

A Process Model of Children's Early Verb Use

Gary Jones (gaj@Psychology.Nottingham.AC.UK)
Fernand Gobet (frg@Psychology.Nottingham.AC.UK)
Julian M. Pine (jp@Psychology.Nottingham.AC.UK)
School of Psychology, University of Nottingham,
Nottingham, NG7 2RD, England

Abstract

The verb-island hypothesis (Tomasello, 1992) states that children's early grammars consist of sets of lexically-specific predicate structures (or verb-islands). However, Pine, Lieven and Rowland (1998) have found that children's early language can also be built around lexical items other than verbs, such as pronouns (this contradicts a strict version of the verb-island hypothesis). This paper presents a computational model (called MOSAIC), which constructs a network of nodes and links based on a performance-limited distributional analysis of the input (mother's speech). The results show that utterances generated from MOSAIC: (1) more closely resemble the child's data than the child's mother's data on which MOSAIC is trained; and (2) can readily simulate both the verb-island and other-island phenomena which exist in the child's data.

Introduction

One of the most influential recent constructivist accounts of early grammatical development is Tomasello's (1992) verb-island hypothesis. According to this view children start producing multi-word speech without knowledge of syntactic categories, such as noun and verb. Instead, children's early language use is based on a "functionally based distributional analysis" (Tomasello, 1992, p.28) of the language they hear. This analysis assigns predicate¹ status to specific words based on their function in sentences. For example, in the sentence "Adam kicks the ball", the roles of Adam and the ball are centred around "kicks", such that Adam is someone who can kick things, and the ball is something that can be kicked. The lexical item "kick" is therefore assigned a predicate role which takes as arguments a "kicker" (Adam) and a "kickee" (the ball).

The notion of "verb-island" arises because most predicates are verbs in adult language and the arguments the predicate takes are specific to that predicate (e.g., "kickers" and "kickees"). Based on this idea, children's early grammar will consist of inventories of verb-specific predicate structures (i.e., verb-islands). For example, the child will use any object which it knows has performed kicking as the antecedent to "kick". Verb-general marking (e.g., knowing that someone who *kicks* can also be someone who *hits*) does not occur until the formation of a verb category.

¹ For Tomasello, a predicate is a lexical item (typically a verb) which forms the main relational structure of a sentence. Arguments are the lexical items to which the predicate relates. Therefore in the sentence "John walks the dog", "walks" is the predicate and "John" and "dog" are the arguments.

In agreement with Ninio (1988), Tomasello argues that children will only start to construct word categories such as noun and verb when they begin to use instances of these categories as the arguments of predicates (e.g., using "ball" as an argument to the predicate "kick"). As verb-islands often use nouns as their arguments, children should form noun categories relatively early in their language development. Verb categories will only be formed later when children begin to use verbs as the arguments of other predicates (e.g., in double-verb constructions such as "Want to + V" or "Can't + V").

The verb-island hypothesis can account for a number of phenomena in children's early multi-word speech. First, it can explain the lexically-specific patterning of children's early verb use. For example, Tomasello (1992) has shown that in the early stages of grammatical development his daughter's ability to generate longer sentences built up piecemeal around particular verbs, and failed to generalise to new verbs which typically entered her speech in very simple structures. Second, it can explain the restricted nature of children's early word order rules. For example, Akhtar and Tomasello (1997) have shown that young children not only fail to generalise Subject-Verb-Object (SVO) word order knowledge from one verb to another, but are also unable to use it as a cue for sentence comprehension with novel verbs. Third, it can explain differences in the flexibility with which children use nouns and verbs in their early multi-word speech. For example, Tomasello and his colleagues have shown that young children will readily use novel *nouns* as arguments in familiar verb structures but tend to restrict their use of novel *verbs* to the structures in which they have heard those same verbs modelled in the input (Akhtar & Tomasello, 1997; Olguin & Tomasello, 1993; Tomasello & Olguin, 1993).

One weakness of the hypothesis is that there are aspects of children's early multi-word speech that do not fit a strict version of the verb-island hypothesis. For example, Pine, Lieven and Rowland (1998) have shown that many children acquire structures based around high frequency items which Tomasello would not define as predicates (e.g., case-marked pronouns such as "I" and "He" and proper-nouns such as "Mummy" and the child's name). Moreover, these pronoun and proper-noun islands not only seem to be functioning as structuring elements in children's speech, but as structuring elements which accept verbs as slot fillers. These data suggest that the lexical specificity of children's early multi-word speech is not always "verb-specificity" or even "predicate-specificity" (because verbs can be slot fillers of

other structures). Verb-island effects may simply be a special case of more general frequency effects on children's acquisition of lexically-specific structures.

This paper presents a computational model called MOSAIC (Model of Syntax Acquisition in Children), which combines naturalistic input (mother's speech) and a performance-limited distributional learning mechanism in order to produce child-like utterances as output. The results will show that MOSAIC is able to: 1) simulate verb-island phenomena that are consistent with children's early multi-word speech; 2) simulate other-island phenomena which exist in children's early multi-word speech but which are problematic for a strict version of the verb-island hypothesis; and hence 3) provide a process-based explanation of why some lexical items come to function as "islands" in the child's grammar and others do not.

The MOSAIC model

MOSAIC is a variant of EPAM/CHREST (De Groot & Gobet, 1996; Gobet, 1996; Gobet & Simon, in press) which creates a discrimination network (a hierarchical structure of nodes which are linked together) based on a given input. Discrimination networks have a root node at the top of the hierarchy, with all other nodes cascading from the root node (see Figure 1 for an example). Nodes are connected to each other by links. This section will describe the basic working of MOSAIC, and then give an example of MOSAIC's learning mechanisms using mother's speech as input.

A general overview of MOSAIC

MOSAIC's discrimination network begins with a root node (which always contains no information). As in other models of the EPAM family (Feigenbaum & Simon, 1984), learning occurs in two steps. The first step involves traversing the network as far as possible with the given input, taking one feature of the input at a time. This is done by starting at the root node and examining all the test links from the root node, selecting the first test link whose test is fulfilled by the first feature in the input (when beginning learning, only the root node will exist and therefore no tests can be fulfilled). The node at the end of the test link now becomes the current node and the next feature of the input is applied to all the test links immediately below this node. The traversal continues until a node is reached where no further traversing can be done (either because the current input feature fulfilled none of the tests of the node's test links, or the current node has no test links below it). Traversing the network in this way is also how information can be output from the network (this will be explained later).

The second step involves adding new information, nodes, and test links. The full input is compared to the information at the final node that was reached by traversal. Based on this comparison, learning can arise in two ways:

1. *Discrimination*. When the input information mismatches the information given at the node (the *image*), a new test link and node are added to the tree below the

node that has just been reached. The new test will relate to the next immediate mismatched feature in the input.

2. *Familiarisation*. When the input information is under-represented by the image (the information given at the node), additional feature(s) from the input are added to the image. The information in the node will contain all information that led to the node during traversal, plus any additional feature(s).

Discrimination therefore creates nodes and test links, and familiarisation creates or modifies the information contained in nodes. The amount of information stored at nodes increases with their distance from the root, because each node contains the accumulation of information of all the nodes that were accessed in traversing to the node.

There are two constraints that are imposed when learning by discrimination and familiarisation. First, before creating a node containing more than one input feature (i.e., a sequence of features), the individual features in the sequence must have been learnt (each input element is said to be a *primitive*). Second, all nodes containing just one input feature are linked to the root node (i.e., all primitives are immediately below the root node; in this way all sequences of input features are below the node which represents the initial feature in the sequence).

Learning can also occur *whilst* traversing the network. MOSAIC compares each node traversed with other nodes in the network to see if they have a similar usage. Similar usage means that there are common test links below each of the two nodes. When this is the case, a lateral link is created between the nodes (this is explained further in the following section).

An example of MOSAIC learning an utterance

The input given to MOSAIC consists of a set of mother's utterances. Each line of input corresponds to a single utterance (delimited by an END marker which signifies the end of the utterance), and each word in the utterance is an input feature. The example utterance "Who came to see you on the train?" will be used as input to illustrate how MOSAIC learns.

The first input feature ("who") is applied to all of the root node's test links in the network. As the network is empty, there are no test links. At this point MOSAIC must discriminate because the input feature "who" mismatches the information at the root node (the root node information is null). The discrimination process creates a new node, and a test link from the root node to the new node with the test "who" (see Figure 1). MOSAIC must then familiarise itself with the input feature, in order to create the "who" information in the image of the node.

When encountering the same input for a second time, the test link "who" can be taken, and the input can move to the next feature, "came". As the node "who" does not have any test links below it, then under normal circumstances discrimination would occur below the "who" node. However, MOSAIC has not yet learnt the input feature

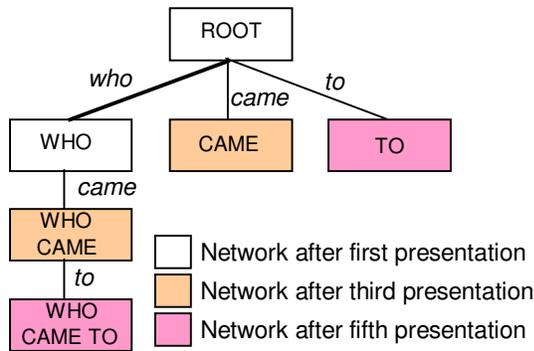


Figure 1: Structure of the MOSAIC net after five presentations of the input “Who came to see you on the train”.

“came”, and so discrimination occurs below the root node. Familiarisation will then fill the image of the new node with “came”. The third time the input is seen, the “who” test link can be taken, and the input can move onto the next feature (“came”). No further test links are available, but the input “who came” mismatches the information at node “who” and so discrimination occurs. A new node “who came” is created (see Figure 1). Familiarisation will fill in the image of the new node.

After a total of five presentations of “Who came to see you on the train?”, the network will have learnt the phrase “Who came to” (see Figure 1). This simple example serves to illustrate how MOSAIC works; in the actual learning phase each utterance is only used once, encouraging a diverse network of nodes to be built.

During traversal of the network, lateral links can be created. A lateral link is a link between any two nodes in a MOSAIC network (excepting the root node). Lateral links are designed to link together nodes which are used in the same manner. Usage is based on the test links that are immediately below a particular node. The way that MOSAIC creates nodes and test links means that all the test links that are below a particular node will consist of the word or words that follow that node in the input (as shown in the previous section). For example, in Figure 2, the words “moves”, “sits”, “walks”, and “chases” must have followed “cat” in the input, meaning sentences such as “cat sits

down” have been seen in the input.

When there is a significant amount of overlap between words or phrases that follow a particular word in the input (i.e., there is significant overlap between the test links that are below two particular nodes) then the two nodes can be linked by a lateral link. The minimum number of test links which must overlap for a lateral link to be created is determined by an *overlap parameter*. Using the network in Figure 2 as an example, “cat” and “dog” will have a lateral link between them when the overlap parameter is set to 3 because at least 3 of the test links below “cat” and “dog” are shared. The next section shows how lateral links are used when generating output from MOSAIC.

Generating utterances from a MOSAIC network

Utterances can be generated from MOSAIC by beginning at the root node and traversing down until encountering a node which contains an END marker (i.e., the last word in the utterance must be one which ended an utterance in the input). Whilst traversing down the network, both test links and lateral links can be taken. To help explain how utterances are generated from the network, test links will be called *rote links* hereafter, and lateral links will be called *generative links*. This is because test links are created from rote learning, and lateral links are created from overlap in node use. When traversing the network, if only rote links are taken then the resulting utterance must have been present in the input (because of the dynamics of the creation of the discrimination network, traversing down from the root node will always produce a phrase that existed as a full utterance or part of an utterance in the input). However, when a generative link is taken, the resulting utterance may never have been seen before in the input.

When generative links exist, MOSAIC can take these links as part of the traversal of the network. For example, in the network shown in Figure 2, the generated utterance could begin with “cat”, take the generative link across to “dog”, and then continue the utterance with any phrase that follows “dog” (i.e., the remainder of the phrase is built up by traversing the nodes below “dog”). This produces novel utterances that were not seen before in the input, such as “cat runs” and “dog moves”. Currently, only one generative link is taken per traversal of the network in order to limit the number of generated utterances (the next section

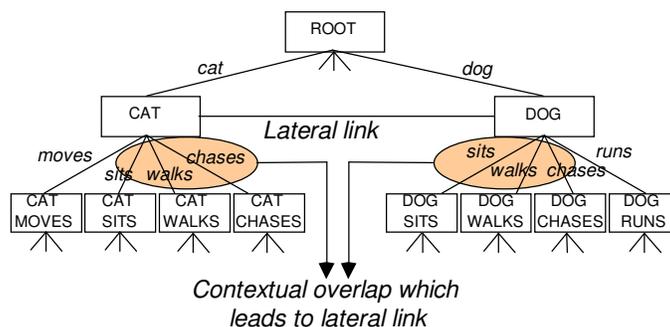


Figure 2: Example of how lateral links are created.

shows that taking only one generative link enables the network to produce over seven generated utterances to every one rote learned utterance).

Modelling verb-island phenomena

The verb-island hypothesis states that children's early language consists of lexical items (typically verbs) existing as predicates, which take other lexical items as arguments. As lexical items such as pronouns cannot, in Tomasello's terms, be predicates, then for flexibility the terms frame and slot filler will be used in place of predicate and argument. A frame is therefore a relational structure of a sentence and the slot fillers to the frame are the lexical items which relate to the frame. For example, the sentence "Daddy moves the chair" has "moves" as the frame and "Daddy" and "chair" as the slot fillers.

The verb-island hypothesis can be confirmed if the language data contain verbs which exist as frames (i.e., verbs which take several different lexical items as slot fillers), and contain *very few other lexical items* which exist as frames. To examine this, the language data will be analysed by extracting verb+common-noun and common-noun+verb sequences. Common-nouns, rather than all lexical items, are examined because: 1) they are the most common category in children's speech; 2) Tomasello (1992) predicts that children form noun categories earlier than verb categories based on their use as slot fillers (i.e., they should be used often as the slot fillers of verb frames); and 3) the analysis is more tractable with only two types of lexical item.

To investigate whether other-island phenomena exists, pronoun+verb and proper-noun+verb combinations will be extracted and analysed. Pronouns are used because a strict version of the verb-island hypothesis does not allow pronouns to act as islands. Also, pronouns occur with high frequency in the child's data and are often followed by a verb (i.e., they may show verbs being used as slot fillers to other frames). Proper-nouns are used for an additional test of other-islands.

Method

Subject data

Three sets of data are compared for the verb-island phenomena: the utterances from one child, Anne; the utterances from Anne's mother; and the utterances from MOSAIC when trained using Anne's mother's utterances as input. The utterances for Anne and her mother were taken from the Manchester corpus (Theakston, Lieven, Pine & Rowland, in press) of the CHILDES database (MacWhinney & Snow, 1990). The corpus consists of transcripts of the mother-child interactions of twelve children over a period of twelve months. The transcripts contain both the utterances

and the syntactic categories (e.g., noun, verb) of all words in the utterances. The child focused on here, Anne, began at age 1;10.7 and completed the study at age 2;9.10. Her starting MLU (Mean Length of Utterance) was 1.62 with a vocabulary size of 180.

For Anne there were 17,967 utterances (i.e., utterance tokens), of which 8,257 utterances were unique (i.e., utterance types). There were 7,331 multi-word utterance types. For Anne's mother, there were 33,390 utterance tokens, 19,358 utterance types, and 18,684 multi-word utterance types. A random sample of 7,331 of Anne's mother's multi-word utterance types were taken to match Anne for quantity of data.

MOSAIC data

MOSAIC was trained on the full 33,390 utterance tokens of Anne's mother in chronological order, one utterance at a time (as a list of words). MOSAIC's overlap parameter was set to 15. The input to MOSAIC did *not* contain any coding information. This means that MOSAIC was not presented with any information about the categories of words (e.g., that "dog" was a noun or "go" was a verb) or about noun or verb morphology (e.g., "going" was seen rather than the morpheme "-ing" attached to the root form of the verb "go").

After MOSAIC had seen all of the input utterances, every possible utterance that could be generated was output. This resulted in 178,068 utterance types (21,510 produced by rote and 156,558 produced by generation). Examples of the utterances generated from MOSAIC are shown in Table 1. The analyses of the data from MOSAIC are based on a random sample of 7,331 (i.e., matching Anne for quantity) of the multi-word utterance types produced by *generation*, because these are the novel utterances that will not have existed as part of the mother's input.

Procedure

The utterances for both the child and mother included the syntactic category for each word in an utterance. The codings for the child's utterances were used to determine the categories of words in the utterances of the child; the codings for the mother's utterances were used to determine the categories of words in the utterances of the mother. Some words (such as "fire") belong to more than one category. In these cases, a category was only assigned if the word was used as that category in at least 80% of the instances in which the word was used. For MOSAIC's utterances, the categories were calculated based on the codings from the mother's utterances.

The three sets of data were analysed in the same way. The method of extracting verb+common-noun combinations is detailed here but the method is the same

for the extraction of common-noun+verb, pronoun+verb, and proper-noun+verb combinations.

Each utterance was searched for a word which was categorised as a verb. The two words following the verb-category word were examined to see if either occurred as a common-noun. If so, the verb+common-noun pair was stored for analysis. Verbs were then converted to their root form (e.g., ‘going’ and ‘goes’ both become ‘go’) and common-nouns to their singular form (e.g., ‘dogs’ becomes ‘dog’), and any duplicate pairs were removed. Analysis was therefore conducted on types, not tokens. The number of slot fillers for a verb is the number of different common-noun types that were paired with that verb.

How well does the output of MOSAIC match the subject data?

Table 2 shows the proportion of pairs that contained a pronoun, a proper-noun, and a common-noun (the label ‘nominal’ refers to the group of all pronouns, proper-nouns, and common-nouns) for Anne, Anne’s mother, and MOSAIC.

The quantitative data show that the utterances from MOSAIC match Anne much more closely than they do Anne’s mother (on whose utterances MOSAIC was trained). For example, there are 378 and 347 pronoun+verb combinations for MOSAIC and Anne respectively, compared with 648 for Anne’s mother. In fact, despite all three datasets having been matched for overall sample size, Anne’s mother produces many more instances of *every* combination shown in Table 2 (e.g., producing over twice as many different nominal+verb combinations [1,204] as Anne [583] and MOSAIC [503]).

The percentage data suggest that the major difference between Anne and her mother in the composition of their nominal+verb and verb+nominal combinations is Anne’s more extensive use of pronouns (59.5% versus 53.8%) and proper-nouns (16.5% versus 11.8%) in nominal+verb combinations. This is a feature of Anne’s data that MOSAIC is partially able to simulate, because MOSAIC has an extensive use of pronouns (75.1%) but not proper-

nouns (8.0%). In fact, the percentage for pronoun+verb combinations for MOSAIC is much greater than for Anne’s mother (75.1% versus 53.8%), suggesting that MOSAIC is actually not learning to produce enough proper-noun+verb and common-noun+verb combinations. One obvious explanation for this problem is that the current version of MOSAIC only scans the data from left to right, meaning that it is unable to acquire structures such as ‘X + gone’ where the verb-island comes *after* the relevant nominal slot.

Verb-islands exist in the data

As explained earlier, the data are expected to show that verbs act as frames (taking lots of different common-nouns as slot fillers) whereas common-nouns are not expected to act as frames. Whether this is true can be examined by looking at the number of common-noun types that follow verb types, and vice versa. We operationalise the concept of an ‘island’ as a lexical item which acts as a frame for at least ten different slot fillers (e.g., a verb type would have to have ten different common-noun types as slot fillers). For example, for Anne, the verb ‘Find’ is an island because it is followed by ten common-noun types (‘Dolly’, ‘Plate’, ‘Seat’, ‘Welly-boot’, ‘Baby’, ‘Ribbon’, ‘Hat’, ‘Duck’, ‘Pen’, and ‘Bird’). Table 3 shows these data for Anne, Anne’s mother, and MOSAIC. This shows that there are many verb-islands for all three sources of data, but very few common-noun islands. In both cases, MOSAIC provides an identical match to Anne for number of islands.

Other-islands exist in the data

Table 3 shows that both pronoun-islands and proper-noun islands exist for Anne, Anne’s mother, and MOSAIC. The pronoun-islands are particularly strong (the mean number of slot fillers for pronouns is more than 20 for all three sets of data) and because pronouns take verbs as slot fillers, these islands are problematic for a strict version of the verb-island hypothesis which predicts that only verbs are initially used as frames. The other-islands, as Table 3 shows, are readily simulated by MOSAIC.

Table 1: Sample of the utterances generated from MOSAIC.

MOSAIC utterance
I forgotten
That’s my toes again
Where’s the magic bag
And she like them
Baby put the sheep in the farmyard
What about the camel
All on the settee
Who can you see on here
He didn’t catch me

Table 2: Number of nominal+verb and verb+nominal combinations accounted for by pronouns, proper-nouns, and common-nouns, based on the 7,331 multi-word utterances from Anne, Anne’s mother, and MOSAIC.

Pair distribution	Anne		Anne’s mother		MOSAIC	
	Nominal+ Verb	Verb+ Nominal	Nominal+ Verb	Verb+ Nominal	Nominal+ Verb	Verb+ Nominal
No. of pairs	583	914	1,204	1,324	503	616
Pronouns	347 (59.5%)	333 (36.4%)	648 (53.8%)	451 (34.1%)	378 (75.1%)	189 (30.7%)
Proper-nouns	96 (16.5%)	44 (4.8%)	142 (11.8%)	109 (8.2%)	40 (8.0%)	47 (7.6%)
Common-nouns	140 (24.0%)	537 (58.8%)	414 (34.4%)	764 (57.7%)	85 (16.9%)	380 (61.7%)

Discussion

The output from MOSAIC more closely resembles the child than the child’s mother, demonstrating that MOSAIC is doing more than just a straightforward distributional

The results presented here show that when combined with naturalistic input, a simple distributional learning mechanism is able to provide an effective simulation of child language data. The simulations show that first, it is possible to model verb-island phenomena as the product of

Table 3: Verb-island data for Anne, Anne’s mother, and MOSAIC (mean=mean number of slot fillers for each frame type; islands=number of frames that have 10 or more slot fillers).

Data source	Mean	Islands	Islands having the most slot fillers
VERB+COMMON-NOUN (frame=verb; slot filler=common-noun)			
Anne	6.24	10	Get, Put, Want, Go, Need, Make
Mother	5.97	13	Get, Put, Want, Need, Have, Find
MOSAIC	9.74	10	Get, Put, Eat, Think, Want, Find
COMMON-NOUN+VERB (frame=common-noun; slot filler=verb)			
Anne	1.51	1	Baby
Mother	2.08	4	Baby, Animal, Dolly, Penguin
MOSAIC	1.57	1	Baby
PRONOUN+VERB (frame=pronoun; slot filler=verb)			
Anne	21.69	10	I, You, He, It, That, They, We
Mother	27.65	11	You, I, He, We, She, They, It
MOSAIC	25.20	12	You, It, That, I, He, We, She
PROPER-NOUN+VERB (frame=proper-noun; slot filler=verb)			
Anne	5.65	3	Anne, Mummy, Daddy
Mother	3.23	3	Anne, Mummy, Daddy
MOSAIC	6.67	2	Anne, Mummy

analysis of its input. In fact, it is a combination of the performance-limitations imposed on the model (e.g., learning one word at a time), and the frequency of occurrence of items in the input, that enable MOSAIC to match the child data. MOSAIC seeks to maximise the information held at nodes in the network, but can only do so for input sequences that occur frequently (e.g., due to limitations in only learning one item at a time). MOSAIC therefore offers a process-based explanation of why some lexical items come to function as ‘islands’ in children’s grammar and others do not: children are maximally sensitive to the high frequency lexical items that exist in their input.

a frequency-sensitive distributional analysis of the child’s input, and, second, that the same mechanism can also simulate other-island patterns which are problematic for a strict version of the verb-island hypothesis.

Acknowledgements

This research was funded by the Leverhulme Trust under grant number F/114/BK.

References

- Akhtar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology*, 33, 952-965.
- De Groot, A. D., & Gobet, F. (1996). Perception and memory in chess: Studies in the heuristics of the professional eye. Assen: Van Gorcum.
- Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science*, 8, 305-336.
- Gobet, F. (1996). Discrimination nets, production systems and semantic networks: Elements of a unified framework. In *Proceedings of the 2nd International Conference on the Learning Sciences*, 398-403. Evanston, IL: Northwestern.
- Gobet, F., & Simon, H. A. (in press). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*.
- MacWhinney, B., & Snow, C. (1990). The Child Language Data Exchange System: An update. *Journal of Child Language*, 17, 457-472.
- Ninio, A. (1988). On formal grammatical categories in early child language. In Y. Levy, I. M. Schlesinger, & M. D. S. Braine (Eds.), *Categories and processes in language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Olguin, R., & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development*, 8, 245-272.
- Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. *Linguistics*, 36, 807-830.
- Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (in press). The role of performance limitations in the acquisition of 'mixed' verb-argument structure at stage 1. In M. Perkins & S. Howard (Eds.), *New directions in language development and disorders*. Plenum.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: CUP.
- Tomasello, M., & Olguin, R. (1993). Twenty-three-month-old children have a grammatical category of noun. *Cognitive Development*, 8, 451-464.