

5

The usage-based theory of language acquisition

Michael Tomasello

5.1 Introduction

The usage-based approach to linguistic communication may be summarized in the two aphorisms:

- meaning is use
- structure emerges from use

‘Meaning is use’ represents an approach to the functional or semantic dimension of linguistic communication. It originated with Wittgenstein (1953) and other pragmatically based philosophers of language, who wanted to combat the idea that meanings are things and instead focus on how people use linguistic conventions to achieve social ends. ‘Structure emerges from use’ represents an approach to the structural or grammatical dimension of linguistic communication. It is implicit in the work on grammaticalization and language change of many historical linguists, and has been made explicit by Langacker (1987, 2000) and other usage-based linguists, who want to combat the idea of a wholly formal grammar devoid of meaning and instead focus on how meaning-based grammatical constructions emerge from individual acts of language use.

Drawing on the work of many other researchers, Tomasello (2003) proposes a usage-based theory of language acquisition. Paralleling the two aphorisms above, the proposal is that children come to the process of language acquisition, at around one year of age, equipped with two sets of cognitive skills, both evolved for other, more general functions before linguistic communication emerged in the human species:

- intention-reading (functional dimension)
- pattern-finding (grammatical dimension)

‘Intention-reading’ is what children must do to discern the goals or intentions of mature speakers when they use linguistic conventions to achieve

social ends, and thereby to learn these conventions from them culturally. Intention-reading – including skills of joint attention – is the central cognitive construct in the so-called social-pragmatic approach to language acquisition (which is most often used in the study of word learning; Bruner 1983, Nelson 1996, Tomasello 1992, 2000d, 2001). ‘Pattern-finding’ is what children must do to go productively beyond the individual utterances they hear people using around them to create abstract linguistic schemas or constructions. As a summary term for such things as categorization, analogy and distributional analysis, pattern-finding is the central cognitive construct in the so-called usage-based approach to the acquisition of grammar (Goldberg 1995, 2006, Tomasello 2000a, 2003).

These theoretical positions on the functional and grammatical dimensions of language use and acquisition are minority positions in the field. Essentially, they represent the view that the pragmatics of human communication is primary, both phylogenetically and ontogenetically, and that the nature of conventional languages – and how they are acquired – can only be understood by starting from processes of communication more broadly. In this chapter I provide a synoptic account of the usage-based approach to language acquisition, in both its functional and grammatical dimensions.

5.2 Prelinguistic communication

In the usage-based view one must always begin with communicative function, and it turns out that human infants communicate in some fairly sophisticated ways before they have acquired any linguistic conventions (see Goldin-Meadow Ch. 9). For example, almost all infants communicate by pointing before they have acquired any productive language, and many also use some kind of iconic or conventionalized gestures as well. Interestingly and importantly, other animal species, including our nearest primate relatives, do not communicate with conspecifics in these ways. This suggests that human pointing and other gestures may already embody forms of social cognition and communicative motivation that are unique to the species, and that are necessary as a first step on the way to linguistic conventions both phylogenetically and ontogenetically (Tomasello *in press*).

The interesting thing about pointing is that there is almost no information *in* the gesture itself; it basically says ‘look in that direction and you’ll know what I mean’. So where does the meaning come from? One can say it comes from context, but this has a very special significance with respect to human communication; specifically, it means mutually understood context. One person could point for another in exactly the same way to exactly the same clock on the wall, for example, and mean everything from ‘what a beautiful clock’ to ‘our friend is late’, depending only on their shared

experience and attention prior to the pointing act. From their earliest communicative pointing, infants understand and produce pointing gestures in the context of some such joint attentional frames or common conceptual ground (Tomasello *et al.* 2007). For example, if an 18-month-old girl is engaged in cleaning up toys with an adult, and the adult points to a toy across the room, she will fetch it and clean it up also – assuming that the adult pointing gesture is relevant to their shared activity. But if another person enters the room and points to the exact same toy in the exact same way at a comparable moment, even though the infant herself has been engaged in cleaning up (with the first adult), she does not interpret this pointing gesture as relevant to her own activity egocentrically and so she does not clean up the toy but instead shares attention to it declaratively or gives it to the new adult (Moll *et al.* in press).

Even young infants do not just communicate about what they understand of the world, but about the shared understandings they have with other potential communicative partners. Infants have the ability to construct such shared understandings – in the form of specific formats, scripts, routines or joint attentional frames in specific interactive contexts – from around the first birthday, and these structure their earliest intentional communication (Bruner 1983, Tomasello 1988). The cognitive aspect of these joint attentional frames comprises precisely those conceptualizations that will later structure young children's complex utterances: agents acting on patients, agents giving things to others, objects being in locations or moving to locations, objects changing states, people in various psychological states and so forth. Importantly, when children communicate in specific instances of such situations or events, they comprehend both their role and the role of the communicative partner. For example, in the diary observations of Carpenter *et al.* (unpublished data) a 14-month-old boy on two different occasions wants his chair pushed up to the dining room table in preparation for mealtime. On one occasion he and his mum are standing next to the table and so he points to the chair; on another occasion he and his mum are standing next to the chair and so he points to the table. This suggests that this child already has some understanding – which he knows he shares with his mum – about preparations for mealtime, where his chair goes at the table, and so forth, that serve as a kind of background topic for the communicative act. He then highlights for his mum, by pointing, the aspect of the situation he wants her to focus on – the one that is new for her – so that she can discern his communicative intention (that the chair be placed under the table in its usual place). On other occasions, with a different joint attentional frame as common ground, it is easy to imagine that this child might point to his chair wanting to be placed in it, or point to the empty space at the table simply to indicate dispassionately that the chair that is normally there is missing (and indeed the Carpenter *et al.* observations include several from prelinguistic children indicating absent referents; see also Lizskowski *et al.* 2007).

In terms of communicative motives, it is well established that infants point for both imperative and declarative motives before language (Bates *et al.* 1979, Carpenter *et al.* 1998). Recent research has even documented that 12-month-old infants point helpfully to inform others of things they are ignorant about. For example, if the mother is searching for something and the child knows where it is, even 12-month-old infants will inform her of this with no desire for the object themselves (Lizskowski *et al.* 2006). The imperative, declarative and informative motives underlying infants' pre-linguistic communication are of course exactly the same motives that will structure their early language in the coming months.

Infants' prelinguistic gestural communication, therefore, already includes a species-unique ability to construct with others various kinds of joint attentional common ground to serve as background topic for the attention-directing act of pointing – comprising such things as agents, locations, objects, etc. – as well as species-unique motives for communicating (declarative and informative) that are the exact same motives with which they will use their earliest language. Indeed, many of young children's earliest uses of language are actually accompanied by pointing or other gestures, and these partition the communicative intention in ways that demonstrate the equivalence of gesture and language from a communicative point of view; for example, the child might point to the door while saying "Daddy" to indicate what he might later indicate with "Daddy leave" or some such (Iverson & Goldin-Meadow 2005). In general, prelinguistic communication paves the way for the acquisition of the 'arbitrary' linguistic conventions that infants use, initially, in exactly the same kinds of situations, for exactly the same kinds of communicative motives, as their early gestures.

5.3 Utterances and words

When we turn to children's early linguistic communication, the most basic unit of linguistic experience, and the one with which children begin, is not the word but the utterance. An utterance is the smallest unit in which a person expresses a complete communicative intention – that is, an intention that another person attend to something within the joint attentional frame and so do something as a result – and it thus corresponds to prelinguistic communicative acts such as pointing. Like an act of pointing, an utterance is used to both direct a recipient's attention to something referentially, and also to express a communicative motive (imperative, declarative, informative and others), typically through some form of emotional expression in the face and/or voice. When the child either comprehends or produces an utterance such as 'Birdie!' (to point it out) or 'Hold!' (to request), he or she understands a full communicative act, comprising both reference and motive – even though the form

is simply a single adult word expressed with a certain emotion. These so-called holophrases are thus already, in a very simple way, composite structures.

When an adult speaks to him or her, then, what the child is attempting to do most urgently is to comprehend the overall communicative intention behind the utterance; what does the adult intend for me to attend to and to do in the joint attentional situation? At the same time, he or she is also attempting to determine the communicative function of particular constituents within the utterance. This is a kind of 'blame assignment' procedure in which the child attempts to determine the functional role of a constituent in the utterance as a whole. This requires that the child determine, to some degree of specificity, the communicative intention of the whole utterance; one cannot determine a novel sub-function without knowing something about the overall function. Presumably, particular utterance constituents such as words are most easily identified – and emerge as independent units – when the same phonological form appears in different utterances over time with some functional consistency. Thus, if the child hears 'There's the ball', 'Gimme my ball', 'The ball's rolling', 'The ball's bouncing', 'I want a ball', 'Throw the ball', 'That ball's Jeffery's', 'Where's your ball?', etc., the word *ball* comes to exist as a potential utterance constituent for future use when the child needs to indicate one of a certain class of objects as one sub-function of an utterance. One thing that facilitates this process is if the adult stresses the key word, as an indication of its referential newness, and its associated referent is indeed new to the situation (Grassman & Tomasello 2007).

As a non-linguistic example, a young girl may see her father use a stapler and understand that his goal is to staple together two pieces of paper. In some cases, the girl may understand also that the sub-goal/function of placing the papers inside the stapler's jaws is to align them with the stapling mechanism inside the stapler, and that the sub-function of pressing down on the stapler is to eject the staple through the two papers – with both of these sub-functions being in the service of the overall goal of attaching the two sheets of paper. The girl does not need to understand all of this to mimic an adult stapling papers with the same stapler over and over again (analogy: child can say "There-ya-go" over and over again without understanding its internal constituents). But to the extent that the girl does not understand these sub-functions, she will be lost when she encounters some new stapler in which the sub-functions are effected by a different means, for example, one whose stapling mechanism does not require pressing down but rather squeezing. Only to the extent that the girl understands the relevant sub-functions, will she be able to adapt to new situations creatively by, for example, adjusting her behaviour to effect the same outcome with the new stapling mechanism. In the same way, the child may hear an adult say "I stapled your papers" and comprehend not only the utterance and its overall communicative intention, but also, for

example, the words *I* and *stapled* and their communicative sub-functions in the utterance (the contributions they are making to the utterance as a whole), along with the phrase *your papers* and its communicative sub-function in the utterance (and the sub-sub-functions of *your* and *papers*). As in the case of the stapler, it is only if the child performs some kind of blame assignment that she will be able to comprehend the constituent linguistic elements in a deep enough way to enable her in the future to use them creatively in novel utterances (Tomasello 2003).

This is the way children learn words. That is, children do not try to learn words directly; they try to comprehend utterances and in doing so they often must comprehend a word in the sense of determining the functional role it is playing in the utterance – and they see commonalities in this functional role across utterances. The lexicon, as it were, is thus only an emergent phenomenon in the sense of Bybee (1998). This is true despite the fact that the process is sometimes obscured in Western middle-class culture because parents and children often establish highly frequent utterance schemas for naming objects (e.g. ‘That’s a ___’, ‘It’s a ___’, ‘Here’s the ___’, etc.). Children understand quite well the overall function of these utterances as well as the function of the open slot, with the new word in the slot always serving to name the new object in the situation. This gives the impression that what children are doing is mapping a single word onto a single object or action, or concept thereof, as in most theories of word learning (e.g. Bloom 2000, Markman 1989). But if ‘mapping’ means simply associative learning, this is clearly not how things work. Children are attempting to understand how the adult is using an utterance (and its constituents as sub-elements) to direct their attention. The process is not one of association or mapping but of intention-reading and blame assignment.

We may use children’s learning of new word in an experiment as an example. Akhtar and Tomasello (1996) had an adult set up a joint attentional game with 24-month-old children in which a novel action was performed always and only with a particular toy character on a particular substrate (e.g. Big Bird on a swing, with other character-action pairings demonstrated as well). She then picked up Big Bird and announced “Let’s meek Big Bird”, but the swing was nowhere to be found – so the action was not performed. Children thus never saw the new word *meek* paired with the corresponding action. But later, when the adult handed them a new toy and told them to ‘Meek it’, they searched for (and found) the swing and used it to swing the new character, thus demonstrating their understanding of the action intended. The only way they could do this was to understand the adult’s intentions with respect to the key objects and actions in this jointly understood situation when she originally said “Let’s *meek* Big Bird.” – and something of the particular intentions behind the use of *meek* – even though she never actually did it. That is to say, the child had to identify the aspect of the adult’s overall communicative intention not

covered by the known parts of the utterance *let's* and *Big Bird* and connect it to the unknown word *meek*. To learn a new word, children must extract it from a larger utterance and connect it with the relevant aspect of the joint attentional frame they share with the adult.

In many ways this process is even clearer for word types other than nouns and verbs for concrete objects and actions. Thus, many function words can **only** be learned through efforts to isolate their functional contribution in some larger and less predictable set of phrases. For example, Tomasello (1987) reports that his daughter learned the preposition *of* from such expressions as *piece of ice*, *piece of bread*, *scared of that*, and *scared of monsters*. It is hard to conceive of any method of acquisition here other than some process of extracting *of* from larger expressions and attempting to discern its function in the overall utterance. Levy and Nelson (1994) make a similar argument about children's earliest uses of causal and temporal terms as *because*, *so*, *since*, *and*, *but*, *before* and *if*. And, of course, there can be no question of mapping or association when what is involved is not learning a word per se, but rather learning which referential term of several to choose for a given referent – for example, *the chair* or *that chair in my room* or *it* – in different communicative situations. Learning to make these pragmatic choices in the conventional way – so-called referential choice – requires children to understand why a person chose one means of expression rather than another, that is, her intentions in making the choice (Matthews *et al.* 2006).

5.4 Schemas and constructions

This communication-based, usage-based way of looking at things means we cannot explain children's acquisition of grammatical competence by starting with individual words, learned in isolation, and then gluing them together with abstract meaningless rules, as in the very common 'words and rules' approach (Pinker 1999). Instead, we must begin with children's comprehension and production of whole, meaningful utterances. We then investigate how children *extract* words (with their functions) from utterances and, at the same time, how they find analogical patterns across utterances (based mainly on communicative function) and thereby *abstract* meaningful grammatical constructions.

A linguistic construction is prototypically a unit of language that comprises multiple linguistic elements used together for a relatively coherent communicative function, with sub-functions being performed by the elements as well. Consequently, constructions may vary in their complexity depending on the number of elements involved and their interrelations. For example, the English regular plural construction (N+s) is relatively simple, whereas the passive construction (NP *was* VERBed *by* NP) is relatively complex. Constructions also vary in their abstractness, from abstract

constructions such as the English plural and passive, to various concrete idioms such as *kick the bucket* and *hold one's breath*. Importantly, even the most abstract constructions are still symbolic, as they possess a coherent, if abstract, meaning in relative independence of the lexical items involved (Goldberg 1995). Thus, we know the general profile of the event when we hear 'The dax got mibbed by the gazzer', even though we know none of the individual content words.

Children begin, as noted above, by producing holophrases – one unit utterances with an intonational contour expressing communicative motive. Their earliest multi-unit utterances soon form schemas or constructions, but ones that are highly concrete, not abstract (i.e. based on particular words and phrases not abstract categories). From the point of view of linguistic form, the utterance-level constructions underlying children's earliest multi-word utterances come in three types: word combinations, pivot schemas, and item-based constructions.

5.4.1 Word combinations

Beginning at around 18 months of age, many children combine two words or holophrases in situations in which they both are relevant – with both words having roughly equivalent status. For example, a child has learned to name a ball and a table and then spies a ball on a table and says, "Ball table". Utterances of this type include both 'successive single-word utterances' (with a pause between them; Bloom 1973) and 'word combinations' or 'expressions' (under a single intonational contour). The defining feature of word combinations or expressions is that they partition the experiential scene into multiple symbolizable units – in a way that holophrases obviously (by definition) do not – and they are totally concrete in the sense that they are comprised only of concrete pieces of language, not categories.

5.4.2 Pivot schemas

Beginning at around this same age, however, many of children's multi-word productions show a more systematic pattern. Often there is one word or phrase that seems to structure the utterance in the sense that it determines the speech act function of the utterance as a whole (often with help from an intonational contour), with the other linguistic item(s) simply filling in variable slot(s) – the first type of linguistic abstraction. Thus, in many of these early utterances one event-word is used with a wide variety of object labels (e.g. 'More milk', 'More grapes', 'More juice') yielding a schema such as 'More ___'. Following Braine (1963), we may call these pivot schemas or constructions (see also Lieven *et al.* 1997, 2003).

Not only are pivot schemas organized only locally, but even within themselves they do not have syntax; that is, 'Gone juice' does not mean something different from 'Juice gone' (and there is no other marking to

indicate syntactic role for elements in pivot schemas). The consistent ordering patterns in many pivot schemas are very likely direct reproductions of the ordering patterns children have heard most often in adult speech, with no communicative significance. This means that although young children are using their early pivot schemas to partition scenes conceptually with different words, they are not using syntactic symbols – such as word order or case marking – to indicate the different roles being played by different participants in that scene.

5.4.3 Item-based constructions

Item-based constructions go beyond pivot schemas in having syntactic marking as an integral part of the construction. For example, children barely two years of age respond appropriately to requests that they ‘Make the bunny push the horse’ (reversible transitives) that depend crucially and exclusively on a knowledge of canonical English word order (e.g. DeVilliers & DeVilliers 1973b, Hirsh-Pasek & Golinkoff 1996). However, the syntactic marking in these item-based constructions is still verb specific, depending on how a child has heard a particular verb being used. Thus, in experimental studies, when children who are themselves producing many transitive utterances are taught a new verb in any one of many different constructions, they mostly cannot transfer their knowledge of word order from their existing item-based constructions to this new item until after their third birthdays – and this finding holds in comprehension as well (Tomasello 2000d, 2003). These findings would seem to indicate that young children’s early syntactic marking – at least with English word order – is only local, learned for different verbs on a one-by-one basis. What little experimental evidence we have from nonce verb studies of case-marking languages (e.g. Berman 1993, Wittek & Tomasello 2005) is in general accord with this developmental pattern.

The main point is that unlike in pivot schemas, in item-based constructions children use syntactic symbols such as morphology, adpositions and word order to syntactically mark the roles participants are playing in these events, including generalized ‘slots’ that include whole categories of entities as participants. But all of this is done on an item-specific basis; that is, the child does not generalize across scenes to syntactically mark similar participant roles in similar ways without having heard those participants used and marked in adult discourse for each verb specifically. This limited generality is presumably due to the difficulty of categorizing or schematizing entire utterances, including reference to both the event and the participant roles involved, into more abstract constructions – especially given the many different kinds of utterances children hear and must sort through. Early syntactic competence is therefore best characterized as a semi-structured inventory of relatively independent verb-island constructions that pair a scene of

experience and an item-based construction, with very few structural relationships among these constructional islands.

5.4.4 Abstract constructions

Between two and three years of age, children begin constructing some more abstract constructions, with fewer particular lexical items necessary. However, despite their abstractness, each of these has a particular function in the sense of the communicative contexts in which it is appropriately used. Examples of some early abstract constructions in English are as follows:

5.4.4.1 Identificationals, attributives, and possessives

Serve to identify an object or to attribute to it some property. Most common for the identification function: *It's a/the X*; *That's a/the X*; or *This's a/the X*. Most common for the attributive function: *It's X*; *That's X*. Most common for the possessive function: *(It's) X's* _; *That's X's/my* _; *This is X's/your* _.

5.4.4.2 Simple transitives and intransitives

Serve to indicate or request an activity or state of affairs. Transitives (NP + V + NP): prototype is a scene in which there are two participants and one acts on the other (e.g. *Daddy cut the grass*). Intransitives (NP + V): prototype is an activity involving a single participant; either an actor does something (e.g. *Mummy smiled*; unergatives) or something happens to something (e.g. *The vase broke*; unaccusatives).

5.4.4.3 Datives, ditransitives, and benefactives

Serve to indicate or request the transfer of objects (and other things) between people. Dative (NP + V + NP to NP): *He gave it to Mummy*. Ditransitive (NP + V + NP + NP): *Daddy sent her a present* or *Daddy told me a story*. Benefactive (NP + V + NP for NP): *She did it for me*.

5.4.4.4 Locatives, resultatives, and causatives

Serve to indicate or request spatial or causal relations. Early locatives include such things as *Put NP in/on/ the NP*, *Take NP off my shirt*, *NP's under the NP*, etc. Resultatives indicate outcomes of actions and include such things *NP eat NP all up*, *NP wash it off*, *NP push it down*, etc. Causatives prototypically involve as a first verb *make*, *let* or *help*, as in *Make NP do it*, *Help NP do it* or *Let NP do it*.

5.4.4.4 Passives and reflexives

Serve to indicate things happening to people or things, who are not active agents. Children's early passives (NP + *be/get* + V + *by* NP) are such things as *Spot got hit by a car* or *Mummy got sick* or *It was taken by a bear*. Reflexives are such things as *I hurt myself*.

5.4.4.5 Imperatives and questions

Many of the above construction types can be used as imperatives to request certain kinds of actions, typically without a subject as in: *Push it here, Smile, Don't do that, etc.* Many of the above construction types can be used as questions to request certain kinds of information. While mature questions are quite complex, two very common formulae early on are: *What NP doing?* and *Where NP (going)?* Slightly later they start with such things as: *How do ..., What are ..., and Where is ...*

The key theoretical point is that when we conceptualize children's early grammatical competence not in terms of abstract computational rules with no semantic content, but rather in terms of constructional patterns conventionally associated with particular semantic content, the acquisition processes needed are not so different from those we need for word learning. The child needs first to see that when the adult produces an utterance that fits a particular linguistic pattern (construction), he or she intends a particular kind of meaning. To see similarities among different utterances, young children need skills of schematization and analogy – skills they also use in other domains of cognitive activity (Gentner & Markman 1997).

5.5 Common objections

More formally oriented theorists object on a number of grounds to this usage-based, item-based approach to child language acquisition. The three most common objections are: (1) it cannot deal with more complex constructions, especially those involving two verbs and syntactic embedding; (2) it does not specify how the generalization/abstraction process is to be constrained, and (3) it does not deal with the so-called 'poverty of the stimulus'.

5.5.1 Complex constructions

Many more formally oriented theorists agree that the kind of account given above works for the very earliest stages of language acquisition – for very simple constructions – but it does not work for more syntactically complex constructions. Recent research has found, however, that complex constructions may not be so different if children's actual productions are looked at carefully (Diessel 2004).

For example, among the more complex constructions in English are sentential complement constructions. The prototype is an utterance like 'I know she hit him' and 'I think I can do it'. Diessel and Tomasello (2001) looked at young English-speaking children's earliest utterances with sentential complements from 2 to 5 years of age. They found that virtually all of them were composed of a simple sentence schema that the child had

already mastered combined with one of a delimited set of fixed phrases containing a complement-taking matrix verb (see also Bloom 1992). The matrix verbs were of two types. First were epistemic verbs such as *think* and *know*. As one example, in almost all cases children used *I think* to indicate their own uncertainty about something, and they basically never used the verb *think* in anything but this first person, present tense form; that is, there were virtually no examples of *He thinks ...*, *She thinks ...*, etc. virtually no examples of *I don't think ...*, *I can't think ...*, etc. and virtually no examples of *I thought...*, *I didn't think ...*, etc. And there were almost no uses with a complementizer (virtually no examples of *I think that ...*). It thus appears that for many young children *I think* is a relatively fixed phrase meaning something like *Maybe*. The child then pieces together this fixed phrase [or one of the other similar phrases like *I hope ...*, *I bet ...*, etc.] with a full proposition, with its function being as a sort of evidential marker (not as a matrix clause that embeds another as in traditional analyses). The second kind of matrix verbs were attention-getting verbs like *Look* and *See*, used in conjunction with full finite clauses. In this case, children used these 'matrix' verbs almost exclusively in imperative form (again almost no negations, no non-present tenses, no complementizers), as in 'See the dog eating a bone,' suggesting again an item-based approach not involving syntactic embedding. (See Brandt *et al.* submitted, for very similar findings in German – even though German subordinate clauses have a different word order from main clauses.)

A second example is relative clauses. Textbook descriptions focus on so-called restrictive relative clauses – e.g. 'The dog that barked all night died this morning' – in which the relative clause serves to identify a noun by using presupposed information (both speaker and listener already know that there was barking all night – that's why it can be used as identifying information). Because relative clauses are a part of a noun phrase argument, they are classically characterized as embedded clauses. Diessel and Tomasello (2000) studied four English-speaking children between ages 1;9 and 5;2 in quantitative detail and made a surprising discovery: virtually all of these children's earliest relative clauses were of the same general form, and this form was not the form typically described in textbooks. Examples would be:

Here's the toy that spins around
That's the sugar that goes in there

What is noteworthy here is: (1) the main clause is a presentational construction (predicate nominal or closely related), basically introducing a new topic using a previously mastered fixed presentational phrase such as *Here's the...*, *That's the ...* ; and (2) the information in the relative clause is not presupposed, as in textbook (restrictive) relative clauses, but rather is new information about the just-introduced referent. Again, the main point is that, when examined closely, even this very complex

construction is firmly based in a set of simpler constructions (copular presentationals) that children have mastered as item-based constructions some time before relative clauses are first acquired and produced. Even in German, where again relative clauses have a different word order from simple main clauses, this same basic acquisition pattern is found (Brandt *et al.* in press)

Finally are questions. A particularly interesting phenomenon is so-called inversion errors. English-speaking children sometimes invert the subject and auxiliary in *wh*-questions and sometimes not – leading to errors such as ‘Why they can’t go?’ A number of fairly complex and abstract rule-based accounts have been proposed to account for these errors, but in a more detailed analysis Rowland and Pine (2000) discovered the surprising fact that the child they studied from age 2 to 4 consistently inverted or failed to invert particular *wh*-word-auxiliary combinations on an item-specific basis. He thus consistently said such incorrect things as *Why I can... ? What she will... ? What you can... ?*, but at the same time he also said such correct things as *How did... ? How do... ? What do ... ?* In a recent experiment, Ambridge *et al.* (2006) elicited inversion errors from 4-year-old English children and confirmed this pattern. Young children do not seem to have an overall rule for forming questions, or even *wh*-questions, but rather they have a collection of more item-based schemas that presumably will become a set of more coherent and abstract constructions later in ontogeny.

5.5.2 Constraining constructions

In all theories of language acquisition, there must be some constraints on children’s linguistic generalizations and abstractions. Classically, a major problem for formal theories is that as the rules and principles are made more elegant and powerful through theoretical analyses, they become so abstract that they generate too large a set of grammatical utterances – and so constraints (e.g. the subadjacency constraint) must be posited to restore empirical accuracy. In usage-based theories children are abstracting as they learn, but they cannot do this indiscriminately; they must make just those generalizations that are conventional in the language they are learning and not others. It is thus clear that any serious theory of syntactic development, whatever its basic assumptions, must address the question of why children make just the generalizations they do and not others.

We may illustrate the basic problem with so-called dative alternation constructions. The situation is that some verbs can felicitously appear in both ditransitive and prepositional dative constructions, but others cannot; for example:

He gave/sent/bequeathed/donated his books to the library.

He gave/sent/bequeathed/*donated the library his books.

Why should the other three verbs be felicitous in both constructions, but *donate* be felicitous only in the prepositional dative? The three verbs have very similar meanings, and so it would seem likely that they should all behave the same. Another example is:

She said/told something to her mother.

She *said/told her mother something.

Again, the meanings of the verbs are very close, and so the difference of behaviour seems unprincipled and unpredictable (Bowerman 1988, 1996). Other similar alternations are the causative alternation (*I rolled the ball*; *The ball rolled*) and the locative alternation (*I sprayed paint on the wall*; *I sprayed the wall with paint*) – both of which also apply only to limited sets of verbs.

One solution is quite simple. Perhaps children only learn verbs for the constructions in which they have heard them. Based on all of the evidence reviewed above, this is very likely the case at the earliest stages of development. But it is not true later in development, especially in the 3-to-5-year age period. Children at this age overgeneralize with some regularity, as documented most systematically by Bowerman (1982b, 1988, see Pinker 1989, for a summary of evidence): ‘Don’t giggle me’ (at age 3;0) and ‘I said her no’ (at age 3;1). It is thus not the case that children are totally conservative throughout development, and so this cannot be the whole answer. A second simple but untrue solution is that when children make overgeneralization errors adults correct them, and so children’s overgeneralization tendencies are constrained by the linguistic environment. But this is not true in the sense that adults do not explicitly correct child utterances for their grammatical correctness with any frequency (Brown & Hanlon 1970). Adults, at least Western middle-class adults, do respond differently to well-formed and ill-formed child utterances (e.g. Bohannon & Stanowicz 1988, Farrar 1992), but this kind of indirect feedback is generally not considered by most theorists sufficient to constrain children’s overgeneralization tendencies, and it is far from consistent.

Given the inadequacy of these simple solutions, three factors have been most widely discussed. First, Pinker (1989) proposed that there are certain very specific and (mostly) semantic constraints that apply to particular English constructions and to the verbs that may or may not be conventionally used in them. For example, a verb can be used felicitously with the English transitive construction if it denotes ‘manner of locomotion’ (e.g. *walk* and *drive* as in ‘I walked the dog at midnight’ or ‘I drove my car to New York’), but not if it denotes a ‘motion in a lexically specified direction’ (e.g. *come* and *fall* as in *‘He came her to school’ or *‘She falled him down’). How children learn these verb classes – and they must learn them since they differ across languages – is unknown at this time. Second, it has also been proposed that the more frequently children hear a verb used in a particular construction (the more firmly its usage is entrenched),

the less likely they will be to extend that verb to any novel construction with which they have not heard it used (Bates & MacWhinney 1989, Braine & Brooks 1995, Clark 1987, Goldberg 1995). And third, if children hear a verb used in a linguistic construction that serves the same communicative function as some possible generalization, they may infer that the generalization is not conventional – the heard construction preempts the generalization. For example, if a child hears ‘He made the rabbit disappear’, when she might have expected ‘He disappeared the rabbit’, she may infer that *disappear* does not occur in a simple transitive construction – since the adult seems to be going to some lengths to avoid using it in this way (the periphrastic causative being a more marked construction).

Two experimental studies provide evidence that indeed all three of these constraining processes – entrenchment, preemption and knowledge of semantic subclasses of verbs – are at work. First, Brooks *et al.* (1999) modelled the use of a number of fixed-transitivity English verbs for children from 3;5 to 8;0 years – verbs such as *disappear* that are exclusively intransitive and verbs such as *hit* that are exclusively transitive. There were four pairs of verbs, one member of each pair typically learned early by children and typically used often by adults (and so presumably more entrenched) and one member of each pair typically learned later by children and typically used less frequently by adults (less entrenched). The four pairs were: *come–arrive*, *take–remove*, *hit–strike*, *disappear–vanish* (the first member of each pair being more entrenched). The finding was that, in the face of adult questions attempting to induce them to overgeneralize, children of all ages were less likely to overgeneralize the strongly entrenched verbs than the weakly entrenched verbs; that is, they were more likely to produce ‘I arrived it’ than ‘I comed it’.

Second, Brooks and Tomasello (1999a) taught novel verbs to children 2.5, 4.5, and 7.0 years of age. They then attempted to induce children to generalize these novel verbs to new constructions. Some of these verbs conformed to Pinker’s (1989) semantic criteria, and some did not. Additionally, in some cases experimenters attempted to preempt generalizations by providing children with alternative ways of using the new verb (thus providing them with the possibility of answering ‘What’s the boy doing?’ with ‘He’s making the ball tam’ – which allows the verb to stay intransitive). In brief, the study found that both of these constraining factors worked, but only from age 4.5. Children from 4.5 showed a tendency to generalize or not generalize a verb in line with its membership in one of the key semantic subclasses, and they were less likely to generalize a verb to a novel construction if the adult provided them with a preempting alternative construction. But the younger children showed no such tendency.

Overall, entrenchment seems to work early, from 3;0 or before, as particular verb island constructions become either more or less entrenched depending on usage. Preemption and semantic subclasses begin to work sometime later, perhaps not until 4 years of age or later, as children learn

more about the conventional uses of verbs and about all of the alternative linguistic constructions at their disposal in different communicative circumstances. Thus, just as verb–argument constructions become more abstract only gradually, so also are they constrained only gradually.

5.5.3 Poverty of the stimulus

The fundamental argument for the existence of an innate universal grammar – and against the kind of item-based, usage-based approach advocated here – 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 behaviouristic learning theory – which enables him or 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). But, as Tomasello (2003) argues, there is no poverty of the stimulus if linguistic competence is conceived not as a set of formal, algebraic rules but rather as a structured inventory of meaningful grammatical constructions, with the child possessing sophisticated learning skills involving categorization, analogy and distributional learning. There is certainly no poverty of the stimulus when it comes to the particular constructions children learn. Each of those listed in the preceding section – e.g. transitives, ditransitives, passives, questions, etc. – are heard by young children many dozens or hundreds of times each and every day for several years before they have mastered them on an abstract level (Cameron-Faulkner *et al.* 2003). And, importantly, the acquisition of these constructions is determined in large measure by the frequency (cue availability) and consistency (cue reliability) with which children hear them – along with their complexity (cue cost) of course (Lieven & Tomasello *in press*). Indeed, relatively precise predictions about age of acquisition may be made crosslinguistically by quantifying these three input variables (Bates & MacWhinney 1989, Chan *et al.* *in press*, Dittmar *et al.* *in press*).

The poverty of the stimulus problem only arises in very abstract arguments against approaches that recognized no kind of structure dependency within utterances (again, presumably behaviourism). Chomsky (1980) gives the following example of question formation in English.

- (1) a. The man is tall.
b. Is the man __ tall?
- (2) 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 examples like (2c) 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 Schulz (2002) find many of just the right kind of examples that children need, such things as:

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

But actually, if one thinks about it for a bit, children do not really need to encounter such sentences at all (Elman 2001). If children understand NPs with relative clauses – if they understand 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; they 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 children understand structure dependence. True. And that is the point. If we allow children to have some notion of meaning or function, then they understand structure of sentences to the extent needed to form a conventional English yes–no question. Modern usage-based theorists are not behaviourists 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.

5.6 conclusions

The usage-based theory of language acquisition makes the fundamental claim that language structure emerges from language use. This applies at the level of individual words, as their communicative function derives from their use, as well as at the level of grammar, as structure emerges from patterns of use of multi-unit utterances. Historically, the structure of a language emerges through processes of grammaticalization. Ontogenetically, children hear individual utterances and then (re-) construct the abstract constructions of a language. All of this is done with general cognitive processes, and universals of linguistic structure derive from the fact that people everywhere have the same set of general cognitive processes. As noted at the outset, 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*, which account for how children constrain their abstractions to just those that are conventional in their linguistic community; 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 Lieven and Tomasello (in press) and Abbot-Smith and Tomasello (2006), 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 all aspects of cue validity – and processes of exemplar-based learning and categorization. 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. Children acquire language first and foremost by understanding how others use language.

Suggestions for further reading

- Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.
- Goldberg, A. (2006). *Constructions at Work*. Oxford: Oxford University Press.
- Lieven, E., & Tomasello, M. (in press). Children's first language acquisition from a usage-based perspective. In N. Ellis (Ed.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Lawrence Erlbaum Associates.

- Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (in press). *Origins of Human Communication*. MIT Press.

