Article

FIRST LANGUAGE

First Language
30(1) 1–23

DOI: 10.1177/0142723709350525
http://fla.sagepub.com

SAGE

The Influence of Frequency and Semantic Similarity on How Children Learn Grammar

Kirsten Abbot-Smith
University of Plymouth

Michael Tomasello
Max Planck Institute for Evolutionary Anthropology

Abstract

Lexically based learning and semantic analogy may both play a role in the learning of grammar. To investigate this, 5-year-old German children were trained on a miniature language (nominally English) involving two grammatical constructions, each of which was associated with a different semantic verb class. Training was followed by elicited production and grammaticality judgement tests with 'trained verbs' and a 'generalization' test, involving untrained verbs. In the 'trained verbs' judgement test the children were above chance at associating particular verbs with the constructions in which they had heard them. They did this significantly more often with verbs which they had heard especially frequently in particular constructions, indicating lexically based learning. There was also an interaction between frequency and semantic class (or the particular verbs). In the generalization judgement test the children were at chance overall. In the elicited production generalization test 75% of the children used the same construction for all items.

Keywords

frequency, miniature language, verb argument structure learning, verb semantics

One of the key issues in language acquisition research is how children eventually come to understand and produce novel sentences. This ability indicates that the child is grammatically productive (i.e., he/she has generalized or abstracted something from his/her input) and is of course essential for developing language mastery. In research on first language grammatical development the 'litmus test' of grammatical productivity is the

Corresponding author:
Kirsten Abbot-Smith, School of Psychology, University of Plymouth, B213 Portland Square, Drake Circus, Plymouth PL4 8 AA, UK; email: kirsten.abbot-smith@plymouth.ac.uk

ability to transfer knowledge of a grammatical pattern or rule to a novel lexical item. Thus, much debate has focused on studies which test children's ability to comprehend or produce sentential constructions such as the English active transitive with novel verbs, as in '*the girl is gorping the boy*' (e.g., Abbot-Smith, Lieven, & Tomasello, 2008; Gertner, Fisher, & Eisengart, 2006).

Such novel verb tests are essential if we are to ensure that the child has attained adult-like grammatical productivity. At the same time, researchers from a wide range of theoretical perspectives emphasize that conservative lexically based learning has been found to play a role even in adult processing of verb argument structure. Thus, syntactic priming effects in adults have been found to be even stronger when the prime and target share the same verb (e.g., Pickering & Branigan, 1998). In addition, adults are quicker at parsing particular verbs in particular sentential constructions dependent on the frequency with which those verbs occur in the target vs. alternative constructions (e.g., Trueswell, Tanenhaus, & Kello, 1993). Similarly, 5- to 6 year-old children and adult speakers of English are more likely to judge, for example, '*the man fell Lisa into a hole*' as incorrect than '*the man tumbled Lisa into a hole*' because '*fall*' is more frequent than '*tumble*' in the language these users are processing on a daily basis (e.g., Ambridge, Pine, Rowland, & Young, 2008; Theakston, 2004). Thus, experienced language users appear to show effects of grammatical 'entrenchment'; that is, the more often a particular verb is heard in a particular grammatical construction, the more the grammatical system is predisposed to consider that construction to be the 'correct' one for that item (e.g., Matthews, Lieven, Theakston, & Tomasello, 2005; Theakston, 2004). The only way to explain how lexical frequency information could play a role in grammatical processing is by assuming that school-age children and adults are storing actual exemplars of previously heard sentences in memory. Thus, it is extremely likely such lexicalist learning also plays a role in grammatical development.

Currently, there are two main theoretical perspectives about how and when conservative lexically based learning plays a role in verb argument acquisition by children (e.g., Pinker, 1989; Tomasello, 2003). According to one view, children's grammatical and lexical systems are independent from the start. Indeed, there is evidence that even before children have learnt much about the lexicon, they are capable of abstracting and generalizing basic sentential constructions, such as the active transitive NP VERB NP (Gertner et al., 2006). Linguistic nativists argue that this early capacity to generalize is constrained by basic (innate) semantics relating to agents, patients and causal actions. Later as the children learn more about the lexicon, they acquire more 'narrow-scope' semantic restrictions on alternation patterns between the active transitive and other sentential constructions, such as the passive (e.g., '*the car got fixed*') and the unaccusative intransitive, such as '*the ball fell*' (e.g., Pinker, 1989). If children have more abstract semantic and syntactic categories from the start, one might predict that children initially generalize sentential constructions more freely and later become more lexically conservative (e.g., Pinker, 1989). This view does, however, allow the opposite prediction: namely that children are initially very lexically conservative in production (because they know that they do not know much about lexical restrictions yet) but show their ability to generalize in comprehension (e.g., Fisher, 2002).

In an alternative view, certain usage-based/construction grammar theorists argue that children's representations of verb argument structure are initially lexically based in that they are closely tied to particular lexical items, especially verbs (e.g., Goldberg, 1999; Tomasello, 2003). These theorists argue that only later do language learners form more abstract semantic-syntactic mappings, which allow them to generalize particular verb argument structure constructions to novel verbs. Even in the adult state, however, these more abstract sentential constructions are represented as the sum or the similarity matrix of the particular stored sentence exemplars which instantiate them (e.g., Langacker, 2000). Whether a particular familiar word is 'entrenched' in a particular grammatical pattern in the mind of the language user depends on the frequency with which that language user has heard that word–grammar pairing in the past. This of course depends on both the relative frequency of the pairing in the input and the amount of experience that language user has had with the target language.

Despite disagreement about when and the degree to which lexicalist conservative learning plays a role, language acquisition theorists from both perspectives agree that semantic analogy may be a key mechanism in the process of forming argument structure generalizations over sentence exemplars heard in the input, at least for some verb classes (e.g., Fisher, 1996; Gentner & Namy, 2006; Goldberg, 1999; Tomasello, 2003). Tomasello (2003, p. 166), for example, argues that children may initially represent the ditransitive construction (NP VERB NP NP) as a set of isolated verb-based schemas such as 'NP GIVE NP NP' and 'NP SEND NP NP'. Gradually they may notice that these schemas share a semantic similarity (the 'transfer' meaning, with NP1 being the 'giver', NP2 being the 'receiver' and NP3 being the 'gift') and shared distributional 'structure'. This may allow the child to extract a form–meaning mapping where the meaning is something like 'GIVER transfers SOMETHING to a RECEIVER'). Then if he/she encounters a new verb which contains these elements of meaning, he/she should be able to extend the form or argument structure to this verb through the process of analogy and structure mapping.

Surprisingly, there have been very few experimental studies which have touched on the role of semantic analogy in the learning of sentential constructions. One is a transitive training study by Abbot-Smith, Lieven, and Tomasello (2004) with English-speaking two-and-a-half-year-olds. Although the factor of semantic similarity between training and test verbs was not directly manipulated in this study, its potential influence could be seen when their results were compared with those of a previous study whose methodology they adapted (Childers & Tomasello, 2001). In Childers and Tomasello both training and test (novel) verbs were of the same semantic class (caused-motion). In Abbot-Smith et al. (2004) only training verbs were of this semantic class. Their test (novel) verbs were verbs of emission, such as '*beep'* or '*flash'*. Thus, if semantic similarity is essential for the proposed process of analogy and structure mapping, one would expect that the two-and-a-half-year-olds in Abbot-Smith et al.'s study should have performed worse on the generalization test than the children the same age in the same condition in Childers and Tomasello's study. In fact, the percentage of two-and-a-half-year-olds who generalized (45%) was exactly the same across both studies. Ninio (2005) interprets this as indicating that semantic similarity is not necessary for generalizing argument structure constructions. Abbot-Smith et al. (2004), however, noted that it is simply not known to date what

counts as semantically similar for children this age. They point out that the children could have interpreted the verbs of emission as having an element of causation which they shared with the training verbs. This is highly plausible as these novel test events involved one hand-puppet acting on the other to cause the second one to emit either sound or light.

A second reason that it is difficult to interpret the findings of Abbot-Smith et al. (2004) is that two-and-a-half-year-olds may already have some kind of representation of the English transitive construction. Therefore these two first language acquisition studies may not have trained these children to generalize the transitive construction per se. Rather, they may merely have given children practice in giving pragmatically appropriate answers to the elicitation questions, since the same elicitation questions were asked during training and testing. This is not merely a problem for these two training studies but for any study which seeks to examine the role of particular generalization mechanisms in first language grammatical development. That is, we can never be sure exactly which aspects of the target grammar have been previously learned nor how well they have been learned. An additional problem which is particular to generalization mechanisms involving semantics is that we cannot easily determine how young toddlers interpret the semantics of events relevant to verbs.

This is one reason why many researchers interested in the ability to learn and generalize grammar have instead taught artificial or miniature grammars to adults or school-age children. Unfortunately, studies of artificial grammar learning without semantics have dominated in this field, and those studies in which grammar had a semantic reference have either not involved event semantics (Meier & Bower, 1986) or have not examined generalization to novel words (Johnston, Blatchley, & Streit Olness, 1990). In addition, it is difficult to extrapolate from most artificial grammar learning studies to real-life first language (L1) grammatical development because artificial grammar learning is essentially second language acquisition (L2). There are differences between first and second language grammatical acquisition, the most noticeable being that most second language learners fail to ultimately attain the level of a native speaker (J. Johnson & Newport, 1989).[1] However, many have argued that there are essential similarities between L1 and L2 in the process of grammatical acquisition. First, the order of acquisition (or degree of difficulty) of grammatical morphemes in L2 has been found to parallel that of L1 and is often clearly determined by the same factors, especially input frequency (e.g., Goldschneider & DeKeyser, 2001). Second, the learning of formulaic phrases as been noted by many in the early stages of both L1 and L2 grammatical acquisition (e.g., Bolander, 1989). Some have also argued that it plays an essential role in both for learning more abstract sentential constructions via a transition involving prefabricated routines with open slots (e.g., Ellis, 2003; Myles, Hooper, & Mitchell, 1998; Weinert, 1995).

That said, there is some evidence that adult (especially tutored) language learners may tend to rely on different learning strategies to those used by children in both L1 and spontaneous L2 (e.g., Yorio, 1989). Certain artificial grammar learning studies have also found differences between adults and children. Hudson Kam and Newport (2005), for example, found that when presented with a language in which nouns occurred with determiners some of the time (varying across between-subjects conditions) adults approximated the same level of inconsistency in their own elicited production. However, 70% of

6-year-olds tested in a similar paradigm were found to over-regularize inconsistent input; that is, input in which nouns were used as bare nouns 40% of the time and with a determiner 60% of the time. These children either used determiners more than 90% of the time or used them less than 10% of the time. Hudson Kam and Newport (2005) interpret their findings as indicating that children impose systematic 'rules' on their input. An alternative explanation is, however, that children might be more inclined to choose one grammatical option (or morpheme, in this case) and stick to it when they have no idea as to the 'correct' response. One indication that this was the case is that the children who apparently 'imposed a rule' were not uniform in which 'rule' they chose. Rather, there were individual differences. Some children decided to ignore the determiners altogether. Some apparently decided that 'this funny language game involves adding the new funny word'. These individual differences among child 'miniature language' learners are interesting in that individual differences also occur in L1 grammatical development (see e.g., Lieven, 2008; Richards, 1990).

Regardless of the reasons, the fact that differences between adult and child learners have been found means that artificial grammar learning studies with children are more likely to reflect the mechanisms of spontaneous grammatical acquisition than are those with adults. Unfortunately, artificial grammar learning studies with children have been far and few between (although see Braine, 1963; Braine, Brody, Fisch, Weisberger, & Blum, 1990; Johnston et al., 1990; MacWhinney, 1983). Although studies such as Hudson Kam and Newport's (2005, Exp 1) are an exception to the former, it is still true that the particular grammatical pattern investigated had no semantic correlates in their study with children (Exp 2). Even their adult study (2005, Exp 1) did not examine generalization to novel words. Therefore it is unclear whether their learners truly learnt to use and grammatically judge particular determiners on the basis of semantics or whether they merely rote-learned noun–determiner combinations on an item-by-item basis.

As far as we are aware, Casenhiser and Goldberg and colleagues are the only group which has used artificial grammar learning with children to examine the roles of event semantics and frequency in grammar learning (Boyd & Goldberg, in press; Casenhiser & Goldberg, 2005). In their first experiment, Casenhiser and Goldberg (2005, Exp 1) taught English-speaking 5- to 7-year-olds (M = 6;4) a novel grammatical construction – SUBJECT-OBJECT-VERB (SOV) – paired with the meaning 'thing appears in a particular location' with five novel verbs. In one condition ('balanced frequency') the learners heard the verbs with relatively low token frequency. That is, they heard eight sentences each for three verbs and four each for two verbs (frequency = 8–8–8–4–4). In the other condition ('skewed frequency') one verb type ('*moop*') was heard with relatively high token frequency (frequency = 16 –4–4–4–4). In the generalization test, the learners heard an untrained novel verb in the SOV construction and had to choose between two possible meanings, one involving appearance and one involving a related action which did not involve appearance. Children in both training conditions chose the appearance interpretation significantly more often than chance (and than the children in the control condition). One conclusion is that the children did this on the basis of semantic similarity between the training and test items. However, an alternative possibility is that the children simply preferred to point at the appearance scene, since that is what they had been trained on. To rule out the latter, in Casenhiser and Goldberg's (2005, Exp 2)

second experiment the test consisted of three SVO sentences as well as three SOV sentences and the choice was always between an appearance scene vs. a scene in which an agent acted on a patient. The children who had been trained on the novel SOV construction were significantly more likely to select the 'appearance' scene for the SOV test sentences than were the control group children. Thus, 5- and 6-year-old children do appear to be able to learn to associate a particular class of verb meanings with *one* novel construction (particularly if there is lexical overlap in the noun phrase between the training and test sentences; see Boyd & Goldberg, in press). However, in real-life learning of a first language children are not just exposed to one construction–meaning pairing at a time, but are exposed to many simultaneously (see e.g., Stoll, Abbot-Smith, & Lieven, 2009).

No previous study has to our knowledge experimentally manipulated semantic similarity between training and test verbs in investigating the ability to generalize *two* argument structure constructions. Therefore, we set out to do exactly this. We choose the English unaccusative intransitive with the form '*NOUN was VERBing*' (e.g., '*Alex was turning*') as one target construction. Because we wanted to ensure that children generalized on the basis of semantic similarity between the unknown and the trained verb, and not merely because the target construction was the only trained construction, we also trained children on an alternative construction – the agentless passive with the form '*NOUN got VERBed*' (e.g., '*Hanna got scared*'). This meant that the children would only have a 50% chance of choosing the target construction in the generalization test. This necessitated that the two constructions be trained with two distinct semantic classes of verbs during training. This is potentially problematic because it is possible that young children see semantic similarities between certain verb classes which are often analysed as linguistically distinct (Abbot-Smith et al., 2004; Brooks & Tomasello, 1999). Therefore, we chose two verb classes which were as semantically distinct as possible, namely verbs of caused motion (e.g., *bend*, *topple*) and verbs of psychological state (e.g., *annoy*, *bore*).

We choose to follow Casenhiser and Goldberg (2005) in training novel grammatical constructions in order to ensure that our learners really were learning the argument structure constructions from scratch during the experiment. This is probably not the case for first language grammar training studies such as that of Abbot-Smith et al.'s (2004) study. We were somewhat adverse to using a purely artificial language and grammar because by 'cleaning up' the natural language input, many artificial grammar learning studies may obscure many factors usually involved in real language learning (see e.g., E. Johnson & Jusczyk, 2001, for evidence that the ability to learn the statistical cues of an artificial language may not be the predominant means by which infants learn real language). On the other hand, we were concerned that by presenting a novel construction in the children's first language they might be encouraged to treat the whole experience as a 'game' rather than 'real language learning'. Therefore, we told our preschool German participants that they were learning a real 'second' language, namely English. This had the additional advantage of making the training task more exciting as 'learning English' is something that 'big' children do (i.e., from the age of 10 years). (German preschool children are not generally exposed to English as all television and films are dubbed into German and all radio programmes are in German, at least in Leipzig where we tested.)

Our hypotheses were as follows. If learners who are just starting to break into a grammatical system tend to prefer lexical conservativism, our participants should just use and preferentially choose (in grammaticality judgement) the argument structure in which they have heard those particular verbs used. This predicts, first, that they should perform well on a test of the trained verbs but at chance when asked to generalize to novel verbs; and second, that they should be more likely to choose the target construction for a trained verb, the more frequently they have heard that particular verb in that particular construction. Alternatively, if children generalize but are constrained by verb class semantics, they should perform above chance on both the trained verbs and the generalization tests (by choosing the '*was __ing*' form for the motion verbs and the '*got __ed*' form for the psychological verbs. A third alternative is that children are generalizers unconstrained by verb semantics. In this case they should choose either the '*was __ing*' form OR the '*got __ed*' form for all items, regardless of frequency of occurrence or semantics.

These hypotheses may need to be qualified in the light of Lieven's (2008) and Richard's (1990) findings on individual differences in syntactic development. First, we may find that some learners tend to search for 'rules' and overextend these, whereas others of the same age are more conservative. Second, some learners may need more time and input to grasp the task than do others.

Method

Participants

The participants were 16 monolingual normally developing German-speaking children – seven girls and nine boys – aged between 5;6 and 5;11 ($M = 5;8$). A further seven children were tested but excluded from the study due to either showing a side bias during the German pointing pre-test (2 children), experimenter error (2), major inattentiveness (1), or because they were absent from kindergarten in the final session (2). At the end of session one, each child completed the morphological productivity subtest of the Sprachentwicklungstest 3–5 (language development test – SETK) for 3- to 5-year-olds (Grimm, 2001), in order to ensure that all were within the normal range for their (first) language development. This was indeed the case.

Importantly, all our participants had no prior exposure to English (except possibly through hearing pop songs). We selected the children by checking with kindergarten teachers and parents that the children had had no prior exposure to English either in the home, on holiday or through taught classes. (Some Leipzig kindergartens did teach English for one hour a week, so they were excluded prior to the study.)

Training Design

Each child was trained on two verb classes – caused-motion and psychological state – each of which was associated with one argument structure. The forms of both were distinct from the German equivalents. The motion verbs were heard in the English unaccusative intransitive in the form '*X was VERBing*' (which would be translated as *X hat sich geVERBt* or *X VERBte sich* in German). The psychological verbs were heard in the

Table 1. The Frequency with which the Verb Types Were Heard during Training and Testing

	Semantic Condition 1: Caused Motion (always heard in intransitive 'NOUN was __ing')	Semantic Condition 2: Psychological State (always heard in passive 'NOUN got __ed')
Training:	**swing (x 4 items)**	**annoy (x 4 items)**
Stories 1 & 3:	bend (x 2)	surprise (x 2)
	slide (x 2)	bore (x 2)
	topple (x 2)	tempt (x 2)
Stories 2 & 4/5:	**swing (x 4 items)**	**annoy (x 4 items)**
	turn (x 2)	scare (x 2)
	bounce (x 2)	embarrass (x 2)
Generalization test	twist (x 2), drop (x 2)	worry (x 2), grieve (x 2)

English (agentless) passive in the form '*X got VERBed*' (which would be translated as *X wurde geVERBt/en* or *X ist geVERBt/en worden* in German). For the learners there was no indication which form might be passive and which form might be unaccusative intransitive and thus the two had equal structural complexity.

For each verb class (motion and psychological), each child was trained on six verb types over five sessions. Each verb was heard in conjunction with a film scene involving two participants. For each of our semantic verb classes, one verb type was heard four times as often as the other verb types; that is, '*swing*' for motion and '*annoy*' for psychological were heard with a total of 16 items each over the five sessions. The other verb types were heard with two different film scene items per session and appeared in two of the four stories (making a total of four film items/sentences per verb over the five sessions). Thus, for each semantic condition, the frequency was skewed in the same fashion as in Casenhiser and Goldberg's (2005) study, but the current study had one extra verb type (namely 16–4–4–4–4–4). The verb types are shown in Table 1.

Training Procedure

First, prior to session one, each child was trained to understand that when he/she was asked to describe something, this required a full sentence (and not just one word). In this phase, the child was shown six pictures of people carrying out various actions and heard the same prompts that he/she would hear during the experiment, namely (the German equivalent of) '*the word for what is happening is called X [e.g., klettern = climbing.]. How do you think you describe what's happening?*' The child was also given corrective feedback, in the (German equivalent of the) form *'that's right! When you describe it, it sounds like this: (e.g., "the boy is climbing")*'. It was hoped that this would facilitate the elicited production tests at the end of sessions four and five.

Then, for the experiment itself, the children were presented with films of real actors involving four characters in a series of events which formed a story. A prerecorded

Table 2. Examples of the Story + Visual Scene Context for One Motion Verb and One Psychological Verb Item (the Second Item Followed the First Consecutively as Part of the Narrative)

Verb	German Narration	Visual Scene
Caused-motion: *swing*	*'Now they are playing on the playground. What they are doing can be done on almost any playground? Yes, that's great, back and forwards, back and forwards. That's how you do what they're doing here.'*	Alex swings Hanna (camera includes full body shot of two participants including swing)
Psych. verb: *bore*	*'They've been doing that for a long time. It's fun for Hannah, but Alex doesn't want to do it anymore. After a while it's silly – over and over always the same thing!'*	Hanna bores Alex (camera zooms to Hanna and Alex's faces: Hanna = neutral and Alex displays a bored facial expression)

narrator told a story in German. At the end of each scene one of two hand-puppets (Susi or Lotte) said the target training sentences in English. Inside both hand-puppets were loudspeakers attached to two separate i-Pod players (held under the table by the experimenter) which alternately emitted the prerecorded training sentences at the appropriate moment (both in female voices). No translation of the target verbs was given in the German narration but it was intended that an approximate meaning of the verb be inducible from the combination of the story's context and the film scene they were currently watching. An example of part of the German story context – translated into English – is given in Table 2.

There were four stories which contained the same verbs (two stories set in a playground and two set in a school gym). Each lasted for one session, except for the last story, which was split across sessions four and five (so that there was enough time for the final tests). Half the children heard the playground stories first and the other half heard the school gym stories first. Each verb was heard with at least two different patient subjects and most were associated with at least two different types of movement/reasons for the psychological state. For example, in story one, '*swing*' was on a swing and in story two '*swing*' was on a rope. Likewise in the playground stories, '*surprise*' was caused by (1) being taken unexpectedly to a playground and (2) seeing another character slide unexpectedly fast; and in the school gym stories the same verb was caused by (1) seeing another character do a cartwheel and (2) slide down stairs on a piece of plastic.

At the end of the text for a particular scene and before the hand-puppet said the sentence, the experimenter told the child to look at the patient of the caused-motion/or experiencer of the psychological event. Then the experimenter said (the German equivalent of) '*the word for what is happening is called in English: X [e.g., swing]. Let's ask Susi/Lotte how one describes that in English – hmm Susi?*' Following this, either the Susi or Lotte hand-puppet used the verb in either the '*X was __ing*' or the '*X got __ed*' form, depending on the semantic class of the particular verb.

Half of the children started with a psychological verb and the other half with a caused-motion verb. From then on psychological verbs were alternated with caused-motion verbs. The particular hand-puppet which said the target sentence for the first training scene was counterbalanced across children. Following this, the other hand-puppet said the target sentence for the second training scene. From then on, a particular hand-puppet would say the target sentences for two consecutive scenes, the other hand-puppet the next two and this was alternated for the rest of the training sessions. Thus, it was never the case that a particular hand-puppet always used the same argument structure for a particular child.

Pointing Practice Test (in German)

At the end of session two, each child participated in a pointing practice test. In this test, each child saw four pictures of a character carrying out a familiar action. The experimenter introduced this test saying (the German equivalent of) '*Now Lotte and Susi will say what you can see in this picture. But one will say it wrong and the other will say it right. After both have said it, you point to the one which said it right, OK?*' They heard one hand-puppet describe the picture using a grammatically correct German sentence (e.g., Lotte: *der Junge reitet* the boy ride.3rd.sg.PRES *the boy is riding*) and other incorrectly (e.g., Susi: *der Junge reiten* the boy ride.3rd.pl.PRES/INFINITIVE *the boy ride*) for the same picture. If the children were confused about the aim of the task, the experimenter corrected them. The verb order for this pointing practice test was counterbalanced according to Latin squares.

Trained Verbs Test

At the end of session four, each child was tested on four of the verb types on which they had been trained, two caused-motion and two psychological. For each verb class one verb type had occurred particularly frequently (with 16 items) during training (FREQUENT: *annoy* and *swing*) and the other two had occurred with four items each overall (INFREQUENT: *bore* and *topple*). In the test, each verb type occurred twice, each time with a different patient–subject. All children heard and saw the same verb–event pairing with the same participants but half saw a caused-motion verb first and the other half a psychological verb. Elicited production always preceded the forced-choice grammaticality judgement test. For the former, the children were given the German equivalent of the following prompt: '*the word for what is happening there is "bore" in English. How do you think you describe in English what's happening?*' For the latter, they heard the same 'pointing' instruction as during the practice pointing task. Following this, one hand-puppet described the scene using the passive '*NOUN got __ed*' and the other using the intransitive '*NOUN was __ing*', whereby regular '*ed*' endings were always used.[2]

The order of the verb film scenes for both the 'trained verbs' test and generalization test was counterbalanced so that half the children started with a motion verb and the other half with a psychological verb. (In addition, half the children who started with motion in training session one also started with motion in the test.) Following this the

verb classes alternated, and the second motion verb was always a different type from the first (likewise for the psychological verbs). The particular verb type which served as the first test verb was counterbalanced across children, so that there were eight possible orders.

We counterbalanced: which hand-puppet spoke first, the side (to the left or right of the child) of the hand-puppet and which was correct. For any one child it was not the case that a particular hand-puppet always used a particular argument structure. In both grammaticality judgement tests, for each child, Lotte and Susi were correct on half the trials each; each spoke first for half the trials and the first speaker was correct for half the trials. This meant that chance was 50% for individual children as well as across the group. Half the children who started with motion in the test also had Lotte as the first speaker in the first test trial and for half of those, the first speaker was also correct. (This was also counterbalanced across whether motion or psychological verbs appeared in the first trial.) The first speaker, who was correct, was fully counterbalanced over the rest of the test trials with the restriction that neither hand-puppet could be correct more than twice in a row and that there was no regular alternating pattern for who was correct. (But the hand-puppet which spoke first alternated across trials.)

Generalization Test

In the generalization test phase (which took place after the training in session five), the story continued with the same characters but with four new verbs: two caused-motion (*drop* and *twist*) and two psychological (*worry* and *grieve*). Each verb occurred twice, each time with a different patient–subject (see Appendix 1 for the generalization grammaticality judgement test sentence stimuli). All children heard and saw the same verb–event pairing with the same participants but half saw a caused-motion verb first and the other half a psychological verb. As for the trained verbs test, elicited production always preceded the grammaticality judgement test and the prompts for both were the same as above. The grammaticality judgement test followed the same procedure as for the trained verbs test and was counterbalanced in the same manner.

Translation Post-Test

After the experiment had been completed on day 5, the children were given a 'translation post-test' in order to determine whether they really did interpret the psychological verbs as psychological verbs and the motion verbs as motion verbs. We had intended to show each child one scene per verb (which would make 16 scenes in total – 12 training verbs and 4 generalization test verbs). However, in piloting we found this was too long for many 5-year-olds. Therefore, the 16 verbs were divided up into four sets of four, so that each child saw four verb scenes (two motion and two psychological verbs). Each child was pre-assigned to one of these four sets (see Appendix 2 for an example of one of these sets). Once a child had completed these four, if he/she showed no signs of tiring, he/she was shown the films from the other sets. For each verb, the experimenter said (as she had done during the experiment) '*look at Alex. The word for what is happening there is called "annoy" in English*'. The difference here was that she then said (the German equivalent

Table 3. (Chronological) Summary of Procedure over the Five Sessions

Session	Tasks
S1	Training on 'word' vs. 'describe' (German – with pictures)
	Training x 20 film scenes
	Standardized German language sub-test
S2	Training x 16 film scenes
	Pointing practice test German
S3	Training x 20 film scenes
S4	Training x 8 film scenes
	Trained verbs test English (scenes with *swing*, *annoy*, *bore*, *topple*)
S5	Training x 8 film scenes
	Generalization test
	Translation post-test

of) '*What is the word for that in German?*' If the child did not reply or gave an uncodable response, the experimenter then asked (the German equivalent of) '*What do you think of when you look at PATIENT (e.g., Alex)?*'

A summary of the whole procedure is presented in Table 3.

Results

Each child had eight test trials for the 'trained verbs' (those which they had experienced during training) and eight test trials for the 'generalization verbs'. For each test trial there was a production opportunity followed by a request to point to the hand-puppet who said the sentence correctly (forced-choice grammaticality judgement). The test trials for the trained verbs came at the end of session four and those of the generalization verbs came at the end of session five (but before the translation post-test). Because of these potential order effects, we analysed grammaticality judgement and production and the trained verbs and the generalization verbs separately.

Grammaticality Judgement: Trained Verbs

There was one verb (with two trials each) per condition. '*Swing*' was frequent motion (heard with 16 items during training), '*topple*' was infrequent motion (heard with four items during training), '*annoy*' was frequent psychological (heard with 16 items during training) and '*bore*' was infrequent psychological (heard with four items during training). In the grammaticality judgement test, the children were basically asked to point to the hand-puppet 'who said it right'. If the child pointed to the hand-puppet which used a motion verb in the '*X was VERBing*' construction, this was counted as 'correct' and was coded as '1'. Likewise, if the child pointed to the hand-puppet which used a psychological verb in the '*X got VERBed*' construction, this was counted as 'correct'. The reverse in

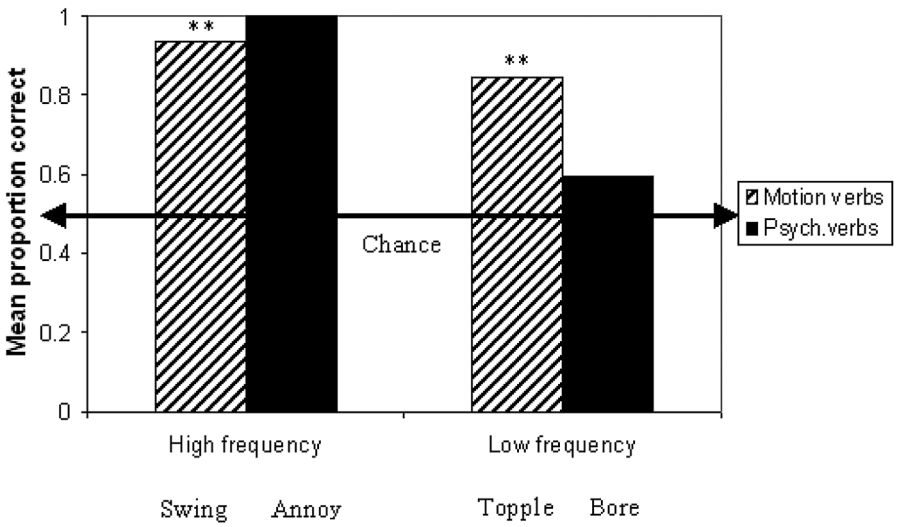

Figure 1.

both cases was counted as 'incorrect' and was coded with zero. Thus, pointing at chance was equivalent to scoring a mean of 1 out of 2 for a particular verb/condition (e.g., '*swing*'). The experimenter wrote down the child's response at the time of testing but both the trained verbs and the generalization grammaticality judgement tests were videotaped and were later coded by a native German-speaking English linguistics student. Reliability between the two coders was 99% (Cohen's kappa = .98).

To examine the effects of semantic class and frequency we carried out a 2 (Semantic class: motion vs. psychological verbs) × 2 (Frequency: frequent vs. infrequent) repeated-measures ANOVA. We found no main effect for semantic class, $F(1,15) = 1.42$, $p = .25$, $\eta_p^2 = .09$, but a main effect for frequency, $F(1,15) = 24.00$, $p < .001$, $\eta_p^2 = .62$. This is illustrated in Figure 1. Clearly the children were much better at associating a particular construction with those verbs ('*swing*' and '*annoy*') which they had heard four times as often (16 times each) as the other verb types (4 times each), suggesting support for the hypothesis that lexical learning plays an important role. However, this needs to be interpreted in the light of a semantics × frequency interaction, $F(1,15) = 5.95$, $p < .05$, $\eta_p^2 = .28$. Since we only tested one verb per cell, it is difficult to determine from this analysis alone whether this interaction is truly due to verb class 'semantics' or whether it is due to (salience) effects of particular verb–event pairings. We return to this issue when we come to our discussion of the generalization and translation tests.

We further investigated this interaction by carrying out post-hoc paired *t*-tests between all the items (correcting with Bonferroni, whereby .05/six comparisons). These revealed significant differences between '*bore*' and '*annoy*', $t(15) = 4.33$, $p = .001$, and a borderline significant difference between '*bore*' and '*swing*', $t(15) = 2.91$, $p = .011$. Thus, the children clearly performed worse with '*bore*' than with either of the two frequent verbs.

Wilcoxon tests confirmed this pattern of results. We also looked to see whether the children pointed above chance to the construction which the children had previously heard with these verbs during training. The answer is that they did with all verbs ($p < .001$, with a Bonferroni correction for four comparisons) apart from '*bore*' ($p = .33$). (We couldn't actually perform a *t*-test for '*annoy*' because the standard deviation was zero; all children performed at ceiling on this verb.) Thus, although the children performed significantly above chance in the grammaticality judgement test for the trained verbs when conflated, $t(15) = 19.05$, $p < .001$, indicating conservative treatment of these verbs on the whole, this was not the case for '*bore*'. For this verb, they either did not learn that the '*X got VERBed*' construction was associated with '*bore*', or they had a tendency to generalize the other construction ('*X was VERBing*') to this verb.

Production: Trained Verbs

The production tests always preceded the grammaticality judgement tests, for both trained verb and generalization. In the production tests the child was asked to describe the test scene 'in English' (e.g., they were asked the German equivalent of '*the word in English for what is happening there is "bore". How do you think you describe in English what's happening?*'). Both the trained verbs and the 'generalization' production tests were videotaped and all child utterances were transcribed by a native speaker of German who was a student of English linguistics. The first author also transcribed 19**%** of the utterances. Reliabilities between the two were 98.95%. The first author then coded the utterances. If the child used either the '*got*' or the '*ed*' morpheme with a psychological verb, this was counted as 'correct' and scored with '1'. If the child used either the '*was*' or the '*ing*' morpheme with a motion verb, this was counted as 'correct' and scored with '1'. The reverse cases were scored with zero. If a child used a mixture of both morphemes (e.g., '*Hanna got boring*'), this was scored as .5. (For the trained verbs test this happened for five out of 66 codable trials). In the trained verbs production test, one child always generalized the '*was __ing*' morphemes to all items (regardless of semantic class) and two children always generalized the '*got __ed*' morphemes to all items (regardless of semantic class). Two other children never used any morphemes, but simply produced two-word utterances of the type '*Hanna swing*'. We excluded these five children from further analyses for this test. For the remaining 11 children, if a child gave a null or uncodable response (e.g., '*Alex wat topple*'), the cell was left empty and excluded from analyses. This was the case for 22 trials. A codable response was given for 66 out of 88 production trials (75%). Because of this, the results for the production tests are given as proportions of codable responses. Figure 2 shows the mean proportion of 'correct' production for each verb separately. Since the number of children (= the denominator for the proportion) differed for each verb, we give the *N* in brackets for each on the x-axis.

As for the grammaticality judgement measure, we again carried out a 2 (Semantic class: motion vs. psychological verbs) × 2 (Frequency: frequent vs. infrequent) repeated-measures ANOVA on the mean proportion of 'correct' responses. We found no main effect for semantic class, $F(1,6) = 0.36$, $p = .57$, $\eta_p^2 = .06$ but we did find a main effect for frequency, $F(1,6) = 15.76$, $p = .007$, $\eta_p^2 = .72$, and the interaction reached borderline significance, $F(1,6) = 5.87$, $p = .052$., $\eta_p^{2p} = .50$. Thus, the results for the production measure tie in with

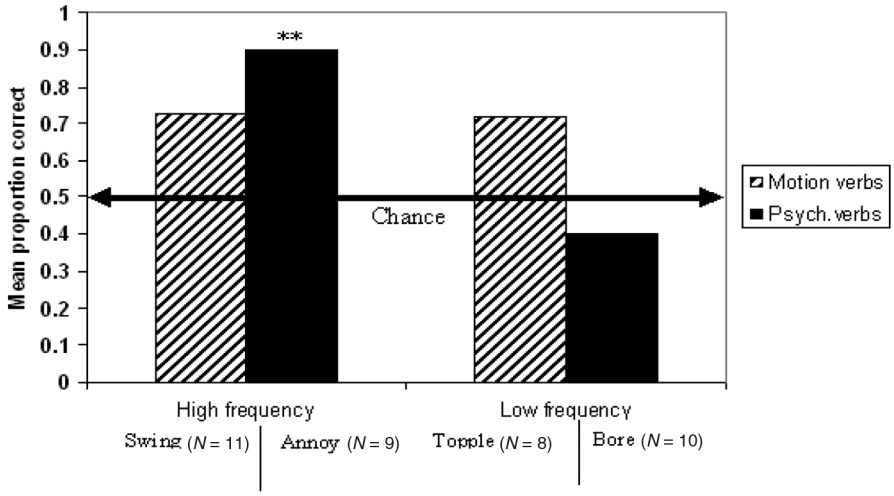

Figure 2.

those from the grammaticality judgement measure in indicating support for the hypothesis that lexical learning plays an important role. In addition, it can be seen from Figure 2 that the trend for production mirrors that for grammaticality judgement, with the children being more likely to use '*annoy*' in the construction in which they had heard it used ('*X got annoyed*') than they were for '*bore*' ('*X got bored*'), $t(8) = 4.26$, $p < .008$, with a Bonferroni correction for six comparisons. (None of the other five comparisons was significant with a paired samples *t*-test. Wilcoxon tests found the same result.) This would imply that for '*bore*' the children actually preferred to use the '*X was boring*' construction, but in fact one-sample *t*-tests found that only production of '*annoy*' differed from chance, $t(8) = 5.29$, $p = .001$, with a Bonferroni correction for four comparisons.

Generalization Test: Grammaticality Judgement

Our main focus of interest with the generalization test was to investigate the 'semantic analogy' hypothesis; that is, that children should be more likely to use construction A ('*X was VERBing*') with motion verbs and construction B ('*X got VERBed*') with psychological verbs, if those semantic classes had been associated with these particular constructions during training. To examine this, we analysed the grammaticality judgement results from the generalization test, which had exactly the same structure as that for the trained verbs, except that 'novel' verbs were used. Chance was again calculated as a mean score of 1 out of 2 for each verb. Figure 3 below shows that the children pointed at chance for '*drop*', '*twist*' and '*grieve*', $p > .16$ for all. This means that they were equally likely to select either of the two constructions for these verbs. This was not the case for '*worry*', where the children pointed above chance to the puppet(s) which used this verb in the '*X got worried*' construction, $t(15) = 1.78$, $p < .05$, one-tailed.[3]

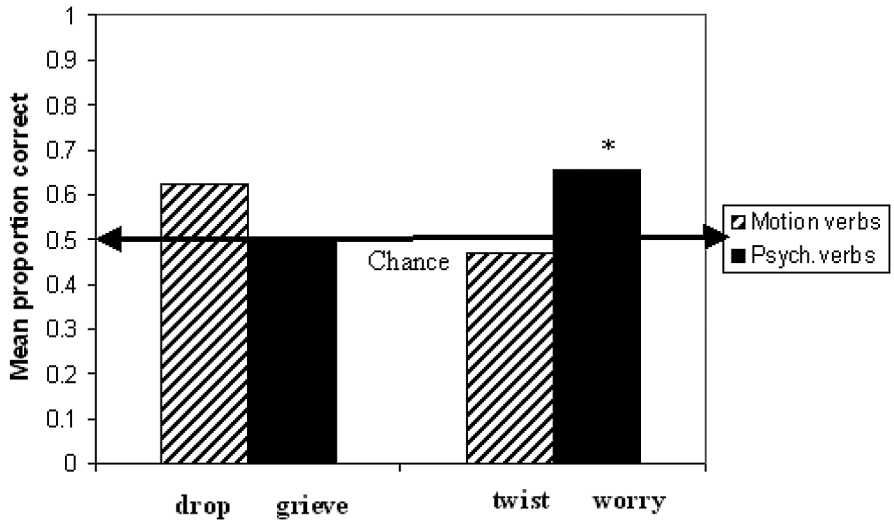

Figure 3.

Generalization Test: Production

In the generalization production test, seven of the 16 children always generalized the '*was __ing*' morphemes to all items (regardless of semantic class) and two of the 16 children always generalized the '*got __ed*' morphemes to all items (regardless of semantic class). Three children never used any morphemes,[4] but simply produced two-word utterances of the type '*Hanna drop*'. Thus, in this production test 75% of our children behaved like the children in Hudson Kam and Newport's (2005) study in that they picked up on one pattern and generalized this. The remaining four children did use a mixture of both morphemes but there was no clear pattern relating to semantic class; the mean proportions of 'correct' responses for these four children were: '*drop*' = .25; '*twist*' = .75; '*worry*' = .75; '*grieve*' = .50.

Translation Test

The finding that the children were (marginally) above chance in choosing the '*X got worried*' construction (over the '*X was worrying*') in the novel verbs forced-choice judgement test is potentially interesting, because the only basis which the children could have done this is through analogy with one or several of the class of 'trained' psychological verbs. One might think that this indicates tentative support for the 'semantic analogy' hypothesis. However, such an interpretation does not explain why the children were not able to carry out semantic analogy for the other verbs. One possibility is that the intended semantics were in fact not always clear. To investigate this latter possibility, we examined the children's responses in the translation post-test.

For the translation post-test, the children's responses were transcribed by a native German speaker and then coded by the first author into one of two categories. A second (blind) coder (a psychology PhD student who had previously been a German translator) also coded 100% of the responses. There was 89% agreement between the two coders, with good agreement between the two (Cohen's kappa = .768). The first category contained 'correct' responses in which the child had (1) clearly focused on the relevant action/psychological state of the relevant participant *and* (2) managed to correctly induce that the word referred to an action (or psychological state respectively). Sometimes the children produced fairly accurate translations (e.g., *das ist ihm peinlich* = 'that's embarrassing for him' for the verb '*embarrass*'; or *das hat sie nicht erwartet* = 'she hadn't expected that' for the verb '*surprise*'). Often, however, responses in this category were of the correct semantic class but could not be considered to be a translation of the intended verb (e.g., *traurig* = 'sad' for the verb '*embarrass*'). The 'other' category contained responses in which the child either (1) did not focus on the relevant action/psychological state of the relevant participant (e.g., *die lachen ihn aus* = 'they are laughing at him' for '*embarrass*') or (2) it was not clear exactly what the child had focused on (e.g., *Alex sagt 'Hör auf!'* = 'Alex says "stop it"!' for '*embarrass*').

For each verb separately we compared the 'correct' vs. 'other' responses using chi-square (exact) tests. Among the trained motion verbs, only for '*slide*' were there significantly more 'correct' than 'other' children, χ^2 (1, $N = 9$) 5.444, $p < .05$. However, this would also be true for '*swing*' – the high frequency motion verb – if there had been any variation (i.e., all children who participated in this trial interpreted this item perfectly as *schaukeln* = 'swing'). Among the trained psychological verbs, only for '*surprise*' were there significantly more 'correct' than 'other' children, χ^2 (1, $N = 11$) = 7.364, $p < .05$. However, many of the non-significant findings here are likely to be due to lack of power. Some children gave null responses (e.g., *weiss nicht* = 'don't know') for some items, and many children were not given all 16 items due to fatigue (see Method / Translation Post-Test section).

The 'correct' responses were in the majority for half the verbs in each verb class (motion: '*swing*', '*slide*', '*bounce*', '*twist*'; psychological: '*tempt*', '*surprise*', '*scare*', '*grieve*'). For the other half of the verbs the numbers of 'correct' vs. 'other' responses were roughly equal. Interestingly, one of the latter verbs was the high frequency psychological verb '*annoy*' (for which all children were perfect at selecting the '*X got annoyed*' form in the judgement test). That is, '*annoy*' was not clearly interpreted as a psychological verb by nearly half the children who gave a response. Instead, these children said things like *Fahrrad fahren* (= 'riding a bike') – where they focused on the action the other character had just carried out or *streitet* (= argues), which is not clearly a psychological state. Thus, the apparent lack of generalization on the basis of the semantic similarity in the current study may be due to the simple reason that the intended meaning of the trained verbs within a semantic class was not always clearly transmitted.

Discussion

The current study investigated the comparative roles of lexicalist frequency-based learning and semantic analogy in the acquisition of verb argument structure. To ensure that we were examining the learning process (and not the effects of additional training on prior

language acquisition) we taught children a miniature language. However, since we were worried about the representativeness of artificial grammars for real language learning, we in fact exposed 5-year-old German children to the vocabulary and two verb argument structures of a real language – English. In the forced-choice judgement test for 'trained verbs', the children most often chose the argument structure in which they had had previously heard those verbs used, with a significant effect of verb frequency. The trained verbs production test reflected the same pattern. However, for both tests, there was an interaction between frequency and particular verb types (or perhaps classes). There was also some evidence of possible item effects in the forced choice judgement test for novel verbs, in which the children were only (marginally) above chance for one verb type (despite there being no evidence of semantic-based generalization overall). In the novel verb production test, the children tended to use one verb argument structure for all items.

The latter finding – that 81% of the children used one or both of the morpheme frames with the 'novel' verbs in the generalization test – indicates that the children did manage to segment the morphological paradigms out of the sentences, and readily used them with verbs they had never encountered in that paradigm. This shows that they formed an independent representation of the structure, and were willing to extend it to new verbs. The behaviour of the children in the generalization production test is very similar to the children in Hudson Kam and Newport (2005). The latter authors interpreted their findings as showing that children (unlike adults) do not follow the distributional probabilities of the input, but impose 'rules' on the language they are learning.

This interpretation does not, however, account for our findings for the trained verbs production test, in which our children tended not to overgeneralize, but rather showed a frequency effect in using the argument structure in which they had heard a particular verb used. These results from the trained verbs test do not support the 'children-as-rampant-generalizers' hypothesis. Instead, there was strong support for our 'lexical conservatism' hypothesis, which predicted, first, that the children should perform well on tests of trained verbs but at chance on tests involving novel verbs and, second, that there should be frequency effects within the trained verbs tests. These frequency dependent findings tie in well with findings from first language acquisition indicating that when children initially start to produce various verb argument structures, they are more likely to do so successfully with verbs which are more frequent in the input (e.g., Matthews et al., 2005). They also fit with findings from second language acquisition indicating that input frequency is a key determinant of order of acquisition (e.g., Ellis, 2003; Goldschneider & DeKeyser, 2001). Last, they also fit with findings that 5-year-old and adult speakers' grammaticality judgements of causative overgeneralizations appear to relate to the relative grammatical entrenchment of the verbs concerned (e.g., Ambridge et al., 2008; Theakston, 2004). Thus the performance of our children in the generalization production test is probably best characterized as the use of a heuristic.

However, frequency effects are clearly not the only thing going on in this study. The results from both grammaticality judgement and production with the trained verbs clearly show an interaction with particular verbs. In addition, the children also showed a trend to be above chance with one novel verb in the grammaticality judgement test. Both of these sets of findings may reflect verb semantics on some level, whether that be semantic verb class or the meaning of particular items or with how well particular items were assimilated. However, the results do not support the predictions of the 'generalization-constrained-by-verb-semantic-class' hypothesis.

There are a number of potential reasons for this. One is that the intended meaning of the various verbs may not always have been clearly transmitted. Support for this possibility can be seen in the results of our translation post-test and from Hudson Kam and Newport (2005), who found that their child learners had to be explicitly taught vocabulary items through direct translation. Teaching through direct translation would, however, be inappropriate for our current research question for two reasons; first, our focus of interest precisely relates to verb semantics (so we could not explicitly teach this) and second, this is not how naturalistic first language acquisition proceeds. It should also be kept in the mind that tests such as our translation post-test rely on metalinguistic awareness and may underestimate children's semantic interpretation of the events.

However, even if the children did correctly interpret the events, it may be that they perceived the caused-motion verbs as belonging to two separate classes, since some involved punctual verbs (e.g., '*drop*', '*slide*') and some involved durative verbs (e.g., '*swing*', '*twist*'). Alternatively, it may be that the children perceived all the events – including the psychological state ones – as belonging to the same class, as they all involved causation. Lastly, it may be that the children had difficulty interpreting the psychological state events. Interestingly, the results of the translation post-test indicate that the children were no worse at correctly interpreting the psychological events than they were at correctly interpreting the caused-motion events (cf. Papafragou, Cassidy, & Gleitman, 2007). Rather, they were particularly good at 'translating' certain psychological events such as '*surprise*' (91% correct), '*scare*' (78% correct) and '*grieve*' (88% correct).

Indeed, it is intriguing how difficult these 5-year-old children found it to translate apparently clear and simple caused-motion events such as '*turn*' (only 45% correct) or '*drop*' (only 33% correct). This supports the point made by theorists from a wide spectrum of perspectives who have argued that learning verb meaning is particularly difficult since the boundaries of events are often unclear (although see Loucks & Baldwin, 2006) and the particular verb chosen depends on speaker perspective (e.g., Gleitman, January, Nappa, & Trueswell, 2007). Focusing on the relevant event is difficult even for adults when they are asked to name the verbs used in a parent–child interaction for which they cannot hear the audio (Gillette, Gleitman, Gleitman, & Lederer, 1999). Clearly much work is needed not only to determine what counts as semantically *similar* for preschool children, but also to first determine the kinds of event meanings children initially derive for verbs (see e.g., Imai, Haryu, & Okada, 2005; Kersten & Smith, 2002).

In conclusion, the current study ties in with previous findings that child and adult learners are (at least partially) constrained by verb-specific frequency when learning verb argument structure. However, further work is needed on how young preschoolers interpret the semantics of individual verbs and whether semantic analogy plays a role in the initial stages of learning the verb argument structures of a language.

Acknowledgements

Many thanks to Claudia Salomo for testing the children, to Eva Leerman for help with filming and editing the stimuli and many suggestions in the preparation phase, to Elena Rossi for organizing the actors, to Dana Langner for transcribing and coding the data, to Kinga Morsanyi for coding reliabilities for the 'translation' test, to Petra Jahn and Ronny Barr for technical assistance and to Anna Theakston, Danielle Matthews, Kevin Durkin and two anonymous reviewers for helpful comments.

A special 'thank you' to the following kindergartens: Andromedaweg 25, Arthur- Nagel- Str. 1b, Delitzscher Str. 141, Demmeringstr. 35, Nordstr. 43, Rudolf- Sack- Str. 12, Yorkstrasse 43b.

Appendix 1: Sentence Stimuli Used in the Generalization Test (One Example Order)

Each pair was heard twice (hand-puppet and order were counterbalanced).
Susi: *Hanna was grieving* + Lotte: *Hanna got grieved*
Lotte: *Luisa was twisting* + Susi: *Luisa got twisted*
Susi: *Alex got worried* + Lotte: *Alex was worrying*
Lotte: *Hanna got dropped* + Susi: *Hanna was dropping*
Susi: *Max was grieving* + Lotte: *Max got grieved*
Lotte: *Max was twisting* + Susi: *Max got twisted*
Susi: *Luisa got worried* + Lotte: *Luisa was worrying*
Lotte: *Alex got dropped* + Susi: *Alex was dropping*

Appendix 2: Set 'C' as Example of Film Clip Items Used during the Translation Test

WORRY: Alex falls to the ground and is in pain. Luisa looks worried.
TURN: Alex turns Hanna on the roundabout (on a playground).
SURPRISE: Hanna surprises Alex by turning a cartwheel. Alex looks absolutely amazed.
BEND: Max stands behind Luisa, holds both her arms and bends her up and down.

Notes

1 However, compare with Flege and Liu (2001), Hakuta, Bialystok, and Wiley (2003) and Piller (2002).
2 This resulted in the children hearing '*NOUN got swinged*' as one of the pairs but recall that this was not ungrammatical to them as they were learning English as a miniature artificial language.
3 This would not be above chance if the Bonferroni correction were applied.
4 Two of these children were the same two who never used morphemes in the trained verbs production test.

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2004). Training 2;6-year-olds to produce the transitive construction: The role of frequency, semantic similarity and shared syntactic distribution. *Developmental Science*, *7*, 48–55.

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development*, *23*, 48–66.

Ambridge, B., Pine, J., Rowland, C., & Young, C. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgments of argument-structure overgeneralization errors. *Cognition*, *106*, 87–129.

Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss*. Cambridge: Cambridge University Press.

Boyd, J., & Goldberg, A. E. (In press). Children's failure to generalise when exposed to the same input as adults and its benefits in language learning. *Lingua*.

Braine, M. D. S. (1963). The ontogeny of English phrase structure – the 1St Phase. *Language*, *39*, 1–13.

Braine, M. D. S., Brody, R. E., Fisch, S. M., Weisberger, M. J., & Blum, M. (1990). Can children use a verb without exposure to its argument structure? *Journal of Child Language*, *17*, 313–342.

Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language*, *75*, 720–738.

Casenhiser, D. M., & Goldberg, A. E. (2005). Fast mapping between a phrasal form and meaning. *Developmental Science*, *8*, 500–508.

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*, *37*, 739–748.

Ellis, N. (2003). Constructions, chunking and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 33–68). Oxford: Blackwell.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, *31*, 41–81.

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

Flege, J., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*, 527–552.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science*, *15*, 297–301.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, *17*, 684–691.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*, 135–176.

Gleitman, L., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, *57*, 544–569.

Goldberg, A. E. (1999). The emergence of the semantics of argument structure constructions. In B. MacWhinney (Ed.), *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum.

Goldschneider, J., & DeKeyser, R. (2001). Explaining the 'natural order of L2 morpheme acquisition' in English: A meta-analysis of multiple determinants. *Language Learning*, *51*, 1–50.

Grimm, H. (2001). *Sprachentwicklungstest für drei- bis fünfjährige Kinder.* Goettingen: Hogrefe.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*, 31–38.

Hudson Kam, C. L., & Newport, E. L. (2005). Regularising unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151–195.

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development*, *76*, 340–355.

Johnson, E., & Jusczyk, P. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, *44*, 548–567.

Johnson, J., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–90.

Johnston, J., Blatchley, M., & Streit Olness, G. (1990). Miniature language system acquisition by children with different learning proficiencies. *Journal of Speech and Hearing Research*, *33*, 335–342.

Kersten, A., & Smith, L. (2002). Attention to novel objects during verb learning. *Child Development*, *73*: 93–109.

Langacker, R. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmerer (Eds.), *Usage-based models of language*. Stanford, CA: SLI Publications.

Lieven, E. (2008). Learning the English auxiliary: A usage-based approach. In H. Behrens (Ed.), *Trends in corpus research: Finding structure in data*. Amsterdam: John Benjamins.

Loucks, J., & Baldwin, D. (2006). When is a grasp a grasp? Characterizing some basic components of human action processing. In K. Hirsh-Pasek & R. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford: Oxford University Press.

MacWhinney, B. (1983). Miniature linguistic systems as tests of the use of universal operating principles in second-language learning by children and adults. *Journal of Psycholinguistic Research*, *12*, 467–478.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, *20*, 121–136.

Meier, R., & Bower, G. (1986). Semantic reference and phrasal grouping in the acquisition of a miniature phrase structure language. *Journal of Memory and Language*, *25*, 492–505.

Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, *48*, 323–363.

Ninio, A. (2005). Testing the role of semantic similarity in syntactic development. *Journal of Child Language*, *32*, 35–61.

Papafragou, A., Cassidy, K., & Gleitman, L. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition*, *105*, 125–165.

Pickering, M., & Branigan, H. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651.

Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics*, *6*, 179–206.

Pinker, S. (1989). *Learnability and cognition: The acquisition of verb-argument structure*. Cambridge, MA: Harvard University Press.

Richards, B. J. (1990). *Language development and individual differences: A study of auxiliary verb learning*. Cambridge: Cambridge University Press.

Stoll, S., Abbot-Smith, K., & Lieven, E. (2009). Lexically restricted utterances in Russian, German and English child-directed speech. *Cognitive Science*, *33*, 75–103.

Theakston, A. (2004). The role of entrenchment in constraining children's verb argument structure overgeneralisations: A grammaticality judgment study. *Cognitive Development*, *19*, 15–34.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 528–553.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, *16*, 180–205.

Yorio, C. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity and loss*. Cambridge: Cambridge University Press.