

Usage-based and emergentist approaches to language acquisition*

HEIKE BEHRENS

Abstract

It was long considered to be impossible to learn grammar based on linguistic experience alone. In the past decade, however, advances in usage-based linguistic theory, computational linguistics, and developmental psychology changed the view on this matter. So-called usage-based and emergentist approaches to language acquisition state that language can be learned from language use itself, by means of social skills like joint attention, and by means of powerful generalization mechanisms. This paper first summarizes the assumptions regarding the nature of linguistic representations and processing. Usage-based theories are nonmodular and nonreductionist, i.e., they emphasize the form-function relationships, and deal with all of language, not just selected levels of representations. Furthermore, storage and processing is considered to be analytic as well as holistic, such that there is a continuum between children's unanalyzed chunks and abstract units found in adult language. In the second part, the empirical evidence is reviewed. Children's linguistic competence is shown to be limited initially, and it is demonstrated how children can generalize knowledge based on direct and indirect positive evidence. It is argued that with these general learning mechanisms, the usage-based paradigm can be extended to multilingual language situations and to language acquisition under special circumstances.

1. Introduction

What is the exact form and content of the uniquely human capacity to learn language? There must be a genetic component in this capacity because every normally developing child is able to learn language, and there must be an environmental component because no one is born with a specific language.

A powerful solution to the acquisition problem is the assumption that innate linguistic structure helps the child overcome the (possible) underspecification of language structure in the input. In this view, the input has to be mapped onto innate linguistic categories, but the categories or principles of core syntax do not have to be learned because they are there right from the beginning. But innate structure, as rich as it may be, cannot account for the acquisition of language-specific properties, such as the lexicon of a language or the inflectional morphology of German or Dutch. These language-specific properties must be derived from the input over a number of years. In contrast, the so-called *emergentist* and *usage-based* approaches to language acquisition, that have become prominent in the past decade, are based on the assumption that language structure can be learned from language use by means of powerful generalization abilities (e.g., Elman et al. 1996; Tomasello 2003a). The rationale of acquisition theory has changed accordingly: if the child has to learn the irregular and peculiar aspects of a language by general learning mechanisms, these mechanisms should suffice to learn the more general and predictable patterns of that language as well.

In Section 2 the theoretical assumptions underlying usage-based approaches to language and language acquisition will be reviewed, and in Section 3 the major lines of empirical research that test and provide evidence for these assumptions will be presented. The paper concludes with a short discussion of the application of usage-based approaches to other situations regarding language learning like multilingualism and language disorders (Section 4) and an outline of the major open issues and fields for new developments (Section 5).

2. Theoretical issues

2.1. Usage-based linguistics

To explain language learning it must be clear what has to be learned. Thus, language acquisition theories are intertwined with linguistic theories on the representation of linguistic structure in adults. The usage-based approach to language acquisition relies on insights from cognitive linguistics, a nonmodular theory that assumes that linguistic structure is tied to the semantics and pragmatics it encodes (e.g., Langacker 1987, 1988, 2000; cf. Kemmer and Barlow 2000 for a summary). The term “usage-based” goes back to Langacker’s assumption that a speaker’s linguistic system is grounded in concrete usage events or utterances (Langacker 1987). This means that the linguistic system is built-up from usage

events of particular symbolic units. With increasing linguistic experience, more abstract linguistic patterns may evolve, but still the assumption is that these more abstract patterns are grounded in usage. Joan Bybee summarizes the usage-based perspective.

While all linguists are likely to agree that grammar is the cognitive organization of language, a usage-based theorist would make the more specific proposal that grammar is the cognitive organization of one's experience with language. [...] [C]ertain facets of linguistic experience, such as the frequency of use of particular instances of constructions, have an impact on representation that we can see evidenced in various ways, for example, in speakers' recognition of what is conventionalized and what is not, and even more strikingly in the nature of language change. The proposal [...] is that the general cognitive capabilities of the human brain, which allow it to categorize and sort for identity, similarity, and difference, go to work on the language events a person encounters, categorizing and entering in memory these experiences.

The result is a cognitive representation that can be called a grammar. This grammar, while it may be abstract, since all cognitive categories are, is strongly tied to the experience that a speaker has had with language. (Bybee 2006: 711)

In usage-based theory there is no level of grammar that is independent of language use, rather, all abstract grammatical rules were at some point induced from concrete and particular usage events.¹ Langacker specifies the mental representation of a usage-based grammar as follows:

The grammar lists the full set of particular statements representing a speaker's grasp of linguistic conventions, including those subsumed by general statements. Rather than thinking them an embarrassment, grammarians regard particular statements as the matrix from which general statements (rules) are extracted. (Langacker 1987: 46)

A number of implications follow from this statement: Cognitive Grammar is nonreductionist and maximalist: it does not strive to reduce language to as abstract a rule system as possible, because particular (lexically specific) and abstract phenomena are the same in kind, namely symbolic form-function units. Parsimony of storage and representation is not the goal of the theory, nor the underlying assumption of how grammar works. Consequently, usage-based grammar is maximalist, because it considers idiosyncratic phenomena, low-level schemas, as well as very productive schemas with general, rule-like properties. Langacker explicitly rejects formalist approaches to grammar which make a division between words and rules that operate words (e.g., Pinker 1999), and considers the division between *rules* and *lists* a fallacy because fully regular as

well as subregular phenomena can be handled in the same fashion without evoking two dichotomous processes (Langacker 2000: 57–60).

But how do the generalization of schemas and regular aspects of a language work? Here, usage-based grammar relies on established learning mechanisms as attested in other cognitive domains as well (Tomasello 1998a: ix–xiv).

2.2. *Generalization in usage-based theories*

2.2.1. *Entrenchment, categorization, and schema formation.* In usage-based linguistics, a few basic psychological processes can account for a linguistic structure (cf. Langacker 2000: 3–5; see Langacker 1987: 485–495 for a glossary of the terms used): A central cognitive phenomenon is *entrenchment*, the fact that repeated encounter of a unit leaves memory traces that stabilize the more often this unit recurs. Entrenchment is involved in psychological processes such as routinization and automatization, and applies to smaller units like words as well as “prepackaged” larger units or constructions, if they can be retrieved without attention to detail. However, repetition alone does not lead to the abstraction of more general information. In order to generalize and form categories, the mind must recognize similarities as well as dissimilarities. It filters out aspects that do not recur, and registers commonalities by comparing stored with new units. New units are *categorized* along those dimensions where similarities with stored units are detected. Through abstraction and generalization, *schemas* are formed. Langacker (1987: 492) defines a schema as a “semantic, phonological, or symbolic structure that, relative to another representation of the same entity, is characterized with lesser specificity and detail”. Schemas unite experiences from an overarching perspective. We can look at one and the same entity from different perspectives and degrees of granularity. For example, a speaker can assess commonalities between usage-events from a coarser perspective even when the items do not have much in common on the surface level: There is a range of low-level schemas that share, for example, a lot of lexical material up to very abstract schemata (e.g., transitive sentences as an abstract category). Schema formation is made possible, for example, by type variation. Whereas high token frequency leads to entrenchment and storage as “chunk”, type frequency leads to the recognition of analogies between constructions (Bybee 2006, see Section 3.2.1 below).

Schema formation and categorization can account for the extraction of very abstract rule-like phenomena that are not tied to specific lexical material.

2.2.2. *Compositionality and emergence.* Since usage-based approaches emphasize the necessity of experience in order to form linguistic and non-linguistic categories, one might conclude that learners cannot go *beyond* their experience. However, since children have powerful generalization mechanisms, they can generalize new structures based on previous experience (O'Grady 2005; cf. Gentner 2003; see Section 3.3 below for examples). Schemas can be integrated to form larger units by means of *composition*. It is characteristic of the new unit that compositionality is only partial, because the composite structure does not just equal the sum of the parts (Langacker 2000: 3–5). In particular, the composite structure typically has new qualities that *emerge*. For example, combining two sounds to a cluster will change the motor patterns necessary for their execution.²

The term *emergentism* emphasizes the idea that qualitatively new and more complex structures can emerge from simpler, basic facts (MacWhinney 1999: ix). Emergence is thus a central component of the human language learning capacity since linguistic knowledge emerges from the child's interaction with the ambient language.

2.3. *Nature and nurture: evolution, species specificity, and modularity*

The main tenet of usage-based and emergentist theories of acquisition is that the linguistic structure is an emergent property of language use, i.e., the child is not innately equipped with specifically linguistic representations (“representational nativism”). Marchman and Thal (2005) summarize the difference as follows:

Both emergentist and nativist approaches share the appreciation that the acquisition of grammar is a very complex and special human accomplishment. In the nativist view, however, children are special because they “have” something (i.e., a domain-specific genetic endowment for particular kinds of representations with particular kinds of computational processes). In an emergentist view, in contrast, children are special because what they have enables them to *do something*, i.e., they construct an impressive system of grammar using domain-general skills. (Marchman and Thal 2005: 144)

The insistence on inductive learning does not mean that usage-based theories assume that children's mind is a blank slate, to be conditioned by behaviorist conditioning practices (Bates 1997; Tomasello 2003a). The very principle of emergentism is that a combination of smaller quantitative differences can produce a new quality, for example, human language as opposed to the communication systems found in other species

(Elman et al. 1996; Tomasello and Rakoczy 2003). Regarding universals, the notion of *emergence* is of particular importance because it allows us to account for universal behavior and structures without having to assume innate representation that is responsible for this universal manifestation. For example, all human beings with unimpaired motor skills eat with their hands rather than their feet. This can be explained without assuming an “eating-with-hands gene” by the assumption that all humans share the same physical endowment (limbs, mouth) and face the same task (the need to eat). To insert food to mouth with hands proves to be the most efficient way to solve this task, although it is by no means the only possible way. Likewise, when standing in line at the supermarket, all lines will tend to be equally long. This is not because humans have a genetic predisposition to make queues match in length. Rather, matching queue length is an emergent result of the individual’s desire to spend as little time as possible waiting in line. Regarding language, emergence can be observed in a number of phenomena: the basic idea is that language structure emerges from the processes of listening and producing speech (MacWhinney 1999: xi). For example, Gupta and Dell (1999) argue that the structural properties of language emerge from serial order and procedural memory.

Gentner (2003) sees the lack of innate linguistic endowment as an evolutionary advantage. She argues that an organism has higher adaptive power if the innate component is as small as possible, and if the organism has powerful generalization abilities instead. While innate specifications at first glance seem to facilitate development, a richly specified innate component makes organisms more vulnerable and less flexible or adaptive, whereas a powerful “generalizer” can easily adapt to new circumstances. In sum, the concept of emergence tries to explain how relatively small genetic and behavioral differences lead to wide-ranging differences in cognitive abilities, including the competence to use a full-fledged linguistic system (see Elman et al. 1996 for a detailed explanation).

If it can be shown that the symbolic units of a language can be derived from a human’s experience with language, innate language-specific representations are not necessary (see also Elman et al. 1996; MacWhinney 1999; Tomasello 2003a):

Just as plausible [...] is the hypothesis that language rests on more general biological predispositions, such as the abilities to create and learn symbols, to form concepts and categories, to process vocal-auditory information rapidly, and to interact and communicate with other people intersubjectively. (Tomasello 1998a: xi)

The human capacity to learn language may be enhanced by humans’ advanced social cognition (e.g., greater tendency to use joint attention

and intention reading, greater tendency to transmit knowledge explicitly, greater skills of pattern recognition, and an advanced ability for inferring, as compared to nonhuman primates, Tomasello 2003a).³

While usage-based approaches are functional approaches to acquisition in the sense that the communicative function of a symbolic unit is a major driving force in acquisition, they do not consider semantics or concepts as being more important than distributional factors. Given the symbolic nature of linguistic units, both contribute to one another (see the discussion of Weinert in this issue for related findings that show that semantic information does not seem to be the driving force for acquisition).

By virtue of being nonreductionist and by assuming a limited set of psychological processes that can explain how structure emerges from usage, usage-based approaches offer a *domain-general* perspective on language, language learning, and language change. A fundamental difference to modern generative versions of grammar is the assumption that all linguistic units are *symbolic*, i.e., pairings of (written or phonetic) form and meaning. This implies that usage-based linguistics is *nonmodular*, since all formal aspects of a unit contribute to its semantics. The non-modularity of linguistic representation also implies that bootstrapping mechanisms in the technical sense where learning is made possible by an interface between levels of representation (see Höhle this issue) do not apply.

2.4. *Language-specific versus domain-general learning mechanisms*

To date, usage-based and emergentist approaches offer a more comprehensive view on acquisition than those based on universal grammar. The main differences concern the assumptions regarding specifically linguistic representations: Whereas Generative Grammar relies on innate universal representations⁴ as the prerequisite for language development (see Eisenbeiß this issue), emergentist approaches see linguistic structure as the result of nonlinguistic cognitive processing (Bates and Goodman 1999). Here, the focus of interest is to explain how different cognitive skills interact to produce a qualitatively new skill, namely complex human language system (e.g., Smith 1999).

It follows that usage-based approaches do not make a distinction between (universal) core syntax and other properties of language. Rather, all linguistic phenomena are considered to be form-function correspondences. They may differ in size (from morphemes to idioms) and in degrees of abstractness or schematicity and productivity. Because of this,

a usage-based approach is ultimately more parsimonious than a UG-approach: First, it does not require a separate innate endowment just for the purpose of language. Moreover, the same mechanisms will account for the acquisition of all of language, i.e., one does not assume separate learning mechanisms like triggering or maturation that solely serve the purpose of activating and setting the language-specific parameters of UG.

2.5. *Mental representation and processing*

Regarding language representation and processing, utterances do not necessarily have to be generated from scratch: speakers of a language may access high-frequent or idiomatic structures as a whole, and they store prefabricated chunks as well as the component parts (Bybee and Scheibman 1999; Dabrowska 2004: 18–22). The proposal that adults store structures on various levels of abstraction — from fully analyzed to fully frozen — relates to previous findings in language acquisition research that children might operate with constructions that are less analyzed than those of adults. They may produce several transitive sentences, but may not be able to produce the same construction with new words (Tomasello 2000a, 2000b). That is, children’s constructions can be correct and error-free, yet formulaic or lexically specific.

MacWhinney (2004: 910–911) argues that item-based processing is the major aspect of data-driven acquisition theory that relies on positive evidence and indirect negative evidence alone. The importance of individual usage events is also prominent in so-called exemplar-based models of linguistics (cf. Bybee 2006). Here, it is assumed that each exemplar is stored. The question then is how is storage affected by experience? Do we store all exemplars in the literal, verbatim way? Or do the stored exemplars accumulate more abstract properties such that we categorize them as instances of a syntactic or morphological schema, or is the concrete lexical material still relevant (see Bybee 2006)? In cognitive linguistics, no such distinction has to be made: Langacker (1987) speaks of the “rule-list-fallacy” of the words-and-rules model (Pinker 1999). The words-and-rules model proposes that for reasons of economy linguistic units are either stored analytically if they can be processed by rules, or holistically if they are irregular or idiomatic. Langacker argues that there is no reason to assume that analytic and holistic storage are mutually exclusive (Langacker 1987: 42; Bybee and Scheibman 1999; see also Jackendoff 1997; Pinker and Jackendoff 2005).

2.6. *Is there poverty of the stimulus?*

The emergentist perspective is further motivated by a certain stagnation in generativist acquisition theory. The widespread consensus about an innate language-specific component is not accompanied by similar scientific agreement about the exact nature of this innate component nor about the exact mechanisms by which it becomes activated (see Eisenbeiß [2002] for a recent survey). But the postulate of innateness can only be tested if there is a very concrete definition of what this knowledge includes. The claim that certain syntactic structures cannot be learned on the basis of positive evidence (i.e., the language the child hears) is a very powerful and complex prediction. For a real poverty-of-the-stimulus situation, three conditions have to be met: (1) the structure under investigation indeed is *inaccessible* from the input, (2) such evidence is *indispensable*, i.e., the structure could not be inferred by other means, and (3) children indeed *acquired* the underlying structure that is postulated (cf. Akhtar et al. 2004: 142). There is a vivid debate on whether there are linguistic phenomena for which the pragmatic and linguistic cues in the input are underspecified (e.g., question formation from embedded clauses, anaphoric reference, or argument structure). These discussions have taken place in a series of target articles and replies. Crain and Pietroski 2002; Pullum and Scholz 2002; and several other papers in a special issue of *The Linguistic Review* dealt with the learnability of question formation in complex sentences. Lidz et al. 2003a proposed poverty-of-the-stimulus claims for children's acquisition of argument structure (see the response by Goldberg 2004 and the reply by Lidz and Gleitman 2004). Likewise, there is an ongoing discussion whether children's early and successful interpretation of anaphoric pronouns is due to innate syntactic knowledge (Lidz et al. 2003b; Lidz and Waxman 2004) or whether it can be inferred on semantic grounds (Akhtar et al. 2004; Tomasello 2004) or Bayesian learning procedures that are sensitive to the absence of certain input patterns (Regier and Gahl 2004).

This discussion shows that the main divide between nativist and emergentist proposals is whether there are aspects of syntax that cannot be learned empirically (see Culicover and Nowak 2003 for a dynamic systems perspective on learning from a Minimalist vantage point). The argument for innateness is not only that a structure cannot be learned in principle, but also that the children display linguistic abilities at a very young age where learning is unlikely (Lidz and Gleitman 2004). Hence, it does not suffice to demonstrate that a computer can learn a complex structure based on positive evidence, but it has to be demonstrated that a two-year-old can do it, too. In this vein, proponents of a usage-based approach are

inspired by recent evidence that even very young infants and children are sensitive to statistical generalizations in the language they are exposed to (e.g., Aslin et al. 1999; Newport and Aslin 2004). This ability certainly plays a role in language-specific learning, but might also suffice to explain those aspects of grammar for which concrete poverty-of-the-stimulus claims have been made (see Pullum and Scholz 2002; Elman 2003; MacWhinney 2004). If this were the case, the usage-based model would offer a more parsimonious account of language acquisition because all of acquisition would be accounted for by only one model and with mechanisms attested in other cognitive domains. So what are the mechanisms for generalizing knowledge from positive evidence?

2.7. Statistical learning

Usage-based acquisition theory does not assume language-particular learning mechanisms but relies on mechanisms known from psychology. Moreover, there are conceptual relationships to computational approaches to language learning, known under headings such as connectionist modeling, probabilistic grammars, and distributional, stochastic, or statistical learning (e.g., Redington et al. 1998; Mintz et al. 2002; Newport and Aslin 2004). The linguistic background for these implementations is provided by stochastic grammars (Bod 2003; Jurafsky 2003; Manning 2003). Probabilistic approaches to syntax include Bayesian statistics of conditional probabilities, i.e., distributional properties, and entropy measures of information processing i.e., assumptions about the amount of information that is processed in a particular variable (see Bod et al. 2003; Manning 2003; Manning and Schütze 1999). Connectionist models of language learning investigate how a neural network can discover structural properties (see Elman et al. 1996; Westermann et al. this issue). The leading hypothesis is that linguistic experience is so rich that it guides learners functionally and stochastically to the structures they learn (see Manning 2003; Gupta and Dell 1999; Klein and Manning 2002, 2004). That is, competence arises from performance (Allen and Seidenberg 1999).

In developmental psychology, the dynamic systems theory (Thelen and Smith 1994; van Geert 1994) explains development with similar concepts: categories are gradient with fuzzy boundaries rather than absolute, and developmental progress does not proceed in a linear growth curve but shows variability in the form of plateaus and sudden developments. Each individual is considered a dynamic system whose progress is best predicted by this individual's prior experience (cf. Hohenberger and Peltzer-Karpf this issue; Hockema and Smith this issue).

3. Empirical evidence

I will review two research traditions in the usage-based approach to language. Since it is assumed that children's early knowledge is item based and becomes (partially) productive only gradually, two types of investigation are critical to such a theory:

- (a) the degree to which children's linguistic competence is limited.
- (b) the processes which allow children to generalize

Three types of empirical studies are conducted to provide support for and test the theory:

First, studies that challenge the claim that children's early performance is adult-like. This type of evidence is taken to prove that children's early categories are not abstract and category general, but — for example — lexically specific. Second, studies that show that children's early usage is input-based, i.e., can be related to positive evidence that children receive. Third, studies that show that children's progress in language development can be explained by general social and cognitive skills and domain-general learning processes. In Section 3.1 I will summarize studies that investigate the constraints in children's early linguistic knowledge. Section 3.2 and Section 3.3 will deal with children's generalization over positive evidence, and their induction of knowledge without positive evidence, respectively.

3.1. *Islands and slot and frame patterns*

That early child language is (partially) formulaic and item-based is one of the cornerstones of usage-based acquisition theories. Observations of this kind first became introduced with Braine's concepts of *limited scope formulae* (Braine 1976), and are also known as *slot and frame patterns* (Pine and Lieven 1997; Lieven et al. 1997; Pine, Lieven and Rowland 1998). These terms indicate that children's internal representation of a construction may differ from that of adults. Although children's utterances look adult like, distributional analyses show that parts of the utterance are likely to represent unanalyzed chunks that act like a frame into which (a limited number of) words or phrases can be inserted.

Tomasello (1992) introduced the notion of "islands" of development. The underlying idea is that first usages of categorically similar verbs (e.g., transitive verbs or ditransitive verbs) do not provide evidence for verb-general categories, i.e., an abstract representation of transitivity or ditransitivity, but that each individual transitive verb may be used in a

different lexical frame, without much overlap between those frames. In the course of development, then, the child recognizes relationships between such constructions such that the sentence schemas become more verb-general and less item-specific.

3.1.1. *Early productivity.* Although it seems to be uncontroversial that some aspects of early child language are lexically specific, there is a debate about what lexical specificity really means in terms of productivity, and how prevalent it is. A controversial and as of yet open issue is whether item-based formulas are the starting point for each aspect of language development, and whether item-based learning is equally relevant crosslinguistically. It is under debate whether children's competence is really as limited as proposed by some authors. Based on results from a diary study on sentences produced with a list of common verbs, Vear et al. (2002) argue that English-speaking children generalize earlier than proposed in the studies by Tomasello and colleagues. The diary data collected by Vear et al. (2002) suggest that not all children go through a verb island stage. However, these results are not incompatible with usage-based learning. It seems plausible that children take the abstract lexical and syntactic knowledge they have already acquired, instead of going through item-specific learning again and again. Abbot-Smith and Behrens (2006) argue that even very complex constructions can be acquired quickly, provided children command the component parts of the new complex construction. For example, the German passive is acquired rather easily because children can build on other auxiliary constructions. Since in spoken language the "by phrase" is almost always omitted, the verbal passive (1) is very similar to present perfect constructions (2). They only differ in the auxiliary.

- (1) verbal passive
das Kind wird gewaschen
'the child is being washed'
- (2) present perfect
das Kind hat gewaschen
'the child has washed (something)'

Interestingly, the resultative passive with the auxiliary *sein* 'to be' (e.g., *das Kind ist gewaschen* 'the child is in a washed state') is acquired earlier than the nonresultative *werden*-passive as in (1) above, presumably because the acquisition of the *sein*-passive is facilitated by the semantically related and formally identical present perfect with the auxiliary *to be* (*das Kind ist gewachsen* 'the child has grown'). Such facilitatory effects of related constructions were also demonstrated by connectionist networks

(Morris et al. 2000). Here, the acquisition of grammatical construction is seen in terms of family resemblance structures where their relationship cannot be reduced to one factor alone. Rather, in a bottom-up fashion more and more similarities between seemingly unrelated constructions are discovered as a function of growing experience with language.

Finally, there is research as to which properties of a language facilitate abstraction. From a crosslinguistic perspective it seems rather strange, for example, that children take relatively long to acquire English morphology although there are only very few paradigms to acquire. Given the syntactic properties of English, namely a relatively fixed word order and little variation in the form of words, it seems that English is more susceptible to the entrenchment of lexically specific patterns because identical strings of words are more likely to occur. That is, English has high token frequency and less type variation. Languages with more variation in both word order and word forms, in contrast, show more type variation and lower token frequency, such that abstraction processes may start earlier. This interdependence of type and token frequency may explain why English-speaking children take relatively long to acquire the little morphology there is, whereas children learning highly inflectional languages seem to perform better and to start earlier (Slobin 1973).

But corpus data alone only give limited evidence about the underlying mental representation: it may well be that a certain collocation is highly frequent and shows little variation or linkage with other construction, but may nevertheless be fully analyzed. Thus, the results from corpus studies must be compared to results from experimental studies that test productivity (e.g., nonce-words experiments, priming studies). Here, results point in the same direction of rather limited competence initially (Tomasello 2000a).

3.1.2. *Social and cognitive prerequisites for language learning processes.*

In its current version, usage-based acquisition theory is the result of convergent evidence from the domain of social-cognitive development, pattern recognition and schema abstraction abilities. Language learning seems to be based on three sources

- (i) Social cognition
- (ii) Constraints on working memory
- (iii) Generalization mechanisms (see Sections 3.2 and 3.3)

(i) Social cognition

Tomasello and Rakoczy (2003) argue that social development is what possibly demarcates the human species from other animals. In their

view, social cognition leads to differences in the various kinds of learning processes. Whereas humans are able to make use of *imitative* learning, where they start off from the goal of the action, animals seem only to be able to use *emulative* learning, where they replicate the action without recourse to the goal of the action. Meltzoff (1995) asked 18-month-old children to imitate an action they just saw, whereby the facial expression of the experimenter suggested that he or she had failed to achieve the results. E.g., the experimenter was pulling at the cap of a pen as if to make the cap come off. However, the cap did not come off. It turned out that children did not imitate what they saw but what they thought the experimenter tried to achieve, e.g., they pulled off the cap of the pen rather than just pulling at the cap. In their review of the literature on social behavior in humans and nonhuman species, Tomasello and Rakoczy (2003) conclude that humans show cultural transmission of knowledge in a way that has not yet been observed in animals.

The species-specific behaviors shown by children are joint attention, and the understanding of others as intentional agents, including the awareness that their mental state may vary from others' (Tomasello and Rakoczy 2003). If we assume that children are good at working out the intention underlying the verbal utterances of others, they have a head start into the linguistic system. It still needs to be worked out in detail, however, how children map their pragmatic knowledge onto linguistic forms. Tomasello argues that children's earliest utterances, typically one-word phrases, represent communicative acts (Tomasello 2003a). I.e., they form a construction (form/function pairing) that can be differentiated and reanalyzed as children proceed.

(ii) Constraints on working memory

One of the puzzles in language acquisition research is why small children with their limited cognitive resources are so good at language learning, a rather complex task, but they fail at other, seemingly simpler tasks. A possible answer is that it is exactly children's limited working memory that focuses their attention to the relationships between words or segments in close proximity (cf. Clark 2003: 413–416 for a summary). This assumption became known as the *starting small hypothesis* (Newport 1990; Elman 1993).

In training experiments with Tamarin monkeys and human adults on sequences of nonsense syllables where regularities held between nonadjacent segments or syllables, Newport et al. (2004: 112) conclude that “for regularities among elements that are at moderate distances from one another, a cognitively limited learner might be more capable than a learner with greater cognitive capacities.”

In a neural network study, Elman (1993) was able to show that the network was able to learn complex structures only if it was trained on shorter sequences before, and when the viewing window was limited. When the model was fed with input that was too long and complex, it failed to generalize (also see Culicover and Nowak 2003).

The connectionist model MOSAIC was employed by Freudenthal et al. (2002) to simulate the stages of acquisition regarding nonfinite and finite forms in English and Dutch. The model managed to acquire verbal inflection in the order of acquisition similar to that attested in child learners, simply by enlarging the model's working memory and by processing the input sentences, all of which were questions from right to left (because the end of the sentence is what children hear last). Consider the Dutch sentence *papa wil koekje eten* 'daddy wants cookie eat-INF'. If one assumes that language processing proceeds incrementally from right to left and that function words are often omitted, the model predicts the acquisition sequence *eat-INF* → *cookie eat-INF* → *daddy cookie eat-INF*. This sequence is a typical developmental sequence attested in Dutch children. In a recent adaptation of the model (Freudenthal et al. 2005, 2006) a primacy bias was added, such that the model also processed utterance initial clusters. This enabled the network to use declarative and imperative sentences as input as well. With these additions, the model could simulate the language-specific differences in optional infinitive phenomena in English, Dutch, German, and Spanish. This suggests that attention to salient positions (the end and the beginning of utterances) combined with a growing working memory enabled the MOSAIC model to generate sequences typically found in early child language, and to simulate language-specific differences in acquisition.

3.2. *Pattern recognition and generalization over positive input*

In order to induce linguistic rules, children need powerful generalization skills. Three components can be distinguished: First, learners must be able to recognize patterns. Second, they must be able to compare patterns and generalize, i.e., to abstract schemas. This type of schema abstraction is still built on generalization over positive evidence. Third, they must command inferencing skills like analogical reasoning such that they generalize schemas that are not attested in the input. In the following I will review two types of learning mechanisms, those based on pattern recognition in positive evidence, and those based on analogical reasoning without positive evidence. The common idea is that learning language consists of pattern recognition and generalization over these patterns such that

language structure can be induced from distributional patterns (Klein and Manning 2002, 2004).⁵

Stochastic models of acquisition became highly influential in the past decade. It turned out that infants are highly efficient pattern recognizers, when exposed to series of nonsense syllables. These pattern recognition skills are skills that can be used to discover linguistic regularities or rules (Newport and Aslin 2004). Most likely, some of these pattern recognition skills are not uniquely human (cf. Newport et al. 2004 for a detailed discussion). Recently, Hudson Kam and Newport (2005) showed that children can even repair inconsistencies in the input patterns, and thus show that children start to regularize based on positive evidence.

These results suggest that children pay close attention to the input patterns they hear. This is also confirmed by studies that analyze the overlap of child and adult language. Behrens (2003, 2005, 2006) studied the distribution of part-of-speech categories in extensive corpora of German child and adult language. She found that (a) the distribution of linguistic categories like part-of-speech distribution was extremely stable over time in the input, and that (b) the child approaches the adult distributional patterns within a year or two, depending on the structure under investigation. Such input orientation is also documented in crosslinguistic studies on children's categorization of spatial relations, where early language specificity is observed (e.g., Choi and Bowerman 1991; see Slobin 2001 for a summary).

Also, children's reliance on positive evidence does not only allow them to acquire the simple structures of language, but also complex ones. Diessel (2004) demonstrated how English-speaking children gradually acquire complement clauses. They start out from a semantically limited use of the main clause in a modifying way (e.g., they used the clause "I think" like an adverbial), and proceed to generalize to different types of syntactic dependencies.

Once patterns are abstracted, priming effects should show this. Priming is based on the idea that since knowledge is connected, the activation of cognitive content activates related content. E.g., hearing the word *sheep* activates lexical entries that are semantically or phonologically related (e.g., goat or shawl). In language production, experienced speakers show priming effects such that related constructions are activated in the mental lexicon. Priming thus indicates that children have built mental networks. Savage et al. (2003) tested children's use of passives in a priming design where children were presented pictures with active or passive sentences. They were asked to repeat the sentence, and then to describe the picture. 3- and 4-year olds showed lexical priming only, i.e., they repeated the verb they heard. 6-year-olds, however, also showed structural priming,

i.e., they used the active or passive structure they heard with new lexical material.

If learning is based on positive evidence, another relevant question is what type of evidence enhances learning. Recently, two training studies were carried out that experimented with different types of learning. Childers and Tomasello (2001) analyzed the differential role of type and token frequency. Ambridge et al. (2006) studied the effect of timing in terms of massed vs. spaced learning.

3.2.1. *The role of type and token frequency.* The interplay between type and token frequency provides the basis for generalization. Type and token frequency play different roles in the process: High token frequency leads to entrenchment by leaving strong memory traces, whereas type variation leads to abstraction (cf. Tomasello 2003: 173–175). In a training study, Childers and Tomasello (2001) tested the effect of different types of input. In three sessions, children were trained on 16 novel transitive actions where a stuffed animal did something to an inanimate object. Each action was commented upon by the investigator with 6 repetitions of a transitive sentence of various types. That is, during the three training sessions, the children heard 288 transitive sentences. Children were assigned to one of five training conditions. A control group played these games without verbal input. In the four verbal conditions, the actions were named with familiar or unfamiliar verbs and the subjects and objects were encoded either with a full NP or with pronouns only (cf. Examples [3a]–[3d]).

- (3) Two-noun condition with (a) known or (b) unfamiliar verb
- a. *Look, the bird is swinging the bathtub. See? The bird is swinging the bathtub.*
 - b. *Look, the dog is hurling the chair. See? The dog is hurling the chair.*

Two-pronoun condition with (a) known or (b) unfamiliar verb

- c. *Look, the cow is pulling the car. See? He's pulling it.*
- d. *Look, the bear is striking the tree. See? He's striking it.*

In the fourth session, children were tested on new transitive actions with pseudoverbs. The success of training was measured by the number of children (out of 10 per group) who produced at least one transitive sentence with the novel verbs. Even by this very lenient criterion of productivity, only two children in the control condition managed to produce a transitive sentence with a novel verb. Children in the noun conditions performed a little better, but did not differ significantly from the control condition. Only the children in the pronoun condition showed a significant

learning effect, regardless of whether they were exposed to familiar or unfamiliar verbs during the training phase (Childers and Tomasello 2001: 742). These results show that modifications in the input properties have an effect in learning abstract syntactic constructions. Only those children that received training were able to generalize. The results also indicate that stability and variation have different roles in development: children in the fixed pronoun condition managed best, although it is unclear whether they had to rely on abstract linguistic knowledge about transitive sentences to do so. Variation in the NPs did not facilitate learning of transitive constructions very much, but such variation in the arguments may help children to learn something about the verb meaning (cf. the notion of syntactic bootstrapping, Gleitman 1990; Naigles 1990; Höhle this issue).

3.2.2. *The effect of massed and spaced learning.* If frequency matters, does distribution in timing matter as well? This is the question underlying the massed versus spaced learning paradigm. In massed training, all evidence is presented at once. In spaced learning, the evidence is distributed over time, i.e., subjects are presented with the data over several sessions. Ambridge et al. (2006) found that various types of distributional learning had a better effect than a massed presentation. Studies of this type can also be used to investigate the minimum number of examples needed to learn, and the optimal distribution of evidence over time.

While these training studies investigate how children generalize based on positive evidence, other researchers focused on learning situations where positive evidence was missing, and where children had to rely on their generalization skills in order to induce the missing link.

3.3. *Generalization over indirect positive evidence*

At first glance, generalization without positive evidence seems to constitute a problem for usage-based approaches because they emphasize that linguistic structure is derived from language use, that is, production data. However, usage-based acquisition can even work on missing evidence, as long as the cognitive resources to deduce missing evidence are available to children. As of yet, evidence for generalization over missing evidence comes from connectionist models.

Elman (2003) argues that the key to acquisition of complex linguistic rules might well be *indirect positive evidence*, which relies on children's ability to generalize across domains. Morris et al. (2000) showed that a simple recurrent network was able to produce a supposedly unlearnable construction, even without positive evidence, but on the basis of simpler

utterances that contained the relevant morphosyntactic information. The construction at stake is question formation from sentences with embedded clauses like *The boy who is smoking is crazy*. Inversion based on simple linear order (the first “is”) would lead to the incorrect question *Is the boy who _ smoking is crazy?* The child must recognize that question formation is structure dependent and must have knowledge that the second “is”, which is part of the main clause, needs to be fronted in order to come to the correct question: *Is the boy who is smoking _ crazy?* Elman’s (2003) network was able to generate such complex questions on the basis of having learned simple utterances that provided

- (i) evidence for agreement, subcategorization, and selectional restrictions
(e.g., *The women smoke _ cigarettes; The boy is funny*)
- (ii) evidence for question formation in simple sentences
(e.g., *Can you read this? Is she friendly?*)
- (iii) evidence for complex sentences and noun phrase constituency
(e.g., *The girl with the cat is nice. I like the teacher who speaks Spanish.*)

The network generated the complex sentence from knowledge of simpler sentences that contained the relevant components of the complex structure.

In developmental psychology, the concept of structure mapping (Gentner 2003) is one that can be applied to the problem of generalizing over missing evidence. Structure mapping is a form of complex analogical reasoning that allows us to draw alignments across domains (Gentner 2003). To my knowledge, this concept has not yet been applied to experiments on language learning.

To sum up, usage-based approaches to acquisition explain language learning by a combination of powerful generalization mechanisms with sophisticated social skills which help the child to overcome the problem of semantic indeterminacy. Due to schema formation and comparisons across schemas, the representations become more abstract, qualitatively new representations emerge.

4. Applications of usage-based approaches in other domains

The vast majority of studies in the emergentist framework is concerned with unimpaired, monolingual first language acquisition. The emphasis is on (a) word learning, (b) morphosyntactic development and (c) discourse patterns (see Tomasello 2003a). Less attention has been paid to

phonological and prosodic development, or to semantics and pragmatics. In principle, usage-based acquisition research should address these domains using the same concepts and methods.

4.1. *Language disorders*

There is extensive work on language development under impairment. Again, the key assumption of emergentist approaches is that phenotype and genotype are in no direct relation, but that rather small aberrations in the genotype can have wide-ranging effects in the phenotype because of an extended developmental trajectory (Karmiloff-Smith 1998, 2004). Likewise, research on children with frontal lesions has shown that the brain shows high plasticity, rather than deterministic localization for higher cognitive functions (Elman et al. 1996).

4.2. *Multilingualism*

To date, there is a lack of studies on multilingualism, be it childhood or adult bilingualism or second language learning. Hernandez et al. (2005) developed a competition model of how children can distinguish and build up multiple language systems based on experience. They argue that children will be able to construe modular multiple grammars based on the competition between cues, and predict entrenchment of units belonging to the different languages to be learned. Ellis (2002, 2008) shows that frequency factors account for several aspects of second language learning as well as language change.

5. Open issues and new directions

There are a number of open issues that need to be addressed in more detail in future research. These open issues include the comparison between child and adult representations, the acquisition of semantics, the construction of meaning in discourse, and the study of individual differences.

5.1. *Child and adult representations*

More comparative research between adults and children needs to be done in order to establish how language is represented, especially regarding

holistic and analytic storage of related forms. Usage-based approaches predict that there should be quantitative differences, because adults are able to store larger sequences, but they can also store items with a higher degree of abstraction because they have had more experience with language.

5.2. *The acquisition of semantics*

Although usage-based theories are *functional* theories of language, there is a lack of hypotheses and studies regarding the acquisition of function or meaning proper. E.g., if one assumes that even prelinguistic children are good at identifying the intentions of others, how does this contribute to the understanding and refinement of the semantics of the utterances that the child processes? A possible answer lies in the study of interactive processes in order to investigate how meaning is established in discourse. In usage-based theory, meaning is a flexible concept with prototype effects (Langacker 2000). Our representations of meaning will change if we hear a construction is used in a different fashion. In order to trace this process, detailed corpus studies or training studies on small scale changes will have to be carried out.

5.3. *Relationship between quantitative and qualitative learning processes*

The success of cluster analyses, probabilistic and connectionist models in learning processes, as well as the experimental research with infants on pattern recognition suggest that language learning can be explained to a large degree by quantitative learning processes. It is as of yet unknown how the registering of quantity information proceeds exactly (cf. the discussion on exemplar based models and growing abstractness of linguistic representations above).

Although researchers in the usage-based paradigm do not propose that child language acquisition is merely stochastic, qualitative aspects of learning situations are not as widely discussed. Küntay and Slobin (2002) criticize the current treatment of corpora because as soon as conversation data are entered into a database, they are only considered in quantitative fashion. However, they point out that real learning takes place in concrete discourse situations and explore the types of interactive patterns that enhance language learning. They identified so-called *variation sets* as one possible context for abstracting linguistic knowledge. Variation sets are characterized as reformulations where the form varies

but the content stays the same. In example (4) a father reformulates a request for information in five different ways:

- (4) Father to son, age 2;3
Who did we see when we went to the store?
Who did we see?
Who did we see in the store?
Who did we see today?
When we went out to shopping, who did we see?
 (Küntay and Slobin 2002)

In terms of function, variation sets typically are control oriented (call for child's action), ideational (provide information), or information querying (prompt for child). Variation in form concerns lexical substitution and rephrasing, addition and deletion of referential terms, and reordering of constituents.

5.4. *Individual differences*

In usage-based approaches, differences in input language matter — both from a crosslinguistic as well as from a language-specific perspective. More research is needed regarding the effect of quantitative and qualitative features of the input on individual children's course of language development and success in language development. In a large cross-sectional study, Hart and Risley (1995) found that children's language development highly correlated with the quantity of the input they received. Here, dynamic system models (e.g., Hockema and Smith this issue) which assume that each learning step depends on the individual state of the system provide a powerful conceptual and methodological tool to understand different learning outcomes.

6. Summary

In cognitive linguistics, constructions are the basic linguistic units. They are defined as form-function pairings of any size, for example morphemes, words, clauses, or sentences. The emphasis on the interrelationship between form and meaning made cognitive linguistics an appropriate framework for those acquisition researchers who emphasize the communicative basis for language acquisition and who assume that pragmatics and semantics are powerful cues in the acquisition process (see Tomasello 1998a, 1998b, Tomasello 2000b, and Diessel 2004: Chapter 2, for a de-

tailed account of the possible links between acquisition research and cognitive linguistics). This claim is supported by findings in linguistic and nonlinguistic domains that the mind can store information in multiple ways, i.e., there is analytic *and* holistic storage (see Section 2.5; see also Weinert in this issue for the importance of redundant encoding in implicit learning processes).

The aim of cognitive linguistics is to be maximalist, i.e., to integrate all factors that are needed to systematically explain all aspects of all linguistic effects (Langacker 1988). Systematicity in the treatment of linguistic phenomena is achieved by establishing the network-like relationships between different constructions (see Section 3.1.1).

Cognitive linguistics is regarded as a perspective on the organization and representation of language that is highly compatible with linguistic facts (see the contributions in Tomasello 1998a and 2003b). In this sense, the theory is eclectic because it relies on *converging evidence* from linguistic and from learning theories. Although this eclecticism makes the usage-based approach less cohesive as a theory in the philosophical sense (see Jordan 2004 and Russell 2004 for extensive discussions of this issue), this eclecticism or integrativeness is seen as the more realistic approach when dealing with complex and multifaceted issues like language use and language learning. As stated above, many cognitive linguists do not strive for economy of representation or elegance of the theory (cf. Tomasello 1995), because there is no evidence that parsimony has psychological reality since the brain allows for multiple ways of storage and processing.

A major tenet of usage-based approaches is that language has to be learned on the basis of experience, because no innate linguistic representations are assumed. Therefore, a number of studies were conducted that showed that children's early competence was limited. Usage-based approaches are inspired by studies on the generalization skills of infants and young children. Thus, the focus in the usage-based approach changes from emphasizing the limited productivity of early child language to the study of the exact learning mechanisms and the stochastic basis for abstracting generalizations.

The converging evidence from developmental and computational studies make data driven learning appear more likely than was expected a decade ago. There is a growing body of evidence that structural properties of language can be induced from, for example, co-occurrence relationships. Also, there is a growing body of evidence that even very young children are very good at pattern recognition and schema abstraction, such that they are able to generalize over inconsistencies in the input and to make use of indirect positive evidence. In this sense, linguistic competence is an emergent property of language use.

The usage-based approach is falsifiable if it can be shown that there are linguistic phenomena which cannot possibly be learned on the basis of (direct and indirect) experience. This challenge is known as the *poverty of the stimulus argument*. However, even if truly “unlearnable” phenomena could be identified, this would not automatically constitute evidence for innateness of the linguistic knowledge underlying this phenomenon. For that, it also needs to be demonstrated that an attested feature of universal grammar does explain the acquisition of such a phenomenon.

Received 25 May 2005

University of Basel

Revised version received

2 May 2007

Notes

- * Correspondence address: English and German Department, University of Basel, Nadelberg 6, 4051 Basel, Switzerland. E-mail: Heike.Behrens@unibas.ch.
1. Langacker (1987) used the term “particular” to refer to actual usage events and their storage. In later acquisition research, the term “concrete” became used (e.g., Tomasello 1998b).
 2. The use of terminology reflects differences in emphasis. E.g., Tomasello (2003) does not mention the term emergentism, but puts forward a *usage-based* theory that strongly emphasises the social roots of language development, whereas the term *emergentism* seems to be preferred in the computational and neurolinguistic literature (see Elman et al. 1996, and the articles in MacWhinney 1999). But the concept of “emergence” is also used in nativist approaches (see Hohenberger and Peltzer-Karpf this issue). The difference lies in the assumptions regarding the initial state of linguistic representation, and the linguistic framework used. Russell (2004) distinguishes between connectionist and other usage-based approaches by placing connectionism into the *empiricist* tradition of thought, whereas he categorizes usage-based approaches as *pragmatist* because of the emphasis on the social foundations of language learning. Both are contrasted with the *rationalist* paradigm that underlies generative theories of language acquisition.
 3. The phrasing is deliberately vague because current comparative research aims at identifying the cognitive skills in other species in order to define the delineation to humans. For example, dogs are better than nonhuman primates on social cognition tasks (Hare et al. 2002), monkeys demonstrate pattern recognition skills (Newport et al. 2004; Hauser et al. 2001), and dogs show word learning strategies similar to those of children (Kaminski et al. 2004).
 4. Russell (2004: 79) points out that Chomsky in the Minimalist Framework changed his view such that instead of linguistic representations, the process “merge” is the basic operation of language. However, in contrast to usage-based theories, this process is considered to be domain-specific to language. Note that in more recent publications, the principle of “recursion” is seen as the central characteristic of human language in the narrow sense, with an ongoing debate of whether this process is domain specific (Fitch et al. 2005).

5. Distributional learning is compatible with nativist approaches as well, because one does not make assumptions about where the categories come from (see Redington et al. 1998). They can be emergent or pre-specified. Likewise, stochastic grammars can be modular or nonmodular (Jurafsky 2003: 90).

References

- Abbot-Smith, Kirsten & Heike Behrens. 2006. How known construction influence the acquisition of other constructions: The German periphrastic passive and future constructions. *Cognitive Science* 30. 995–1026.
- Akhtar, Nameera, Maureen Callanan, Geoffrey K. Pullum & Barbara C. Scholz. 2004. Learning antecedents for anaphoric 'one'. *Cognition* 93. 141–145.
- Allen, Joseph & Mark S. Seidenberg. 1999. The emergence of grammaticality in connectionist networks. In Brian MacWhinney (ed.), *The emergence of language*, 115–151. Mahwah, NJ: Lawrence Erlbaum.
- Ambridge, Ben, Anna L. Theakston, Elena V. M. Lieven & Michael Tomasello. 2006. The distributed learning effect for children's acquisition of an abstract linguistic construction. *Cognitive Development* 21. 174–193
- Aslin, Richard N., Jenny R. Saffran & Elissa L. Newport. 1999. Statistical learning in linguistic and nonlinguistic domains. In Brian MacWhinney (ed.), *The emergence of language*, 359–380. Mahwah, NJ: Lawrence Erlbaum.
- Bates, Elizabeth. 1997. Origins of language disorders: A comparative approach. *Developmental Neuropsychology* 13. 447–476.
- Bates, Elizabeth & Judith Goodman. 1999. On the emergence of grammar from the lexicon. In Brian MacWhinney (ed.), *The emergence of language*, 29–79. Mahwah, NJ: Lawrence Erlbaum.
- Behrens, Heike. 2003. Verbal prefixation in German child and adult language. *Acta Linguistica Hungarica* 50. 37–55.
- Behrens, Heike. 2005. Wortartenwerb durch Induktion. In Clemens Knobloch & Burkhard Schaefer (eds.), *Wortarten und Grammatikalisierung: Perspektiven in System und Erwerb*, 177–198. Berlin and New York: Walter de Gruyter.
- Behrens, Heike. 2006. The input-output relationship in first language acquisition. *Language and Cognitive Processes* 21. 2–24.
- Bod, Rens. 2003. Introduction into elementary probability theory and formal stochastic language theory. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 11–37. Cambridge, MA: MIT Press.
- Bod, Rens, Jennifer Hay & Stefanie Jannedy. 2003. Introduction. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 1–10. Cambridge, MA: MIT Press.
- Braine, Martin D. S. 1976. Children's first word combinations. *Monographs of the Society for Research in Child Development* 41(1) [Serial No. 164].
- Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82. 711–733.
- Bybee, Joan L. & Joanne Scheibman. 1999. The effects of usage of degrees of constituency: The reduction of "don't" in English. *Linguistics* 37. 575–596.
- Childers, Jane B. & Michael Tomasello. 2001. The role of pronouns in young children's acquisition of the English transitive construction. *Child Development* 37. 739–748.
- Choi, Soonja & Melissa Bowerman. 1991. Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition* 41. 83–121.

- Clark, Eve V. 2003. *First language acquisition*. Cambridge: Cambridge University Press.
- Crain, Stephen & Paul Pietroski. 2002. Why language acquisition is a snap. *The Linguistic Review* 19. 163–183.
- Culicover, Peter W. & Andrzej Nowak. 2003. *Dynamical grammar: Minimalism, acquisition, and change*. Oxford: Oxford University Press.
- Dąbrowska, Ewa. 2004. *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Edinburgh: Edinburgh University Press.
- Diessel, Holger. 2004. *The development of complex sentence constructions in English: A usage-based approach*. Cambridge: Cambridge University Press.
- Eisenbeiß, Sonja. 2002. *Merkmalsgesteuerter Grammatikerwerb: Eine Untersuchung zum Erwerb der Struktur und Flexion der Nominalphrase* [Feature-driven structure building: A study on the acquisition of noun-phrase structure and inflection]. Düsseldorf: University of Düsseldorf dissertation.
- Ellis, Nick C. 2002. Frequency effects in language acquisition. *Studies in Second Language Acquisition* 24. 143–188.
- Ellis, Nick C. 2008. The dynamics of language use, language change, and first and second language acquisition. *Modern Language Journal* 92(2). 232–249.
- Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48. 71–99.
- Elman, Jeffrey L. 2002. Generalization from sparse input. In Mary Andronis, Erin Debenport, Anne Pycha & Keiko Yoshimura (eds.) *Proceedings of the 38th Meeting of the Chicago Linguistics Society*, vol. 2, 175–200. Chicago: Chicago Linguistics Society.
- Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi & Kim Plunkett. 1996. *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA.: MIT Press.
- Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. Discussion: The evolution of the language faculty: Clarifications and implications. *Cognition* 97. 179–210.
- Freudenthal, Daniel, Julian Pine & Fernand Gobet. 2002. Modelling the development of Dutch optional infinitives in MOSAIC. In Wayne D. Gray & Christian D. Schunn (eds.), *Proceedings of the 24th Meeting of the Cognitive Science Society*, 328–333. Mahwah, NJ: Lawrence Erlbaum.
- Freudenthal, Daniel, Julian Pine & Fernand Gobet. 2005. Simulating optional infinitive errors in child speech through the omission of sentence-internal elements. In Bruno G. Bara, Lawrence Barsalou & Monica Buchiarelli (eds.), *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, 708–713. Mahwah, NJ: Lawrence Erlbaum.
- Freudenthal, Daniel, Julian Pine & Fernand Gobet. 2006. Unifying Cross-linguistic and within-language patterns of finiteness marking in MOSAIC. In Ray Sun & Naomi Miyake (eds.), *Proceedings of the 28th annual meeting of the Cognitive Science Society*, 232–236. Mahwah, NJ: Lawrence Erlbaum.
- van Geert, Paul. 1994. *Dynamic systems of development: Change between complexity and chaos*. New York: Harvester Wheatsheaf.
- Gentner, Dedre. 2003. Why we are so smart. In Dedre Gentner & Susan Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and cognition*, 195–235. Cambridge, MA: MIT Press.
- Gleitman, Lila R. 1990. The structural sources of verb meaning. *Language Acquisition* 1. 3–55.
- Goldberg, Adele. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Science* 7. 219–224.
- Goldberg, Adele E. 2004. But do we need universal grammar? Comment on Lidz et al. (2003). *Cognition* 94. 77–84.

- Gupta, Prahlad & Dell, Gary D. 1999. The emergence of language from serial order and procedural memory. In Brian MacWhinney (ed.), *The emergence of language*, 447–481. Mahwah, NJ: Lawrence Erlbaum.
- Hare, Brian, Michelle Brown, Christina Williamson & Michael Tomasello. 2002. The domestication of social cognition in dogs. *Science* 298. 1634–1636.
- Hart, Betty & Todd R. Risley. 1995. *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- Hauser, Marc D., Elissa L. Newport & Richard N. Aslin. 2001. Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition* 78. B53–B64.
- Hernandez, Arturo, Li Ping & Brian MacWhinney. 2005. The emergence of competing modules in bilingualism. *Trends in Cognitive Science* 9. 220–225.
- Hudson Kam, Carla J. & Elissa L. Newport. 2005. Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development* 1(2). 151–195.
- Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.
- Jordan, Geoff. 2004. *Theory construction in second language acquisition*. Amsterdam: John Benjamins.
- Jurafsky, Dan. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 39–95. Cambridge, MA: MIT Press.
- Kaminski, Juliane, Josep Call & Julia Fischer. 2004. Word learning in a domestic dog: Evidence for “fast mapping”. *Science* 304. 1682–1683.
- Karmiloff-Smith, Annette. 1998. Development itself is the key to understanding developmental disorders. *Trends in Cognitive Science* 2. 389–398.
- Karmiloff-Smith, Annette. 2004. Bates’s emergentist theory and its relevance to understanding genotype/phenotype relations. In Dan I. Slobin & Michael Tomasello (eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates*, 219–236. Mahwah, NJ: Lawrence Erlbaum.
- Kemmer, Suzanne & Michael Barlow. 2000. Introduction: A usage-based conception of language. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, vii–xxviii. Stanford, CA: CSLI.
- Klein, Dan & Christopher Manning. 2002. Natural language grammar induction using a constituent context model. In Thomas G. Dietterich, Suzanna Becker & Zoubin Ghahramani (eds.), *Advances in neural information processing systems 14 (NIPS 2001)*, vol. 1, 35–42. Cambridge, MA: MIT Press.
- Klein, Dan & Christopher Manning. 2004. Corpus-based induction of syntactic structure: Models of dependency and constituency. *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004)*.
- Küntay, Aylin C. & Dan I. Slobin. 2002. Putting interaction back into child language: Examples from Turkish. *Psychology of Language and Communication* 6. 5–14.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar, vol. 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- 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. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford: CSLI.
- Lidz, Jeffrey, Henry Gleitman & Lila R. Gleitman. 2003. Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition* 87. 151–178.

- Lidz, Jeffrey & Lila R. Gleitman. 2004. Yes, we still need universal grammar. *Cognition* 94. 85–93.
- Lidz, Jeffrey & Sandra Waxman. 2004. Reaffirming the poverty of the stimulus: A reply to the replies. *Cognition* 93. 157–165.
- Lidz, Jeffrey, Sandra Waxman, Jennifer Freedman. 2003. What infants know about syntax but could not have learned: Experimental evidence for syntactic structure at 18 months. *Cognition* 89. B65–B73.
- Lieven, Elena V. M., Julian M. Pine & Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24. 187–219.
- MacWhinney, Brian (ed.). 1999. *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum.
- MacWhinney, Brian. 2004. Multiple solutions to the logical problem of language acquisition. *Journal of Child Language* 31. 883–914.
- Manning, Christopher D. 2003. Probabilistic syntax. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 289–341. Cambridge, MA: MIT Press.
- Manning, Christopher D. & Hinrich Schütze. 1999. *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Marchman, Virginia & Donna Thal. 2004. Words and grammar. In Dan I. Slobin & Michael Tomasello (eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates*, 141–164. Mahwah, NJ: Lawrence Erlbaum.
- Meltzoff, Andrew. 1995. Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology* 31. 838–850.
- Mintz, Toben H., Elissa L. Newport & Thomas G. Bever. 2002. The distributional structure of grammatical categories in speech to young children. *Cognitive Science* 26. 393–424.
- Morris, William C., Garrison W. Cottrell & Jeff L. Elman. 2000. A connectionist simulation of the empirical acquisition of grammatical relations. In Stefan Wermter & Ray Sun (eds.), *Hybrid neural symbolic integration*, 175–193. Berlin: Springer.
- O’Grady, William. 2005. *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Lawrence Erlbaum.
- Naigles, Letitia R. 1990. Children use syntax to learn verb meaning. *Journal of Child Language* 17. 357–374.
- Newport, Elissa L. 1990. Maturation constraints on language learning. *Cognitive Science* 14. 11–28.
- Newport, Elissa L. & Richard N. Aslin. 2004. Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48. 127–162.
- Newport, Elissa N., Marc D. Hauser, Geertrui Spaepen & Richard N. Aslin. 2004. Learning at a distance II: Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology* 49. 85–117.
- Pine, Julian M. & Elena V. M. Lieven. 1997. Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics* 18. 123–138.
- Pine, Julian M., Elena V. M. Lieven & Caroline F. Rowland. 1998. Comparing different models of the English verb category. *Linguistics* 36. 807–830.
- Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.
- Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: What’s so special about it? *Cognition* 95. 201–236.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19. 9–50.
- Redington, Martin, Nick Chater & Steven Finch. 1998. Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22. 425–469.

- Regier, Terry & Susanne Gahl. 2004. Learning the unlearnable: The role of missing evidence. *Cognition* 93. 147–155.
- Russell, James. 2004. *What is language development? Rationalist, empiricist and pragmatist approaches to the acquisition of syntax*. Oxford: Oxford University Press.
- Savage, Ceri, Elena V. M. Lieven, Anna Theakston & Michael Tomasello. 2003. Testing the abstractness of children's linguistic representation: Lexical and structural priming of syntactic constructions in young children. *Developmental Science* 6. 557–567.
- Slobin, Dan I. 1973. Cognitive prerequisites for the development of grammar. In Charles A. Ferguson & Dan I. Slobin (eds.), *Studies of child language development*, 175–208. New York: Holt, Rinehart & Winston.
- Slobin, Dan I. 2001. Form-function relations: How do children find out what they are? In Melissa Bowerman & Steve Levinson (eds.), *Language acquisition and conceptual development*, 406–449. Cambridge: Cambridge University Press.
- Smith, Linda B. 1999. Children's noun learning: How general learning processes make specialized learning mechanisms. In Brian MacWhinney (ed.), *The emergence of language*, 277–303. Mahwah, NJ: Lawrence Erlbaum.
- Thelen, Esther & Linda B. Smith. 1994. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Tomasello, Michael. 1992. *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 1995. Language is not an instinct: Book review of Steven Pinker. *The language instinct: How the mind creates language*. *Cognitive Development* 10. 131–156.
- Tomasello, Michael. 1998a. Introduction: A cognitive-functional perspective on language structure. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, vii–xxiii. Mahwah, NJ: Lawrence Erlbaum.
- Tomasello, Michael. 1998b. The return of constructions: Review Essay on: Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. *Journal of Child Language* 25. 443–484.
- Tomasello, Michael. 2000a. Do you children have adult syntactic competence? *Cognition* 74. 209–253.
- Tomasello, Michael. 2000b. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11. 61–82.
- Tomasello, Michael. 2003a. *Constructing a language: A usage-based account of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, Michael (ed.). 2003b. *The new psychology of language, Volume 2: Cognitive and functional approaches to language structure*. Mahwah, NJ: Lawrence Erlbaum.
- Tomasello, Michael. 2004. Syntax or semantics? A reply to Lidz et al. *Cognition* 93. 139–140.
- Tomasello, Michael. & Hannes Racoczy. 2003. What makes human cognition unique? From individual to shared to collective intentionality. *Mind & Language* 18. 121–147.
- Vear, Donna, Letitia R. Naigles, Erika Hoff & Eliane Ramos. 2002. *Grammatical flexibility in early verb use*. Ms. University of Connecticut.