# Constructions emerging : a usage-based model of the acquisition of grammar

Beekhuizen, B.F.

# CHAPTER 1

# Introduction

When I utter the sentence *John kissed Mary*, anyone having a sufficient command of English will understand that I make an assertion, namely that an event took place in which some person whom we both know, named *John*, engaged in the act of kissing another person, again known to both of us, named *Mary*. In comprehension, language users connect an observable signal, in this case the string of sounds produced when uttering *John kissed Mary*, to an unobservable conceptualization of the situation and the speaker's communicative intent. In production, the reverse process takes place: given a conceptualization of a situation and a communicative intention, the speaker tries to figure out which observable signals to use in order for the hearer to arrive at the desired conception of the situation and communicative intent.

At the heart of linguistics is the question how observable signals, such as speech or sign, are connected to conceptualization. Various theoretical frameworks have been developed to account for the connection. Whereas there is widespread agreement between various recent frameworks concerning the question how words work (everyone harks back to de Saussure's (1916) idea of a word being a symbol, i.e., a conventional pairing of a signifying form and a signified meaning), the paths separate when it comes to the question how words are combined into larger units, such as phrases (*the red ball*, *deeply enlightened*) and whole clauses (*The red ball seems deeply enlightened*). Generative grammar, from the 1950s (Chomsky 1957) up to its various present-day incarnations (Chomsky 1993), argues that at the core of the human ability to form complex linguistic representations is a cognitive mechanism for structure-building that is autonomous, i.e., whose properties cannot be de-

rived from other cognitive domains. Construction grammar, the theoretical perspective constituting the starting point of this dissertation, provides a different perspective: the cognitive representations responsible for comprehending and producing complex utterances are **constructions** – like words, conventional pairings of an (observable) signal and a signified meaning (e.g. Goldberg 1995). Because the content and structure of these representations fully comes from other cognitive domains, a language user's grammar is not an autonomous cognitive system. Furthermore, if the representations responsible for building complex linguistic structure are pairings of an observable signal and an inferred meaning, they are qualitatively the same as regular words, and as such, according to construction grammar, all linguistic representations can be regarded as constructions.

The magnificent task faced by an infant being born into a community of speakers, is to figure out how the connections between the observable signals and the meanings work. Again, various theoretical frameworks differ in how they conceive of this task: in the Generative tradition, rooted in rationalist thought, the language learner's task is to deduce which properties from among a finite set of possibilities the grammar of her community's language has (Baker 2001). The constructivist approach, on the other hand, argues that children build up their inventory of linguistic representations in a bottom-up way, and without any preconceptions concerning grammar-wide regularities.

In the literature, one often finds a comparison of the frameworks, where empirical data is presented as being suggestive for the truth of the one and the falsehood of the other framework. To my mind, this approach is unwarranted given the state of any current theoretical framework. Because of their informal nature, any datapoint presented by an adherent of one framework can be brought into accordance with another framework or simply be dismissed as non-data (because it is not part of the 'core' of language, or because it is a 'theory-internal matter'). Of course, this possibility exists even with highly explicit and formalized theories in other fields of research, but it seems that challenges to any linguistic theory can be resolved too easily. This is not to lament the state of linguistics: independently, various theories have internally developed themselves to an enormous extent.

A more productive strategy would be to devote one's energy to the maturation of the theoretical framework by scrutinizing the theoretical construct on which it rests. This dissertation should be read as an attempt to do so. In it, I employ formalization and computer simulation as tools for achieving a deeper understanding of the theoretical framework of construction grammar. Results from the modeling work presented in this dissertation that confirm the line of reasoning of the theoretical vantage point are interesting and provide a basic sanity check of the relation between the formalized model and the informal theory. More insightful, however, are the cases where we fail to simulate a phenomenon. In those cases, it is either the model that is not faithful to the theory, or there are gaps in the theory. It is in the recognition of these gaps that theoretical progress can be made.

## 1.1 Early grammar

Before we turn to the more theoretical issues, let us have a brief look at what kinds of phenomena this dissertation will be occupied with. Children's early linguistic productions differ in several ways from the utterances adults produce. Their utterances are typically shorter, and markers of grammatical categories such as tense and number are mostly omitted (Brown 1973, 98-99). The main phenomenon of interest in this research, however, is the realization of argument-structure patterns over development. The event expressed by a verb has certain roles, which are linguistically realized as the arguments of that verb. Some examples of deviations from the language of adults are shown below.

(1) open drawer (Kathryn 2;0[1], opening a drawer, Bloom, Lightbown & Hood (1975))

(2) I made (Eric, 1;1 1, just reassembled a train, Bloom et al. (1975))

(3) put truck window (Adam 2;3, Brown (1973))

(4) pick up (.) puzzle up (Adam 2;6, wants to pick the puzzle up, Brown (1973))

(5) Adam fall toy (Adam 2;3, dropped a toy, Brown (1973))

(6) eat Benny now (Ben, between 1;7 and 2;6, wants to eat Sadock (1982), cited in O'Grady (1997, 61))

(7) the bridge knock down (Aran 2;4, knocked the bridge down; Manchester corpus, cited in Marcotte (2005))

Examples (1)-(4) show how elements of the argument-structure patterns that are grammatically obligatory for adult speakers of English are omitted. We find subjects and objects being left unexpressed, for instance in examples (1) and (2), but also obligatory prepositions, as in example (3). As Bloom et al. (1975) note, children often produce patterns expressing different aspects of the event they want to express in a sequential way. Example (4) is such a case: to express an event that would be expressed by an adult as *I want to pick up the puzzle*, or *I want to pick the puzzle up*, the child uses two structures, separated by a short pause, seemingly because he was not able to integrate both under a single syntactic constellation. Some basic findings with regards to the argument omissions are that more arguments are expressed over time (Tomasello 1992, 244) and that subjects more frequently omitted than objects (Bloom et al. 1975).

In examples like (5)-(7), we see cases of children's productions where not only are elements omitted, but also rules applied that are not in line with the adult usage, so-called errors of commission. In a case like example (5), the

[1] I adopt the conventional <year>;<month> notation here for the child's age.

verb *drop* would be used in the transitive frame, instead of *fall*, which cannot be used in a transitive syntactic frame. In example (6), the child's wish would be expressed with a pre-verbal subject and *want to*, as in *Benny wants to eat now*, but we find the child changing the word order of the subject and the verb. From this piece of behavior it is hard to glean what the underlying representation might have been: is the child applying a grammatical rule at all, and if so, which one? Example (7) is the mirror image of example (5): here, the noun expressing the patient role is expressed pre-verbally, in the position where subjects are typically found. Some verbs, such as *roll*, allow for the alternation whereby the semantic patient role is expressed as the syntactic subject, but *knock down* is not among them. Again the question is: which representations underlie this production?

From a developmental perspective, the occurrence of both errors of omission and commission in the same time span (roughly, Roger Brown's Stage I) is interesting, because the underlying representations and mechanisms produce both of them. As acquisitionists, we thus face the puzzle of how the learner both under- and overshoots the target. Despite decades of work on these early productions, a comprehensive account of the representations and cognitive mechanisms leading to these productions, and their development over time, has not been satisfactorily given.

## 1.2 Theoretical background

The constructivist theory of language acquisition, briefly introduced earlier, constitutes the starting point for our understanding of the phenomena I just outlined. Construction grammar centrally makes a representational claim: all linguistic representations are pairings of signifying elements, prototypically phonological form, and signified meaning or conceptualization. This holds for words, but also for grammatical patterns. A complementary addition to the representational claims of construction grammar is the usage-based perspective. This perpective holds that linguistic representations are qualitatively and quantitatively grounded in language use (Langacker 1988). The qualitative grounding is taken to mean that the representations consist of the cognitive reflections of usage events. Importantly, only representational content derived from the processed usage events can be part of the linguistic representations. That is to say: the language learning child has no preconception of contentive elements that are expected to be part of linguistic representations. An important consequence is that there is no place in a usage-based constructivist theory of language for universal syntactic categories, such as 'noun' and 'verb'.[2] At the core of linguistic representations should be elements of the observable signal (sound structure for spoken languages) and elements of the

[2]The rejection of universal distributional or syntactic categories follows not only from Langacker's (1988) ideas about the grounding of linguistic representation in usage. The most explicit rejection is given by Croft (2001) on the basis of language-typological arguments.

conceptualization of the world that these signaling elements are assumed to refer to.

Given this perspective, the language-acquiring child faces the task of building up an inventory of representations allowing her to communicate successfully with the other members of her community. To do so, she has several, highly general mechanisms at her disposal. Primarily, these involve mechanisms of understanding other people's intentionality, which emerge around the first birthday, and a set of abilities to recognize patterns (Tomasello 2003). The former allow the child to recognize that the speaker has communicative intentions with the speech signal he is producing and roughly what this communicative intention encompasses. The latter enable the learner to discover patterns of regularity in the signal and inferred communicative intent. These patterns are discovered, furthermore, in a bottom-up and gradual way. If we assume that the child has no preconceptions of the content of the linguistic representations, any abstraction over the processed usage events can be expected to arrive through a comparison of the structure of various usage events. Abstraction in the inventory of linguistic representation therefore only emerges after multiple comparable usage events have been observed.

## 1.3 Computational cognitive modeling

The past decade has seen a large number of publications on computational models of the acquisition of grammar from a usage-based constructivist perspective. We find many learning models taking semantics into account, in line with the constructivist tenet that linguistic representations consist of pairings of form and meaning throughout. This does not only include work explicitly being framed as being constructivist, such as the dissertation of Chang (2008) and Alishahi & Stevenson's (2008) clustering approach to argument-structure constructions, but also modeling research in other frameworks such as Combinatorial Categorial Grammar (e.g. Kwiatkowski 2011). Distributional learners, taking only distributional properties of the phonological form into account, have shown how multi-word units, which are assumed to play a large role in language acquisition on the usage-based perspective, are extracted (McCauley & Christiansen 2014*a*), how the integration of the constructivist view and a connectionist, sub-symbolic representation can be achieved (Borensztajn 2011), how grammatical patterns assumed to be unlearnable can be learned from usage data (Bod 2009), and how early patterns of language production can be modeled (Freudenthal, Pine & Gobet 2010). Finally, highly interesting work on the appropriate analysis of corpus data of children's early production have sparked interesting, and perhaps even productive back-and-forths between adherents of the usage-based perspective and generativist researchers (Lieven, Salomo & Tomasello 2009, Yang 2011).

The development of computational models is interesting for several reasons. First of all, in developing a computational model, the researcher is forced

to operationalize the concepts of a linguistic theory at a level of precision that allows a computer to run it as a piece of software. This means that the modeler has to make design choices in the representations and algorithms of the model that are not specified in the theory. The observed consequences of these design choices can then be 'fed back' into the theory. For this reason, it are not only the success stories that are interesting in modeling. In fact, the rhetorics of success can be misleading: when we regard modeling as an extension of theorizing, a model's failure is essentially the exclusion of a logically possible variant of the theory. Through successive failures, we can gradually delimit the range of potentially successful theoretical options. Unfortunately, modeling failures are rarely shown.

Secondly, modeling allows us to observe the interaction of various components of the model. Rather than isolating the phenomenon of interest, as is often done in experimental studies, modeling allows us to observe what the effects are if multiple components of a theory interact in processing the data (cf. Beekhuizen, Bod & Verhagen 2014). This does not mean that all models do so, or should do so, but it is a possibility of the method that I believe is worth exploring. A further consequence of this approach is that we can work towards comprehensive models, that is, models being able to fully comprehend and produce utterances.

Finally, modeling allows us an evaluation of the theory both on a wide level and a narrow level. We can model both how an operationalization of the theory behaves across the board (e.g., in understanding or producing novel utterances), but also how the operationalized theory behaves in a simulation of a particular experimental set-up or in the case of a rare event. Both lines of inquiry are important: with linguistic theory often focussing on rare events (crossing dependencies, sentences like *Pat sneezed the foam off the capuchino*), the more global behavior of the theory is often left out of consideration.

## 1.4 Goals of this research

It is to the background of the modeling inquiries presented in section 1.3 that the line of research reported on in this dissertation starts. At the time when I started it, there were several issues that I thought to be insufficiently addressed in existing work. Four of them constitute the central theoretical issues of this dissertation: a call for greater **comprehensiveness** of usage-based computational models, a scrutiny of the **conception of learning**, a plea for more **naturalism** when the acquisition of meaning is concerned, and the reassessment of the **starting-small** position.

As such, the research presented in this dissertation should be regarded primarily as a theoretical exercise. It involves reinterpreting, scrutinizing, synthesizing, operationalizing, and, finally, evaluating ideas from the usage-based approach to grammar. The conceptual work, to my mind, is an important part of the endeavour of computational cognitive modeling. On an epistemological

note: this does not suggest that the conceptual work has any kind of primacy over the more empirically oriented work: both operate in a heuristic loop, informing and enhancing each other.

### 1.4.1 Providing a comprehensive model

The usage-based approach to language is a non-modular approach, meaning that what in componential theories are called 'the lexicon' and 'the grammar' are not distinct entities, both conceptually and cognitively. Because of this, it is not possible to fully isolate the acquisition of lexical constructions ('words') from that of grammatical constructions. At the very least, the mechanisms involved in them operate in lockstep: it is unthinkable (from any theoretical perspective, really) that the child waits until she has acquired an adult-like lexicon before she starts to figure out the grammatical rules. Starting from a usage-based perspective, even the weaker form (the child waits until it has a lexicon of some size before it starts learning the grammar) is not satisfying. As both are symbolic representations, roughly the same set of mechanisms have to be involved in acquiring them from situated utterances, and as such it has to be possible for a learner to pick them up at the same time. Except for a model that is, interestingly enough, not framed as a usage-based model (Kwiatkowski 2011), none of the models listed in section 1.3 does this. Models that do incorporate meaning start with many lexical form-meaning pairings already acquired (Chang 2008, Alishahi & Stevenson 2008). The first goal of this dissertation therefore, is to develop a computational model of language acquisition that starts with no representational symbolic content and that acquires lexical and grammatical constructions at the same time.

The comprehensiveness of a model bears on another issue too. If language users are able to form utterances on the basis of some conceptualization and a linguistic inventory and understand language on the basis of the utterance and the same linguistic inventory, we have to model the task of language use on both ends. That is to say: we want a model to be able to produce utterances from a conceptualization of some situation and to comprehend utterances with its inventory of linguistic representations to arrive at an interpretation. Again, many computational models model only part of this task and do not show how they can account for comprehension on the basis of an utterance, and production on the basis of only a meaning to be expressed.

### 1.4.2 The conception of learning

When modeling the acquisition of grammar, cognitive modelers often find inspiration in learning algorithms developed in computational linguistics or artificial intelligence. Even if they turn out to be descriptively adequate, it remains important to reflect upon the conception of learning they encompass. Usage-based theorizing adheres to a general empiricist approach to learning, in which categories are induced from the input rather than deduced given

a pre-existing hypothesis space, as rationalist approaches have it. This does not mean that there is no hypothesis space: given the nature of our cognition, certain representations are possible, whereas other, logically possible ones, are not. This 'design space', however, is not very informative, and it is mainly the learner's exploration of the space through language use that is of interest for usage-based theorists. The learning algorithms used in usage-based models typically reflect this (e.g., unsupervised clustering in Alishahi & Stevenson (2008) and Bayesian Model Merging in Chang (2008)).

Another question is whether learning involves a rational decision making process. Linguistic production and comprehension arguably involves such processes: the speaker selects from his inventory a number of representations that allow him to produce of comprehend an utterance. Whether the learning process is also affected by (subconscious) decision making, is another issue. If we follow Langacker (1988) in the argument that learning is merely the effect of processing linguistic usage events, are there any cognitive mechanisms that only affect the learning taking place after the processing of the usage events but not the processing itself? This is not a question that I believe has been answered yet, but for the research presented here, I believe it to be the best null hypothesis to start from the idea that all learning takes place in the processing of usage events and the reinforcement of representations used in processing.

### 1.4.3 Starting small

One of the central tenets of construction grammar, as well as the usage-based perspective, is the idea that language users compute language with symbolic representations that are 'bigger' than a single word. This focus has led to the view in language acquisition that children often first acquire larger, unstructured wholes, 'chunks' of linguistic material, which they subsequently start breaking down into their component parts.

There is definitely a lot to be said for this 'starting-big' approach. However, the wholes-to-parts account of language acquisition may be overemphasized in current usage-based approaches. Especially when we look at early production, it seems that, besides the use of chunk-like structures, there is also a gradual build-up of used representations taking place. If children produce increasingly more arguments with a verb, for instance, I believe it is likely that they are extending existing representations with additional valency roles, rather than learning the larger unit first as a chunk, then breaking it down, and only then recognizing its similarity with earlier acquired representations.

### 1.4.4 Naturalism in meaning

A final issue that is central to this thesis is naturalism in the acquisition of any linguistic structure involving conceptual structure, or meaning. Many modeling approaches, but psycholinguistic studies as well, make simplifying assumptions concerning the question how available the conceptualization of the

situation is to the child. If children are able to understand the communicative intention of the speaker, they must derive the content of this intention from some source. This source is often thought to be the situational context, but if we assume the situational context to provide the learner with a source for her meanings, we face a number of problems. The situational context, after all, is a confusing place if you want to learn meaning. Even granted that the child has joint-attentional skills that narrow down the set of conceptualizations that the speaker can possibly be trying to convey, there are still often many conceptualizations of the same situation possible, and often many situations co-present with the utterance that are likely to be expressed. On the other hand, some situations may simply be absent from the consideration of the learner. If we want to simulate the acquisition of meaningful structures, we have to make a realistic assessment of the overwhelmingness of other possible conceptualizations and the frequency of the absence of the situation which the speaker is trying to express from the learner's consideration.

## 1.5 A note on notation

In this thesis, I adopt Langacker's (1987) Cognitive Grammar shorthand formalism for describing constructions. I briefly describe this notation at the outset, because I will use it throughout the dissertation. The formalism works as follows. Conceptual structure is represented in small capitals and phonological structure in italics. As a conceptual structure in the case of the model to be developed often has many features (e.g., the feature set behind the meaning of *daddy* being {ENTITY, ANIMATE, MALE, PERSON, FAMILY-MEMBER, FATHER}), I use only the most concrete, or otherwise most recognizable feature to represent the set of features (so *daddy* would signify FATHER). A **unit**, or construction, is a linguistic representation of a number of constituents where each constituent is given between square brackets.

If, in a constituent, a phonological and a conceptual structure are present, the phonological structure can be said to **symbolize** the conceptual structure, which is represented with a slash sign. The whole of constituents is delimited by square brackets as well. If the unit signifies conceptual structure beyond the meaning in the constituents, this meaning will be represented after the formal structure, marked with a pipe '|'.[3] In many cases the non-compositional meaning is not directly relevant, and will be omitted for the sake of space. Examples are given in (8) and (9).

(8) [ BALL / *ball* ]

(9) [ [ HUMAN ] [ GRAB / *grab* ] ] | GRAB(GRABBER(HUMAN))

Using this formalism, we can also describe complex analyses or constructs. Suppose the slot of the construction in (9) is filled (by means of what Lan-

[3] Here, my notation differs from Langacker's.

gacker calls 'elaboration') with a construction [ FATHER / *daddy* ]. In that case we denote this slot-filling operation with an arrow after the slot, followed by the unit filling the slot, as in example (10).

(10) [ [ HUMAN ]→[ FATHER / *daddy* ] [ GRAB / *grab* ] ] |
GRAB(GRABBER(FATHER))

Another feature of Langacker's notation is the use of parentheses for extensions of known units. I will employ this notation in several cases. To give an example: the model to be presented allows for the concatenation of two partial analyses of an utterance without any construction governing both of them. We find a case of this concatenation operation in (11), where the partial analysis in (10) is concatenated with the unit given in (8).

(11) ( [ [ HUMAN ]→[ FATHER / *daddy* ] [ GRAB / *grab* ] ] [ BALL / *ball* ] ) |
GRAB(GRABBER(FATHER),GRABBEE(BALL))

Similarly, the notation in parentheses will be used for the so-called 'bootstrapping' of unknown words into open slots of constructions. When a word is bootstrapped into a schematic position of a construction, a novel representation is created that links the hypothesized meaning of that word to its form. An example of the notation of the bootstrapping process is given in (12), where the unknown word *epidemiologist* is bootstrapped into the slot of the construction we saw before in examples (9) and (10).

(12) [ [ HUMAN ]→( EPIDEMIOLOGIST / *epidemiologist* ) [ GRAB / *grab* ] ] |
GRAB(GRABBER(EPIDEMIOLOGIST))

## 1.6 Overview of the dissertation

The core of this dissertation is the model, the Syntagmatic-Paradigmatic Learner (SPL), presented in chapter 3. I set out the theoretical and empirical issues which I believe are worth studying using a computational model in chapter 2. After this theoretical core, we first look at the issue of the acquisition of meaning in chapter 4. This chapter does not constitute a modeling experiment, but rather an observational study into the observable sources of meaning. The results of this study are used in the subsequent three chapters that report several aspects of simulation experiments performed with the model. In chapter 5, first, we look at the model's ability to comprehend utterances in a noisy situational context. Chapter 6 presents a 'look under the hood': what is the nature of the representations acquired over time by the model, and what learning mechanisms are involved in doing so. In chapter 7, finally, I discuss the results of several production studies: what happens if we give the model a situation and ask it to express it.