Graded representations in the acquisition of English and German transitive constructions

Kirsten Abbot-Smith a,*, Elena Lieven b, c, Michael Tomasello b

a School of Psychology, University of Plymouth, B213 Portland Square, Drake Circus, Plymouth PL4 8AA, United Kingdom
b Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
c School of Psychological Sciences, University of Manchester, United Kingdom

Abstract

English and German children aged 2 years 4 months and 4 years heard both novel and familiar verbs in sentences whose form was grammatical, but which mismatched the event they were watching (e.g., ‘The frog is pushing the lion’, when the lion was actually the ‘agent’ or ‘doer’ of the pushing). These verbs were then elicited in new sentences. All children mostly corrected the familiar verb (i.e., they used the agent as the grammatical subject), but there were cross-linguistic differences among the two-year-olds concerning the novel verb. When English 2-year-olds used the novel verb they mostly corrected. However, their most frequent response was to avoid using the novel verb altogether. German 2-year-olds corrected the novel verb significantly more often than their English counterparts, demonstrating more robust verb-general representations of agent- and patient-marking. These findings provide support for a ‘graded representations’ view of development, which proposes that grammatical representations may be simultaneously abstract but ‘weak’.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Grammatical development; Transitive; German; English

Many textbooks on child development describe grammatical development as proceeding very rapidly. Indeed, even two-year-olds are able to describe scenes in which an agent (or ‘doer’ of the action) acts on a patient (or one affected by the action) by using transitive sentences, that is, ones with two noun phrases as in ‘Ben pushed the cat’. However, it is difficult to tell from production and comprehension studies involving familiar verbs whether young children are productive with these sentence-types (i.e., whether they can apply the ‘rules’ generally to all appropriate verbs
in an adult-like manner) or whether they represent them with lower-level schemas dependent on familiar lexical items, such as ‘push’ (Tomasello, 1992).

Therefore, the most widely accepted method of testing grammatical productivity is the ‘wug’ test, in which children are taught a novel word and are required to generalize their knowledge of a grammatical pattern to this new item. Many production and act-out comprehension studies have found that English-speaking children have difficulty demonstrating productivity with novel verbs in active transitive sentence types until around the age of 2 years 6 months (Chan & Meints, 2005; Tomasello & Brooks, 1998). In stark contrast, Gertner, Fisher, and Eisengart (2006) recently found that when English-speaking children aged 1 year 9 months hear a transitive sentence such as ‘The boy is gorping the girl’ with a novel verb they are more likely to look at the ‘correct’ video (e.g., the one in which the boy acts on the girl) than the ‘incorrect’ video (e.g., the one in which the girl acts on the boy).

Such measure-dependent results have also been found within the same task for the same children in the ‘weird word order’ paradigm in which children are taught verbs in either grammatical or ungrammatical word orders and their own productions are later elicited for new scenes (Akhtar, 1999; Matthews, Lieven, Theakston, & Tomasello, 2005). Here, English-speaking children aged 2 years 8 months frequently matched the word order in which they heard the novel or rare familiar verbs used, even when these were ungrammatical word orders, which would appear to indicate that they had not completely acquired a non-verb-specific representation of the transitive. However, these children also often ‘avoided’ using the novel verb they had heard used ungrammatically; that is, in this condition they were less likely to use the novel verbs in sentences altogether (whether grammatically or ungrammatically) than they were in the condition in which they had heard novel verbs used in grammatical word orders. This avoidance would appear to indicate that they do have some verb-general knowledge of SV(O) word order (Akhtar, 1999; Abbot-Smith, Lieven, & Tomasello, 2001).

One type of explanation for this method-dependence in empirical findings is to claim that previous novel verb elicited production studies have been inadequate in revealing underlying child knowledge of the transitive because these tasks are essentially too difficult for young two-year-olds (Fisher, 2002; Naigles, 2002). This ‘inadequate task’ view treats child acquisition as basically all-or-none in that the evidence from preferential-looking studies is argued to indicate that the English transitive is ‘acquired’ at 1–9. The failure of most English-speaking toddlers to produce novel transitives until around 2–6 is explained either in terms of some sort of deficit in a supporting subsystem separate from the linguistic system, such as memory, or in terms of incomplete acquisition not of the transitive construction but of the alternation rules relating to other constructions (Fisher, 2002; Naigles, 2002). Among other things, all previous production studies have ‘primed’ the usage of the novel verb in an alternative construction (Fisher, 2002; Naigles, 2002).

An alternative explanation comes from Abbot-Smith and Tomasello (2006). Their account is one of a number of usage-based theories, which assume that that the basic level units of grammar are sentence-types (or constructions) such as the ‘transitive’, which the child gradually abstracts out of learned sentence exemplars (Goldberg, 2006). Abbot-Smith and Tomasello (2006) argue that as the child processes sentence-scene pairings such as ‘The man rolled the ball’ and ‘Tom dropped the doll’, these stimulate similar activation patterns because of their similarities both in terms of meaning and sentence form. When the child has only learned a small number of exemplars (for instance of the transitive), it will not yet be obvious which semantic and distributional features are crucial because there will not yet have been enough overwriting of the more idiosyncratic elements. This would predict that initially sentences with familiar relational elements such as
familiar verbs should be much easier to produce and comprehend, because they show the greatest degree of relational similarity to previously learned exemplars. However, that is not to say that the child does not have any access to a more abstract (verb-general) representation at this stage. The exemplars are stored in a similar fashion, so the presentation of a novel utterance (comprehension) or novel scene (production) which shares crucial similarities with this category of sentence-scene exemplars will activate the group as a whole and consequently also the more abstract category, which is basically the ‘sum’ of these similarities (Shanks, 1997).

Abbot-Smith and Tomasello’s (2006) account draws heavily on the graded representations account, which was originally formulated to explain differences between looking and reaching measures within the domain of object permanence development (Munakata, McClelland, Johnson, & Siegler, 1997). This account proposes that representations are graded in strength, with only strong representations allowing clean signaling to other parts of the cognitive system (Munakata, 2001). Their strength depends on the number of exemplars which the child has previously experienced in the input. If a child’s representation is still ‘weak’, it is accessible but because its signal is noisy it may only show up with dependent measures which do not tax executive functions mechanisms and memory systems. Some evidence for this approach comes from Munakata et al.’s (1997) connectionist simulation where a single mechanism accounted for early success in looking-time tasks and only later success in reaching. This view also receives some support from a connectionist model by Chang, Dell, and Bock (2006) in which they simulated the discrepancy between production and looking measures in early acquisition of the English transitive construction.

In all usage-based approaches (including the current ‘graded representations’ view) the input plays a crucial role, both in terms of the amount of input the child has previously processed and in terms of the relative learnability of the subject and object markers of a particular language. In English, the subject is marked by word order; that is, it is almost always the noun phrase preceding the verb (e.g., ‘the ball’ in both ‘The ball hit the boy’ and ‘The ball was hit by the boy’). Many other languages mark the subject using case-marking. This involves using a different word or word-form depending on which noun phrase is the agent. There are many ways in which particular markers or ‘cues’ (e.g., word order versus case-marking) to semantic roles might be easier to learn than others and this issue has not yet been sufficiently empirically resolved (but see Kempe & MacWhinney, 1998; Devescovi, D’Amico, & Gentile, 1999; Slobin & Bever, 1982; Lindner, 2003). One simple proposal is that all things being equal, languages (or constructions within particular languages) which mark semantic roles using two cues (e.g., both word order AND case-marking) should be easier to learn than those in which there exists only one cue. This is basically the ‘coalitions-as-prototypes’ view proposed by Bates and MacWhinney (1987).

Thus, both the ‘inadequate task’ view and ‘graded representations’ accounts would agree that English speaking children have some kind of non-lexically specific representation of the action-verb active transitive construction from early on. They would disagree, however, as to why. While innate linking theorists assume some basis in a priori knowledge, many emergentist theorists note that word order in English action-verb active transitives has high cue availability and 100% reliability (Bates & MacWhinney, 1987) and children hear them extremely frequently. British mothers, for example, have been found to use around 70 declarative full active transitives on average per hour (Cameron-Faulkner, Lieven, & Tomasello, 2003).

Since neither account makes precise predictions as to exactly when we should begin to see evidence for an abstract transitive representation in English-speaking children, we need to find another way to tease the two apart. One way would be through a cross-linguistic comparison examining two-year-olds. Here, the ‘inadequate task’ view would predict that if the poor performance
of young English-speaking two-year-olds in production tasks is solely due to memory difficulties, incorrect interpretation of the novel verb’s meaning, or other performance limitations, children the same age learning other languages should perform equally poorly. In contrast, usage-based theories would predict that we should see cross-linguistic variation among young two-year-olds in their elicited production depending on relative learnability of the formal cues to agent and patient. More specifically, with constructions in which the agent is marked with two or more cues, young 2 year-olds may find it easier to overcome the processing difficulties involved in producing a transitive with a novel verb, than with constructions in which the agent is marked with only one cue.

Therefore we compared English active transitives, in which semantic roles (e.g., agent versus patient) are marked by one cue—word order, with German. In German, word order and case-marking collaborate in marking the same noun phrase as subject in 68% of active transitive sentences in child-directed speech. This coalition could be considered the German active transitive ‘prototype’ (Dittmar, Abbot-Smith, Lieven, & Tomasello (in press). Case-marking is highly available in active transitives in child-directed speech. It is only ambiguous in the 11% of active transitives in which both the subject and object are either feminine, neuter or plural (see Dittmar et al., in press). In such sentences, subject-first word order has almost 100% reliability (MacWhinney, Bates, & Kliegl, 1984). Case-marking is also 100% reliable within active transitives; that is it is THE overriding grammatical marker. When masculine lexical forms such as der (the/himasculine-accusative) or den (the/himasculine-nominative) are present, they must refer to the subject and object, respectively, regardless of word order (Lindner, 2003; Szagun, 2004).

To compare English and German two-year-olds we adapted the ‘weird word order’ method to deal with criticisms that previous elicitation methods have required children to overcome the tendency to use the novel verb in the alternative argument structure in which they heard it used. For example, in Akhtar’s (1999) study, the children heard the novel verbs used in either NOUN-NOUN-VERB or VERB-NOUN-NOUN form but a ‘correct’ response involved using these verbs in the NOUN-VERB-NOUN (NVN) form. Thus, these previous studies may have syntactically primed the children to use the novel verbs in the non-target structure (see Branigan, 2007, for a review of syntactic priming). The current method eliminated this potential flaw because the form in which the target verb was heard was the same as that of the ‘correct’ responses we attempted to elicit; that is, both involved NVN form. The ‘weirdness’ in the current method instead comes from the fact that the children heard sentences in which the grammatical markers were linked ungrammatically to semantic roles. In English there was one grammatical marker—word order. This means that the ‘weird linking’ arose from the first noun phrase being mapped on the patient and the second noun phrase being mapped onto the agent (e.g., ‘The frog is pushing the lion’, where the lion is the agent and the frog is the patient). In German there were two grammatical markers – word order plus case-marking – which collaborated in marking the first noun phrase as subject. This meant that the German sentences also had subject-first word order. In addition, the determiner of the first noun phrase was marked nominative (which is used to mark grammatical subjects) and the second noun phrase was marked accusative (which is used to mark grammatical objects). Thus, in German word order and case-marking cues collaborated in both being ‘weirdly linked’ to the semantic roles. Therefore the only difference between the two languages was that the English linguistic models contained one grammatical marker whereas the German linguistic

---

1 The principle behind this method bears some resemblance to previous studies by Marantz (1982) and Pinker, Lebeaux, & Frost (1987) which were carried out with older children. The current task is, however, quite different from both of these.
models contained two. In both cases, the form of the models constituted the prototypical active transitive form for that particular language.

In the elicitation trials the children could use (a) ‘grammatical’ linking, (b) ‘ungrammatical linking’ or c) they could avoid using the target verb in a sentence altogether (i.e., if they have difficulties in actively correcting). All theoretical perspectives discussed here would predict that the young 2-year-olds of both languages should tend to not use ungrammatical linking, since children have some kind of non-lexically specific representation of this construction from early on. The differences are as follows. An ‘inadequate task’ view would predict that both English and German young 2-year-olds should have equal difficulties in demonstrating productivity with the novel verb; that is, equal difficulties in actively correcting it to grammatical linking. Usage-based theories, on the other hand, one would predict that German young 2-year-olds should be more likely to use the novel verb in correct linking for the ‘prototypical’ transitive construction than their English counterparts, because the German construction has both case-marking and word order markers. In addition, the ‘graded representations account’ would predict that while English young 2-year-olds may have difficulty correcting the novel verb to grammatical linking, they should show sensitivity to its ungrammaticality by avoiding its usage in argument structure; that is, by using it in verb-argument structure less often than the familiar verb on the one hand, and less often than their German age counterparts on the other.

1. Method

1.1. Participants

92 children were included in the final sample. 46 were monolingual German children and the other 46 were monolingual British English-speaking children. The German children were recruited from 17 kindergartens in Leipzig, a medium-sized German city. The number of kindergartens was due to the difficulty of recruiting large numbers of children from any one kindergarten due to the narrow age ranges required. The English children were recruited from five kindergartens in Manchester, a medium-sized English city, and through the Max Planck Child Study Centre participant database. For this database, parents are recruited from all over the Greater Manchester area through advertisement in newspapers, parenting magazines and, for example, kindergartens and doctors’ surgeries. After registering with the Child Study Centre, parents are called in when their child is required for a study.

There were two sessions and our criterion for inclusion in the study was merely completion of both. There were 30 2-year-old children in each language group; 16 English boys and 14 girls (mean = 2 years 4 months, range = 2-2–2-6) and 17 German boys and 13 girls (mean = 2 years 4 months, range = 2-2–2-6). We also tested 4-year-olds as a control group. For both languages there were 16 four-year-olds with 8 boys and 8 girls in each age group, with a mean age of 4 years (German range = 3-10–4-2; English range = 3-10–4-2). A further 64 children did not complete both sessions due to either fussiness (13 German 2-year-olds, 24 English 2-year-olds, 2 English 4-year-olds), refusing to speak (two 2-year-olds), bilingualism (1 child), absence in the second session (6 children), experimenter or parent error (9 children) or technical failure (6 children). In addition, the 4-year-olds in both languages completed standardized language tests, on the basis of which two 4-year-olds were excluded because they fell below the normal range. This ‘drop-out rate’ is similar to that found in other recent behavioral studies of the acquisition of verb-argument constructions in young English-speaking 2-year-olds (Fernandes, Marcus, DiNubila, & Vouloumanos, 2006; Abbot-Smith et al., 2001).
1.2. Design and materials

The children were shown films on a laptop, all depicting causative scenes in which an agent (a hand puppet) caused a change of location to a patient (an animal soft-toy). There were two within-subjects experimental verb conditions: familiar verb and novel verb. The familiar verb was ‘pushing’ (or schubsen ‘push/shove’ in German) and the novel verb was ‘domming’ (or dommen in German). In the novel game, the patient toy animal hung on the end of a string attached to the top of a pole. The novel action involved the agent hitting the patient so that it swung around the pole and then fell off. All of the hand puppets and toy animals took masculine gender in German (in which case-marking is unambiguous) and were familiar to young 2-year-olds.

There were two sessions. In each session there were two ‘blocks’ for the familiar verb and two for the novel verb. Each ‘experimental verb’ block contained a verb-learning phase, followed by a verb-argument models phase, and finally an elicitation trial. Each phase of each ‘block’ was associated with different agent and patient ‘scenes’, so that any sentences the child produced during the elicitation trial could not involve imitation. The familiar and novel verb ‘blocks’ were presented in an alternating fashion, interspersed with single ‘filler’ scenes (so that child X, for example, was shown familiar-novel-familiar-novel in session one). Two of the filler verbs involved caused change of location, namely ziehen or pulling and holen or getting. The other two filler verbs involved caused change of state, namely waschen or washing and anmalen or painting.

The order in which the experimental conditions were shown was fully counterbalanced across children in each age group; within each age group, half the children saw the novel before the familiar condition in session one and vice versa. The filler scenes always occurred first and then between the experimental conditions in both sessions. The order of which filler scenes were shown in which slots was counterbalanced according to Latin squares. There were thus 16 filler verb orders, one for each child in both 4-year-old groups. The remaining 2-year-olds were randomly distributed across various filler verbs orderings, but most importantly, one half of them saw the novel before the familiar condition as session 1 and for the other half vice versa.

1.3. Procedure

Each child participated in two sessions that lasted between 20 and 25 min each, within a one week period. There were three experimenters; two were female native speakers of German (who each tested half the German 2-year-olds and half the German 4-year-olds) and one was a female native speaker of British English. All the German children were tested individually in their kindergartens and no experimenter effects were found. Half the English 4-year-olds and 14 of the 30 English 2-year-olds were also tested in kindergartens and the rest were tested in a child lab in a British Psychology Department. All instructions, verb-argument models and elicitation questions heard by the English children were the translations of the German as given below.

As can be seen from Table 1 below, in each session the child watched two experimental verb blocks interspersed with individual filler scenes. At the beginning of each filler scene, the experimenter asked “Wer macht was?” (or ‘Who is doing what?’). During the filler scene the action was described using correct linking three times (e.g., ‘The AGENT is VERBing the PATIENT’). Then after the action was over, the experimenter asked “Wer hat was gemacht?” (‘Who did what?’). Then the experimenter described the filler action correctly twice using the past tense. Following this, if the child had neither spontaneously answered the question nor imitated the entire utterance, the experimenter asked the child to imitate the utterance by saying for example “Kannst du das sagen: Der Hase hat den Elefanten gewaschen.” (‘Can you say that: the rabbit washed the elephant’).
Table 1
Order of filler and experimental verb block scenes within a given session for a particular child

<table>
<thead>
<tr>
<th>Condition</th>
<th>Materials</th>
<th>Child saw (e.g.)</th>
<th>Child heard</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Block 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filler: ‘pulling’ + 1st agent–patient pair</td>
<td>Film scene</td>
<td>Boy pulls dinosaur</td>
<td>Grammatical linking (\times 5) Elicitation questions</td>
</tr>
<tr>
<td>'Pushing'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 1: verb-learning</td>
<td>Live scene</td>
<td>Experimenter pushes dog</td>
<td>Verb in citation form (\times 6)</td>
</tr>
<tr>
<td>Phase 2: verb-argument models</td>
<td>Film scene 1</td>
<td>Tiger pushes rooster</td>
<td>Verb in citation form (\times 3)</td>
</tr>
<tr>
<td>Phase 3: elicitation</td>
<td>Film scene 2</td>
<td>Tiger pushes elephant</td>
<td>Verb in ungrammatical linking (\times 5) Elicitation questions</td>
</tr>
<tr>
<td>Filler: ‘pulling’ + 2nd agent–patient pair</td>
<td>Film scene</td>
<td>Tiger pulls dog</td>
<td>Grammatical linking (\times 5) Elicitation questions</td>
</tr>
<tr>
<td>'Domming'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 1: verb-learning</td>
<td>Live scene</td>
<td>Experimenter ‘domms’ dog</td>
<td>Verb in citation form (\times 6)</td>
</tr>
<tr>
<td>Film scene 1</td>
<td>Film scene 1</td>
<td>Man ‘domms’ hedgehog</td>
<td>Verb in citation form (\times 3)</td>
</tr>
<tr>
<td>Phase 3: elicitation</td>
<td>Film scene 2</td>
<td>Man ‘domms’ dog</td>
<td>Verb in ungrammatical linking (\times 5) Elicitation questions</td>
</tr>
<tr>
<td>Film scene 3</td>
<td>Film scene 3</td>
<td>Lion ‘domms’ bird</td>
<td>Elicitation questions</td>
</tr>
</tbody>
</table>

Each experimental block had three phases; verb-learning, verb-argument models and elicitation. Table 1 above shows the ordering within one session with examples of what the children saw and heard (or what was elicited) for each phase. The number of times the children heard particular sentence types is given in brackets in the Table.

1.3.1. Phase 1: verb-learning
In each session, for both experimental conditions each child saw three scenes (one live and two filmed) which were described using the verb in citation form. The ‘live’ verb-learning phase came at the beginning of the session and involved the experimenter describing what she was doing to the patient (e.g., “Guck, was ich jetzt mit dem Hund mache. Das heißt dommen” ‘Look what I am doing to the dog? That is called domming’). The first film scene of each block of each experimental block was also described in the citation form (e.g., “Guck mal, was der AGENT mit dem armen PATIENT macht. Das heißt dommen.” ‘Look what the AGENT is doing to the poor PATIENT. That is called domming’). Over both sessions, for each experimental verb, each child heard a total of 24 verb models in the citation form.

1.3.2. Phase 2: verb-argument models
In the next phase – ‘verb-argument models’ – the children saw a different film scene. First, the scene was introduced by the experimenter saying: “Guck was der AGENT mit dem armen PATIENT macht/gemacht hat!” (‘Look what the AGENT is doing/did to the poor PATIENT!’). Then the scene was described by a hand puppet using the verb with ungrammatical linking (e.g., Der PATIENT dommt gerade den AGENT ‘The PATIENT is domming the AGENT’). Three of
these models were in the present tense and two were in the past tense. Over the two sessions, for each experimental verb each child saw four scenes which were described with a total of 20 verb-argument models with ungrammatical linking.

1.3.3. Phase 3: elicitation

At the end of each experimental block came another film scene with new characters during which elicitation questions were asked. The elicitation questions always followed the following pattern. Before the action, the experimenter first asked the child to name both animals. If the child did not name one of the animals, the experimenter asked her to imitate its name. Then the experimenter said “Und jetzt sagst du mir, was passiert.” (‘And now you tell me what happens!’). Then, during the action, the experimenter asked “Wer macht was?” (‘Who is doing what?’). If the child did not answer, or only named the action, the experimenter asked “Wer VERBt wen?” (‘Who is VERBing who?’). If the child merely gave an animal-only response, the experimenter asked “Was ist mit dem X?” (‘What is happening with the NAMED ANIMAL?’). After the action, the experimenter made the hand puppet ask the past tense version of the three questions above. In addition, if the child still had not answered any of the questions using the verb together with one argument, at the end the experimenter asked “Und was ist passiert?” (‘And what happened?’). If, during any of these elicitation questions or during the verb-argument models, the child used a non-target verb to refer to the target action, she was told the name of the verb (e.g., Dieses Spiel heißt nicht Xen, sondern dommen ‘This game is not called Xing, it is called ‘domming’).

Over both sessions for each experimental verb each child saw a total of four scenes during which elicitation questions were asked (at least one present tense and one past tense question for each).

1.4. Post-tests

All children completed a general language post-test at the end of the second session. The German 4 year-olds completed the morphological productivity subtest of the SETK for three to five year-olds (Grimm, 2001). The English 4 year-olds completed the repeating sentences subtest of the Clinical Evaluation of Language Fundamentals—Preschool (CELF-P) for 3–7-year-olds (Semel, Wiig, & Secord, 1992).

The parents of the English 2-year-olds were asked to complete the British version of the MacArthur CDI. All the parents of the lab-tested 2-year-olds and six parents of the kindergarten 2-year-olds did so. The MacArthur CDI has been normed for British-English children between 1-0 and 2-1 (Hamilton, Plunkett, & Schafer, 2000) but has not yet been normed for British 2-year-olds in our age range of 2-2 and 2-7. The current English 2-year-olds scored a mean of 93% for comprehension and 86% (S.D. = 0.13) for production of the vocabulary test items. For the German 2-year-olds we instead used the vocabulary comprehension sub-test of the “Sprachentwicklungstest” (language development test) for 2–3-year-olds (SETK-2) in order to at least provide some check that no 2-year-old was markedly below the others in their language group in terms of vocabulary development. The current German 2-year-olds scored a mean of 83% (S.D. = 0.91) on this test.

1.5. Scoring

All utterances produced by the 2-year-olds were transcribed from mini-disc separately by two native speakers of the particular language, and any utterance on which the transcribers did not agree was checked by a third (blind) transcriber. If the third transcriber did not agree with either of
the two, the utterance was discarded from the final sample. As the utterances produced by the 4-year-olds were much easier to understand, one transcriber transcribed half their data and the other transcriber transcribed the other half, with 15% overlap between the two. All utterances in the final sample which related to the target events were then coded by a German linguistics Masters student (who spoke fluent English) for construction and case-marking type and whether they involved grammatical, ungrammatical or ambiguous linking. Reliability checks were performed by the first author on 15% of the same data, with very high agreement between the two coders (Cohen’s Kappa = 0.9637 for German and 0.9408 for English).

2. Results

Each child had eight elicitation opportunities per verb (four film scenes with at least one present and past tense question each). However, it should be noted that some of the children did not use the target verb at all in some elicitation opportunities and others gave multiple responses within one elicitation opportunity. Fig. 1 shows that the children gave three types of responses; grammatically linked sentences, ungrammatically linked sentences and ‘other’ responses. The ‘other’ responses were ones in which the children did not use the target verb together with at least one noun phrase in a semantically unambiguous manner. These responses were made up of ‘noun-only’ responses, ‘target-verb only’ responses or intransitives or transitives using non-target verbs. Grammatically linked sentences were those in which the semantic roles of agent and patient were mapped on the grammatical markers of the particular language. In English the grammatical marker is word order, so if the agent was marked by preverbal position, this was coded as grammatical, whereas if it was marked by post-verb position, this was coded as ungrammatical. This meant that in English, ungrammatical linking was defined as either a full transitive with PATIENT-VERB-
AGENT order or a PATIENT-VERB\(^2\) (objectless) transitive or a VERB-AGENT (subjectless) transitive. In German the overriding grammatical marker is case-marking. Thus, in German if the agent was marked nominative, this was coded as ‘grammatical’ and if it was marked accusative it was coded as ‘ungrammatical’ (and vice versa for the patient). In fact, the German children hardly ever used word order and case-marking in conflict (as in “Den Frosch dommt der Löwe” ’the + accusative frog is domming the + nominative lion’) It is the frog that the lion is domming). Rather, the nominative-marked noun phrase was almost always also the first noun phrase. For sentences in the case-markers were omitted or phonetically unclear, word order was coded as the grammatical marker, as this is the default in adult German sentences without case-marking (Lindner, 2003).

2.1. Mean frequency

We examined the mean frequency of response types (shown in Fig. 1 above) by carrying out a 2 (Language) × (2 (Verb Condition) × 3 (Response type)) mixed factorial ANOVA for each age group separately. For the 4-year-olds we found main effects for Verb Condition \((F(1,30) = 4.657, p < .05, \eta^2_p = .134)\) and Response type \((F(2,60) = 39.811, p < .001, \eta^2_p = .570)\), plus an interaction between these two \((F(2,60) = 9.706, p < .001, \eta^2_p = .244)\). For the 2-year-olds these were also significant (Verb Condition = \(F(1,58) = 8.936, p < .005, \eta^2_p = .134\); Response Type = \(F(2,116) = 130.676, p < .001, \eta^2_p = .693\); Verb × Response Type = \(F(2,116) = 32.904, p < .001, \eta^2_p = 362\)). However, unlike the 4-year-olds, the 2-year-olds also showed a borderline main effect for Language \((F(1,58) = 3.889, p = .053, \eta^2_p = .063)\) and a significant interaction between Language and Response Type \((F(2,116) = 4.879, p < .01, \eta^2_p = .078)\).

We investigated the latter further by comparing the German and English 4-year-olds for each response type in both conditions. No significant cross-linguistic differences were found for any condition or response type for this older age group. We did the same for the 2-year-olds (with a Bonferroni correction for six comparisons). We found that the German 2-year-olds were significantly more likely to correct the novel verb to grammatical linking \((M = 1.77)\) than were the English 2-year-olds \((M = 0.60)\) \((t(58) = 2.966, p < .008)\). The only other comparison to approach significance was also in the novel verb condition, where the English 2-year-olds tended to use an ‘other’ response \((M = 13.33)\) more often than their German age counterparts \((M = 9.77)\) \((t(58) = 2.162, p = .035)\). The same results were found when we carried out these cross-linguistic using Mann–Whitney U-tests. This provides initial support for our hypothesis that young German 2-year-olds have more robust representations of the prototypical transitive than do English children the same age.

Of course, one might counter-argue that if the English 2-year-olds were more verbally reticent than the German children, the probability of them producing a target verb in argument structure would be lower. However, Fig. 1 shows that the English 2-year-olds in fact gave MORE verbal

\(^2\) It might be argued that ‘PATIENT-VERB’ utterances might have been intended as (grammatical) unaccusative intran-
sitives (e.g. ‘the ball dropped’). Our coders agreed that this was probably not the case in our experimental scenario as the children used such utterances following questions like ‘Who is pushing who?’ and they themselves only ever heard these actions described with sentences containing two noun phrases (preceded by the lead-in ‘Look what the agent is doing to the patient’). If we excluded PATIENT-VERB utterances, we would also have to exclude AGENT-VERB utterances, which are indisputably grammatical. Moreover, the number of PATIENT-VERB utterances was small (e.g. the English two-year-olds produced only one such utterance in the novel verb condition and two such utterances in the familiar verb condition).
responses over the two conditions ($M = 26.63$) than their German age counterparts ($M = 21.33$) ($t(58) = 1.972, p = .053$). Thus, the probability of them producing a sentence with a novel verb together with a noun phrase should actually be greater than that for the German 2-year-olds. Nonetheless, since there was some variability in mean frequency of usage between verb conditions and age groups, the results can be interpreted more clearly by examining mean proportions of grammatical and ungrammatical linking within Age Group and Verb Condition and Language.

2.2. Mean proportions

Since we were using proportional data, we first carried out arcsine transformations on all dependent variables. We then carried out a 2 (Age) $\times$ 2 (Language) $\times$ 2 (Verb Condition) mixed factorial ANOVA on mean proportions for grammatical, ungrammatical, and other responses separately. For corrections to grammatical linking we found a main effect for Verb Condition ($F(1,88) = 45.490, p < .001, \eta^2_p = .341$—overall the familiar verb was corrected to grammatical linking more often than the novel verb) and for Age Group ($F(1,88) = 54.450, p < .001, \eta^2_p = .382$—overall the 4-year-olds corrected to grammatical linking more often than the 2-year-olds). There were no other main effects or interactions for grammatically linked responses. For ungrammatical linking there was a significant Verb $\times$ Age Group interaction ($F(1,88) = 5.871, p < .05, \eta^2_p = .063$) but no other main effects or interactions (presumably because numbers of ungrammatical usage were very small in all conditions and groups). For ‘other responses’ we found main effects for Verb Condition ($F(1,88) = 50.017, p < .001, \eta^2_p = .362$) and Age Group ($F(1,88) = 76.511, p < .001, \eta^2_p = .465$); the 2-year-olds used more ‘other responses’ overall than did the 4-year-olds. However, this must be seen within the context of both a Verb $\times$ Language interaction ($F(1,88) = 4.956, p < .05, \eta^2_p = .053$) and a Verb $\times$ Age Group interaction ($F(1,88) = 5.115, p < .05, \eta^2_p = .055$). This Verb $\times$ Language interaction for ‘other’ responses is relevant to the ‘graded representations hypothesis’ relating to avoidance of using the novel verb in a sentence. Because this point is so crucial to our experimental hypothesis, we asked whether young English 2-year-olds avoided using the novel verb more often than their German age counterparts. The answer was that they did ($t(58) = 3.608, p < .005$).

It might be argued, however, that our data are not suited to parametric analysis since certain age groups in certain conditions tended to fall into predominately one value. Therefore we carried out 12 Mann–Whitney $U$-tests on the non-transformed proportional data comparing the two language groups, for each age group, verb condition and response type separately. No cross-linguistic differences were found for the 4-year-olds in any condition. No cross-linguistic differences were found for the 2-year-olds for the familiar verb condition. In the novel verb condition, however, the German 2-year-olds used grammatically linked responses significantly more often than the English 2-year-olds ($U = 248, n_1 = 30, n_2 = 30, p < .005$). For ungrammatically linked responses this comparison reached borderline significance ($U = 368, n_1 = 30, n_2 = 30, p = .055$). This is because in this condition, the English 2-year-olds used ‘other’ responses significantly more often than the German two-year-olds ($U = 212, n_1 = 30, n_2 = 30, p < .001$); i.e., they avoided using the novel verb in a sentence more often than did their German age counterparts (Fig. 2).

2.3. Numbers of children avoiding usage of the target verbs in sentences

Thus, our crucial finding is that in terms of actively demonstrating their knowledge of grammatical linking by actively generalizing to a novel verb, the current English 2-year-olds were
clearly not as competent as the German children the same age. To demonstrate this further, for each verb condition and age group separately we categorized the 2-year-olds in terms of whether they (1) used that verb in a sentence (with at least one noun or pronoun), or (2) always avoided using the target verb in a sentence. The results are shown in Fig. 3 below. The difference between
the language groups in terms of the number of 2-year-olds who never used the novel verb in a sentence was found to be significant ($\chi^2(1, N=60) = 15.429, p < .001$). No significant cross-linguistic differences were found for the 4-year-olds (who always used the verb in a sentence) or for the 2-year-olds in the familiar verb condition ($p = 1.0$, and $p = .604$ for usage and correction, respectively). The current study thus supports the hypothesis that German 2-year-olds develop more robust abstract representations of active-transitive sentences earlier than English 2-year-olds, at least for those in which case-marking and word order converge.

2.4. Evidence that English 2-year-olds have acquired non-verb-specific transitive linking

The English 2-year-olds were clearly not willing to use the novel verb in ungrammatical or grammatical linking. The majority (66%) never used verb-argument responses in the Novel Verb Condition at all. Instead, their largest response type in this condition was overwhelmingly ‘other verb’ (over 40% of responses) and noun-only utterances (over 40% of responses, with just over half referring to the agent). Their next largest category was ‘unanalyzable’ (circa 10%), which mostly involved usage of the novel verb alone. Thus, most English 2-year-olds apparently avoided using the novel verb in argument structure altogether despite simultaneously having much greater difficulties than the German 2-year-olds in correcting to grammatical linking. This fits with avoidance behavior found in studies using the weird word order paradigm and supports a graded representations view.

In contrast to weird word order studies, however, when the current English 2-year-olds used the novel verb in verb argument structure, they hardly ever used it in ungrammatical linking (see Fig. 4 below). To further investigate this, we re-analyzed the mean proportions. This time we calculated the proportions out of grammatically linked plus ungrammatically linked responses only (i.e., we excluded the ‘other’ responses from the denominator). We again used arcsine transformations and first carried out a 2 (Language) × 2 (Age Group) × 2 (Verb Condition) × 2 (Response Type) mixed factorial ANOVA. This time we included only those 2-year-olds who used the target verb in argument structure in both verb conditions ($N = 10$ for English and $N = 22$ for German). The only
effects were a main effect for Response Type \( (F(1,60) = 163.461, p < .001, \eta^2_p = .731) \), where the children used grammatical linking significantly more often than ungrammatical linking, and an interaction Age Group \( \times \) Response Type \( (F(1,60) = 4.173, p < .05, \eta^2_p = .065) \), because the 4-year-olds were significantly less likely to use ungrammatical linking (and more likely to use grammatical linking) than the 2-year-olds. We followed up this interaction by carrying out two 2 (Language) \( \times \) 2 (Verb Condition) mixed factorial ANOVA for grammatical and ungrammatical linking responses separately for the 2-year-olds only. No significant effects were found (for all \( p > .411 \)). That is, there were no differences between verb conditions nor were there any differences between the German and English 2-year-olds regarding proportions of ungrammatically or grammatically linked sentences, when we only include those children who used the target verbs in a sentence. Rather, the ten English 2-year-olds who used the novel verb in argument structure corrected to grammatical linking in 77% of their responses. This is strikingly greater than Akhtar’s (1999) and Matthews et al.’s (2005) older 2-year-olds. In those studies, when the 2-year-olds did use the novel or low frequency real verbs which they had heard ungrammatically, they were willing to use them ungrammatically themselves in over 50% of their verb-argument responses. While there are numerous differences between the weird word order studies and the current weird linking method, one possible explanation for this difference is that the current method may have syntactically primed the use of NVN form, freeing up enough processing capacity for the children to focus on linking the correct semantic roles to the correct noun slots.

2.5. Evidence against potential criticism

One potential argument against our interpretation of the cross-linguistic differences in young two-year-olds might be that all the German 2-year-olds were all tested in kindergartens, whereas nearly half of the English 2-year-olds were tested in the child lab. One might hypothesize that young children who have had exposure to kindergarten schooling might have more advanced language skills. Therefore, we compared novel verb productivity between our English 2-year-olds who were tested in the lab \( (N = 16) \) with those who were tested in kindergartens \( (N = 14) \). In actual fact we found the reverse pattern; only two of the English kindergarten 2-year-olds ever used the novel verb together with a noun phrase, and both of these always used it incorrectly. From this, one would predict that had all the German and English children been tested in kindergartens, the cross-linguistic differences in the 2-year-old age group would have in fact been even stronger.

A second potential criticism is that the children might have used ungrammatical linking because they thought that was the ‘game’. In this case, we might expect a higher use of ungrammatical linking in session two than session one. \( t \)-Tests were carried out comparing session 1 with session 2 for both age groups, both languages, both Verb Conditions, and each Response Type separately. No significant differences were found (for all \( p > .33 \)), indicating that no age or language group was more likely to use grammatical linking more often in session 2, or more likely to use ungrammatical linking in session 1.

3. Discussion

The current study used a new “weird linking” elicitation method to investigate the acquisition of semantic role-to-grammar linking in action-verb transitives. If young 2-year-olds operate solely on the basis of verb-specific representations, we should expect them to use the novel (but not the familiar verb) in ungrammatical linking. This was not the case. The English 2-year-olds were rarely
willing to use the novel verb in the ungrammatically linked argument structure in which they had heard it used. Instead, when they did use the novel verb in argument structure, they corrected it in 77% of their verb-argument responses. However, the majority of the time, the English 2-year-olds appeared to avoid using the novel verb in argument structure altogether. In contrast, the German 2-year-olds were significantly more likely to correct a novel verb to grammatical linking than were their English age counterparts.

Our finding that German 2-year-olds perform well in the novel verb elicitation task throws doubt on the ‘inadequate task’ account of why English 2-year-olds have difficulty producing correct transitives with novel verbs in the current and previous production tasks. This account would predict that 2-year-olds should universally show the same performance difficulties, regardless of the language they are learning. That is, if the poor performance of young English 2-year-olds in the current and previous novel verb elicited production tasks is due to difficulties with lexical access, or with interpreting the meaning of the novel verb or other performance difficulty, German children the same age should show the same difficulties when carrying out the same task. This was not found to be the case.

However, while there is no disputing the good performance of the German 2-year-olds, there are a number of potential alternative accounts of the current cross-linguistic differences in regards to willingness to produce the novel verb in a sentence. Firstly, it is impossible to be absolutely sure that the German and English 2-year-olds had attained the same language level since there currently exist no identical language tests which have been standardized for the two groups. However, both groups were within the normal range on the vocabulary tests we used. A second possibility might be to argue that the British 2-year-olds were more verbally reticent than were the German 2-year-olds. In fact, we found the British 2-year-olds to be significantly LESS verbally reticent than the German children (see section 2.1). A related worry might be that the British 2-year-olds were shyer about using a novel word at all (perhaps for cultural reasons). However, they were not reticent about using the novel verb in isolation (e.g., “dommed” in response to “What happened with the lion?”); such utterances make up circa 10% of their verbal responses (see section 2.4). In addition, previous ‘weird word’ order studies have also not found British 2-year-olds to be reticent in using novel verbs in sentences per se, but rather only when they have heard that particular verb used in an ungrammatical sentence (Abbot-Smith et al., 2001).

A third counter-argument is that the English 2-year-olds were significantly worse as a group because the half that we tested in the child lab felt less secure than those – like the German children – who were randomly selected and tested in urban nurseries. In fact only two of the English kindergarten 2-year-olds ever used the novel verb together with a noun phrase, and both of these always used it incorrectly. From this one would predict that had all the German and English children been tested in the same venue, the cross-linguistic differences would be even greater.

The differences we found between British 2-year-olds tested in kindergartens and British 2-year-olds tested in the lab do, however, indicate that there is considerable developmental variability in this age group. There are a number of potential reasons for this (at least British) difference between the kindergarten and child lab ‘populations’. One is that British children felt less comfortable being tested in a kindergarten, and it is indeed possible that there are also cross-cultural differences in how comfortable children feel being in kindergarten at all. Another possibility is that this difference in fact relates to socio-economic status (SES) as volunteers for child labs tend to be parents from middle-class backgrounds. While we do not think that the British kindergarten/lab disparity impacts on our cross-linguistic findings for the reasons given above, it does indicate the importance of future work developing uniform standardized cross-linguistic vocabulary tests and cross-cultural SES criteria appropriate for very young pre-schoolers.
There are two further potential factors, however, which might be thought to have impacted on our cross-linguistic findings. That is, one might worry about the relative availability of an alternative verb for describing the novel action. In fact, neither language has a precise word for ‘swing someone around a pole so that he falls off’. While it is true that the English 2-year-olds tended to give more ‘other verb’ responses than did their German counterparts, the same is also true for ‘noun-only’ responses. The actual ‘other’ verbs used in the novel condition were almost identical across the two languages; the most common in both were _schubsen_ ‘push’ and _runterschmeissen_ ‘throw down’. Thus, it appears that the English 2-year-olds were applying a range of alternative non-target responses in order to avoid using the novel verb in a sentence.

The last possibility that needs addressing is that more British than German 2-year-olds were excluded from the study. However, if one were to argue that excluding the ‘fussy’ (i.e., inattentive) children led to the elimination of the less-advanced children, one should predict that the English 2-year-olds should perform better as a group than the German 2-year-olds (because more English 2-year-olds were excluded). In fact, we found the opposite pattern.

We therefore argue that our cross-linguistic findings suggest that the difficulties English 2-year-olds have in demonstrating productivity with the transitive are not due solely to insurmountable performance limitations but to input-related properties of the particular language and construction type that they are learning. That is, there is something about German prototypical active transitive constructions (in which word order and case-marking collaborate) which makes them easier to learn than English active transitive constructions (in which the agent and patient are marked solely through word order). This leads young German two-year-olds to develop of a more robust knowledge of linking in ‘prototypical’ active transitive constructions earlier, at least in sentences without pronouns. Of course a truly prototypical sentence would have a pronoun subject (Cameron-Faulkner et al., 2003). Even in English, pronouns are often case-marked. In fact, if we had tested English 2-year-olds on a truly prototypical active transitive with case-marked pronouns, we predict that they might have performed as well as the German 2-year-olds. However, it is of course impossible to use pronouns in the linguistic models using the current methodology since one would not know which animal each pronoun refers to.

This raises the question of exactly which aspect of the German input led to this effect. One possibility is simply that the German children are getting two cues to the agent whereas the English children are merely getting one, as suggested by Bates and MacWhinney’s (1987) ‘coalitions-as-prototypes’ proposal. However, there are a number of other (non-mutually exclusive) possibilities, one of which is that the case-markers serve as positional makers highlighting the structural positions. Hearing the contrastive usage of case-marked lexical items such as ‘der’ versus ‘den’ might be likely to lead German children to search for a meaning distinction from early on due to the pragmatic principle of contrast (Clark, 1990). That is, the child will assume that there must be a reason for why adults sometimes use ‘der’ and at other times use ‘den’ when referring to the same object. Since German case-marking is more strongly associated with subject-before-object word order, it may highlight the fact that structural positions in German are relevant for encoding who did what.

Of course, German children must eventually learn to dissociate the word order cue from case-marking in the process of learning the adult system of heuristics for semantic role assignment in action-verb active transitives. In adult German, when case-marking is present, this must be followed, but when case-marking is entirely absent, agent-first word order is most likely. Thus, innate linking rules would not really assist the child very far because how such grammatical roles are marked varies greatly both within languages as well as between languages. In addition to compelling evidence from child acquisition against innate linking (Bowerman, 1990), there is recent evidence that children can learn to link phrasal patterns and meanings, even when such
mappings are supposedly disallowed by universal grammar (see Casenhiser & Goldberg, 2005). Innate linking theories would, moreover, presumably also have difficulty accounting for the far better performance of the German over the English 2-year-olds in the current task.

Despite the better performance of the German 2-year-olds, the current English 2-year-olds were clearly not willing to use the novel verb in ungrammatical linking. Thus, the current findings do not support purely lexicalist accounts of initial grammatical development, such as early versions of the verb-island hypothesis (Tomasello, 1992). This theory claims that representations of the transitive are initially entirely based around familiar verbs (e.g., English children know that the ‘pusher’ comes before the verb ‘push’, and the ‘pushee’ follows it), and that this knowledge is not generalised to unfamiliar verbs until around the age of three years old. This would predict that English 2-year-olds should be willing to use the novel verb in ungrammatical linking, which is clearly not the case.

Instead, we found measure-dependent results for the English 2-year-olds, depending on whether we only examine actual production of grammatically linked transitives or whether we take avoid-ance behavior into account. This fits with similar findings of avoidance of ungrammaticality using the weird word order method (Akhtar, 1999; Abbot-Smith et al., 2001). This avoidance behavior would appear to indicate that the English children had acquired some kind of non-lexically specific representation of the transitive construction, but had difficulty demonstrating this by producing transitives with the novel verb. Thus, proponents of the ‘inadequate task’ view are clearly correct that high task demands may make it difficult for 2-year-olds to demonstrate what they have already learned about the grammar of their language.

The controversial question is, however, not whether 2-year-olds have performance limitations but rather the importance of these performance limitations. According to the ‘inadequate task’ explanation, performance limitations result from irrelevant tasks demands which need to be eliminated in order to tap the underlying grammatical competence which they obscure. In contrast, in the graded representations account, task difficulties are not something which always need to be eliminated. Rather, the level of task difficulty allows us to see how robust the child’s representation is. In this view, cognitive representations are embedded in specific processes underlying overt behavior (Munakata, 2001). Successful performance depends on an interplay between degree of automaticization and processing capacity (Case, 1985; Siegler, 2000) and strength of representation, which depends on the input (Munakata et al., 1997; Huttenlocher, Vasilyeva, & Shimpi, 2004). Even if a particular behavioral response (such as the production of two- or three-word utterances) has not been completely automaticized, but the transitive representation is strong enough (as with our German 2-year-olds), the child will still have enough processing capacity to successfully carry out the task.

The reverse situation also leads to successful performance; namely even a ‘weak’ transitive representation (such as that of the English-speaking children aged 1 year 9 months in (Gertner et al., 2006) will be adequate to successfully carry out the task if a particular behavioral response HAS been adequately automaticized. Looking has of course been automaticized by 1- and 2-year-olds. The same is probably true of pointing, although this remains to be clarified for English-speaking children under 2–6. If, however, a transitive representation is weak AND the required response has not been sufficiently practiced, then a child may ‘fail’ this task due to processing capacity

---

3 Fernandes et al. (2006) claim that they found above chance performance with the active transitive construction for English-speaking 2; 4-year-olds using a pointing task. However, it is impossible to verify their claim since they conflate the results from their 2; 4-year-old group with those of the 2; 10-year-olds.
limitations but ‘pass’ on another measure, which does not impede on processing capacity to the same extent. This would appear to describe the behavior of the current English 2-year-olds.

In the current study we have focused on the representational strength of the link between agent and patient roles and the particular construction-dependent marking in English and German. We found that performance appears to depend on both the particular task used AND the particular language the child is learning. On the basis of this, we argue that the transitive representations of young pre-school children are graded in strength depending on both the amount of input they have previously processed and the relative salience of perceptual cues to the semantic roles. This graded representations interpretation fits well with a growing body of researchers in syntactic acquisition who argue that emergent knowledge may be represented at a more abstract level but nonetheless be further strengthened through continued experience of processing the input (Huttenlocher et al., 2004; Narasimhan, Budwig, & Murty, 2005; Chang et al., 2006).

Acknowledgements

Many thanks to Silke Brandt and Eva Leermann for testing the German children and for very helpful suggestions; to Anna Roby for testing the English children, to Anne Alexander, Katja Fuchs, Ellie O’Malley, Anna Roby and Tim Manning for transcribing; to Silke Brandt for coding; and to Colin Bannard, Edith Bavin, Heike Behrens, Tanya Behne, Peter Bryant, Miriam Dittmar, Letitia Naigles, Anat Ninio and five anonymous reviewers for helpful comments on earlier drafts. Thank you to Roger Mundry for help with the statistical analyses. A special thank you to the staff and caregivers of the following English kindergartens: Brighter Beginnings, Quays Day Nursery, Ruperts Day Nursery, Rainbow Day Nursery, St. Gilberts Primary School; and German kindergartens: Eisenacher-Straße, Gundermannstraße, Gustav-Freytag-Straße, Hans Marchwitzer-Straße, Hans-Otto-Straße, Holsteinstraße, Kurt-Kresse-Straße, Kurt-Tucholsky-Straße, Leipziger Straße, Lichtenbergweg, Nordstraße, Richard-Lehmann-Straße, Spinnereistraße, Sternwartenstraße, Theodor Neubauer-Straße, Tschaikowskistraße, and Zum Kleingartenpark.

References


