Remi van Trijp*

Cognitive vs. generative construction grammar: The case of coercion and argument structure

DOI 10.1515/cog-2014-0074
Received October 6, 2014; revised February 19, 2015; accepted March 16, 2015

Abstract: One of the most salient hallmarks of construction grammar is its approach to argument structure and coercion: rather than positing many different verb senses in the lexicon, the same lexical construction may freely interact with multiple argument structure constructions. This view has however been criticized from within the construction grammar movement for leading to overgeneration. This paper argues that this criticism falls flat for two reasons: (1) lexicalism, which is the alternative solution proposed by the critics, has already been proven to overgenerate itself, and (2) the argument of overgeneration becomes void if grammar is implemented as a problem-solving model rather than as a generative competence model; a claim that the paper substantiates through a computational operationalization of argument structure and coercion in Fluid Construction Grammar. The paper thus shows that the current debate on argument structure is hiding a much more fundamental rift between practitioners of construction grammar that touches upon the role of grammar itself.

Keywords: cognitive-functional language processing, language formalization, computational modeling, Fluid Construction Grammar

1 Introduction

In 1995, Adele E. Goldberg threw a large pebble in the pond of linguistics with her book *Constructions: A construction grammar approach to argument structure*, in which she forcefully argues against a lexicalist approach to argument realization. Instead of implementing all constraints in the lexicon, Goldberg posits the existence of argument structure constructions, which are quite similar to lexical constructions in the sense that they are also mappings between meaning/function and form. Different patterns of argument realization are then the result of the free combination of lexical constructions with various argument structure

Corresponding author: Remi van Trijp, Sony Computer Science Laboratory Paris, 6 rue Amyot, 75005 Paris, France, E-mail: remi@csl.sony.fr

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

constructions. The constructional account is especially appealing in the case of coercion, where grammatical constraints seem to be violated, as illustrated in example (1).

(1) *He hurried forward to help Aunt Petunia* ***negotiate*** *a weak-kneed Dudley over the threshold while avoiding stepping in the pool of sick.*
(JK Rowling, *Harry Potter and the Order of the Phoenix*, emphasis added)

The example involves Harry Potter's uncle and aunt, who are trying to carry their big and heavy son into their house after he was attacked by a magical creature. The author's choice of the verb *negotiate* is interesting because it is normally not associated with a caused-motion semantics. Native speakers of English are however already familiar with a similar sense of the verb from utterances such as *she carefully negotiated the road with her car*, where it means "to successfully travel over or through a difficult route or obstacle". The use of the verb in example (1) is thus most likely the semantic extension of simple motion to the meaning "X causes Y to move over or through Z". Lexicalist approaches can only account for such examples by adopting an additional verb sense in the lexicon or through a derivational lexical rule. In a constructional account, on the other hand, this new sense of *negotiate* is analyzed as *coercion by construction* in which the Caused-Motion Construction imposes its argument structure onto the verb's participant structure. This analysis is illustrated using a Goldbergian diagram of argument structure in Figure 1.

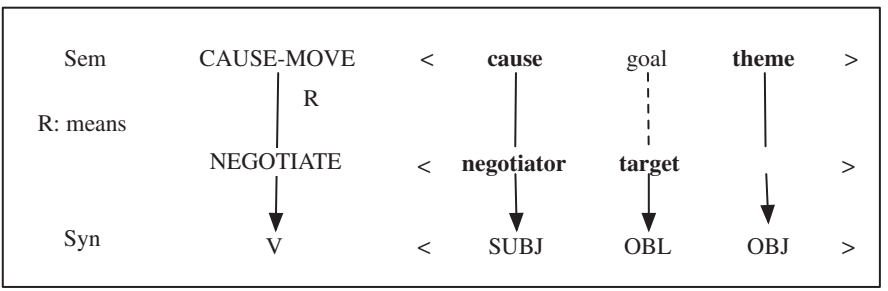


Figure 1: This Goldbergian diagram illustrates the combination of the verb *to negotiate* with the Caused-Motion Construction. The verb's participant structure already contains two obligatory participant roles (a negotiator and a target). The Caused-Motion Construction then imposes its obligatory Theme role onto the verb's participant structure.

But when you throw a pebble in the water, there is always a ripple effect. While many researchers have embraced the notion of argument structure constructions, their exact status and how they should interact with other constructions is

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

a matter of heavy debate both within and outside of the construction grammar community (among others Boas 2008a, 2008b; Croft 1998, 2003; Goldberg 1995, 2006; Goldberg and Jackendoff 2004; Iwata 2008; Kay 2005; Kay and Michaelis 2012; Levin and Rappaport Hovav 2005; Müller 2006; Müller and Wechsler 2014; Nemoto 1998). The most important criticism is that Goldberg's approach leads to problems of *overgeneration*, as shown in example (2) from Kay (2005: Ex. 17b):

(2) **He bragged her to sleep*.

However, none of the participants in the debate ever spell out what overgeneration is supposed to mean. The next section therefore lays bare all of this concept's hidden consequences, and by doing so, will reveal a much more fundamental issue that is at stake here: what is the role of grammar?

2 Grammar as a generative competence model

Overgeneration is a concept that emerged in the tradition of generative grammar. Chomsky (1965) explains a generative grammar as follows:

> To avoid what has been a continuing misunderstanding, it is perhaps worthwhile to reiterate that a generative grammar is not a model for a speaker or hearer. It attempts to characterize in the most neutral possible terms the knowledge of the language [...]. When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. (Chomsky 1965: 9)

The key to the above citation is that a generative grammar is a process-neutral competence model, which means that the words *generate* and *derivation* should not be understood in their intuitive sense, but rather in the sense of a formal theory of a language as a set of expressions. Within the family of construction grammars, Sign-Based Construction Grammar (SBCG; Boas and Sag 2012; Michaelis 2013) most outspokenly continues the process-neutral tradition:

> [G]iven that linguistic knowledge is process-independent, there should be no bias within a grammatical theory – whether overt or hidden, intentional or inadvertent – toward one kind of processing, rather than another. (Sag and Wasow 2011: 368)

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

2.1 The problem of overgeneration

The best-known example of a generative grammar is a *phrase structure grammar* (or *context-free grammar*; Chomsky 1956, 1957), as illustrated in example (3).

(3)

S ➔ NP VP	N ➔ *man*
NP ➔ DN	N ➔ *ball*
VP ➔ V	N ➔ *napkin*
VP ➔ V NP	V ➔ *sneezed*
VP ➔ V NP PP	V ➔ *kicked*
PP ➔ P NP	P ➔ *off*
D ➔ *the*	...

The arrows suggest that a phrase structure grammar needs to be applied from left to right, but this is not true: it simply consists of a set of *declarative rules* that do not specify how they should be used for processing. Indeed, computational linguists have devised various processing strategies that apply these rules in a top-down (i.e., left-to-right) or bottom-up fashion (i.e., right-to-left), or a mixture of both (Jurafsky and Martin 2000: Ch. 10). In the simplest case, a *recognition* algorithm can be used for testing whether a sentence is *accepted* by the grammar.

The problem of overgeneration is that the grammar of example (3) accepts sentences such as **the man sneezed the ball*. This sentence is ungrammatical, yet it satisfies all the constraints of the grammar. The job of the linguist is then to further refine the model in order to ensure that the grammar accepts all and only the grammatical sentences of English. For example, Chomsky's (1965) model added subcategorization and selection restrictions to the (1957) apparatus of phrase structure grammars and transformations in order to improve the model's accuracy in terms of this "all-and-only" requirement.

2.2 Preventing overgeneration leads to undergeneration

Researchers who have criticized Goldberg's (1995) argument structure constructions for "overgenerating" (see for instance Kay 2005; Morita 1998; Nemoto 1998; Boas 2003; Iwata 2002; Müller and Wechsler 2014) are thus concerned that such broad-coverage constructions make the grammar too permissive with respect to which utterances it accepts. As a solution, these scholars propose stronger constraints on lexical constructions, either calling their approach "lexical-constructional" (e.g., Iwata 2008), or arguing in favor of a lexicalist account (e.g., Müller and Wechsler 2014).

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

Fortunately, several lexicalist theories, such as Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag 1994) and Sign-Based Construction Grammar (Boas and Sag 2012), have made their proposals formally explicit, which makes it possible to examine more clearly whether overgeneration can indeed be avoided by lexicalizing a grammar. Let us look at HPSG as an illustration. There are two important differences between HPSG and traditional phrase structure grammars that are important for the current discussion. First, instead of treating non-terminal symbols such as VP and V as atomic categories, HPSG makes fine-grained distinctions in its representation of linguistic knowledge through *feature structures* (i.e., sets of attributes and values; Kay 1979). Secondly, HPSG implements *type definitions* that specify which structures are admissible by a grammar and thereby tackle the problem of overgeneration (Jurafsky and Martin 2000: 438).

For instance, suppose that a generative linguist wants to develop a formal competence model of the English *verb* category that captures the fact that English verbs can appear in different forms, and that they can take agreement marking. At the same time, the model should disallow features that are not appropriate for English verbs, such as the nominal feature CASE. In order to achieve this goal, the linguist then proceeds by implementing a *type definition*, which consists of the name of the type and of the constraints that all feature structures of this type should satisfy. The name of the type can be any symbol, such as *type-1*, but since the type is supposed to constrain feature structures that model verbs, the linguist decides to call it *verb*. Example (4) illustrates the constraints that feature structures of type *verb* should adhere to.

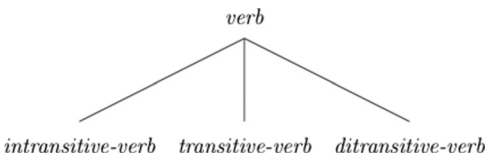
(4)
$$\begin{bmatrix} \textit{verb} & \\ \text{AGREEMENT} & \textit{agr} \\ \text{VERB-FORM} & \textit{verb-form} \end{bmatrix}$$

As can be seen, there are two "appropriate" features associated with feature structures of type *verb*: AGREEMENT and VERB-FORM. Note that their values are typed as well (indicated in italics). The value of the AGREEMENT feature must satisfy the constraints of a type called *agr*, which may for example specify that an appropriate value for AGREEMENT is a feature structure that contains the features NUMBER and PERSON (whose values are also typed). Likewise, the value of the feature VERB-FORM is typed.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

Types are organized in a *type hierarchy*, where subtypes inherit the constraints of their supertypes, as shown in example (5).

(5)



The type hierarchy then enables the grammar writer to model for instance the verb *sneeze* using a feature structure of type *intransitive-verb*. It would thus inherit all the appropriateness conditions of the *intransitive-verb* type, the *verb* type, and so on. Now all that the linguist needs to do to prevent *sneeze* from behaving as a transitive verb (thereby accepting utterances such as **she sneezed the ball*) is to restrict the rule VP ➔ V NP so that it accepts only transitive verbs.

However, type restrictions on lexical entries now lead to the problem of *undergeneration*. For instance, if *kick* is typed as a *transitive-verb*, the grammar can no longer account for instances of multiple argument realization as illustrated in example (6), taken from Goldberg (1995: 11). That is, the grammar would either dismiss the utterance for being ungrammatical, or it would wrongly analyze it as a transitive one. This problem could in principle be solved by listing all the possible senses of a word as a separate entry in the lexicon, but that solution comes at the loss of generalization and would cause spurious ambiguity in parsing. Moreover, the grammar cannot handle cases of coercion, such as *the man sneezed the napkin off the table*, which was still permitted by the phrase structure grammar of example (3).

(6) a. *Pat kicked Bob black and blue.*
b. *Pat kicked his foot against the chair.*
c. *Pat kicked Bob the football.*
d. *The horse kicks.*
e. *Pat kicked his way out of the operating room.*

2.3 Preventing undergeneration leads to overgeneration

Enter lexical rules. Lexical rules can best be thought of as *valency affecting operations* (Keenan and Faltz 1985) that are able to take the valency of a word that was fixed by a type constraint, and then add, delete or shift its values. The

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

most famous lexical rule is the Passive Rule, which changes the default valency values of an active verb form into those of a passive verb.

But here's the rub for critics of Goldbergian argument structure constructions: lexical rules are overly powerful. Even massively so. Carpenter (1991) has already formally proven that lexical rules (in the style of HPSG or categorial grammars) lead to unrestricted generative capacity:

> The inevitable conclusion is that if we want a natural and effectively decidable lexical rule system for categorial grammars or head-driven phrase structure grammars, then we must place restrictions on the system given here or look to state lexical rules at completely different levels of representation which themselves provide the restrictiveness desired [...]. The common assumption that lexical rules can perform arbitrary operations [...] is simply not restrictive enough to yield an effective recognition algorithm. (Carpenter 1991: 311–312)

Another inevitable conclusion is that the "all-and-only" requirement of generative grammar leads to a vicious circle: introducing strong constraints on the lexicon (such as a type system) helps to prevent overgeneration, but then leads to the problem of undergeneration. In order to solve this problem, lexical rules are added that in turn lead to... massive overgeneration. At this point, lexicalists will undoubtedly object to these arguments by saying that there exist computational implementations of lexicalist grammars that make efficient use of lexical rules. Such implementations indeed exist, but they reinforce my point rather than weaken it: they acknowledge the fact that lexical rules are very powerful because they had to tackle the problem head on. Here are some examples of the solutions that have been proposed:

- Using a strategy to postpone and control the application of lexical rules (e.g., Bouma and van Noord 1994).
- Encoding the possibilities of lexical rule application as systematic covariation in lexical entries (Meurers and Minnen 1997). More concretely, Meurers and Minnen translate general lexical rules as written by linguists into a more restricted set of rules, and then extend lexical entries with explicit information about which lexical rules can apply to them. Coercion can still be achieved using probabilistic methods (1997: footnote 27).
- Using probabilistic methods (Briscoe and Copestake 1999).

Just to be clear, none of this proves that a lexicalist approach is wrong, but it does show that the criticism of overgeneration cannot be used in favor of such an approach. What the computational implementations of lexical rules demonstrate quite clearly is that overgeneration is only a problem if the "all-and-only" assumption of generative grammar is taken to its logical conclusion. However, if one cares about how language users are able to process language in a robust

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

and efficient way – in short, if one cares about language usage – it is much better to allow overgeneration *in principle* and to find strategies for avoiding this from happening *in practice*. In fact, the most successful language technologies that exist today are statistical models that "massively overgenerate analyses, but most of those analyses receive very low numerical scores, and so will be outranked by correct analyses" (Penn 2012: 165).

3 Grammar as a problem-solving model

By deconstructing all the implications of the criticism of overgeneration, we have hit upon a fundamental insight: if a grammar is strictly treated as a generative competence-model, linguists are doomed to run in circles to find a model that neither under- nor overgenerates – regardless of whether they use lexical rules or constructions. The only way out is to adopt a different notion of what a grammar should capture, and to bring processing into the loop.

This alternative view of grammar is often called *usage-based* or *cognitive-functional* in the literature. In this paper, I will use the term *problem-solving model* (borrowed from the field of Artificial Intelligence) to emphasize that the grammar is not a processing-independent competence model, but instead plays an active role in the communicative acts and social interactions of language users. More specifically, a problem-solving grammar needs to "solve" the following two problems:

- *Production:* How to verbalize the meanings conceptualized by the language user in such a way that communicative success is maximized and cognitive effort is minimized?
- *Comprehension:* How to extract the intended meaning from a speaker's utterance despite the ubiquity of ambiguity?

Each usage event potentially changes the grammar and its structure (Bybee 2010). That is, the grammar will adhere as much as possible to the conventions of a speech community, but restructure itself for future interactions and remain open-ended in order to learn novel expressions (*acquisition*) or to expand when the speaker's expressive needs call for it (*innovation*). Grammaticality is considered to be a gradient concept, which means that a problem-solving model is inherently probabilistic.

As soon as we adopt the problem-solving model, the criticism of overgeneration becomes void even when using powerful generalizations in the form of Goldbergian argument structure constructions. I will substantiate this claim in the remainder of this paper through a computational implementation of argument structure and coercion in Fluid Construction Grammar (FCG; Steels 2011, 2012a), which refines an earlier implementation by van Trijp (2011). An

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

important disclaimer is that FCG is not a linguistic theory but rather an open-source software tool that can be used for the computational verification of constructional analyses. Throughout this paper, I will use non-formal descriptions for convenience's sake, with pointers to the formal details in the Appendix. Interested readers can also check an online video tutorial on implementing argument structure constructions at www.fcg-net.org/tutorial/, where they can also download a sample grammar and the FCG-system for testing the approach.

3.1 Verbs and argument structure constructions

Goldberg (1995) assumes a frame semantics approach (Fillmore 1976) for handling verbal meaning, which includes an event's *participant roles* (Goldberg 1995: 43). Participant roles are verb-specific and thus distinct from the more general *argument roles* of argument structure constructions. For instance, *break* has at least two participant roles: a breaker, and something that becomes broken. Likewise, the verb *hit* has at least two participant roles: a hitter, and something being hit.

(7) break < breaker, **broken** >

(8) hit < **hitter**, target >

According to Goldberg (1995: 44), verbs *lexically profile* some of their participants, which are obligatory roles. The FCG implementation does not adopt this notion of obligatory roles, because this would require additional mechanisms for deleting or backgrounding these roles in order to process utterances in which they do not occur. As we will see further below, it is sufficient to let argument structure constructions decide on which roles need to be expressed and which can remain implicit in the message.

However, the implementation does highlight one role as the most salient role, which is indicated in boldface in examples (7) and (8). In the case of English, the salient role implements a distinction between accusative verbs (i.e., verbs whose subject in active-intransitive clauses aligns with the subject of active-transitive clauses) and ergative verbs (i.e., verbs whose subject in active-intransitive clauses aligns with the object of active-transitive clauses). This distinction is illustrated in examples (9) and (10) for the verbs *to hit* and *to break* (also see Fillmore 1970). The verb *to hit* behaves like an accusative verb because its subject is preferentially mapped onto the actor in both active intransitive and transitive clauses. The verb *to break*, on the other hand, behaves as an ergative verb because its subject is preferentially mapped onto the undergoer in active intransitive utterances.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

(9) *The storm hit.*

(10) *The window broke.*

In line with usage-based approaches to language, argument structure constructions may range from verb-specific and verb-class-specific constructions to fully schematic constructions (Boas 2008a; Croft 2003). A fully schematic Active-Intransitive Construction is able to abstract away from (classes of) verbs by simply mapping the *Salient role* onto the syntactic function *Subject*, as shown in Figure 2. A fully schematic *Active-Transitive Construction*, shown directly underneath the intransitive construction, performs a similar mapping between an actor and undergoer onto subject and object.

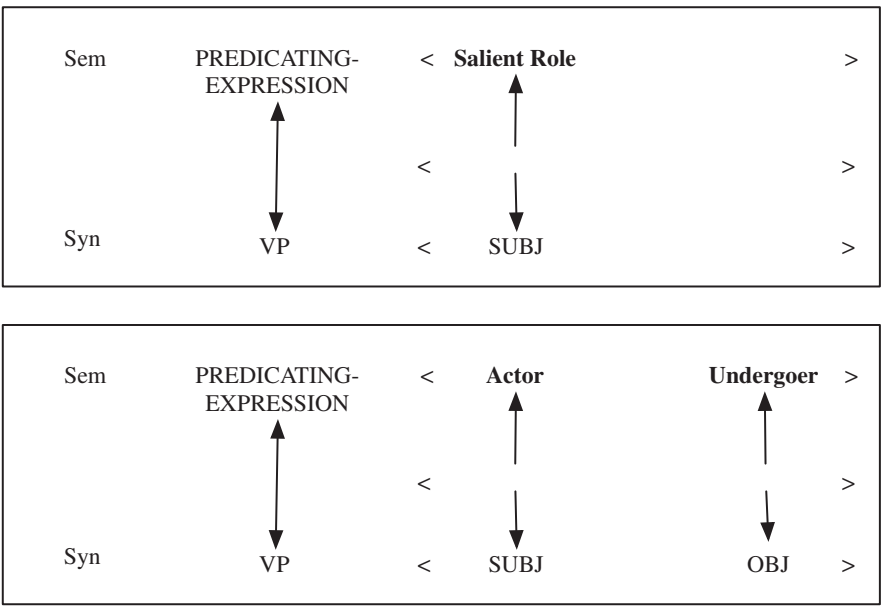


Figure 2: This Figure shows a diagrammatic representation of the Active-Intransitive Construction on top, and the Active-Transitive Construction below. Their formal representations are provided in Appendixes A4 and A5. Instead of directly referring to a verb's participant structure, these fully schematic constructions use more abstract argument roles such as actor, undergoer and salient role.

3.2 Fusion and valency

So how do verbal constructions interact with argument structure constructions? Goldberg (1995: 50) proposes that "the participant roles of the verb may be

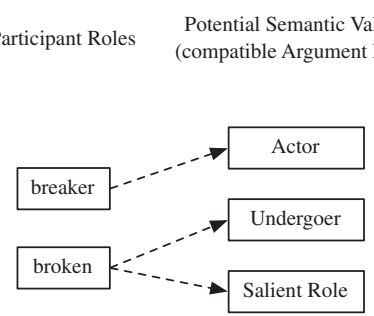
Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

semantically fused with the argument roles of the argument structure construction." Unfortunately, Goldberg has never offered any specific details of how this process of *fusion* might work.

In the FCG implementation, fusion is operationalized as a categorization task for which two different scenarios are possible. The first scenario involves categorization "on the fly" and applies in the case of novelty and language acquisition. Suppose that a learner has already acquired the transitive construction, but now encounters the verb form *broke* for the first time in this distribution. The classification task consists of figuring out whether and how the participant roles of the verb fit in the argument roles of the construction, thereby using all the information at the language user's disposal (including past experiences and discourse context). Recent agent-based models have already demonstrated how *analogical reasoning* can be exploited for testing the semantic compatibility of participant and argument roles, and how the same processes cannot only account for learning, but also for innovation and language change (van Trijp 2015).

Crucially, both argument structure constructions and verbal constructions are affected by the process of fusion, which leads to the second scenario that involves a conventionalized interaction between such constructions. More specifically, the on-the-fly categorizations may become an entrenched part of a lexical construction by extending the construction with a *potential semantic valency* that stores which argument roles are compatible with which participant roles. Example (11) illustrates the combinatorial potential of *to break*.

(11)

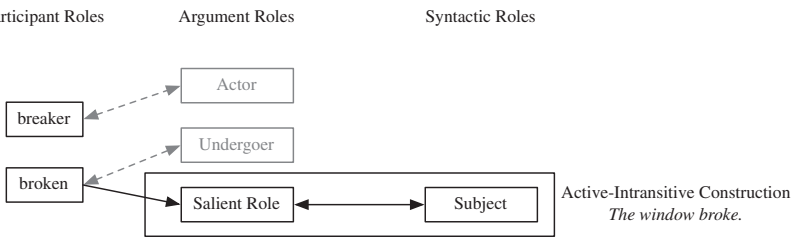


Once the lexical construction keeps information about which argument roles it is compatible with, fusion is achieved directly through routine processing. In the FCG-implementation, this means that an argument structure construction will *select* the roles that it needs from the verb's combinatorial potential, and map them onto a functional structure. Other potential combinations are simply

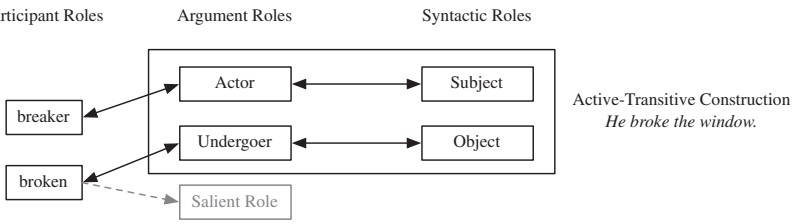
Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

ignored. In parsing, the opposite direction is followed. Examples (12) and (13) illustrate these processes for the Active-Intransitive and the Active-Transitive Constructions. Example (14) shows how a Passive-Transitive Construction selects the undergoer and maps it onto the subject role.

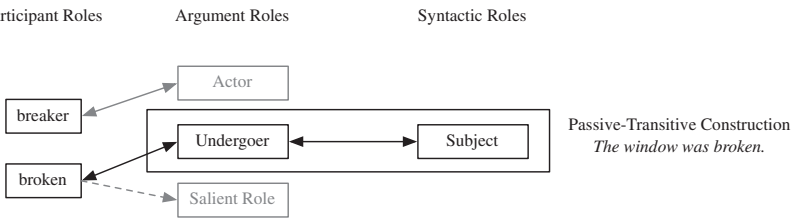
(12)



(13)



(14)



3.3 Coercion by construction

A third scenario is what happens when there is a mismatch between the frame semantics of a verb and the argument roles of the argument structure construction, as in the well-known example *He sneezed the napkin off the table* (Goldberg 1995: 55). The verb *sneeze* is typically defined as an intransitive verb, so the lexical construction of the verb form *sneezed* includes one participant role (the sneezer) and one compatible role in its potential valency (also see Appendix A3):

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

(15) a. sneeze < **sneezer** >
b. sneezer ➔ actor

The Caused-Motion Construction, on the other hand, expresses the abstract meaning "X causes Y to move Z" in which Z is a locative phrase (Goldberg 1995: Ch. 7). The construction thus needs to assign three argument roles, but only one can be found in the verb's potential valency. On-the-fly fusion through analogical reasoning is impossible in this case because there are not enough participant roles to work with.

So how can *sneezed* nevertheless interact with the Caused-Motion construction? Goldberg (1995: 54–55) suggests that an argument structure construction is able to impose its argument structure through coercion-by-construction. Technically speaking, this proposal requires a grammar formalism capable of specifying which parts of a construction must minimally be supplied by other constructions, and which parts can either be fused with existing information or be added by the construction itself. The FCG-system meets this requirement (Steels and De Beule 2006): in all of the graphical representations of constructions in the Appendix, "obligatory" information is always shown above a dotted line, and contributable information is shown below.

The distinction between obligatory and contributable information makes it possible to carefully control the conditions under which coercion is allowed in production. More specifically, in production, *sneezed* can only be coerced into the semantic frame of a Caused-Motion construction if the following conditions are met:

1. The speaker has conceptualized the scene in terms of the CAUSED-MOTION frame. As can be seen in Appendix 6, the Caused-Motion Construction requires this meaning to be supplied before it will apply.
2. The speaker wishes to overtly express all of the frame elements of the CAUSED-MOTION frame (i.e., its actor, undergoer and theme).
3. The speaker has conceptualized the sneeze-event as the cause of the motion.

These requirements are actually quite strict and will automatically prevent the grammar to produce unattested sentences unless the speaker performs a conceptual blend in which *sneeze* is conceptualized as a verb of caused motion. As a consequence, the construction can be more lenient in what it minimally requires from the verb. As can be seen in Appendix 6, the construction only demands the verb to supply the actor role (in the feature SEM-VALENCE above the dotted line). The other roles can either be merged with existing roles (which happens when the construction combines with verbs of caused-motion such as *push*), or

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

they can be contributed by the construction itself. Since the FCG-system does not impose any type constraints, the contribution of additional argument roles can be achieved by simply unifying the potential valency of the verb with the argument roles of the Caused-Motion Construction.

The bonus of this more lenient approach becomes visible in comprehension. The construction can be applied whenever a clause can be identified that contains a verb, subject, object and locative phrase. One anonymous reviewer raised an important concern about this kind of flexibility, namely the risk that the grammar comes up with incomprehensible analyses of any given input. This risk would indeed be problematic if the grammar would evaluate analyses solely based on syntactic (or formal) criteria. A problem-solving model, however, is guided by *semantic goal tests* that filter the hypothesis space based on interpretability (a suggestion also made by the reviewer), such as checking whether the current analysis is compatible with the discourse context, whether additional inferences can be made based on ontological knowledge, and so on (for an overview of the methods of computational semantics, see Bos 2011). The FCG-system provides hooks to perform those tests at each step of a processing task. As I will explain in the following section, however, reducing the search space through semantic goal tests will often not suffice, so the grammar also needs to be coupled to a probabilistic model.

3.4 Constructional competition

The model that I have presented so far has the implicit assumption that every construction expresses its own distinct semantics. However, we know that this assumption is false: there is abundant variation in a speech community and the distinction between constructions may not always be transparent to language users. Indeed, historical linguists have documented quite clearly how argument structure constructions may collapse or merge with one another as a result of constructional competition (Barðdal 2008).

Recent corpus studies have provided countless examples of utterances that make sense from a semantic point of view, but which only occasionally occur because native speakers overwhelmingly prefer an alternative expression. Examples (16–18), taken from Goldberg (2011: ex. 1–6), show instances of such intelligible, yet unconventional utterances paired to their more conventional counterparts.

(16) a. *??She explained her the news.*
b. *She explained the news to her.*

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

(17) a. *??She considered to go to the farm.*
b. *She considered going to the farm.*

(18) a. *??She saw the afraid boy.*
b. *She saw the boy who was afraid.*

Corpus evidence is probably the strongest argument against any criticism of overgeneration: there are always examples to be found that violate the restrictions that linguists try to impose on their grammars. Many linguists therefore now accept that a language user's linguistic knowledge can only be properly modeled by incorporating usage data (Bybee 2010). Moreover, the fact that unconventional utterances are infrequent or even unattested in the corpora at our disposal does not mean that linguists can simply discard them. One should never forget that corpora only contain what people have already said or written, and not what they would or would not say; and that language has a Zipfian distribution in which the majority of examples only occur once, no matter how big the corpus is (Kay 2011).

A major challenge for linguistics is therefore to account for the creativity with which language users produce and comprehend novel utterances, and explain how they are able to handle uncertainty in language processing. Fortunately, there is already a significant state-of-the-art that can be incorporated in linguistic theory for modeling the kinds of uncertainty faced by speakers and listeners. First, the listener must be able to disambiguate utterances by choosing among hundreds or even thousands of possible analyses. For this task, several measures from probability theory (Charniak 1993; Jurafsky 2003) and information theory (Gibson 1998; Hale 2003; Jaeger and Tily 2011) can be used. The speaker, on the other hand, already knows the meaning and must therefore choose between alternative ways of expressing that meaning and estimate whether the listener will be able to retrieve the intended message. The best strategy for this task is to stick close to the conventions of the speech community, for which methods from statistical physics and opinion dynamics can be used (Steels 1995; Loreto et al. 2011). Such measures have already been successfully implemented in Fluid Construction Grammar in the context of experiments on language evolution (see e.g., Steels 2012b; van Trijp and Steels 2012).

In the case of the FCG-formalism, statistical information can be incorporated in several ways. First, every construction has an "attributes" slot in which information can be kept such as a construction's frequency and degree of entrenchment. Next, the FCG-system also allows the construction inventory to be organized using networks (Wellens 2011). In these networks, constructions are the *nodes* and the many kinds of relations that may exist between these

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

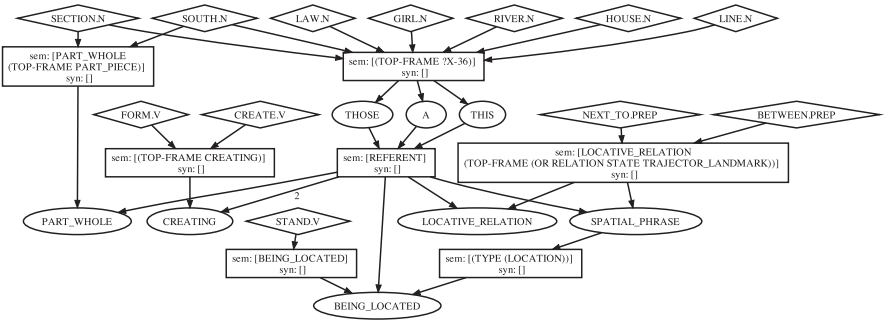


Figure 3: This Figure shows a partial *priming network* for production (adopted from Wellens 2011: Figure 8) that was trained using FrameNet data (Baker et al. 1998) Priming networks capture conditional relations between constructions and co-occurrence data, which can be exploited for robust and efficient processing. In this Figure, lexical constructions that were automatically extracted from the FrameNet database are shown as diamonds. Lexical and grammatical constructions that were manually added are shown as ovals, and the dependency relation between constructions are shown in boxes.

nodes are the *edges*. One example is a *priming network* (see Figure 3) in which the FCG-system automatically creates and maintains edges between constructions that are dependent on each other for application. Such networks make processing more efficient because they prioritize the application of certain constructions depending on which constructions were applied so far. In the simplest case, these networks can thus function in the same way as N-gram models by keeping statistical information about sequences of constructional application on the edges of the network. On top of a layer for routine processing, the FCG-system also has a meta-layer consisting of diagnostics and repairs for handling unseen input (Steels 2011).

4 Conclusion

The construction grammar approach to argument structure has been criticized for allowing overgeneration. This criticism plays a prominent role in the current debates on argument structure, so it is important to examine its importance for the further development of construction grammar as a theory of linguistics.

When tracing back the history of the concept of overgeneration, however, it becomes clear that this criticism is only valid if one assumes that grammar is a generative competence model that is capable of generating all and only the grammatical utterances of a language. I have then exposed the hidden consequences of

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

this assumption by surveying formal evidence proving that lexicalist accounts, in the style put forward by scholars who are critical to argument structure constructions, massively overgenerate themselves. Computational implementations demonstrate that the lexicalist approach can only avoid this problem by giving up on the sharp distinction between competence and performance that is advocated in generative grammar, and instead work with probabilistic methods or careful control strategies. Moreover, decades worth of research in computational linguistics have taught us that the most robust models of language allow overgeneration but filter out improbable analyses using statistical methods.

I argued that once we adopt this alternative model of grammar, which I called a problem-solving model, it is perfectly feasible to implement the constructional account. I have substantiated this claim through an operational implementation of argument structure and coercion in Fluid Construction Grammar, which works for both production and comprehension. The FCG implementation gets rid of the most common design features of lexicalist grammars, such as its type system and lexical rules, and instead empowers individual constructions to freely combine with each other and to decide whether the conventions of the language need to be stretched.

Obviously, this paper does not settle the debate on lexicalist versus constructional approaches to argument structure, for which evidence from other disciplines is needed. It has, however, revealed that a rational debate is currently hindered by the fact that the construction grammar community is split between researchers operating within the paradigm of generative linguistics, and researchers who work within a usage-based approach. If we want to move forward as a field and keep both sides of the paradigmatic rift communicating with each other, future discussions must include more transparency about the underlying assumptions and stronger evidence for substantiating a particular argument. That is, the generativist construction grammarians will have to provide evidence that their competence models are neutral to, but compatible with language processing. The usage-based construction grammarians in their turn have to become much more explicit about how competence and performance interact with each other. Both sides will require an interdisciplinary dialogue between linguists and psychologists, who have a firm understanding of the empirical evidence, and computational linguists, who have the necessary experience in devising representations and processing models.

Acknowledgements: The research reported in this paper has been funded by and carried out in the Sony Computer Science Laboratory Paris. The author would like to thank the editors of *Cognitive Linguistics* and the anonymous reviewers for their

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

invaluable feedback and help in improving this manuscript. He also wishes to thank his colleagues, particularly Luc Steels, Katrien Beuls, Miquel Cornudella Gaya and Paul Van Eecke for their continuous support and for creating such a superb working environment.

References

Baker, Collin F., Charles J. Fillmore & John B. Lowe. 1998. The Berkeley FrameNet project. *Proceedings of the 17th International Conference on Computational Linguistics*. Morristown, NJ: ACL.

Barðdal, Jóhanna (ed.). 2008. *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.

Boas, Hans C. 2003. *A constructional approach to resultatives*. Stanford: CSLI Publications.

Boas, Hans C. 2008a. Determining the structure of lexical entries and grammatical constructions in construction grammar. *Annual Review of Cognitive Linguistics* 6, 113–144.

Boas, Hans C. 2008b. Resolving form-meaning discrepancies in construction grammar. In Jaakko Leino (ed.), *Constructional reorganization*, 11–36. Amsterdam: John Benjamins.

Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based construction grammar*. Stanford: CSLI Publications.

Bos, Johan. 2011. A survey of computational semantics: Representation, inference and knowledge in wide-coverage text understanding. *Language and Linguistics Compass* 5(6). 336–366.

Bouma, Gosse & Gertjan van Noord. 1994. Constraint-based categorial grammar. *Proceedings of the 32nd Annual Meeting of the Association of Computational Linguistics*, 147–154. Las Cruces: ACL.

Briscoe, Ted & Ann Copestake. 1999. Lexical rules in constraint-based grammars. *Computational Linguistics* 25(4). 487–526.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Carpenter, Bob. 1991. The generative power of categorial grammars and head-driven phrase structure grammars with lexical rules. *Computational Linguistics* 17(3). 301–313.

Charniak, Eugene. 1993. *Statistical language learning*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague & Paris: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Croft, William. 1998. Event structure in argument linking. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: Lexical and compositional factors*, 21–63. Stanford: CSLI Publications.

Croft, William. 2003. Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in language studies: Studies in honour of Günter Radden*, 49–68. Amsterdam: John Benjamins.

Fillmore, Charles J. 1970. The grammar of hitting and breaking. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 120–133. Waltham, MA: Glinn and Company.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

Fillmore, Charles J. 1976. Frame semantics and the nature of language. In Steven R. Harnad, Horst D. Steklis & Jane Lancaster (eds.), *Origins and evolution of language and speech*, 20–32. New York: New York Academy of Sciences.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

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

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

Goldberg, Adele E. 2011. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22(1). 131–154.

Goldberg, Adele E. & Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80(3). 532–568.

Hale, John T. 2003. The information conveyed by words in sentences. *Journal of Psycholinguistic Research* 32(2). 101–123.

Iwata, Seizi. 2002. Does MANNER count or not? Manner-of-motion verbs revisited. *Linguistics* 40(1). 61–110.

Iwata, Seizi. 2008. *Locative alternation: A lexical-constructional approach*. Amsterdam: John Benjamins.

Jaeger, T. Florian & Harry Tily. 2011. On language 'utility': Processing complexity and communicative efficiency. *WIREs: Cognitive Science* 2(3). 323–335.

Jurafsky, Dan. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*. Cambridge, MA: MIT Press.

Jurafsky, Dan & James H. Martin. 2000. *Speech and language processing. An introduction to natural language processing, computational linguistics, and speech recognition*. New Jersey: Prentice Hall.

Kay, Martin. 1979. Functional grammar. *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society*, 142–158. Berkeley: BLS.

Kay, Martin. 2011. Zipf's law and l'arbitraire du signe. *Linguistic Issues in Language Technology* 6. http://elanguage.net/journals/lilt/article/view/2584 (accessed 17 February 2015).

Kay, Paul. 2005. Argument structure constructions and the argument-adjunct distinction. In Mirjam Fried & Hans C. Boas (eds.), *Grammatical constructions: Back to the roots*, 71–98. Amsterdam: John Benjamins.

Kay, Paul & Laura A. Michaelis. 2012. Constructional meaning and compositionality. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 3, 2271–2296. Berlin: Mouton De Gruyter.

Keenan, Edward L. & Leonard M. Faltz. 1985. *Boolean semantics for natural language*. Dordrecht: D. Reidel.

Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge: Cambridge University Press.

Loreto, Vittorio, Andrea Baronchelli, Animesh Mukherjee, Andrea Puglisi & Francesca Tria. 2011. Statistical physics of language dynamics. *Journal of Statistical Mechanics: Theory and Experiment*. doi:10.1088/1742-5468/2011/04/P04006

Meurers, Detmar & Guido Minnen. 1997. A computational treatment of lexical rules in HPSG as covariation in lexical entries. *Computational Linguistics* 23(4). 543–568.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51

Michaelis, Laura. 2013. Sign-based construction grammar. In Thomas Hoffman & Graham Trousdale (eds.), *The Oxford handbook of construction grammar*. Oxford: Oxford University Press.

Morita, Joe. 1998. Some notes on prepositional resultatives. *Tsukuba English Studies* 17. 319–340.

Müller, Stefan. 2006. Lexical or phrasal constructions? *Language* 82(4). 850–883.

Müller, Stefan & Stefan Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76.

Nemoto, Noriko. 1998. On the polysemy of ditransitive *save*: The role of frame semantics in construction grammar. *English Linguistics* 15. 219–242.

Penn, Gerald. 2012. Computational linguistics. In Ruth Kempson, Tim Fernando & Nicholas Asher (eds.), *Philosophy of linguistics*, 143–174. Amsterdam: North Holland.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago & Stanford: University of Chicago Press/CSLI Publications.

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, 359–377. West Sussex: Wiley-Blackwell.

Steels, Luc. 1995. A self-organizing spatial vocabulary. *Artificial Life* 2(3). 319–332.

Steels, Luc (ed.). 2011. *Design patterns in Fluid Construction Grammar*. Amsterdam: John Benjamins.

Steels, Luc (ed.). 2012a. *Computational issues in Fluid Construction Grammar*. Heidelberg: Springer Verlag.

Steels, Luc (ed.). 2012b. *Experiments in cultural language evolution*. Amsterdam: John Benjamins.

Steels, Luc & Joachim De Beule. 2006. Unify and merge in Fluid Construction Grammar. In Paul Vogt, Yuuga Sugita, Elio Tuci & Chrystopher Nehaniv (eds.), *Symbol grounding and beyond: Third international workshop on the emergence and evolution of linguistic communication (EELC2006)*, 197–223. Berlin: Springer Verlag.

van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Luc Steels (ed.), *Design patterns in Fluid Construction Grammar*, 115–145. Amsterdam: John Benjamins.

van Trijp, Remi. 2015. *The evolution of case grammar*. Berlin: Language Science Press.

van Trijp, Remi & Luc Steels. 2012. Multilevel alignment maintains language systematicity. *Advances in complex systems* 15(3–4). 1250039.

Wellens, Pieter. 2011. Organizing constructions in networks. In Luc Steels (ed.), *Design patterns in Fluid Construction Grammar*, 181–201. Amsterdam: John Benjamins.

Bereitgestellt von | De Gruyter / TCS
Angemeldet
Heruntergeladen am | 08.04.16 13:51