Kent Academic Repository

Exemplar-learning and schematization in a usage-based account of syntactic acquisition[1]

KIRSTEN ABBOT-SMITH AND MICHAEL TOMASELLO

Abstract

The early phases of syntactic acquisition are characterized by many input frequency and item effects, which argue against theories assuming innate access to classical syntactic categories. In formulating an alternative view, we consider both prototype and exemplar-learning models of categorization. We argue for a 'hybrid' usage-based view in which acquisition depends on exemplar learning and retention, out of which permanent abstract schemas gradually emerge and are immanent across the summed similarity of exemplar collections. These schemas are graded in strength depending on the number of exemplars and the degree to which semantic similarity is reinforced by phonological, lexical, and distributional similarity.

1. Introduction

If we are interested in the nature of linguistic representations, an important source of information is how such representations are built up during the process of language acquisition. In the domain of syntax, there are basically two types of theoretical account of how this happens. Firstly, linguistic formalists/ 'representational' nativists assume that many of the most important linguistic representations are given biologically and do not need to be built up – they 'just' need to be connected (linked) to the particular language being learned. The assumption is that underlying linguistic competence is a symbolic computational system that processes the relationships between abstract variables, which are basically all-or-none in nature (e.g., Chomsky 1995; Marcus 1998).

Thus, syntactic categories in this approach are classical categories that are defined by the presence or absence of certain defining features.

Secondly, theorists from a more usage-based or emergentist perspective believe that an individual's linguistic representations are built up through mechanisms that are not limited to the linguistic domain, which act on individual usage events of linguistic comprehension and production (e.g., Tomasello, 2003; Elman et al. 1996; MacWhinney 1999; Goldberg 1999, 2006). Experience with each exemplar contributes to linguistic competence, though the precise nature of this contribution is characterized differently in different accounts.

In the current article we argue, first, that approaches that rely on a priori, classical categories in their descriptions of linguistic competence are incompatible with the many asymmetries, frequency, and item effects observed during the early acquisition of a variety of grammatical structures. In contrast, a usage-based account (e.g., Tomasello 2003) expects exactly these kinds of acquisition patterns and effects. The key issue in a usage-based approach is how children make generalizations beyond item-based schemas and constructions. Therefore, our main attempt here is to specify a model of generalization/abstraction processes in early grammatical development. We believe that one crucial set of mechanisms in grammatical generalization are categorization processes. A major point of contention within the categorization literature in general is whether learners develop abstractions that supersede (and essentially efface) the experienced exemplars, as in a 'pure' version of prototype theory or, alternatively, whether generalization occurs solely through online analogical comparison to a set of previously learned exemplars. We opt in the end for a kind of hybrid model comprising both abstractions and the retention of the exemplars of which those abstractions are composed.

2.   Item effects and the inadequacy of classical categories

The central idea behind linguistic nativism (a form of representational nativism) is that children just need to 'link in' to the a-priori categories and this results in a rapid acquisition of syntactic rules. Thus, the assumption of linguistic nativists is that if a child applies a symbolic rule to one member of a category, she should apply it to all other members equally and at the same time (e.g., Radford 1990: 61; Marcus 1998: 250). However, this prediction is not borne out by the data from acquisition. A large number of studies have found that in the initial stages of multiword speech of children's production and comprehension of linguistic constructions tends to be based around particular words which frequently occur in these in the input (see Tomasello 2003, for a review). One especially clear example is the acquisition of WH-questions. Four year-old English children have been found to make many more non-inversion

errors when asking questions beginning with *What do …?* than with *What does …?* (Ambridge et al. 2005). That is, English-speaking children are not learning "*WH-word AUXILIARY SUBJECT … ?*" as an ordering of abstract categories of elements, but rather they are learning patternings of particular words such as *what* and *does*. Moreover, children's rate of error is determined mainly by the frequency of particular WH-words-auxiliary combinations (e.g., *what can …?, where is …?*) in the linguistic input and whether they themselves have previously frequently produced these particular item based-patterns (Rowland and Pine 2000; Rowland 2005).

The most extensive evidence for initial item-based usage comes from English-speaking children's acquisition of SVO word order in the transitive construction involving action verbs. English-speaking children around 24 months, who use active transitive utterances frequently every day with familiar verbs that they have heard being used as transitives, have difficulty applying the transitive SVO pattern to novel verbs (e.g., Tomasello and Brooks 1998; see Tomasello 2000, 2003, for a review), and they are no better in act-out comprehension tasks (Chan et al. 2005).

Perhaps even more telling is the fact that English two-year-olds pass through a stage in which they even inconsistently categorize novel causative verbs within the same task. For example, in the 'weird word order' paradigm of Akhtar (1999), children heard novel verbs modeled for novel events. One was modeled in canonical English SVO order and others in ungrammatical SOV or VSO orders. When the children had heard a novel verb in either the SOV or VSO frames, the four-year-olds corrected to SVO 96 % of the time – showing that children this age indeed do apply the SVO pattern to all members of the causative verb category. But the two-year-olds (mean age of 32 months) wavered between using the novel verbs in the weird orders they had heard and correcting them to SVO (circa 50 % of the time). Younger two-year-olds show even greater willingness to use ungrammatical orders with the novel verbs. Using intransitives (SV and VS order for self-initiated actions), Abbot-Smith, Lieven and Tomasello (2001) found that English 28-month-olds used ungrammatical word orders in 69 % of their utterances with novel verbs but in only 34 % of their utterances with a familiar verb.

These 'weird word order' studies also found evidence for pronoun-specific usage. When the children in Abbot-Smith et al.'s study matched the non-canonical order, they virtually never used a pronoun (see also Akhtar 1999; Matthews et al. 2005). Conversely, when they corrected familiar and/or novel verbs to canonical order, they used pronouns much more often than nouns. Similarly, Dodson and Tomasello (1998) found that when children aged between 29 and 37 months used arguments with a novel verb which they had only heard used without arguments, these arguments were pronouns 90 % of the time. In contrast, when children heard a novel verb used with arguments, they used pro-

276　*Kirsten Abbot-Smith and Michael Tomasello*

nouns in their own production significantly less – only 67 % of the time (see also Childers and Tomasello 2001). Thus, English-speaking two-year-olds are apparently aware that pronouns referring to an agent/subject precede and do not follow the word denoting the action (i.e., verb, predicate) and find it easier to access such patterns to generate novel utterances.

Crucially, in these item-based learning phases, the particular items to which children apply a particular syntactic pattern correctly tend to be those which are heard in this pattern most frequently in the input. An illustration of this can be seen in the 'weird word order' study by Matthews et al. (2005) in which English 33-month-olds corrected highly frequent transitive verbs such as *push* and *pull* to SVO word order significantly more often than transitive verbs of medium frequency (e.g., *shove* and *drag*) or of low frequency (e.g., *ram* and *tug*) (see Matthews, Theakston, Lieven and Tomasello, submitted, for similar findings for French children). This mirrors findings for the acquisition of English past tense morphemes, where it appears that the more frequently a particular irregular verb is heard with a particular past tense form (higher token frequency), the more it becomes *entrenched* and consequently the more protected it is from error (e.g., Marchman 1997; Maslen et al. 2004). On the flip side, the higher type frequency with which young children hear a particular morpheme or syntactic construction, the more likely they appear to be to (over)generalize the pattern. MacWhinney (1978), for example, notes that German children tend to overgeneralize the past tense suffix associated with the verb class with the highest type frequency (*-t*, e.g., *gespielt* 'played') and not the one associated with the verb class with the highest token frequency (*-en*, e.g., *geschlafen* 'slept').

This perspective on the acquisition of regular versus irregular morphology is of course quite different from the symbolic-rule (regular morpheme) versus rote-learned lexical item (irregular morpheme) view held by linguistic nativist researchers such as Pinker (1998) and Marcus (1998), who claim that the rule (the regular form) should be applied in all default contexts regardless of its frequency. In fact, the type frequency of the morpheme or syntactic pattern is not the only determinant of the likelihood with which it will be generalized to a particular item. Rather, the phonological (e.g., Köpcke 1998) or semantic (e.g., Ramscar 2002) similarity between the novel item and other items with which that particular pattern has previously been learned has been found to play a crucial role in the generalization of both so-called regular and irregular morphemes even for adults (see Dabrowska 2004: Chapter 8, for more in-depth discussion).

These findings of initial item-based and frequency-based development in the application of syntactic generalizations appear to conflict with 'symbolic rule + abstract variable' approach of most linguistic nativist theorists. It is true that certain recent linguistic nativist theories emphasize the role of lexical learning. However, what most apparently mean by this is that lexical learning is required

to trigger the acquisition of a particular feature, such as finiteness (e.g., Clahsen, Eisenbeiss and Penke 1996). Once triggering has occurred, however, the rule should be applied universally to all items in the category. This would not predict item-based usage or frequency effects within a particular category once the child has started to apply the rule correctly for certain members of that category. Admittedly, other recent linguistic nativist theories refer to the acquisition of features which are associated with lexemes, such as the copula or the auxiliary DO (e.g., Santelmann et al. 2002). But even this type of formalist account would not predict early lexical specificity in the sense of syntactic rules based around particular forms of one and the same lexeme (e.g., '*do* and not *does*) or particular WH-word+person/tense auxiliary forms and not others. Usage-based theorists, on the other hand, argue explicitly that children initially pass through item-based phases and that frequency affects linguistic representations at all points in development. To take one example, Tomasello (1992) argues that young English-speaking two-year-olds initially know how to mark the verb participant roles for familiar verbs (e.g., English children know that the 'pusher' comes before the verb *push*, and the 'pushee' follows it), but this knowledge is not generalized to other verbs.

3. Psychologically plausible models of categorization in grammatical acquisition

Emergentist and usage-based theorists argue that children gradually generalize over these item-based patterns through a variety of mechanisms including categorization mechanisms (e.g., Tomasello 2003; Goldberg 2006). That categorization is involved in some areas of grammatical acquisition is indisputable – all theories agree that children have to learn grammatical categories. But there are at least two issues here on which the usage-based view differs from that of linguistic nativists. Firstly, in usage-based theory syntactic categories are certainly not classical categories. Secondly, categorization is also involved in the generalization of morphemes or sentence-level constructions to novel words (e.g., Langacker 2000).

To date usage-based theory on how categorization works in grammatical development has not been fleshed out to a great degree. Categorization in the non-linguistic domain has been extensively investigated, which is why we look to these studies for a potential model for grammatical categorization. Currently, theoretical views within the literature on non-linguistic categorization can be divided into two extremes: at the one end, the 'pure' prototype formation view and at the other a 'pure' exemplar category/instance-learning model. The concept of a prototype was developed through psycholinguistic research by Rosch and colleagues (e.g., Rosch 1973; Rosch and Mervis 1975), partly as a reac-

tion to the classical view of categories as being definitively bounded/linearly separable and where individual items within that category must all contain the same attributes or features. These authors established the following empirical phenomena, which do not fit with the classical categories view, regarding how people categorize naturally-occurring or man-made objects. Firstly, not all members of such categories share all the same features. Rather, the members have a family-resemblance relationship to each other, whereby one may share one or several features with another member but not with all members. Secondly, not all members of these categories have equal status. Instead, some items are more typical of a particular category than are others, in that they contain more of the features associated with that particular category. Because of this, these categories are graded with fuzzy boundaries so that they are not necessarily linearly separable from other categories. That is, there is no obvious way to determine to which category marginal members belong simply by comparing features or behavior.

A number of linguists have since presented evidence that syntactic categories also exhibit the same prototype phenomena (e.g., Croft and Cruse 2004; Taylor 1995). We do not have the space to survey this evidence here. However, one example is the distinction between passive participles and adjectives in English and German. Classic linguistic tests for adjectival status include the possibility of using the prefix *un-* , the possibility of occurring in pre-noun position, and semantic reference to the state of an entity. Tests for transitive participle status include the possibility of occurring in an active transitive and co-occurrence with a *by*-phrase. However, many adjectives fail one or more of the 'adjective' tests (e.g., **the un-black/un-alone dog*) and some transitive participles pass both but can also be used with a *by*-phrase (e.g., *the un-washed boy*; *the boy is being washed by his mother*). In German, *un*-prefixation and a *von*-phrase (*by*-phrase) occasionally occur with the same participle within the same stative passive utterance (see, e.g., Lenz 1993).

To account for such phenomena, many researchers in the non-linguistic categorization literature have argued that people abstract a mental representation, which has the typical functional and perceptual features associated with that particular category. Although this schema is abstracted over the features shared by learned exemplars, in the 'pure' versions of prototype theory, the memory traces for these learned exemplars decay with time, so that their more individual characteristics are lost and not used for the categorization of new items. Instead, new items are categorized by direct comparison with the abstract prototype. Some evidence for this view comes from experiments in which adults and infants appear to find 'prototypical' items, which they have never seen, as familiar as – or occasionally even more familiar than – actual training items, especially after a delay of a week (e.g., Strauss 1979; Homa, Sterling and Trepel 1981).

An alternative body of opinion views categorization as instance learning (see Chandler 2002, for a review). Instance learning theorists point to a variety of evidence for faster processing of familiar exemplars even when the familiar items are atypical (e.g., Brooks 1987). Secondly, the processing of new instances cannot be accounted for entirely in terms of comparison with a completely abstract prototype, because learners are sometimes more likely to respond accurately to new exemplars when these share a greater number of features with a particular learned exemplar which is itself actually quite distant from the prototype than they are to respond accurately to new exemplars which are closer to the prototype but which share less features with any particular learned exemplar (e.g., Whittlesea 1987). These two characteristics of non-linguistic categorization clearly tie in with findings that early usage of sentential constructions are based around items such as frequent pronouns or particular WH-word + auxiliary forms.

Even more crucial evidence in favor of exemplar models are a number of automated purely instance-learning computational models which can simulate both empirical prototype phenomena and account for the evidence of instance learning (e.g., Nosofsky 1992; Hintzman 1986; Skousen 1992). Such models can simulate prototype effects by, for example, retrieving training items in parallel and 'summing' the similarity between the features (e.g., Medin and Schaffer 1978; Nosofsky 1992). From this, many have concluded that it is not necessary to postulate the abstraction of a permanent prototype representation at all. The same conclusion has been reached by a few researchers in grammatical acquisition who have claimed that child grammatical productivity occurs solely through online analogical comparison to a set of previously learned exemplars (e.g., Ninio 2003).

4.   Towards a usage-based model

In view of the prevalence of item-based effects and frequency effects in syntactic acquisition (and which remain to some degree in adult language usage, see Dabrowska 2004), exemplar models of categorization are more attractive than a 'pure' prototype-abstraction model in which the extraneous details of original instances are completely lost. Furthermore, such exemplar-learning models are perhaps better able than pure prototype models to explain patterns of family resemblance in syntactic and morphological categories where there is no central tendency (e.g., Bybee 1995). However, we would question the assumption that more abstract prototype categories are only generalized online and leave no permanent representational change. Even in exemplar models every time an exemplar is comprehended, its representation must change in some way, even if this merely involves registering frequency. If the comprehension or produc-

tion of a novel utterance involves 'summing over' similar sets of exemplars, the frequency with which a set is called upon probably also leaves a trace. Therefore, if the mutual similarities of a particular collection of exemplars (such as transitive sentences) are 'summed over' regularly, we believe this is highly likely to permanently change the user's linguistic representations in some way equivalent to the formation of some kind of more abstract representation.

A resolution to the drawbacks of both 'pure' prototype and 'pure' exemplar-learning models is a hybrid in which much of the extraneous details of original instances are retained but where some kind of more abstract schema is gradually formed on the basis of these. It may of course prove difficult to empirically differentiate such a hybrid model from certain 'pure' exemplar-learning models. The chief focus of our current concerns, however, is a more serious problem which plagues all categorization models. That is, how does one determine exactly what the similarity features are which are required for the categorization of new items? This is not such a problem if we restrict ourselves to examining grammatical categorization based on perceptual properties. Indeed, this may be all that is required for the acquisition of some areas of grammar, such as grammatical gender or the formation of English irregular past tense classes (although see Ramscar 2002, for evidence that semantics may even be involved here). Perceptual properties such as prosody must also contribute to the acquisition of question versus declarative constructions. Moreover, if there are salient perceptual cues associated with particular grammatical constructions, this may assist children to acquire the constructions more quickly (see e.g. Braine 1963). Some evidence supporting this comes from Childers and Tomasello (2001), who trained English-speaking two-year-olds with a variety of transitive sentences containing the pronouns *he* and *it* as in *He's swinging it!*; *He's turning it!*. This led to twice the number of children extending this construction to a novel verb than did training with variable nouns only in the slots.

However, such perceptual and distributional cues are clearly not sufficient for target-like acquisition of argument structure constructions. When it comes to functional/semantic categories or other relational categories even exemplar models implicitly invoke a similarity metric, which means that they must assume an underlying similarity structure (e.g., Hampton 1997). One potential solution is to propose that people construct relational and semantic categories in order to make sense of the world and in order to communicate with one another. Thus, even when children are processing and learning their very first exemplars, they do so through particular non-linguistic and communicative biases which constrain how they view the world and interpret the formal grammatical patterns they are hearing. The first kind of constraint involves communicative function. From around nine months, humans are possibly unique in their ability to interpret the communicative intentions of others (e.g., Tomasello 2003); for

example, they are well-equipped to interpret whether a particular utterance is a question, imperative or declarative. The second kind of constraint involves relational similarity within these functions; that is, whether the utterance refers to an event, a process or state. Thirdly, the child's interpretation of the utterance is likely to be constrained by attentional and perceptual biases relating to cues such as self perpetuated and/or goal directed action and causality which infants are sensitive to long before they begin to comprehend multiword utterances (e.g., Carpenter, Akhtar and Tomasello 1998; Csibra et al. 1999; Cohen and Oakes 1993). Other aspects of events and processes which are likely to be particularly salient or of interest to young children are those involving movement, location, state change, possession and time (e.g., Slobin 1985).

Naturally, merely arming ourselves with these functional and relational constraints – and allowing both exemplar storage and schema formation – does not give us a full account of the data from grammatical acquisition. If we assume exemplar retention, why do toddlers not rote-learn everything? Further, the item-based effects do not relate to actual exemplars, as two-year-olds appear easily able to generalize new lexemes to the 'slots'. For example, young two-year-olds are able to create novel sentences whereby they insert novel names for objects into the slots of a familiar verb, such as *put wuggy on there* (e.g., Tomasello et al. 1997).

We argue that the reason for item-based rather than purely rote-learning phases is the following. Every utterance a child hears and processes has a lasting effect on linguistic representations; i.e., it is never completely 'lost' in that the representations completely revert to the state prior to having heard the exemplars. However, a single specific token is unlikely to be recalled due to insufficient phonological encoding and to interference from the many other utterances which have been processed before and are processed subsequently. Therefore, individual sentences will only be rote-learned if they are used by the caregivers with extremely high frequency, such as *What's that?* in English. The majority of sentences that children hear are not exactly the same but overlap to a very high degree in terms of certain lexical items, such as *can you X*? or *X put Y here* (see Stoll, Abbot-Smith, and Lieven, submitted). Taking the latter example, if a child hears and processes a large number of sentences involving the verb *put* such as *shall we put your picture here?*, *put it here!*, *you want to put that on there?*, these are likely to activate very similar patterns during processing, because the event is essentially the same and the phonological form *put* is the same, so the exemplars are thus – to express it in connectionist terms – stored as 'neighbors in state space'. However, the form *your picture* will be overwritten or interfered with by the forms *it* and *that* from the other two exemplars. Thus the remaining representation will essentially be something like *put X*, where X stands for an object. From a process such as this, children could learn to generalize *put* to new objects and essentially show the familiar

verb-based learning phases found for early child English and French.

One potential problem with this view is, if we assume that one-year-olds find elements of events such as causation and movement highly salient, why do they not skip the item-based phase and immediately construct more general schemas based on concepts such as "AGENT ACTION OBJECT", as proposed by several semantic primitives theorists in the 1970s and 1980s (e.g., Slobin 1985; Schlesinger 1988). We believe there are several reasons for this. Firstly, if a child has only learned a relatively small number of exemplars concerning a particular event or set of events, the crucial semantic features which correlate with the formal features will not be immediately obvious. Moreover, there will not yet have been enough overwriting of the more idiosyncratic elements of the events concerned which are not directly relevant for the learning of argument structure. In terms of being able to demonstrate productivity in a novel verb experiment, initially sentences with familiar relational elements such as familiar verbs should be much easier to categorize, because they show the greatest degree of relational similarity to previously learned exemplars. As more and more similar exemplars are learned, however, gradually the relational commonality between these events and relational lexemes will be reinforced, the non-common elements being overwritten or interfered with. Through this language learners reach a point in which novel verbs may be categorized (in this case: generalized to) with equal success if the relational features of the event they refer to match the summed similarity of the exemplar category set. Our view is thus very close to that of Langacker (2000: 7) who claims that the more abstract schema which emerges from the process of exemplar-learning (e.g., English NVN = 'agent verb patient') is not stored separately from the individual learned exemplars, but is rather 'immanent' in the sum of their similarities. Exactly how abstract this 'summed similarity' (or schema, to use shorthand) is will depend on the nature of the exemplars, the frequency with which new items need to be categorized and the semantic or analogical distance between the new instances and the previously learned exemplars.

Thus, the strength of a particular 'more abstract' prototype schema depends on its level of entrenchment, which in turn depends on exactly how many exemplars of that schema the child has processed (e.g., Bybee 1995; Tomasello 2003). The strength of the schema will gradually increase until it reaches the point where it is strong enough to support a particular behavior or productivity measure. Thus, if we see that at a certain point in development a child is able to demonstrate productivity with a particular construction where she was not able to do so before, we must assume that the schema for that particular construction has been gradually strengthening for some time beforehand. If we think about this, this means that this schema must have been abstract in some sense for some time before that child demonstrated productivity with it. Indeed, if we take to its logical conclusion claims made by usage-based grammarians that an

abstract schema is immanent in its instantiations (e.g., Langacker 2000), then a potential abstract schema is latent in the first few item-based constructions that share semantic and structural similarities. That is, from this point on, item-based argument structure constructions are not really 'based solely around lexical items', but are (weakly) interconnected with others which are semantically and functionally related.

This idea that more abstract grammatical representations grow in strength is extremely reminiscent of a theory developed by Munakata et al. (1997) to explain discrepancies between looking and reaching measures within the domain of non-linguistic object permanence development. Munakata et al. argue that the knowledge underlying infants' behavior is best viewed as graded in strength, depending on the number of exemplars which the child has previously experienced. 'Strong' representations allow cleaner signaling to the executive functions mechanisms and memory systems involved in reaching tasks which require young infants to coordinate motor movement with remembering where a hidden object is located.

Similar task-dependent results have been found for the acquisition of the English active transitive construction. That is, English-speaking children have difficulties producing this construction with a novel verb until around 2½ years on average. In preferential-looking studies, however, English-speaking children aged 21 months are more likely to look at to the correct video out of two which have reverse agent and patient roles when they hear a transitive with a novel causative verb (e.g. Gertner, Fisher and Eisengart, in press). Such measure-dependent results have also been found within the same task for the same two-year-olds in 'weird word order' studies. That is, when only actual usage (using grammatical versus ungrammatical orders) was measured, the 28-month-olds in Abbot-Smith et al. (2001)'s study tended to use the novel verbs in the orders in which they had heard, even if these were ungrammatical. However, these children also often 'avoided' using the novel verb they had heard used ungrammatically; in this condition they were four times much more likely to describe the action using a noun-only or verb-only utterance or to select an alternative verb (such as *The cow's doing it*) than in the condition where they had heard the novel verb used grammatically (see also Abbot-Smith, Lieven, and Tomasello 2005, for similar findings). This indicates that English 28-month-olds are not willing to use even novel verbs in ungrammatical linking, which would be surprising, if their representations of the transitive were entirely based around familiar verbs. English children at this age appear to have a 'weak' verb-general representation of the active transitive construction which allows them to avoid using ungrammatical orders and to look at the matching picture in preferential-looking, but is not strong enough to assist them in overcoming the burden of remembering the new word and coordinating production of a multi-word utterance.

Supportive evidence for this 'graded representations' claim comes from a connectionist simulation that uses a single mechanism to account for early success in looking-time tasks and only later success in production tasks (Chang, Dell and Bock 2006). The model learns meaning-form mappings during acquisition that are deployed incrementally during the production or processing of a sentence. The incremental nature of syntax in the model interacts with the way the tasks are evaluated. In production tasks, a complete transitive sentence is the dependent measure, while in preferential-looking tasks, children look back and forth between the two pictures and the percent of preference for the matching picture is the dependent measure. If a child has a weak representation for one position in the sentence, they might not produce a complete transitive, but their matching ability at other positions might be enough to yield an overall matching preference. The model provides a demonstration that this account is consistent with the developmental production and preferential looking data.

One potential criticism of the 'graded representations' account of task-dependent grammatical productivity might be that it is indistinguishable from an alternative 'performance limitations' account. In the latter, some researchers have argued that one- and two-year-olds show a better performance in looking tasks because elicited production and act-out tasks are inherently too difficult for children this age (e.g., Fisher 2002). In addition, Naigles (2002) argues that one reason why children have performed better in the preferential-looking studies is that most of these studies have only required children to notice differences in form and match this with extremely rough differences in meaning. The novel verb elicitation studies (and perhaps also the act-out comprehension studies), on the other hand, might require the child to understand the meaning of the novel verb more exactly for the child to be able to extend the transitive frame to it.

A usage-based 'graded representations' account would make the following predictions which would not be made by a 'performance limitations' account. Firstly, we would predict cross-linguistic variation among young two-year-olds in their elicited production of grammatical constructions depending on the perceptual salience and clarity of form-meaning mapping of the formal cues to the semantics of the construction concerned. This appears to be the case. In German, in which case-marking rather than word order is the major determinant of semantic role interpretation, 28-month-olds correct novel verbs to canonical linking significantly more often than their English age counterparts, demonstrating more robust verb-general representations of agent- and patient-marking (e.g., Abbot-Smith et al 2005). The same may even be true of the act-out task; the poor performance of English children before age 2½ cannot be solely due to memory difficulties since Turkish children of the same age can correctly act out transitive sentences around 80 % of the time at 24 months (e.g., Slobin and Bever 1982). Secondly, the graded representations account

would not only predict age effects but also pure exemplar-learning effects. Indeed, when English-speaking 2½-year-olds hear a large number of exemplars of the transitive construction over four days, their performance on novel verb productivity tests improves significantly (see Childers and Tomasello 2001).

5.   Conclusion

In sum, to account for the phenomena of syntactic acquisition we have proposed a hybrid categorization model with instance learning and retention, but with permanent schema abstraction if this is accessed frequently enough to categorize new instances. The schema abstraction for syntactic constructions is based on a combined perceptual categorization of surface form, on the one hand, and functional-semantic categorization of communicative function, on the other. The former is constrained by human skills of vocal-auditory information processing and statistical learning (e.g., Aslin, Saffran and Newport 1998), whereas the latter is constrained by human skills of intention-reading and analogy making when parsing utterances.

However, it is abundantly clear that we know very little about how all of this actually works in practice. There are very few experimental investigations of the abstraction/generalization process, and its relation to input and input frequencies of various types, and so there remain many unanswered questions regarding exactly which elements of events and constructions are most salient to children in the initial phases of grammar learning and how they use these to make generalizations. A great deal of research will be required to determine exactly what features of events and linguistic constructions children find so similar that they are naturally led to generalize.

Department of Comparative and Developmental Psychology
Max Planck Institute for Evolutionary Anthropology

References

Abbot-Smith, Kirsten, Elena Lieven and Michael Tomasello (2001). What pre-school children do and do not do with ungrammatical word orders. *Cognitive Development* 16: 679–692.

— (2005). The acquisition of linking in English and German basic transitive constructions. Ms.

Akhtar, Nameera (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language* 26: 339–356.

Ambridge, Ben, Caroline Rowland, Anna Theakston and Michael Tomasello (2005). Comparing different accounts of inversion errors in children's non-subject wh-questions: 'what experimental data can tell us?' Ms.

Aslin, Richard, Jenny Saffran, and Elissa Newport (1998). Computation of conditional probability statistics by 8-month-olds. *Psychological Science* 9: 321–324.

Braine, Martin (1963). On learning the grammatical order of words. *Psychological Review* 70 (4): 323–348.

Brooks, Lee R. (1987). Nonanalytic cognition. In *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, U. Neisser (ed.), 141–174. Cambridge: Cambridge University Press.

Bybee, Joan (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455.

Carpenter, Malinda, Nameera Akhtar and Michael Tomasello (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development* 21: 315–330.

Chan, Angel, Kerstin Meints, Elena Lieven, and Michael Tomasello (2005). Young children's comprehension of English word order in act-out and inter-modal preferential looking tasks. Ms.

Chandler, Steve (2002). Skousen's analogical approach as an exemplar-based model of categorization. In *Analogical Modeling*, Royal Skousen, Deryle Lonsdale and Dilworth B. Parkinson (eds.). Philadelphia: John Benjamins.

Chang, Franklin, Gary Dell and Kay Bock (2006). Becoming syntactic. *Psychological Review* 113 (2): 234–272.

Childers, Jane and Michael Tomasello (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology* 37 (6): 730–748.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Clahsen, Harald, Sonja Eisenbeiss and Martina Penke (1996). Lexical learning in early syntactic development. In *Generative Perspectives on Language Acquisition,* Harald Clahsen (ed.), 129–159. Amsterdam: John Benjamins.

Cohen, Leslie and Lisa Oakes (1993). How infants perceive a simple causal event. *Developmental Psychology* 29: 421–433.

Croft, William and D. Alan Cruse (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Csibra, Gergely, George Gergely, Szilvia Biro, Orsolya Koos and Margaret Brockbank (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition* 72 (3): 237–267.

Dabrowska, Ewa (2004). *Language, Mind and Brain: Some Psychological and Neurological Constraints on Theories of Grammar*. Edinburgh: Edinburgh University Press.

Dodson, Kelly and Michael Tomasello (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language* 25: 605–622.

Elman, Jeff, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi and Kim Plunkett (1996). *Rethinking Innateness*. Cambridge, MA: MIT Press.

Fisher, Cynthia (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition* 82: 259–278.

Gertner, Yael, Cynthia Fisher and Julie Eisengart (in press) Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*.

Goldberg, Adele (1999). The emergence of the semantics of argument structure constructions. In *The Emergence of Language*, Brian MacWhinney (ed.), 197–212. Mahwah, N.J.: Lawrence Erlbaum.

— (2006). *Constructions at Work: The Nature of Generalization in Language.* Oxford: Oxford University Press.

Hampton, James A. (1997). Psychological representation of concepts. In *Cognitive Models of Memory*, Martin A.Conway (ed.), 81–109. London: Psychology Press Publishers.

Hintzman, Douglas L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93: 411–428.

Homa, Donald, Sharon Sterling and Lawrence Trepel (1981). Limitations of exemplar-based generalization and the abstraction of category information. *Journal of Experimental Psychology: Human Learning and Memory* 7: 418–439.

Köpcke, Klaus-Michael (1998). The acquisition of plural marking in English and German revisited. *Journal of Child Language* 25: 293–319.

Langacker, Ron (2000). A dynamic usage-based model. In *Usage-Based Models of Language*, Michael Barlow and Susanne Kemmer (eds.), 1–63. Stanford: CSLI Publications.

Lenz, Barbara (1993). Probleme der Kategorisierung deutscher Partizipien. *Wuppertaler Arbeitspapiere zur Sprachwissenschaft* 9: 33–77.

MacWhinney, Brian (1978). *The Acquisition of Morphophonology*. (Monographs of the Society for Research in Child Development, Vol 43). Chicago, IL: University of Chicago Press.

— (1999). The emergence of language from embodiment. In *The Emergence of Language,* B. MacWhinney (ed.). Mahwah, N.J.: Lawrence Erlbaum.

Marchman, Virginia (1997). Children's productivity in the English past tense: the role of frequency, phonology and neighbourhood structure. *Cognitive Science* 21 (3): 283–304.

Marcus, Gary (1998). Rethinking eliminative connectionism. *Cognitive Psychology* 37: 243–282.

Maslen, Robert, Anna Theakston, Elena Lieven and Michael Tomasello (2004). A dense corpus study of past tense and plural overgeneralizations in English. *Journal of Speech, Language, and Hearing Research* 47: 1319–1333.

Matthews, Danielle, Elena Lieven, Anna Theakston and Michael Tomasello (2005). The role of frequency in the acquisition of English word order. *Cognitive Development* 20: 121–136.

Matthews, Danielle, Anna Theakston, Elena Lieven and Michael Tomasello (2005). French children's use and correction of 'weird word order's: a constructivist account. Ms.

Medin, Douglas, and Schaffer, Marguerite M. (1978). Context theory of classification learning. *Psychological Review* 85: 207–238.

Munakata, Yuko, James McClelland, Mark H. Johnson and Robert Siegler (1997). Rethinking infant knowledge: toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review* 104 (4): 686–713.

Naigles, Letitia (2002). Form is easy, meaning is hard: resolving a paradox in early child language. *Cognition* 86: 157–199.

Ninio, Anat (2003). No verb is an island: Negative evidence on the verb island hypothesis. *Psychology of Language and Communication* 7: 3–21.

Nosofsky, Robert M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology* 4: 25–53.

Pinker, Steven (1998). Words and rules. *Lingua* 106: 219–242.

Radford, Andrew (1990). *Syntactic Theory and the Acquisition of English Syntax: The Nature of Early Child Grammars of English*. Oxford: Blackwell.

Ramscar, Michael (2002). The role of meaning in inflexion: why the past tense does not require a rule. *Cognitive Psychology* 45: 45–94.

Rosch, Eleanor (1973). Natural categories. *Cognitive Psychology* 4: 328–350.

Rosch, Eleanor and Carolyn Mervis (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7: 573–605.

Rowland, Caroline (2005). Explaining errors in children's questions. Ms.

Rowland, Caroline and Julian Pine (2000). Subject-auxiliary inversion errors and wh-question acquisition: What children do know? *Journal of Child Language* 27: 157–181.

Santelmann, Lynn, Stephanie Berk, Jennifer Austin, Shamitha Somashekar, and Barbara Lust (2002). Continuity and development in the acquisition of inversion in yes/no questions: dissociating movement and inflection. *Journal of Child Language*, 29, 813–842.

Schlesinger, Izchak (1988). The origin of relational categories. In *Categories and Processes in Language Acquisition*, Yonata Levy, Izchak Schlesinger and Martin Braine (eds.), 121–178. Hillsdale, N.J.: Lawrence Erlbaum.

Skousen, Royal (1992). *Analogy and Structure*. Dordrecht: Kluwer.

Slobin, Dan (1985). Cross-linguistic evidence for the language-making capacity. In *The Cross-Linguistic Study of Language Acquisition*, Vol. 2, Dan Slobin (ed.), 1157–1256. Hillsdale, NJ: Lawrence Erlbaum.

288   *Kirsten Abbot-Smith and Michael Tomasello*

Slobin, Dan and Tom Bever (1982). Children use canonical sentence schemas: A cross-linguistic study of word order and inflections. *Cognition* 12: 229–265.

Stoll, Sabine, Kirsten Abbot-Smith and Elena Lieven (2005). Lexical specificity in Russian, German and English child directed speech. Ms.

Strauss, Mark S. (1979). Abstraction of prototypical information by adults and 10 month-old infants. *Journal of Experimental Psychology: Human Learning and Memory* 5: 618–632.

Taylor, John (1995). *Linguistic Categorization: Prototype in Linguistic Theory.* Oxford: Oxford University Press.

Tomasello, Michael (1992). *First Verbs: A Case Study of Early Grammatical Development.* Cambridge University Press

— (2000). Do young children have adult syntactic competence? *Cognition* 74: 209–253.

— (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Cambridge, MA: Harvard University Press.

Tomasello, Michael and Patricia Brooks (1998). Young children's earliest transitive and intransitive constructions. *Cognitive Linguistics* 9 (4): 379–395.

Tomasello, Michael, Nameera Akhtar, Kelly Dodson, and Lauren Rekau (1997). Differential productivity in young children's use of nouns and verbs. *Journal of Child Language* 24: 373–387.

Whittlesea, Bruce W. (1987). Presence of specific experiences in the representation of generic knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13: 3–17.