features influence syntactic processing only at a later stage (cf. Frazier 1985). However, these assumptions have been challenged by a large body of research indicating that the language user’s experience with individual words has an immediate impact on sentence processing (cf. MacDonald et al. 1994; Trueswell et al. 1994). Let us consider a few examples that illustrate the influence of lexical expressions on sentence comprehension.

One structure that has played an important role in research on sentence processing is the reduced relative clause. The sentence in (8) is a frequently cited example from a seminal study by Bever (1970), which is very difficult to understand (out of context). With no further information given, most people misinterpret the verb *raced* initially as a simple past tense form and recognize only later that it serves as passive participle of a reduced relative clause.

(8) The horse raced past the barn fell.

Garden path sentences of this type have been interpreted as evidence for the importance of structural heuristics in sentence processing (cf. Frazier 1985); however, it is easy to show that the ambiguity of this structure hinges on the ambiguity of the verb *raced*. If *raced* is replaced by *kept* the sentence is much less ambiguous, if at all (cf. example 9).

(9) The horse kept past the barn fell.

This observation has given rise to the “lexical guidance hypothesis” of sentence processing, which holds that syntactic ambiguities are commonly resolved by information about particular words. Specifically, the hypothesis suggests that the language user’s knowledge of words includes statistical information about their syntactic distribution that can influence online sentence processing (cf. MacDonald and Seidenberg 2006).

For instance, given that reduced relative clauses have a passive meaning, it is a plausible hypothesis that these structures are easier to process with verbs that are frequently used in passive voice than with verbs that are primarily used in active voice. Trueswell (1996) tested this hypothesis in an experiment in which people had to read two different types of stimuli: (i) reduced relative clauses including verbs such as *select* that are frequently used in passive voice, and (ii) reduced relative clauses including verbs such as *search* that usually occur in active voice. In accordance with the lexical guidance hypothesis this study found that reduced relative clauses with frequent passive verbs (cf. example 10) cause fewer difficulties than reduced relative clauses with frequent active verbs (cf. example 11).

(10) The recipe selected by the judges did not deserve to win.

(11) The room searched by the police contained the missing weapon.

Another lexical aspect that affects the interpretation of reduced relative clauses is the transitivity of the verb (cf. MacDonald 1994). Many verbs can occur in both transitive and intransitive clauses, but are biased towards one of them. The verb *open*, for instance, is much more frequent in transitive than in intransitive clauses (e.g. *He open the door* vs. *The door opens*), whereas *walk* is primarily used in intransitive clauses and only rarely in transitives (e.g. *He walked the dog* vs. *He walked in the park*). MacDonald (1994) showed that reduced relatives are easier to understand with verbs that are primarily used in transitive clauses than with verbs that tend to occur in intransitive clauses. Since reduced relatives are transitive by definition, they force the listener to select the transitive sense of a verb, which can lead to comprehension difficulties if the verb is more frequent in intransitive clauses.

Both the frequency effect of active and passive uses and the frequency effect of transitive and intransitive uses are based on distributional patterns of individual verbs, supporting the hypothesis that sentence processing is "guided" by the language user's experience with particular words that are statistically biased to occur in certain structural positions.

Let us consider a second example to elaborate this point. A prepositional phrase at the end of a sentence is often ambiguous between an interpretation as attribute or verb modifier. Most people interpret the final PP in (12) as a modifier of the preceding NP (cf. 'The teacher pointed at the student who had a pen'); but it can also function as a modifier of the preceding verb (cf. 'The teacher pointed with the pen at the student').

(12) The teacher pointed at the student with the pen.

Interestingly, Ford et al. (1982) observed that the attachment site of a prepositional phrase is crucially influenced by the verb. Some verbs favor the interpretation as noun modifier, whereas other verbs favor the interpretation as verb modifier. For instance, the verbs *discuss* and *keep* give rise to very different interpretations: 90 percent of the participants in the Ford et al. study interpreted the PP in (13) as a noun modifier, i.e. they attached the final PP to the preceding NP, and only 10 percent selected the interpretation as verb modifier. By contrast, 95 percent of the participants interpreted the PP in (14) as an immediate constituent of the verb phrase, and only 5 percent adopted the interpretation as attribute.

(13) The woman discussed [the dogs [on the beach]].

(14) The woman [kept [the dogs] [on the beach]].

Since the test items were identical except for the verb, it must have been the verb that caused the different responses. But what distinguishes the two verbs *discuss* and *keep*? One factor that seems to be important is their meaning: *keep* has a common locational meaning that favors the interpretation of the final PP as a locational adverbial, which is less plausible in the case of *discuss*. However, while semantic factors are important, Ford et al. argue that the main factor influencing the comprehender's choice of a particular attachment site is frequency.

If we look at the frequency distribution of the two verbs we find that they tend to occur in different constructions (or different subcategorization frames): *discuss* occurs primarily with a single NP complement, whereas *keep* tends to occur with two immediate constituents. In corpus data, three out of four instances of *discuss* occur with a single NP complement, whereas four out of five instances of *keep* occur with both an NP complement and a locational PP (cf. Jurafsky 1996), supporting Ford et al.'s hypothesis that the language user's experience with distributional patterns of individual verbs plays an important role in syntactic ambiguity resolution.

Interestingly, there is evidence that syntactic processing is not just affected by the distribution of word forms (or lemmas), but also by the language user's experience with particular word senses (cf. Hare et al. 2003). As can be seen in examples (15) and (16), some transitive verbs can occur with both nominal complements and complement clauses.

(15) I believe the story.
(16) I believe the story is not true.

A number of studies have shown that the online interpretation of the postverbal NP varies with the subcategorization preference of the preceding verb. A verb such as *believe*, which is frequently used with a sentential complement, favors the interpretation as subject of a complement clause, whereas a verb such as *confirm*, which is primarily used with nominal complements, favors the interpretation as direct object (cf. Garnsey et al. 1997). However, many complement-taking verbs are polysemous and their subcategorization preferences vary with the verb sense. For instance, if the verb *find* denotes the successful end of a search, it usually occurs with an NP complement (cf. example 17); but if it indicates a psychological state, it usually takes an S complement (cf. example 18) (cf. Roland and Jurafsky 2002; Wiechmann 2008).

(17) Searching for hours, we eventually found the key in the garage.
(18) I found (that) this movie is pretty boring.

Hare et al. (2003) showed that the online interpretation of these sentences is influenced by the subcategorization preferences of individual verb senses. Specifically, they found that NP complements cause prolonged reading times with verbs that typically occur with S complements given a particular verb sense and vice versa. It seems that our knowledge about verbs includes very fine-grained statistical information about the syntactic distribution of different verb forms and different verb senses (MacDonald et al. 1994; Trueswell et al. 1994).

Taken together these findings indicate that syntactic processing is influenced by distributional information about specific word forms and word senses. While the lexical biases of constructions are always motivated by semantic and pragmatic criteria, they also reflect the language user's lifelong experience with particular words and constructions. Interestingly, when we look at the beginning of this experience in child language development, we find an almost deterministic connection between lexemes and constructions.

2.3. Language acquisition

The earliest utterances children produce consist of isolated words or chunks of unanalyzed expression (e.g. *Find-it; What-s-that*). However, after a few months, children begin to combine individual words to multi-word utterances. Interestingly, the first two- and three-word utterances children produce are often organized around specific lexical expressions; i.e. they are "mixed constructions" (Tomasello 2000) consisting of a relational term and an open slot that can be filled by other expressions (cf. Braine 1976; Bowerman 1976; Tomasello 1992; Lieven et al. 1997). Braine (1976) characterized the relational term as a "lexical pivot" and referred to the whole pattern as a "pivot schema" or "pivot construction" (cf. Table 1).

Table 1. Pivot schemas (examples adopted from Braine 1976: 7)

All done __	__ off	More __
All done milk	Boot off	More car
All done now	Light off	More cereal
All done juice	Pants off	More cookie
All done outside	Shirt off	More fish
All done pacifier	Water off	More hot

Each pivot schema is an isolated construction defined by the pivot word. It usually takes several months until children begin to use a particular pivot word across constructions. Tomasello and colleagues examined this development with experimental methods (see Tomasello 2000 for a review). In one of their experiments they taught 2-year-old children nonce verbs in two different constructions (cf. Brooks and Tomasello 1999). One group of children heard the nonce verbs in the transitive construction, the other group heard them in the passive construction. The two constructions were used with reference to the same scenes (e.g. showing a puppet pulling a wagon up a ramp). After training, the experimenter asked the children two different questions to elicit either an active sentence ('What is X doing?') or a passive sentence ('What is happening to X?') including one of the nonce verbs.

What Brooks and Tomasello found is that 2-year-old children are somewhat reluctant to extend the use of novel verbs from active sentences to passive sentences or vice versa. In fact, some children used the nonce verbs exclusively in the construction in which they had heard them during training. However, when Brooks and Tomasello repeated the experiment with 4-year-olds, all children used them in both constructions; that is, the older children were able to "detach" individual verbs from the syntactic contexts in which they had heard them and extended their use to novel constructions; but the younger children seemed to interpret the nonce verbs as an integral part of an item-specific construction in which lexemes and constructions are inextricably related. Parallel experiments with other constructions yielded similar results, supporting the hypothesis that children's early constructions are tied to particular pivot words (cf. Tomasello and Brooks 1998; see also Diessel 2013 for a review).

2.4. Language change

Finally, there is a large body of research indicating that syntactic change proceeds in an item-specific fashion (see Diessel 2011b, 2012 for some discussion of the parallels between language change and language acquisition). The best evidence for this comes perhaps from research on grammaticalization. Grammaticalization is an item-specific process whereby individual lexemes, or strings of lexemes, become reanalyzed as grammatical markers in the context of particular constructions. Early research on grammaticalization has concentrated on the lexical side of this process; however, soon it became clear that grammaticalization is not just a lexical phenomenon, but involves both lexemes and constructions (see Bybee 2010, Hilpert 2013, and Traugott and Trousedale 2013 for recent construction-based research in grammaticalization).

For instance, Krug (2000) argued that English has developed a new category of secondary modals that are embedded in a particular construction organized around a few verbs, namely the contracted forms of 'want to', 'have to', and 'going to', i.e. *wanna* /ˈwɒnə/, *gonna* /ˈgɒnə/, and *gotta* /ˈgɒtə/. Semantically, these verbs are similar to primary modals (e.g. *can, may, must*); but they have different syntactic properties. In contrast to primary modals, the secondary modals cannot be negated by postverbal *not* and do not occur at the beginning of polar questions (cf. Table 2).

Table 2. Primary and secondary modals in negative sentences and questions

	Primary modals	Secondary modals
Negation	can not may not must not	*wanna not *gonna not *gotta not
Questions	Can I help you? May I help you? Must I help you?	*Wanna you help me? *Gonna you help me? *Gotta you help me?

Note that the secondary modals originate from different source constructions: *want* is (originally) a main verb that can occur with both nominal and verbal complements (e.g. *want coffee, want to leave*), and *gonna* and *gotta* are based on periphrastic expressions consisting of two verbs and an infinitive (e.g. *is going to leave, has got to do*). Interestingly, although these verbs originate from different source constructions they have undergone parallel changes. All three verbs have merged with the infinitive marker *to* and have converged on the same phonetic form (i.e. they constitute a "product-oriented schema"; Bybee 1985). As Krug observed, the target forms consist of two CV syllables, carry stress on the first syllable, and end with a schwa. Interestingly, there are other, semantically related verbs that have developed similar contracted forms (e.g. *have to* /ˈhæftə/, *try to* /ˈtraɪtə/, *need to* /ˈnitə/), indicating that the emergence of secondary models involves a phonetic template (i.e. /ˈCVCə/) that can be seen as the structural core of a new modal construction organized around particular verbs.

A second piece of evidence for the lexically-specific nature of syntactic change comes from research on analogy. It is well-known that analogical extensions in morphology are lexically particular. Other things being equal, lexical expressions with high token frequency are less likely to be leveled by analogy than infrequent expressions. This explains why high token frequency correlates with morphological irregularity (cf. Bybee and Hopper 2001).

The same phenomenon occurs with syntactic patterns, i.e. constructions. For instance, a number of studies have shown that the development of the periphrastic *do*-construction in questions and negative sentences proceeded in a piecemeal, item-based manner (Tottie 1991; Bybee and Thompson 1997; Krug 2000). In Old and Middle English, negation was commonly expressed by a postverbal negative marker (e.g. *I have no answer*); however, this pattern was gradually replaced by the periphrastic *do*-construction (e.g. *I don't have an answer*). The development occurred with different verbs at different times, but did not affect the present day auxiliaries and modals, which are still used with a postverbal negative marker today (e.g. *is not, must not*). Apart from modals and auxiliaries, there are only a few other high-frequency verbs such as *know* and *make* that have preserved the old negation pattern in frequent collocations (e.g. *They know not what they do*) (e.g. Tottie 1991; Bybee 2010: 71).

Item-specific relics of an older grammatical pattern are very common. Another example is the English verb-first conditional construction consisting of a clause-initial auxiliary followed by the subject and the main verb (cf. example 19).

(19) Had she called me, I would have been here a lot sooner.

In Present Day English, the verb-first conditional construction is a lexically specific pattern that occurs with only three verbs in the subjunctive, i.e. *had, were*, and *should*. However, in earlier periods this construction was productive: it occurred with a wide range of verbs, similar to verb-first conditional constructions in modern German (cf. Diessel 1997). While the use of a subjunctive verb form is semantically motivated, there is no obvious semantic reason why the verb-first conditional construction of Present Day English occurs with only three verbs. The current restriction to *had, were*, and *should* is a lexical property of the conditional construction that cannot be explained by general semantic or pragmatic principles.

2.5. Interim summary

To summarize the discussions thus far, we have seen that adult language abounds with lexical prefabs and idiomatic expressions, and that the slots of more schematic constructions are often statistically biased to include particular words that influence sentence comprehension and production. Moreover, we have seen that in child language constructions are commonly organized around particular relational terms, and that in language history syntactic change proceeds in a piecemeal, item-based fashion similar to lexical diffusion in sound change.

3. Lexical specificity in grammar: Towards an explanation

In the remainder of this paper, I will look at the relationship between words and constructions from a more theoretical perspective. I begin with the exemplar model, which is often used to account for the redundant representation of constructions at different levels of abstractness, and then I consider the psychological notion of automatization and its effect on the formation and organization of lexically specific constructions.

3.1. The exemplar model

The exemplar model was developed as a general psychological theory of (non-linguistic) categorization and learning (cf. Nosofsky 1988; see also Murphy 2002: chap 3-4). Usage-based linguists often refer to the exemplar model in order to explain the emergence of linguistic categories and constructions from concrete words and utterances (cf. Abbot-Smith and Tomasello 2006; Bybee 2006, 2010; Goldberg 2006).

In the exemplar model every piece of information, i.e. every token, encountered in experience leaves a trace in memory. Over time, tokens with similar or identical features reinforce each other creating clusters of overlapping tokens known as exemplars (cf. Figure 1). The whole token cluster can be interpreted as an emergent category that functions as an “attractor”, i.e. a cognitive reference point for the classification of future tokens (cf. Nosofsky 1988, 1992; see also Murphy 2002: chap 4).

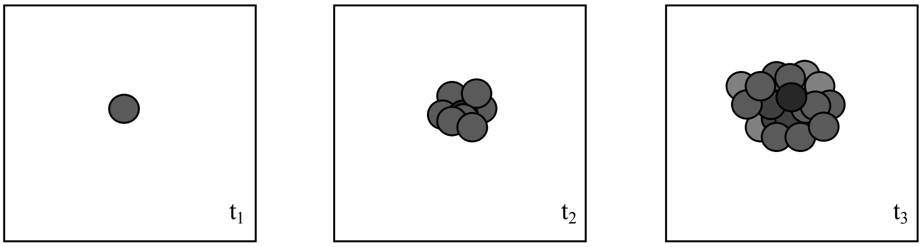

Figure 1. Emergence of a token cluster

In linguistics, the exemplar model has been especially influential in research on phonetics and phonology (cf. Johnson 1997; Bybee 2001; Pierrehumbert 2001). There are different exemplar accounts for the emergence and representation of phonological elements (see Weiser 2006 for a review); but they all assume that similar phonetic tokens in experience are grouped together as an exemplar category. On this view, a phoneme such as English /ɛ/ is emergent from many slightly different phonetic tokens of /ɛ/ that a language user encounters in experience.

The exemplar approach to phonetics and phonology has been extended to the analysis of other linguistic entities, notably to words and constructions (cf. Bybee 2006; Goldberg 2006). On this account, words and constructions are emergent from the language user’s experience with lexical tokens. Of course, words and constructions have both form and meaning, and not just form, as do phonemes; but the cognitive mechanism for the emergence of words and constructions is assumed to be the same as that for phonemes. As for words, Bybee (2010) describes this as follows:

> … each of the phonetic forms of a word that are distinguishable are established in memory as exemplars; new tokens of experience that are the same as some existing exemplars are mapped on to it, strengthening it. Then all the phonetic exemplars are grouped together in an exemplar cluster which is associated with the meanings of the word and the contexts in which it has been used, which themselves form an exemplar cluster … [Bybee 2010: 19]

Constructions exist at different levels of generality and abstractness (see above). Low-level constructions are lexically particular; but higher-level constructions, often referred to as "constructional schemas" (Langacker 2008), are defined over purely structural properties that are not immediately associated with particular lexemes. However, it is a central assumption of the usage-based approach that constructional schemas are emergent from recurrent strings of lexical expressions with similar or identical distributional and morphological properties. Evidence for this hypothesis comes from usage-based research on the acquisition of constructions (e.g. Tomasello 2003; Goldberg 2006) and psychological research on statistical grammar learning (e.g. Gómez and Gerken 1999; Marcus et al. 1999).

For instance, Goldberg (2006) argued that schematic verb-argument constructions are learned as generalizations over item-specific constructions organized around individual verbs that are accompanied by nominal expressions with similar semantic and syntactic properties. In accordance with this view, Gómez and Gerken (1999) showed that young children are able to "extract" syntactic categories (or schemas) from the (statistical) analysis of distributional regularities in strings of lexical expressions. In one of their studies, they exposed 12-month-old infants to an artificial language consisting of a set of monosyllabic nonce words (e.g. *VOT, PEL, JIC*) that appeared in different sentence types (or different constructions) generated by a finite-state grammar. The sentence types were distinguished by linear order and the number of words they include. Each word the children learned occurred in certain structural positions in one or more sentence types. After training, i.e. after the infants had listened to sentences from this language for a few minutes, they were tested under two conditions (cf. Table 3).

Table 3. Sample sentences generated by different grammars, adopted from Gómez and Gerken (1999: 114)

Set A	Set B
VOT PEL JIC	PEL TAM RUD RUD
PEL TAM PEL JIC	VOT JIC RUD TAM JIC
PEL TAM JIC RUD TAM RUD	VOT JIC RUD TAM RUD
REL TAM JIC RUD TAM JIC	VOT PEL JIC RUD TAM
VOT PEL PEL JIC RUD TAM	PEL TAM PEL PEL PEL JIC

In condition 1 (Set A), they were exposed to sentences generated from the same finite-state grammar as the ones used during training, but with different words; that is, each word the children had learned during training was replaced by a novel nonce word with the same distributional properties. And in condition 2 (Set B), the infants were exposed to sentences from a different finite-state grammar with the same novel nonce words as in condition 1. Using the head-turn preference procedure, Gómez and Gerken found that the infants of their study noticed the difference between the two grammars, i.e. they noticed that the sentences of the two grammars involve different word orders or sentence types, although they had not heard any of the words at test before, indicating that children as young as one year of age are able to abstract beyond specific words and to acquire abstract syntactic categories (or schemas).

Similar results were obtained in a study with seven-month-old infants by Marcus et al. (1999) and in a series of experiments with both child and adult language learners conducted by Newport, Aslin, Saffran, and colleagues (e.g. Thompson and Newport 2007; Wonnacott et al. 2008; Aslin and Newport 2012). What all of these studies suggest is that children (and adults) are very sensitive to transitional probabilities—a cognitive capacity that enables language learners to segment the speech stream, i.e. to identify boundaries between words and phrases (e.g. Saffran et al. 1996), and to recognize distributional regularities in strings of lexical expressions (see Aslin and Newport 2012 for a recent review).

In Section 2 we saw that constructions in early child language are often tied to particular relational terms and that children are initially reluctant to abstract away from these terms. In accordance with this view, research on statistical grammar learning has shown that children's generalizations can be constrained by the occurrence of specific phonetic forms. For instance, Gerken (2005) exposed nine-month-old infants to two different types of strings consisting of two identical syllables at the beginning of each string and a different syllable at the end (cf. Marcus et al 1999). One group of children was familiarized with strings that all ended in the same syllable, e.g. *-di* (condition 1), and the other group of children was familiarized with strings that ended in different syllables, e.g. *-di, -je, -li, -we* (condition 2) (cf. Table 4).

Table 4. Stimuli used in Gerken (2005: 68)

Condition 1	Condition 2
le-le-**di**	le-le-**di**
wi-wi-**di**	wi-wi-**je**
ji-ji-**di**	ji-ji-**li**
de-de-**di**	de-de-**we**

Although both sets of stimuli instantiate an AAB pattern, children performed differently under the two conditions. In condition 2, they generalized to an abstract AAB schema (cf. Marcus et al. 1999); but in condition 1 they restricted their generalization to an item-specific pattern including the syllable *di* (i.e. AA*di*), i.e. they did not generalize to structures without -*di*, supporting the hypothesis that children's early constructions are often tied to specific lexical material.

Taken together, these studies present strong evidence for exemplar learning, i.e. they support the hypothesis that schematic syntactic representations are grounded in the language user's experience with strings of lexical tokens, but without necessarily discarding these tokens after a schema has been established. On the contrary, there is evidence that the syntactic patterns children learn remain connected to (the strings of) lexical expressions from which they evolve.

One question that is controversial in this approach is whether schematic representations of syntactic structure are permanently stored in memory or whether they are only constructed online as particular tokens are processed (cf. Abbot-Smith and Tomasello 2006; Goldberg 2006; Ibbotson and Tomasello 2009). This controversy is related to an ongoing debate in cognitive psychology about the relationship between the exemplar model and prototype theory (see Rips et al. 2012 for a summary of this debate; see also Murphy 2002: chap 1-3).

A prototype is an abstract concept defined over weighted and connected features, which (some) cognitive psychologists have characterized as a schematic "summary representation" (Murphy 2002: 42). A number of studies have argued that categorization involves both clusters of concrete exemplars and schematic prototypes that are permanently stored in memory (cf. Ross and Makin 1999); but this assumption is controversial. There are computational models of exemplar learning in which generalizations across similar tokens are based on temporary concepts that are constructed online as the model seeks to classify new tokens but without creating an enduring schema (cf. Nosofsky and Zaki 2002). However, in contrast to these computational models most usage-based linguists assume a mixed model of categorization in which linguistic tokens, i.e. lexically particular constructions, are stored together with permanent schemas (cf. Abbot-Smith and Tomasello 2006; Goldberg 2006).

In sum, exemplar theory provides a simple but efficient learning mechanism that is consistent with the usage-based hypotheses that linguistic generalizations (e.g. constructional schemas) are emergent from the language user's experience with lexical tokens, i.e. words and utterances, and that linguistic structure is represented (or computed) at different levels of generality and abstractness. However, in what follows I argue that an important aspect of language, notably grammar, is not adequately modeled in this approach: Language is sequential and unfolds in time and this affects the way in which lexemes and constructions are related. Since the exemplar model has not been designed for the analysis of sequential information, the model is not fully appropriate to explain the interaction between lexemes and constructions. In the remainder of this paper, I argue that exemplar theory needs to be augmented by a processing mechanism for sequential information, i.e. automatization, in order to provide an adequate account for the emergence and organization of lexically specific constructions.

3.2. Automatization

Like exemplar learning, automatization is determined by (token) frequency. The psychological notion of automatization is related to Langacker's notion of entrenchment. Langacker does not refer to psychological research on automatization, but his definition of the term is consistent with this research. Here is what Langacker says about automatization and entrenchment:

> *Automatization* is the process observed in learning to tie a shoe or recite the alphabet: through repetition or rehearsal, a complex structure is thoroughly mastered, to the point that using it is virtually automatic and requires little conscious monitoring. In Cognitive Grammar parlance, a structure undergoes progressive *entrenchment* and eventually becomes established as a unit. [Langacker 2008: 16]

In the cognitive literature, the notion of entrenchment is often used in a broader sense than in this quote. Many cognitive linguists use the term entrenchment in the general sense of activation strengthening or lexical reinforcement; but this use of the term obscures the sequential nature of entrenchment (in Langacker's sense of the term).

The psychological notion of automatization is unequivocal in this regard. It is closely related to the distinction between 'automatic' and 'controlled processes', which was introduced at the end of the 19th century by the psychologist William James and has become a central topic of research on memory and attention in cognitive science (cf. Anderson 2005; Logan 1988; Schneider and Shiffrin 1977; Shiffrin and Schneider 1977).

The distinction between automatic and controlled processes concerns the way in which sequential information is learned and organized. Automatic processes occur without conscious control and effort; they are fast and can be performed in parallel to other tasks. Using a computer keyboard, for instance, is an automatic process that most people perform without monitoring their finger movements and parallel to other tasks (e.g. watching the computer screen). Controlled processes, by contrast, require attention and monitoring and cannot be so easily combined with other tasks. Entering new values into an electronic database, for instance, is a controlled activity that requires attention and conscious control (see Schneider and Chein 2003 for discussion).

Automatization is the cognitive mechanism whereby controlled processes are transformed into automatic processes (cf. Figure 2). Almost all sequential activities start off as controlled processes, but are then often transformed into automatic processes through repetition or practice. This is a very common cognitive phenomenon involved in many everyday tasks. Automatization enables people to perform complex sequential activities with little effort (cf. Logan 1988), but is also a common source for certain types of mistakes, i.e. slips, that occur for lack of attention or lack of conscious control (cf. Schneider and Chein 2003).

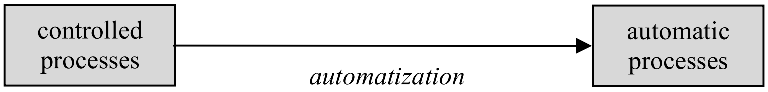

Figure 2. The transformation of controlled processes into automatic processthrough automatization

Language is a sequential medium that is crucially influenced by automatization. Linguists often look at language from a top-down perspective as if all linguistic elements of an utterance are simultaneously present; but language unfolds in time. All linguistic elements, e.g. phonemes, morphemes, words, phrases, and clauses, occur in sequence and are therefore subject to automatization. If we repeatedly process the same string of linguistic elements, e.g. the same sequence of speech sounds, words, or phrases, within a particular period of time, automatization creates associative links between them. The strengths of these links can be expressed in terms of transitional probabilities and other statistical measures that have been explored in corpus and computational linguistics (e.g. Manning and Schütze 2000: chap 5-6) and psychological research on statistical grammar learning (e.g. Saffran et al. 1996; see above).

Like exemplar categorization, automatization is driven by token frequency; i.e. both mechanisms are usage-based. However, in contrast to exemplar categorization and learning, automatization involves sequential information and creates processing units rather than non-sequential token clusters or exemplars (cf. Figure 3). This is an important difference between the two mechanisms, which is immediately relevant for the analysis of grammar in the usage-based approach (see below).

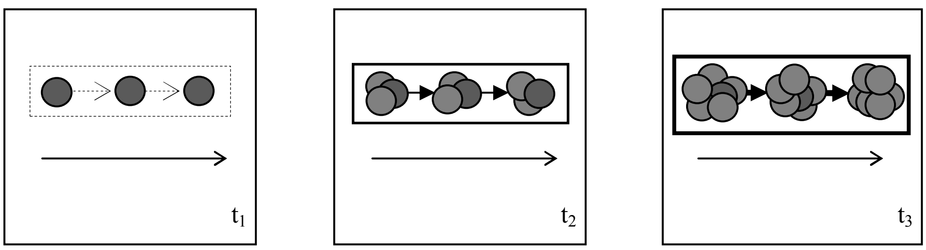

Figure 3. Emergence of linguistic unit

Automatization weakens the boundaries between sequentially organized elements and is the driving force behind univerbation and phonetic reduction. It underlies the formation of phrasal categories and the emergence of new clausal constructions (see Bybee 2010: chap 3, who uses the notion of "chunking" instead of "automatization").

Closely related to the notion of automatization is the notion of 'chunking', which the psychologist George Miller (1956) introduced in a seminal paper on memory (see Cowan 2005: chap 3 for a summary). According to Miller, the maximal number of items that can be held in working memory is seven, plus/minus two; but this number can be increased through chunking, which Miller defines as a cognitive process whereby individual items are grouped together to chunks in long-term memory. Related research on memory in chess by Chase and Simon (1973) demonstrated that chunks are organized in hierarchies, i.e. one chunk can be embedded into another one, so that an individual item in working memory is often linked to several related chunks in long-term memory, increasing the overall number of items that can occur in working memory (see Cowan 2005: chap 3).

The psychological notions of chunking and automatization complement each other; they concern the same cognitive process, but focus on different aspects of it: In the psychological literature, the term automatization is primarily used to characterize the way in which controlled processes are transformed into automatic processes through repetition or practice, and the term chunking is primarily used to characterize the way in which automatized sequences are stored and organized in memory (see also Newell 1990: chap 3-4). In the linguistic literature, the two terms are often used interchangeably and with little or no reference to their use in cognitive psychology (cf. Langacker 2008; Bybee 2010).

3.3. The emergence of processing units

One can think of automatization as a processing mechanism of online language use; but this mechanism has long-term effects for language development. The cognitive result of automatization in language development is the gradual emergence of a 'unit' (or 'chunk of memory'). In Cognitive Grammar, a 'unit' is a technical term that refers to the result of automatization:

> The term "unit" is employed in a technical sense to indicate a thoroughly mastered structure, i.e. one that a speaker can activate as a preassembled whole without attending to the specifics of its internal composition. A unit can therefore be regarded as a cognitive routine. (Langacker 1987: 494)

A good example of a linguistic unit is the word (cf. Langacker 2008: 16-17). A word is formed from a sequence of speech sounds and may consist of several morphemes bound together as a processing unit that leaves very little room for sequential variation: the components of words, i.e. speech sounds and morphemes, are usually arranged in fixed orders. Phrases and clauses are more variable in this regard, they often allow for (some) linear rearrangements; however, like words, phrases and clauses are processing units in Langacker's sense of the term, i.e. they are "cognitive routines" that constitute "preassembled wholes".

Since automatization is a gradual process driven by frequency, the units of speech vary on a continuum. Other things being equal, the more frequent a particular string of linguistic elements is processed the stronger is the cohesion of the emerging unit (see Bybee 2002, 2010: chap 3 and chap 8). The most cohesive (meaningful) unit is the morpheme, which usually does not allow for any sequential variation. Larger units, e.g. phrases and clauses, are more variable in this regard and therefore less cohesive. However, if the elements of larger sequential units occur with high transitional probabilities they may turn into fixed expressions. This is immediately obvious in the case of lexical prefabs; but in what follows I argue that automatization also affects more abstract grammatical patterns, i.e. constructions (or constructional schemas).

Lexical prefabs are automatized sequences of multiple words that have developed into routine expressions. The development is driven by frequency or repetition: Processing a string of lexical expressions strengthens the associative links between individual words and morphemes. The more often a string of words is processed, the stronger is its status as a (lexical) unit (cf. Figure 4).

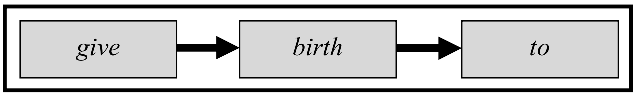

Figure 4. Lexical prefab

The same processing mechanism accounts for the emergence of more schematic (and productive) grammatical patterns. In Section 2 we have seen that verb-argument constructions are often associated with particular words. The ditransitive construction, for instance, is strongly associated with the verb *give*; but *give* also occurs in the *to*-dative construction. The two constructions, i.e. the ditransitive and the *to*-dative, involve the same participant roles, i.e. a recipient and theme; but the roles occur in different orders. In the ditransitive construction the recipient precedes the theme (e.g. *give the girl the thing*), but in the *to*-dative it is the other way around (e.g. *give the thing to the girl*) (if we disregard that in some British varieties of English the ditransitive can also appear with the reverse order of recipient and theme; cf. Section 2.1). Since *give* is more frequent in the ditransitive than in the *to*-dative (see above), the sequence *give*-RECIPIENT has a higher transitional probability than the sequence *give*-THEME, suggesting that *give*-plus-RECIPIENT is more a cohesive unit than *give*-plus-THEME (cf. Figure 5).

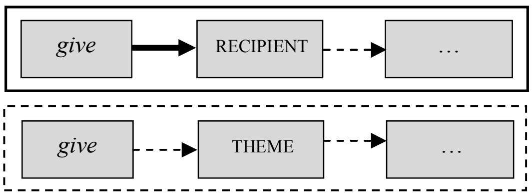

Figure 5. AUTOMATIZED DITRANSITIVES: ASSOCIATIVE LINKS BETWEEN *give* AND DIFFERENT THEMATIC ROLES

Parallel to the relationship between verbs and thematic roles, one can analyze the relationships between verbs and other semanto-syntactic categories. For instance, since *give* is more frequently combined with a personal pronoun referring to an animate being (e.g. *give me* ...) than with a full lexical noun referring to an object (e.g. *give the thing* ...), the strings *give*-plus-PRONOUN and *give*-plus-ANIMATE are more strongly automatized, and therefore more cohesive, than the strings *give*-plus-NOUN and *give*-plus-INANIMATE (cf. Figure 6 and Figure 7).

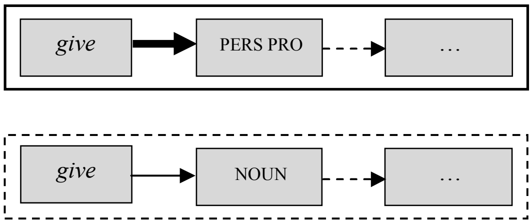

Figure 6. AUTOMATIZED DITRANSITIVES: ASSOCIATIVE LINKS BETWEEN *give* AND DIFFERENT NP TYPE

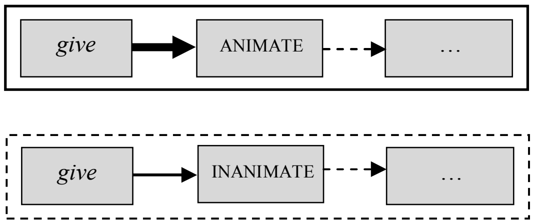

Figure 7. AUTOMATIZED DITRANSITIVES: ASSOCIATIVE LINKS BETWEEN *give* AND DIFFERENT SEMANTIC REFERENTS

Of course, automatization does not only affect the combination of a verb and its adjacent categories. For a full account, one has to consider the transitional probabilities between all linguistic elements in a particular processing unit.

One piece of evidence supporting the hypothesis that the lexical biases of verb-argument constructions are the result of automatization comes from research on phonetic reduction. Most studies on phonetic reduction are concerned with strings of lexical expressions; but Gahl and Garnsey (2004) conducted an experiment in which they examined the occurrence of phonetic reduction in strings defined over (abstract) categories. Specifically, they investigated the pronunciation of sentences including verbs such as *confirm* and *believe* that can occur with different syntactic types of complements. In Section 2 we saw that individual verbs are often statistically biased to occur with a particular type of complement and that these biases influence sentence processing. Building on this finding, Gahl and Garnsey hypothesized that speakers are more likely to reduce sentences that are consistent with the statistical bias of a particular verb than sentences that deviate from the statistically dominant pattern. Using a wide range of phonetic measurements, they examined the pronunciation of 21 verbs before NP and S complements. In accordance with their hypothesis they found that “NP-biased verbs” (i.e. verbs that tend to occur with NP complements, e.g. *confirm*) are more often phonetically reduced before nominal complements (e.g. *He confirmed the rumor*) than before complement clauses (e.g. *He confirmed the rumor is true*); whereas “S-biased verbs” (i.e. verbs that tend to occur with S complements, e.g. *believe*) are more strongly reduced if they precede a complement clause (e.g. *He believed the story is true*) rather than a simple noun phrase (e.g. *He believed the story*).

If we assume that phonetic reduction is caused by automatization, as Bybee (2001) and others have claimed, the results of this research suggest that automatization does not only affect strings of concrete lexical expressions, but also strings involving categories. Thus, we may conclude that automatization is the driving force for the development of both (i) lexical prefabs consisting of strings of lexical expressions (e.g. *give birth to*), and (ii) constructional schemas (or item-specific constructions) consisting of a relational category (or relational term) and a semanto-syntactic category (e.g. *give*-RECIPIENT).

Note that such item-specific constructions are similar to children’s early pivot schemas. As pointed out in Section 2, the earliest multi-word utterances children produce are commonly organized around a relational term, e.g. a verb, that is associated with a slot for non-relational expressions with particular semantic and/or structural properties. In the literature, verb-argument constructions are commonly defined over abstract syntactic categories; but the evidence reviewed in Section 2 has shown that the use and development of argument structure in adult language cannot be separated from particular words. I suggest therefore that adult grammar includes a level of cognitive organization at which verb-argument constructions are tied to particular verbs similar to verb-argument constructions in early child language. On this account, adult grammar includes a large inventory of item-specific constructions organized around individual verbs (with particular senses) that are associated with certain structural positions and/or semanto-syntactic categories (see Dąbrowska 2009 for a related proposal).

Let me emphasize, however, that this analysis does not question the existence of fully schematic verb-argument constructions (in adult grammar). There is evidence that adult speakers are able to abstract away from particular verbs and that their knowledge of argument structure includes fully schematic syntactic patterns (cf. Goldberg 2006). However, in addition to constructional schemas, adult grammar includes an enormous amount of verb-specific information about argument structure that reflects the language user's experience with frequent strings of individual verbs and their associated categories.

4. Conclusion

To conclude, the paper has surveyed linguistic and psycholinguistic research on lexical specificity in grammar. The results of this research support the usage-based hypothesis that syntactic structures, i.e. constructions, are stored and processed across a continuum of abstractness and that lower level constructions are associated with particular lexemes. There is evidence from a wide range of studies that syntactic structures are commonly realized by lexical prefabs and idiomatic expressions, and that the structural positions of verb-argument constructions are statistically biased to include particular nouns and verbs.

In the usage-based literature, these phenomena are commonly explained by the exemplar mechanism of category learning; but the current paper has argued that the "bare" exemplar model is not sufficient to explain how lexemes and constructions are related. What is missing in this model is a cognitive mechanism for the processing of sequential information. Drawing on research from cognitive psychology, I have argued that the exemplar model of language learning needs to be combined with a linear processing mechanism, i.e. automatization, in order to account for the connection between words and structure in usage-based construction grammar.

References

Abbot-Smith, Kirsten and Michael Tomasello. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23: 275-290.

Anderson, John R. 2005. *Cognitive Psychology and its Implications*. [6th edition]. New York: Worth Publishers.

Aslin, Richard N. and Elissa L. Newport. 2012. Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science* 21: 170-177.

Beckner, Clay and Joan Bybee. 2009. A usage-based account of constituency and reanalysis. *Language Learning* 59: 27-46.

Bell, Alan, Jason Brenier, Michelle Gregory, Cynthia Girand, and Daniel Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60: 92-111.

Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and Development of Language,* 279-352, New York: Wiley.

Bloom, Lois. 1973. *One Word at a Time*: *The use of single word utterances before syntax*. The Hague: Mouton.

Bowerman, Melissa. 1976. *Early Syntactic Development: A cross-linguistic study with special reference to Finnish.* Cambridge: Cambridge University Press.

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

Brooks, Patricia and Michael Tomasello. 1999. Young children learn to produce passives with nonce verbs. *Developmental Psychology* 35: 29-44.

Bybee, Joan. 1985. *Morphology*. *A study on the relation between meaning and form*. Amsterdam: John Benjamins.

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

Bybee, Joan. 2002. Sequentiality as the basis of constituent structure. In Talmy Givón and Bertram F. Malle (eds.), *The Evolution of Language out of Pre-Language*, 109-132. Amsterdam: John Benjamins.

Bybee, Joan. 2006. From usage to grammar: The mind's response to repetition. *Language* 82: 711-733.

Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Cambridge: Cambridge University Press.

Bybee, Joan. 2010. *Language, Usage, and Cognition*. Cambridge: Cambridge University Press.

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

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

Chase, William G. and Herbert A. Simon. 1973. Perception in chess. *Cognitive Psychology* 4: 55-81.

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

Cowan, Nelson. 2005. *Working Memory Capacity*. New York: Psychological Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press.

Dąbrowska, Ewa. 2009. Words as constructions. In Vyvyan Evans and Stéphanie Pourcel (eds.), *New Directions in Cognitive Linguistics*, 201-224. Amsterdam: John Benjamins.

Diessel, Holger. 1997. Verb-first constructions in Germa. 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. 2004. *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Diessel, Holger. 2007. Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology* 25: 108-127.

Diessel, Holger. 2011a. Review article of 'Language, usage and cognition' by Joan Bybee. *Language* 87: 830-844.

Diessel, Holger. 2011b. Grammaticalization and Language Acquisition. In Bernd Heine and Heiko Norrog (eds.), *Handbook of Grammaticalization*, 130-141. Oxford: Oxford University Press.

Diessel, Holger. 2012. Language Change and Language Acquisition. In Alexander Bergs and Laural Brinton (eds.), *Historical Linguistics of English: An international handbook*. Vol. II. 1599-1613. Berlin: Mouton de Gruyter.

Diessel, Holger. 2013. Construction Grammar and First Language Acquisition. In Graeme Trousdale and Thomas Hoffmann (eds.), *The Oxford Handbook of Construction Grammar*, 347-364. Oxford: Oxford University Press.

Ellis, Nick C. 1996. Sequencing in SLA. Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18: 91-126.

Ellis, Nick C. 2002. Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 43-188.

Erman, Britt and Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20: 29-62.

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.

Ford, Marylin, Joan Bresnan, and Ronald M. Kaplan. 1982. A competence-based theory of syntactic change. In Joan Bresan (ed.), *Mental Representations and Grammatical Relations*, 727-796. Cambridge: MIT 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.

Gahl, Susanne and Susan M. Garnsey. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80: 748-775.

Garnsey, Susan M., Neal J. Pearlmutter, Elizabeth Myers, and Melanie A. Lotocky. 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 7: 58-93.

Gerken, LouAnn. 2006. Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition* 98: 67-74.

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at Work. The nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, Adele E. and Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80: 532-568

Gómez, Rebecca L. and LouAnn Gerken. 1999. Artificial grammar learning by 1-year-olds leads to specific abstract knowledge. *Cognition* 70: 109-135.

Gries, Stefan and Anatol Stefanowitsch. 2004. Extending collexeme analysis. *International Journal of Corpus Linguistics* 9: 97-129.

Hare, Mary, Ken McRae, and Jeff L. Elman. 2003. Admitting that admitting verb sense into corpus analyses makes sense. *Language and Cognitive Processes* 19: 181-224.

Haspelmath. Martin. 2012. Explaining the ditransitive person-role constraint. *Constructions* 2: 1-72.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in allomorphy, word-formation and syntax*. Cambridge: Cambridge University Press.

Ibbotson, Paul and Michael Tomasello. 2009. Prototype constructions in early language acquisition. *Language and Cognition* 1: 59-85.

Johnson, Keith. 1997. Speech perception without speaker normalization. In Keith Johnson and John W. Mullennix (eds.), *Talker Variability in Speech Processing*, 145-165. San Diego: Academic Press

Jurafsky, Daniel. 1996. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20: 137-194.

Jurafsky, Daniel. 2003. Probabilistic modelling in psycholinguistics: Linguistic comprehension and production. In Rens Bod, Jennifer Hay, and Stefanie Jannedy (eds.), *Probabilistic Linguistics*, 39-95. Cambridge: MIT Press.

Krug, Manfred. 2000. *Emerging English Modals. A corpus-based study of grammaticalization*. Berlin: Mouton de Gruyter.

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. 1987. *Foundations of Cognitive Grammar.* Vol. II. *Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 2008. *Cognitive Grammar. A basic introduction*. Oxford: Oxford University Press.

Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago: University of Chicago Press.

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.

Lieven, Elena V.M., Heike Behrens, Jennifer Spears, and Michael Tomasello. 2003. Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30: 333-370.

Logan, Gordan D. 1988. Towards an instance theory of automatization. *Psychological Review* 95: 492-527.

MacDonald, Maryellen C. 1994. Probabilistic constrainst and syntactic ambiguity resolution. *Language and Cognitive Processes* 9: 157-201

MacDonald, Maryellen C., Neal J. Pearlmutter, and Mark S. Seidenberg. 1994. Lexical nature of syntactic ambiguity resolution. *Psychological Review* 191: 676-703.

MacDonald, Maryellen C. and Mark S. Seidenberg. 2006. Constraint satisfaction accounts of lexical and sentence comprehension. In Matthew J. Traxlor and Morton Ann Gernsbacher (eds.), *Handbook of Psycholinguistics*, 581-611. [2nd edition]. London: Elsevier.

Manning, Christopher D. and Hinrich Schütze. 2000. *Foundations of Statistical Natural Language Processing*. Cambridge, Mass: MIT Press.

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

McEnery, Tony and Andrew Hardie. 2012. *Corpus Linguistics: Method, theory, and practice*. Cambridge: Cambridge University Press.

Miller, George A. 1956. The magical number seven, plus or minus two. Some limitations on our capacity for processing information. *Psychological Review* 63: 81-97.

Murphy, Gregory. 2002. *The Big Book of Concepts*. Cambridge: MIT Press.

Newell, Allen. 1990. *Unified Theories of Cognition*. Cambridge: Harvard University Press.

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

Nosofsky, Robert M. 1992. Exemplars, prototypes, and similarity rules. In Alice F. Healy, Stephen Michael Kosslyn, Richard M. Shiffrin, (eds.), *Essays in honor of William K. Estes.* Vol. II. *From learning processes to cognitive processes*, 149-167. Hillsdale, NJ: Lawrence Erlbaum.

Nosofsky, Robert M. and Safa R. Zaki. 2002. Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28: 924-940.

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

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.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 137-158. Amsterdam: John Benjamins.

Rips, Lance J., Edward E. Smith, and Douglas L. Medin. 2012. Concepts and categories: memory, meaning, and metaphysics. In Keith J. Holyoack and Robert G. Morrison (eds.), *The Oxford Handbook of Thinking and Reasoning*, 177-209. Oxford: Oxford University Press.

Roland, Douglas and Daniel Jurafsky. 2002. Verb sense and verb subcategorization probabilities. In P. Merlo and S. Stevenson (eds.), *Sentence Processing and the Lexicon: Formal, computational, and experimental perspectives*. 325-346. Amsterdam: John Benjamins.

Ross, Brian H. and Valery S. Makin. 1999. Prototype versus exemplar models. In Robert J. Sternberg (ed.), *The Nature of Cognition*, 205-241. Cambridge Mass: MIT Press.

Saffran, Jenny R., Elissa L. Newport, and Richard N. Aslin. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35: 606-621.

Schneider, Walter and Richard M. Shiffrin. 1977. Controlled and automatic human information processing I: detection, search, and attention. *Psychological Review* 84: 1-66.

Schneider, Walter and Jason M. Chein 2003. Controlled and automatic processing: behavior, theory, and biological mechanisms. *Cognitive Science* 27: 525-559.

Shiffrin, Richard M. and Walter Schneider. 1977. Controlled and automatic human information processing II: perceptual learning, automatic attending, and a general theory. *Psychological Review* 84: 127-190.

Siewierska, Anna and Willem Hollmann. 2007. Ditransitive clauses in English with special reference to Lancashire dialect. In M. Hannay, and Gerald D. Steen (eds.), *Structural-functional Studies in English Grammar in Honor of Lacklan Mackenzie*, 83-102. Amsterdam: John Benjamins.

Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Spivey-Knowlton, Michael J. and Julie C. Sedivy. 1995. Resolving attachment ambiguities with multiple constraints. *Cognition* 55: 227-267.

Thompson, Sandra A. and Paul Hopper. 2001. Transitivity, clause structure, and argument structure: Evidence from conversation. In Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 27-60. Amsterdam: John Benjamins.

Thompson, Susan P. and Elissa L. Newport. 2007. Statistical learning of syntax: The role of transitional probabilities. *Language Learning and Development* 3: 1-42.

Tomasello, Michael. 1992. *First Verbs. A case study of early grammatical development.* Cambridge: Cambridge 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 first language acquisition*. Cambridge: Harvard University Press.

Tottie, Gundel. 1991. 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.

Traugott, Elizabteh Closs and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Trueswell, John C. 1996. The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language* 35: 566-585.

Trueswell, John C., Michael K. Tanenhaus, and Susan M. Garnsey. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 33: 285-318.

Weiser, Jana. 2006. *Exemplartheorie: Verschiedene Ansätze im Vergleich*. Diplomarbeit. Universität Stuttgart.

Wiechmann, Daniel. 2008. Sense-contingent lexical preferences and early parsing decisions: Corpus-evidence from local NP/S-ambiguities. *Cognitive Linguistics* 19: 447-464.

Wonnacott, Elizabeth, Elissa L. Newport, and Michael Tanenhaus. 2008. Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology* 56: 165-209.

Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.