

Beyond formalities: The case of language acquisition

MICHAEL TOMASELLO

Abstract

Generative grammar retained from American structural linguistics the 'formal' approach, which basically effaces the semantic and pragmatic dimensions of grammar. This creates serious problems for an account of language acquisition, most especially the problem of how to link universal grammar to some particular language (the linking problem). Parameters do not help the situation, as they depend on a prior linking of the lexical and functional categories of a language to universal grammar. In contrast, usage-based accounts of language acquisition do not posit an innate universal grammar and so have no linking problem. And if children's cognitive and social skills are conceptualized in the right way, there is no poverty of the stimulus in this approach either. In general, the only fully adequate accounts of language acquisition are those that give a prominent role to children's comprehension of communicative function in everything from words to grammatical morphemes to complex syntactic constructions.

1. Introduction

In the 1960's Chomskian generative grammar effected a revolution in linguistics. But in many ways it was only a technical revolution. American structural linguists were looking for formal distributional patterns in languages, and Chomsky (following Harris) created some especially powerful formal tools for doing this, namely, syntactic transformations. Transformations opened up new avenues of inquiry and enabled the discovery of such things as island constraints – arguably one of the important advances in 20th-century linguistics.

American structural linguists were not interested in psychological reality, and so they explicitly eschewed any account of how people use language to

express meaning and communicate – that was considered a problem for psychology. In apparent contrast, Chomsky argued that linguistics is a branch of cognitive psychology, and that universal grammar is an innate faculty of the human mind. Curiously, however, in this psychologizing of linguistics, communicative function was still considered irrelevant for characterizing grammatical constructions – with attempts in this direction, for example by generative semantics, being explicitly rejected. In this respect, Chomskian linguistics may be seen as the theoretical successor to American structural linguistics, since both ignore for principled reasons the semantic and pragmatic dimensions of grammatical constructions.

The failure to break with structural linguistics on this crucial point has determined much of the history of generative grammar (Langacker 1987, Lakoff 1990). Of particular importance is the account of language acquisition. From the beginning, Chomsky stressed that the question of how human beings acquire linguistic competence is the key empirical phenomenon – because it is here that we most clearly discern ‘the poverty of the stimulus’ and the way that universal grammar (UG) must fill in the missing information. But a crucial challenge is presented by the fact that children may be born into any one of several thousand different languages – which vary from one another in myriad ways – and so UG must also have a way of ‘linking’ itself to particular languages. The initial idea in the early 1980’s was that the linking problem could be solved with a smallish set of innate bi-valued parameters. But specifying exactly what these parameters are has turned out to be extremely difficult, and moreover, it has become apparent in recent years that parameters are of no theoretical help until the linking problem is solved – one cannot set the head direction parameter until one can identify heads in the particular language being acquired. And the linking problem derives, ultimately, from the exclusive focus on the formal dimension – to the neglect of the functional dimension – of grammar.

In this article I argue that generative grammar is not an adequate theory of language acquisition – either with or without parameters – mainly because of the linking problem. I propose an alternative theoretical account that explicitly recognizes the functional dimension of grammatical constructions and the cognitive and communicative sources from which constructions emanate historically and ontogenetically. This account does not rely on any form of universal grammar and so the linking problem does not arise.

2. The problem of linking

Each animal species eats only certain things, and so they are born knowing how to identify those things perceptually. In some animal species individuals

are born alone and live mostly solitary lives, but when it is time to mate they know how to perceptually identify a conspecific of the opposite sex. There is no problem with organisms having adaptive specializations that contain pre-existing "knowledge" of something they are likely to encounter in their environments.

The problem is that language cannot work in this way. It cannot work in this way because each of the world's 6000-plus natural languages is different perceptually, and so any universal that might be a candidate for pre-existing knowledge looks different in different languages. For example, let us say that you have an innate idea of "subject of a sentence" or "clausal head". You now encounter the following utterance (from Slave, an Athabaskan language of northern Canada; Rice 1989, as cited in Dryer in press):

(1) *T'eere li ráreyiht'u.*

What is the subject of this sentence? Or the clausal head? Does it help if you are told that the English translation involves 'girl', 'dog', and 'hit'? The general problem is formulated by Fodor (2001: 765) in this way:

For UG to be truly helpful, it should supply innate sentence structures *and* fix their relation to surface words strings ... As long as there is substantial crosslinguistic variation with respect to how innately defined structure is overtly lexicalized, there will be ambiguities of string-to-structure alignment.

The only way out of the impasse is function. Thus, the only explicit theory addressing the problem is Pinker's (1989) theory of semantic bootstrapping. Pinker proposed the following: (a) a list of key syntactic categories innately given to all human beings, (b) a list of key experiential functions innately given to all human beings, and (c) a set of innate linking rules to connect the two. In the case of 'subject of a sentence', as one instance, the process would work in the following way. 'Subject' is innately linked to 'agent of an action'. Thus, if the child saw a girl hit a dog and heard someone say "The girl hit the dog", she would know through perception and her general causal cognition that the girl is the agent of the action; then, her innate linking rule would automatically connect agent to subject. Because she notices the linguistic form associated with the subject, the child can also now recognize future exemplars of 'sentence subject' on the basis of this form alone (say a particular word order configuration or a particular case marker), even if they are not agents. Thus, the English-speaking child will eventually have to deal with experiential subjects that are not agents (as in *John saw Mary*) and even passive sentences in which subjects are not agents and agents are not subjects.

However, in the specific case of 'sentence subject' it is almost certain that Pinker's proposal is not correct. On general theoretical grounds it has been

known for some time that in ergative languages the notion of ‘subject’ does not operate like it does in English and other accusative languages, and so a direct connection to agent is not possible (and some languages probably do not have subjects at all; Croft 2001). Moreover, even if there were some solution to this problem, there are many languages that are so-called split ergative: some of its constructions are ergative while others are accusative based on such things as person or tense (DeLancey 1981; van Valin 1992). In general terms, Slobin (1997) has made a persuasive case that there is much too much variability across languages for any static and innate look-up table to function in the way it would need to solve the problem of linking (see also Braine 1992). Empirically, the problem with Pinker’s proposal is that children do not always begin with agentive subjects. Many early subjects come from such non-causal utterances as *I like it*, *Maria have it*, *I see it*, and *It has a hole* in which there is no agent of an action at all, and sometimes there are even direct violations in early speech such as *Pete hurt by car* (patient = subject, agent = oblique; Tomasello 1992). And there are no other syntactic roles or grammatical relations – that is, other than sentence subject – for which there is any detailed analysis or supportive evidence at all.

So what could possibly work? The only realistic possibility is that children construct from their experience with a particular language some kinds of grammatical categories, based on the function of particular words and phrases in particular utterances – followed by generalizations across these. Then they could look to see if any of these categories corresponds to anything in UG. But then the notion of UG becomes superfluous. If the child is constructing grammatical categories anyway, then what added value is provided by UG?

3. Do parameters help?

On first blush, one might imagine that the notion of parameters would be helpful in solving the linking problem. But actually the opposite is the case: parameters are of no help until the linking problem is solved since parameters presuppose the existence of certain lexical and functional categories. The point is made quite nicely by Mazuka (1996: 24–25):

Setting a Head Direction parameter by analyzing the syntactic structure of the input involves a paradox. The Head Direction parameter is supposed to determine the order in which the head and complement should appear in the language the child is acquiring. But, for a child to set this parameter, she must first find out which units are the heads and the complements in the sentence she hears. If her linguistic skills are sophisticated enough to know which are heads and complements, she will also know which order they came in. If she already knows

which order the head and the complements come in a sentence, there is no need to set the parameter.

The hard part is thus recognizing ‘heads’ and ‘complements’ in a particular language, and this difficulty is logically prior to any act of parameter setting. That this recognition is far from trivial is highlighted by Fodor (2001: 761):

It would have been more convenient for children if natural language parameters were all concerned with surface facts ... Instead ... the relation between word strings and their parametric generators is opaque at best.

It is also telling that there is no agreement among generative grammarians about what kinds of parameters there might actually be, that is, what kinds of things might be parameterized. Again, Fodor (2001: 734) is especially helpful:

It might turn out that there are 20 parameters or 30 or 100 and more ... On one recent estimate ... there would be at least 32 parameters controlling the landing site for verb movement, perhaps multiplied by the number of possible verb forms (finite/infinitive/past participle, etc.).

And in point of fact the two major proponents of a theory of parameters, Mark Baker (2001) and Janet Fodor (2003) provide lists with very little overlap (see Table 1).

Empirically, there are only a few truly cross-linguistic tests of parameters, and they lead to uniformly negative conclusions. For example, perhaps the best-studied parameter is the null subject (pro-drop) parameter. It specifies that the following features should go together in a language (i.e., because they all result from the setting of the parameter): the possibility of null thematic subjects in tensed clauses, null nonthematic (expletive) subjects, subject inversion, and *that*-trace violations. But as Newmeyer (1998: 357–359) shows, once we get away from European languages on the basis of which this parameter was formulated, the correlations simply do not hold. The other widely cited parameter is the head-direction parameter, and it fares no better in cross-linguistic perspective. In a typological survey of a sample of 100 languages, Dryer (in press: 57) demonstrates that there are many languages where, again, the hypothesized correlations (in this case concerning the placement of heads in various types of phrases and clauses) simply do not hold. And to the extent that these correlations do hold, they are best explained by general processing considerations (Hawkins 1994).

It should also be pointed out that, a priori, the proposal of innately pre-specified parameters of variation in a cognitive skill is an extremely implausible biological mechanism. To my knowledge, no one has proposed anything remotely similar for any other cognitive or social skills that human beings possess. For example, such skills as music and mathematics are, like language,

Table 1. Lists of hypothesized parameters from the two most well-known theorists of parameters, with possible overlaps listed on the same line

Baker (2001)	Fodor (2003)
Polysynthesis	
Adjective Neutralize	
Head Directionality	
Ergative Case	
Topic Prominent	Obligatory Topic
Subject Side	Subject Initial
Verb Attraction	
Subject Placement	
Serial Verbs	
Null Subject	Null Subject
	Object Final
	Complementizer Initial
	V to I Movement
	I to C Movement
	Question Inversion
	Affix Hopping
	Topic Marking
	Wh- Movement Obligatory
	Pied Piping
	Null Topic

unique to humans and universal among human groups, with some variations. But no one has to date proposed anything like Universal Music or Universal Mathematics, and no one has as yet proposed any parameters of these abilities to explain cross-cultural diversity (e.g., +/–mathematical variables, which some but not all cultures use in such things as algebra – or certain tonal patterns in music). Perhaps these empirical and logical problems with parameters are the reason that many contemporary theories (e.g., Minimalism; Chomsky 1995) no longer make use of them.

So parameters as theoretical constructs have not worked empirically, and the immediate prospects for progress do not seem promising. Parameters assume a range of formally specified lexical and functional categories, whose cross-linguistic status is problematic at best (Croft 2001), and whose linking to particular languages in the cases where they do apply seems to be pretty much impossible (unless we want to posit an absolutely incredible number of innate linking rules). One possibility, then, is that the search for formal parameters is a mistake, and we would be much better served by trying to figure out how children construct the particular grammatical categories of their particu-

lar language, with the question of universality being an empirical one – with what language universals there are resulting from universal processes of human cognition, communication, and vocal-auditory processing.

4. Is there really a poverty of the stimulus?

The fundamental argument for the existence of an innate universal grammar is the argument from the poverty of the stimulus. Chomsky has made this clear in a number of places, and it has recently been reiterated by Crain and Pietroski (2001). The problem is that the argument is formulated in terms of a formal generative grammar as adult endpoint and a child who has available only behavioristic learning theory – which enables her only to string words together in a Markov chain (with no understanding of phrasal organization or any other structure-function correlations), making blind associations and inductive inferences in the process (with no conceptual understanding of linguistic function at all).¹ But, as Tomasello (2003) argues, there is no poverty of the stimulus if (1) language is conceived as a set of symbolic instruments for directing the intentional and mental states of others (with complex linguistic constructions being one kind of form-function pairing for doing this; Goldberg 1995); and (2) children are given credit for possessing the cognitive and pragmatic skills necessary for learning such meaningful linguistic symbols and constructions, including such things as categorization, analogy, statistical learning, competition among structures, and so forth.

An excellent example of the problem appears in arguments for so-called structure dependency. Chomsky (1980) gives the following example of question formation in English.

- (2) a. *The man is tall.*
b. *Is the man __ tall?*
- (3) a. *The man who is smoking is tall.*
b. **Is the man who __ smoking is tall?*
c. *Is the man who is smoking __ tall?*

The idea is that forming simple questions such as “Is the man tall?” could be done on the basis of either of two hypotheses: move the first-occurring auxiliary to the front or move the auxiliary from the main predicate to the front. To differentiate between these two hypotheses children supposedly need to see

1. In their recent textbook, Crain and Lillo-Martin (1999) have section headings on classical and instrumental conditioning (which have basically been dead in the study of language acquisition for close to half a century), whose inadequacies in accounting for language acquisition they then go on to demonstrate.

examples like 3c in which the subject NP contains a relative clause with an auxiliary (which did not move to the front). Chomsky (1980: 40) has famously claimed that children almost never hear such sentences. But in an analysis of some written corpora and corpora of child directed speech, Pullum and Scholz (2002) find many of just the right kind of examples that children need, such things as:

- (4) *Can those who are leaving early __ sit near the door?*
 (5) *Is the boy who was crying __ still here?*
 (6) *Could those who are coming __ raise their hands?*

But actually, if one thinks about it for a bit, the child does not really need to encounter such sentences at all. If the child understands NPs with relative clauses – if she understands that the whole phrase is used to make one act of reference – then there would never be any temptation to extract an auxiliary from it; she would simply understand that that unit stays together as one functional unit. It may be said that this is simply another way of stating that the child understands structure dependence. True. And that is the point. If we allow the child to have some notion of meaning or function, then she understands the structure of the sentence to the extent needed to form a conventional English yes-no question. Modern functionalist theorists are not behaviorists who believe the child works with unstructured linear strings, but rather they are cognitivists who believe in structure – just not of the purely formal kind.

A related argument may be made about island constraints. For example, the so-called subjacency constraint posits, among other things, the impossibility of questioning an element embedded too far down and in the wrong way in a sentence structure. For example, 7b seems to be a question form of 7a, but the parallel pair in 8 does not work in the same way.

- (7) a. *Bob believes that Jill married Sam.*
 b. *Who does Bob believe that Jill married __?*
 (8) a. *Bob believes the rumor that Jill married Sam.*
 b. **Who does Bob believe the rumor that Jill married __?*

Subjacency precludes the possibility of moving a WH-word out of an embedded clause which is part of a complex NP (*the rumor* + clause), as in (8b). Subjacency is supposed to demonstrate the argument from the poverty of the stimulus because there is no relevant evidence available to the child in the input (Chomsky 1986).

But van Valin (1998) recasts the constraint by focusing on the role of information structure in question formation: *The element questioned must be in a clause within the potential focus domain of the sentence.* The potential focus

domain may be determined by various kinds of tests including a negation test. Thus, with respect to (7), if we ask “Does Bob believe that Jill married Sam?”, we may say “No, Jeff” (i.e., he believes she married Jeff), which shows that the final NP (Sam) is in the potential focus domain. We may not do the same with (8). If we ask “Did Bob spread the rumor that Jill married Sam?”, and we answer “No, Bill”, this can only mean that Bill is the one who spread the rumor. In general, van Valin (1998: 232) explains the restriction in this way:

Questions are requests for information, and the focus of the question signals the information desired by the speaker. It makes no sense, then, for the speaker to place the focus of the question in a part of a sentence which is presupposed, i.e. which contains information which the speaker knows and assumes that the hearer knows or can deduce easily [e.g., in an object complement, as in (8) – MT]. The content of adverbial clauses and restrictive relative clauses is normally presupposed, and consequently constructing questions with the focus in one of these structures generates a pragmatic contradiction.

We do not need to go into details here. It is enough to simply note that if one takes function into account – in this case pragmatic notions of topic and focus – there is no poverty of the stimulus since even the most abstract syntactic principles from generative grammar may be understood as emanating from the language users’ sensitivity to the communicative function of the pieces of language she is using.

Other examples could be adduced. For example, the binding principles are supposed to require the purely formal syntactic operation c-command – because again the child is supposed to have no relevant linguistic experience – but a much more natural account based on the notion of cognitive domains in discourse is provided by Van Hoek (1997, 2003) – an account which explains not only intra-sentential constraints, as the c-command account, but also inter-sentential constraints in discourse as well. Again, the details are not important for current purposes, but what is important is that there is no poverty of the stimulus if we assume that children are cognitive beings seeking to make sense of the linguistic communication addressed to them.

5. What is the alternative?

So if humans are not born with an innate universal grammar, then how do they construct the abstract dimensions of grammatical structure and why are there universals in this process? The key in usage-based approaches is to explicitly recognize the semantic and pragmatic functions of grammatical constructions, that is, to recognize that linguistic constructions are nothing other than complex linguistic symbols with a form and function (Langacker 1987, 1991;

Goldberg 1995). They are constructed by communities of people historically, via processes of grammaticalization, and re-constructed by individual children ontogenetically.²

In this view, there is no need to posit a specific genetic adaptation for grammar because processes of grammaticalization can actually create grammatical constructions out of concrete utterances – and grammaticalization is a cultural-historical process not a biological one. Thus, it is a historical fact that the specific items and constructions of a given language are not invented all at once, but rather they emerge, evolve, and accumulate modifications over historical time as human beings use them with one another and adapt them to changing communicative circumstances (Croft 2000). Through various discourse processes (involving various kinds of pragmatic inferencing, analogy making, etc.) loose and redundantly organized discourse structures congeal into more tightly and less redundantly organized constructions (see Traugott and Heine 1991, Hopper and Traugott 1993, for some recent research). Universals of language structure emerge not from an innate universal grammar, but rather from the simultaneous interaction of universals of human cognition, communication, and vocal-auditory processing in the process of grammaticalization.

In acquisition, children are confronted not with abstract grammatical principles but with particular, language-specific utterances. For example, the English passive construction conveys an abstract meaning packaged in a certain information structure: an entity has something done to it (with the agent either not expressed or expressed obliquely). Because semantic and pragmatic function are part of the definition of the construction, children are presented with no linking problem and there is no poverty of the stimulus. Indeed, children can learn to use this or any construction appropriately and productively only by witnessing events in the world and attempting to discern the speakers' communicative intention in referring to these events with a certain pattern of linguistic symbols – in much the same way they pair a simple word with some type of environmental scene(s). Given that each language has its own lexical and grammatical conventions, there is really no alternative to this basic account of how children acquire the productive use of the constructions of their language.

2. Some people may doubt that cultural-historical processes can create abstract structures such as those embodied in the grammatical constructions of modern-day languages. But, although the analogy is clearly not perfect, there are many highly abstract structures in modern mathematics that could only have been created by cultural-historical processes since they are not universal among cultures (e.g., those of algebra and calculus). Again, there are many dis-analogies between language and mathematics (which is more closely related, both logically and historically, to written language). The only point is that abstract symbolic systems can be created by groups of human beings working together over historical time in the domain of mathematics, and so perhaps they can also be created in similar yet different ways in the domain of language.

Children then find generalizations that apply across some subset of grammatical constructions and end up with a structured inventory of grammatical constructions, organized in an inheritance hierarchy similar to that used to describe knowledge representation in other branches of cognitive science (Goldberg 1995). Importantly, there are many subsets of constructional patterns in every natural language, including ones that are highly idiosyncratic as compared with the rest of the language (e.g., the ‘let alone’ construction, Fillmore et al. 1988; the ‘nominal extraposition’ construction, Michaelis and Lambrecht 1996; the ‘WXDY’ construction, Kay and Fillmore 1999). The constructional approach attempts to account for all different kinds of constructions, from simple words to abstract constructions that are canonical or idiosyncratic. And indeed, the existence of idiosyncratic constructions – and also ‘mixed’ constructions that are partially canonical and partially idiosyncratic – create insurmountable problems for generative theories that posit two sets of acquisition processes: one for the lexicon and other unruly linguistic structures (words) and another for those that conform with the core requirements of universal grammar (rules) (cf. Pinker 1999). There are many constructions that do not fall neatly into either of these categories. One can also make the argument that if the irregular constructions of a language are acquired by means of general cognitive processes, then there is no reason why the more regular aspects cannot be acquired in the same way.

Universals of linguistic structure derive from the fact that children acquire all linguistic symbols of whatever type with one set of general cognitive processes. Tomasello (2003) argues that we may segregate these general cognitive processes into the two overall headings of: (1) intention-reading, comprising the species unique social cognitive skills responsible for symbol acquisition and the functional dimensions of language, and (2) pattern-finding, the primate-wide cognitive skills involved in the abstraction process. More specifically, these two kinds of general cognitive abilities interact in specific acquisition tasks to yield four specific sets of processes:

- *Intention-Reading and Cultural Learning*, which account for how children learn conventional form-function pairings, including everything from words to complex constructions;
- *Schematization and Analogy*, which account for how children create abstract syntactic constructions (and syntactic roles such as subject and direct object) out of the concrete utterances they have heard;
- *Entrenchment and Preemption (Competition)*, which account for how children constrain their abstractions to those that are conventional in their linguistic community (Brooks and Tomasello 1999; Brooks et al. 1999); and
- *Functionally Based Distributional Analysis*, which accounts for how children form paradigmatic categories of various kinds of linguistic constituents (e.g., nouns and verbs).

Together these processes account for how children construct a language, that is, a structured inventory of linguistic constructions, from the language they hear being used around them. Further insights into how these processes work in detail are given in Tomasello (2003), mainly in the form of patterns of linguistic input that facilitate these processes, for example, type frequency for analogy, token frequency for entrenchment, statistical patterns leading to paradigmatic categories, and so forth. Tomasello (2003) also argues that connectionist accounts – at least in their current form in which almost everything is based on distributional analysis with no account of communicative function – are not sufficient to account for language acquisition.

The same general processes also apply to complex constructions such as sentential complements and relative clauses, which are often organized early in development around specific item-based frames and develop into more abstract constructions only gradually (Diessel and Tomasello 2000, 2001). In terms of constraints, as children attempt to produce creative utterances involving some kind of merging of complex constructions, a number of constraints apply based on such things as semantic and information structure coherence (as in the van Valin example concerning subadjacency and question formation above).

6. Conclusion

The theory of generative grammar has produced some insights into language structure. But those insights have been of a formal nature only, that is, they have been described formally with no explanation of why they are the way they are – except to say that that's the way that universal grammar is. But it has been four decades now since *Aspects of a Theory of Syntax*, and there is less agreement than ever before about precisely what is and what is not in universal grammar – with different accounts focusing on everything from nouns and verbs, to the empty category principle, to linking rules, to c-command, to recursion (Tomasello in press). New discoveries about the structure of non-European languages at one point led to the modification that universal grammar should include parameters. But there is no more agreement on parameters than there is on other aspects of universal grammar, with many modern generative grammarians abandoning the idea of parameters altogether.

As a theory of language acquisition, generative grammar has thus come to an impasse. It has come to an impasse because the hypothesis of an innate universal grammar is characterized in purely formal terms, and so it is impossible to link it up with real linguistic constructions in real natural languages – either with or without parameters. And it is simply not helpful to point to some formal formulation of a complex sentence and say that the child has no evidence for this in her language – that there is a poverty of the stimulus. Quite often

it is the case that children can comprehend these complex sentences without ever having heard ones exactly like it – because they know something of the structure of sentences in terms of the function that is played by different parts (e.g., a complex NP containing a relative clause, or some kind of complement clause as a complex argument). There is no poverty of the stimulus when we credit children with having some functional sense.

The overall point is that if we conceive of language not in purely formal terms in the tradition of American structural linguistics and generative grammar, but rather in cognitive and functional terms as in usage-based approaches such as Langacker (1987, 1991), Goldberg (1995), and Croft (2001), then the child has no problem of linking, she has no need for parameters, and there is no poverty of the stimulus. And if we characterize language in this more functional way, what is required for its acquisition is not an innate universal grammar, but rather language learners who have cognitive and pragmatic abilities that go well beyond the abilities of straw man behaviorism (blind association and induction). With these more powerful cognitive and pragmatic abilities language users grammaticalize linguistic constructions over historical time and reconstruct them during ontogenetic time. My view is thus that it is past time for us to take the final step beyond American structural linguistics to recognize that the fundamental reality of language is the use of linguistic forms, including grammatical constructions, to convey intended communicative functions. Without this recognition, language acquisition cannot be fully explained.

Max Planck Institute for Evolutionary Anthropology, Leipzig

References

- Baker, Mark (2001). *The Atoms Of Language: The Mind's Hidden Rules Of Grammar*. New York, N.Y.: Basic Books
- Braine, Martin D. (1992). What sort of innate structure is needed to “bootstrap” into syntax? *Cognition* 45 (1): 77–100.
- Brooks, Patricia and Michael Tomasello (1999a). How young children constrain their argument structure constructions. *Language* 75: 720–738.
- Brooks, Patricia, Michael Tomasello, L. Lewis and K. Dodson (1999b). Children's overgeneralization of fixed transitivity verbs: The entrenchment hypothesis. *Child Development* 70: 1325–1337.
- Chomsky, Noam (1980). Rules and representations. *Behavioral and Brain Sciences* 3: 1–61.
- (1986) *Knowledge of Language*. Berlin: Praeger.
- (1995). *Minimalism*. Cambridge: MIT Press.
- Crain, Stephen and Diane C. Lillo-Martin (1999). *An Introduction to Linguistic Theory and Language Acquisition*. Oxford: Blackwell Publishers
- Crain, Stephen. and Paul Pietroski (2001). Nature, nurture and Universal Grammar. *Linguistics and Philosophy* 24: 139–185.
- Croft, William (2000). *Explaining Language Change: An Evolutionary Approach*. London: Longmans.

- (2001). *Radical Construction Grammar*. Oxford: Oxford University Press.
- DeLancey, Scott (1981). An interpretation of split ergativity and related patterns. *Language* 57: 626–657.
- Diessel, Holger and Michael Tomasello (2000). The development of relative constructions in early child speech. *Cognitive Linguistics* 11: 131–152.
- (2001). The acquisition of finite complement clauses in English: A usage based approach to the development of grammatical constructions. *Cognitive Linguistics* 12: 97–141.
- Dryer, Matthew (in press). Word order. In *Linguistic typology (2nd edition)*, T. Shopen (ed.). Cambridge: Cambridge University Press.
- Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64: 501–538.
- Fodor, Janet Dean (2001). Setting syntactic parameters. In *The Handbook of Contemporary Syntactic Theory*, M. Baltin and C. Collins (eds.), 730–767. Oxford: Blackwell
- Fodor, Janet Dean (2003). Evaluating models of parameter setting. Handout at LSA Institute Workshop on UG Principles and Input. Michigan State University.
- Goldberg, Adele (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Hopper, Paul J. and Elizabeth Closs Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Kay, Paul and Charles Fillmore (1999). Grammatical constructions and linguistic generalizations. *Language* 75: 1–33.
- Lakoff, George (1990). The Invariance Hypothesis: Is abstract reason based on image schemas? *Cognitive Linguistics* 1: 39–74.
- Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar, Volume 1*. Stanford, CA: Stanford University Press.
- (1991). *Foundations of Cognitive Grammar, Volume 2*. Stanford, CA: Stanford University Press.
- Mazuka, Reiko (1996). How can a grammatical parameter be set before the first word? In *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, J. Morgan and K. Demuth, (eds.), 24–25. Mahwah, NJ: Lawrence Erlbaum
- Michaelis, Laura and Knud Lambrecht (1996). Toward a construction-based model of language function: The case of nominal extraposition. *Language* 72: 215–247.
- Newmeyer, Frederick J. (1998). *Language Form and Language Function*. Cambridge, MA: MIT Press.
- Pinker, Steven (1989). *Learnability and Cognition: The Acquisition of Verb-Argument Structure*. Cambridge, MA: Harvard University Press.
- (1999). *Words and Rules*. New York: Morrow Press.
- Pullum, Geoffrey K., and Barbara C. Scholz (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 9–50.
- Slobin, Dan I. (1997). *The Crosslinguistic Study of Language Acquisition: Vol. 4 and Vol. 5. Expanding the Contexts*. Mahwah, NJ: Lawrence Erlbaum
- Tomasello, Michael (1992). *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press
- (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press
- (in press). What kind of evidence could refute the UG hypothesis? *Studies in Language*.
- Traugott, Elizabeth Closs and Bernd Heine (1991). *Approaches to Grammaticalization, Volumes 1 and 2*. Amsterdam: John Benjamins.
- Van Hoek, Karen (1997). *Anaphora and Conceptual Structure*. Chicago: University of Chicago Press.

- (2003). Pronouns and point of view: Cognitive principles of Coreference. In *The New Psychology of Language Vol. II*, M. Tomasello (ed.), 169–194. Mahwah, NJ: Erlbaum.
- Van Valin, Robert D. (1992). An overview of ergative phenomena and their implications for language acquisition. In *The Crosslinguistic Study of Language Acquisition, Volume 3*, D. I. Slobin (ed.), 47–87. Hillsdale, NJ: Erlbaum.
- (1998). The acquisition of WH-Questions and the mechanisms of language acquisition. In *The New Psychology of Language: Cognitive and Functional Approaches*, Michael Tomasello (ed.), 334–378. Mahwah, NJ: Erlbaum.