

The role of entrenchment in children's and adults' performance on grammaticality judgment tasks

Anna L. Theakston*

Department of Psychology, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Received 1 March 2003; received in revised form 1 August 2003; accepted 1 August 2003

Abstract

Between the ages of 3 and 7 years, children have been observed to produce verb argument structure overgeneralization errors (e.g., *Don't giggle me*; Bowerman, 1982, 1988; Pinker, 1989). A number of recent studies have begun to find evidence that the precise distributional properties of the input may provide an important part of the explanation for why children retreat from overgeneralization errors (Brooks & Tomasello, 1999; Brooks, Tomasello, Dodson, & Lewis, 1999). The current study evaluates the role of entrenchment (Braine & Brooks, 1995) in constraining argument structure overgeneralization errors using a grammaticality judgment task. The 5-year-olds, 8-year-olds, and adults were presented with examples of argument structure errors containing high and low frequency verbs matched for semantic class and asked to indicate whether, or the extent to which they found the sentences to be grammatical. The data show that across all groups, sentences with argument structure errors containing low frequency verbs were judged to be significantly more grammatical than those containing high frequency verbs. These findings provide further support for the entrenchment hypothesis and suggest that verb frequency plays an important and continuing role in determining a speaker's choice of verb argument structure.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Overgeneralization errors; Entrenchment; Verb argument structure

1. Introduction

A question that has received a lot of attention in the language acquisition literature is, how children are able to become productive in their use of verb argument structure, and yet learn to avoid producing overgeneralization errors such as *Don't giggle me* (Bowerman, 1982)

* Tel.: +44-161-275-2600; fax: +44-161-275-2588.

E-mail address: theaksto@psy.man.ac.uk (A.L. Theakston).

where, in this case, the intransitive verb *giggle* is used in a transitive construction. From the age of around 3 years, up to and beyond 7 years of age, children have been observed to produce these kinds of errors in their spontaneous speech (e.g., Bowerman, 1982, 1988; Pinker, 1989). Moreover, studies using a grammaticality judgment methodology have shown that children accept these kinds of errors as grammatical (Hochberg, 1986; Mazurkewich & White, 1984). These findings suggest that overgeneralization errors are not always one-shot innovations or performance errors, but rather may reflect more persistent influences on children's early grammatical representations at any given point in development. Given the apparent lack of consistent direct negative evidence in the form of parental recasts that might inform children that these kinds of utterances are ungrammatical (e.g., Brown & Hanlon, 1970), researchers are still a long way from fully understanding the mechanisms that allow children to retreat from these kinds of errors to end up with an adult-like grammar.

A number of explanations have been put forward that attempt to account for how children avoid making overgeneralization errors (e.g., Braine, 1971; Mazurkewich & White, 1984; Randall, 1990). Pinker (1989) suggests that the verbs that may appear in specific syntactic constructions belong to narrowly defined semantic classes. Thus, children will make argument structure overgeneralizations until they have correctly established the meaning of a verb and the narrow semantic class to which it belongs. In a recent experimental study, Brooks and Tomasello (1999) taught children two novel verbs. One verb had a meaning that was consistent with a verb class that can be used both transitively and intransitively, whereas the second verb had a meaning consistent with a verb class that had fixed transitivity. Results showed that from 4.5 years of age, the children tended to observe the transitivity status of the verb from a fixed transitivity class, but extended use of the verb from a bitransitivity class to constructions not modeled in the input. However, 2-year-olds did not observe the semantic constraints and in fact produced a larger number of transitivity violations with fixed transitivity verbs than with bitransitive verbs. These findings suggest that narrowly defined semantic classes may help children to constrain their argument structure productions from around 4.5 years of age (see also Brooks & Zizak, 2002), but that at earlier stages of development children may use alternative mechanisms to constrain productivity.

One weakness of the semantic classes account of the retreat from overgeneralization errors is that it does not make clear predictions about when children might be expected to retreat from making overgeneralization errors with individual verbs. Researchers working within a constructivist framework have begun to show that there may be a much closer relation than has previously been assumed between the language children hear and their early acquisition of verb argument structure. For example, children's use of intransitive or transitive constructions with bitransitive verbs is closely related to their mothers' choice of construction (Theakston, Lieven, Pine, & Rowland, 2001), children's use of the verb *Go* in a number of constructions with a number of meanings is related to their mothers' use of the verb with different constructions and different meanings (Theakston, Lieven, Pine, & Rowland, 2002), and the order of acquisition of individual verbs is influenced by the frequency and syntactic diversity of those verbs in the input (Naigles & Hoff-Ginsberg, 1998). Similarly, children's use of Wh-question syntax is related to the frequency of individual Wh-word + auxiliary combinations in the input (Rowland & Pine, 2000, also see Van Valin, 2002). Given the apparent relation between the language children hear and their early language use, this suggests that researchers should look to the properties of the input

to discover possible explanations for children's early overgeneralization errors, and for the reasons why these errors gradually disappear over the course of development.

In fact, two recent explanations for the retreat from overgeneralization errors are consistent with a constructivist input-driven account of early language development. The first explanation, preemption, suggests that children will be more likely to produce argument structure errors if they have not yet acquired an alternative construction to express their intended meaning. Thus, children will be prone to make errors such as *Don't giggle me* if they have not yet acquired the periphrastic causative construction with the verb *giggle* and therefore cannot produce the sentence *Don't make me giggle* to express their intended meaning while maintaining the transitivity status of the verb (Braine & Brooks, 1995, see also Bates & MacWhinney's competition model, 1987; Clark's principle of contrast, 1987). Experimental evidence provides support for the role of preemption in constraining children's argument structure errors. Brooks and Tomasello (1999) taught children aged 2.5, 4.5, and 6–7 years two novel verbs that were either modeled in a patient-focused intransitive construction (*The sock's tammimg*), or in an agent-focused transitive construction (*The man's meeking the cow*). Half of the children were also presented with preempting structures that would allow them to describe a transitively modeled verb from the perspective of the patient (*The cow's getting meeked*), or an intransitively modeled verb from the perspective of the agent (*The boy's making the sock tam*) while maintaining the transitivity status of the verb. Results showed that only the 6- to 7-year-olds were able to use the presence of a preempting structure to avoid making errors that violated the transitivity status of the verb (see also Brooks & Zizak, 2002). Thus, although preemption may play a role in the retreat from overgeneralization errors, this strategy does not appear to be available to children until relatively late in development. It is also unclear whether children can make use of some preempting structures earlier in development than others. From a constructivist perspective, we might predict that there will be a relation between the availability of individual preempting structures at different stages of development and their frequency in the input, but such a relation has not yet been demonstrated empirically. Moreover, the availability of a specific preempting structure might operate at the level of the individual verb, or at the level of the construction, or at both levels at different stages of development.

The second explanation, entrenchment, is more directly related to the distributional properties of the input (Braine & Brooks, 1995). In this account, the likelihood that children will produce an overgeneralization error with a particular verb is closely related to its frequency in the input. The more often children hear a particular verb used in a particular construction or range of constructions in the input, the less likely they are to overgeneralize use of that verb to a novel construction not modeled in the input. In other words, the familiarity of individual verbs is seen to play an important role in determining the likelihood that children will produce overgeneralization errors. Over the course of development as children are exposed to more linguistic input, the familiarity of individual verbs is expected to increase and thus the number of errors found in children's speech will decrease toward levels in the adult grammar. The direct prediction from this approach, therefore, is that children will be more likely to make argument structure errors with verbs that have lower frequency in the input and are acquired later in development than with verbs that are of higher frequency in the input and acquired earlier in development.

A recent experimental study has provided support for the entrenchment hypothesis. Brooks, Tomasello, Dodson, and Lewis (1999) modeled four actions for 3- to 4-year-olds, 5-year-olds, and 8-year-olds. Each action was associated with a pair of verbs, one of which was a high frequency, early acquired verb (e.g., *come*), and the other a low frequency, late-acquired verb (e.g., *arrive*). Age of acquisition and verb frequency were determined using parental checklists of children's language use and frequency estimates based on use in samples of child-directed-speech. The verbs had fixed transitivity, and were modeled in either the transitive or intransitive construction as appropriate to the verb. The children were then questioned to elicit use of the verbs in the non-modeled, ungrammatical construction, i.e., the transitive construction for intransitive verbs and vice versa. The prediction was that children who heard a particular action described with a low frequency, late-acquired verb would produce more overgeneralization errors than children who heard the action described using a high frequency, early acquired verb. The results confirmed this prediction. At all ages, the children were more likely to violate the transitivity status of the lower frequency verbs than the higher frequency verbs, suggesting that the familiarity of individual verbs may indeed play an important role in constraining argument structure generalizations. Although there was no effect of age on the absolute numbers of errors made by the children, the errors made by the younger children constituted a greater proportion of their overall utterances with individual low frequency, late-acquired verbs than in the older group. These results provide some support for the suggestion that as linguistic experience increases with age, the likelihood that children will make errors with later acquired verbs decreases.

The current study focuses on the role of entrenchment in constraining children's verb argument structure overgeneralizations, but adopts a different methodology in an attempt to provide further support for an input-driven approach to constraining children's productivity. The current study differs from previous studies examining the role of entrenchment in the following ways:

- A grammaticality judgment methodology was adopted rather than an elicitation task to evaluate the effects of verb frequency on children's judgments of argument structure errors. In production tasks, it is difficult to elicit argument structure errors (the children in Brooks et al.'s (1999) study produced just 2.2-argument structure errors on average with error-prone late-acquired verbs). This reflects the fact that children are generally conservative and reluctant to use a verb in constructions that they have not heard used with that verb. A grammaticality judgment task reduces the performance demands on the children by testing their comprehension rather than production of argument structure errors. This methodology may, therefore, overcome some of the problems associated with children's tendency to produce verbs in the same way as they have heard them, and provide greater insight into children's underlying verb representations.
- Adults were included in the study as well as 5- and 8-year-olds. The aim was to establish whether the kinds of processes operating in children's judgments of grammaticality might continue to play a role in adult language use. From a usage-based perspective, adult linguistic representations are thought to be dynamic rather than static, and continually influenced by the frequency of use of individual lexical items and constructions (e.g., Bybee, 1998; Croft, 2001). We might therefore expect that even adult judgments of grammaticality will reflect the frequency of use of individual verbs. Frequency effects

have been found to be important in many aspects of adult language processing (e.g., Bock, 1986; Bybee, 1995; Ellis & Schmidt, 1998; Hasher & Chromiak, 1977; Kuiper, 1996; Marslen-Wilson, 1990; Trueswell, 1996), but it has not yet been shown that the same frequency effects may play a role in both adult's and children's use of verb argument structure.

- A larger number of high and low frequency verbs were included in the current study than in Brooks et al.'s (1999) study of entrenchment. The aim was to investigate in greater detail the relation between verb frequency and the extent to which individual verbs are judged to be grammatical when presented in sentences containing argument structure errors.

One additional area of interest in the current study concerns the interaction between verb frequency and semantic class in the adult data. According to Pinker's (1989) theory, verb argument structure is determined by the narrow semantic class to which a verb belongs. It is therefore unclear why verb frequency might have an effect on judgments of grammaticality as membership of a verb class is absolute rather than relative. Although Brooks et al. (1999) found an effect of verb frequency on children's production of verb argument structure within specific semantic classes, it is possible that the children in the study had not yet formed adult-like representations of verb meaning, and therefore made errors that violated the semantic class for unfamiliar low frequency verbs. However, if a frequency effect is found for adults who are expected to operate with fully defined semantic classes, this would suggest that narrowly defined semantic categories may be constructed, at least in part, on the basis of the frequency of use of individual verbs in individual constructions. Membership of a specific semantic class may be organized around prototypes where some less frequent verbs are viewed as more marginal members of the class rather than in absolute terms.

2. Study 1: children's grammaticality judgments

2.1. Method

2.1.1. Participants

There were two groups of participants in this study, 5- and 8-year-olds. Fifty-nine children with a mean age of 5.8 (range = 5.1 to 6.2) formed the younger group of children. Fifty-five children with a mean age of 8.5 (range = 7.1 to 9.2) formed the older group of children. The children were recruited from two primary schools in the Greater Manchester area and were monolingual English speakers.

2.1.2. Materials

Twenty-four sentences were selected for use with the children plus a further two practice sentences. The sentences were selected as follows.

- *Group 1*: six sentences containing argument structure errors with high frequency verbs.
- *Group 2*: six sentences containing argument structure errors with low frequency verbs matched for semantic class to the high frequency verbs in Group 1.
- *Group 3*: six fully grammatical sentences.
- *Group 4*: six ungrammatical sentences.

Groups 1 and 2. The 12 test sentences contained argument structure errors of the kind produced by children from around 3 years of age and continuing until age 7 or older (see [Appendix A](#)). Six were selected from examples of argument structure errors produced by children that are documented in the literature ([Bowerman, 1978, 1982; Pinker, 1989](#)), and were chosen on the basis that their intended meaning could be understood out of context. Each of the six verbs used in these sentences was then paired with another verb from the same semantic class ([Levin, 1993](#)) that had either higher or lower frequency, but would still be known to the children, to create the remaining six test sentences. Thus, Group 1 consisted of six sentences containing argument structure errors with high frequency verbs, and Group 2 consisted of six sentences containing argument structure errors with low frequency verbs matched for semantic class.

To determine the relative frequency of use of each verb, the Manchester corpus ([Theakston et al., 2001](#)) was searched for the frequency of each verb in the combined sample of 12 British-English mothers' child-directed-speech (see [Table 1](#)). The CHILDES input frequency corpus ([MacWhinney, 2000](#)) was also searched for comparison purposes, but as this corpus is not tagged for morphological category, the frequencies for lexical items that can be used as both nouns and verbs (e.g., *brush*) are likely to be inaccurate. Verbs that appeared in these samples of maternal input to young children were used in the test sentences as these verbs are likely to be known to the younger children in the current study.

Groups 3 and 4. Two groups of sentences were used as control sentences to determine whether individual children were able to rate unambiguously grammatical or ungrammatical sentences according to their syntactic structure. Group 3 sentences were fully grammatical simple declaratives (e.g., *The cat drank the milk*). Group 4 sentences were ungrammatical and contained word order errors and/or additional or omitted words or phrases (e.g., *Mum shopping went, Jump the cat the mouse*). Both groups of sentences contained verbs that are familiar to young children and were of a roughly similar length to the test sentences.

In addition to the four groups of sentences described above, there was one practice grammatical control sentence and one practice ungrammatical control sentence. These were

Table 1
Verb frequency in the Manchester corpus

High frequency verbs	Frequency	Low frequency verbs	Frequency
Brush	95	Shampoo	1
Come	6766	Arrive	24
Disappear	49	Vanish	2
Fall	1406	Tumble	3
Fill	81	Line	27
Laugh	72	Giggle	3
Pour	76	Dribble	18
Shout	100	Whisper	28
Spill	86	Drip	21
Squash	148	Squeeze	66
Stay	467	Remain	2
Touch	142	Stroke	43
Mean	790.7	Mean	19.8

included to ensure that the children understood how they were expected to respond prior to hearing any of the test sentences containing argument structure errors. Each sentence was written on a card to allow the children to select the sentences in random order.

2.1.3. Procedure

The procedure used with the two groups of children was virtually identical. The children were introduced to two investigators and told that they were going to play a game. One investigator played the game with the child while the second investigator recorded the child's responses. The purpose of the game was to help a toy animal decide whether various sentences sounded grammatical or not.

The younger children were told that when they heard a sentence, if they thought it sounded okay they should move the animal to a green card showing a tick, but if they thought the sentence sounded a bit silly, they should move the animal to a card showing a red cross. The older children were simply asked to point to the red or green card. A forced choice method was adopted to increase the likelihood that the youngest children would understand the requirements of the task, and to minimize the performance demands on them.

The investigator then read the children the first practice sentence that was grammatically acceptable and asked them whether they thought the sentence sounded okay or silly. If the children judged the sentence correctly, the investigator moved on to the second practice sentence that was ungrammatical. If they judged the first sentence incorrectly, the investigator explained the task to them again and indicated the correct answer before moving on. This procedure was followed for the first two sentences only. The remaining 24 cards were then laid out in front of the children. They were asked to select cards for the investigator to read out one at a time, thus randomizing the order of presentation of the sentences. The children were not given feedback after the two practice sentences.

3. Results

To determine which children should be included in the analyses, the children's individual scores on the control sentences were examined. As all of the children responded to the two practice sentences, their judgments of these sentences were included in the results. Children who either judged four or more of the seven ungrammatical control sentences to be grammatical, or who judged four or more of the seven grammatical control sentences as ungrammatical were excluded from the analyses. These children were not able to rate unambiguous sentences according to their syntactic structure, and therefore we cannot draw any conclusions from their ratings of sentences containing argument structure errors. Ten children from the younger age group were excluded on these grounds. The remaining 5-year-olds judged a mean of 6.41, S.D. = 0.81, of the grammatical control sentences to be grammatical, and a mean of 0.80, S.D. = 0.98, of the ungrammatical sentences to be grammatical. None of the older children were excluded from the analyses. The 8-year-olds judged a mean of 6.89, S.D. = 0.31, of the grammatical control sentences to be grammatical, and a mean of 0.07, S.D. = 0.33, of the ungrammatical sentences to be grammatical. These data show that using the current methodology, both groups of children were capable of judging the grammaticality of individual sentences correctly.

3.1. Number of sentences in each group judged grammatical

First, the 5-year-olds responses were examined to determine whether there were differences between the sentences containing high and low frequency verbs in the number of sentences out of a possible six in each group that were judged to be grammatical by the children. A Wilcoxon test revealed that the children were significantly more likely to accept sentences containing low frequency verbs, $M = 2.86$, $S.D. = 1.55$, than sentences containing high frequency verbs, $M = 1.98$, $S.D. = 1.31$, $Z(48) = -4.65$, $P < .01$. The data from the 8-year-olds were then examined to determine whether a similar pattern of responses could be observed. A Wilcoxon test revealed that the 8-year-olds were significantly more likely to accept sentences containing low frequency verbs, $M = 1.81$, $S.D. = 1.20$, than those containing high frequency verbs, $M = 1.00$, $S.D. = 0.79$, $Z(54) = -4.47$, $P < .01$.

Both the 5- and the 8-year-olds were more likely to accept sentences containing low frequency verbs than those containing high frequency verbs. A further comparison was conducted to determine whether there were changes in the children's response patterns with age. Two Mann–Whitney tests were conducted to compare the number of sentences containing high and low frequency verbs that were accepted by the two groups of children. Results show that the 5-year-olds were significantly more likely to accept sentences containing both high and low frequency verbs than the 8-year-olds, sentences containing high frequency verbs $Z(103) = -3.99$, $P < .001$; sentences containing low frequency verbs $Z(103) = -3.43$, $P < .01$.

3.2. Number of children judging each sentence grammatical

The children's responses were then analyzed to determine whether there were differences in the number of children judging the sentences containing the low and high frequency verb in each pair as grammatical. A Wilcoxon test on the data from the 5-year-olds was significant showing that a larger number of the children judged sentences with low frequency verbs to be grammatical, $M = 23.33$, $S.D. = 9.85$, than judged sentences with high frequency verbs to be grammatical, $M = 16.00$, $S.D. = 9.47$, $Z(5) = -2.02$, $P < .05$. A second Wilcoxon test revealed the same pattern of results in the data from the 8-year-olds, M number of children accepting sentences with low frequency verbs = 16.67, $S.D. = 14.51$, M number of children accepting sentences with high frequency verbs = 9.17, $S.D. = 11.94$, $Z(5) = -2.21$, $P < .05$.

3.3. Individual verbs

The overall pattern in the children's data shows that the children tend to accept ungrammatical sentences containing low frequency verbs more than they accept ungrammatical sentences containing high frequency verbs, both with respect to the mean number of sentences containing low and high frequency verbs accepted by each child, and in terms of the mean number of children accepting each sentence with low and high frequency verbs. The children's data were then examined to determine whether their responses to each individual pair of sentences containing high and low frequency verbs followed the overall pattern.

Table 2

The proportion of 5- and 8-year-olds accepting the sentences containing each high and low frequency verb as grammatical

High frequency verbs	% children accepting sentence	Low frequency verbs	% children accepting sentence
5-year-olds			
Come	14.3	Arrive	46.9
Disappear	44.9	Vanish	59.2
Fall	22.4	Tumble	28.6
Laugh	18.4	Giggle	39.6
Pour	32.7	Dribble	30.6
Shout	65.3	Whisper	81.6
Mean	33.0		47.8
8-year-olds			
Come	1.8	Arrive	12.7
Disappear	27.3	Vanish	50.9
Fall	3.6	Tumble	18.2
Laugh	1.8	Giggle	9.1
Pour	3.6	Dribble	16.4
Shout	56.4	Whisper	74.5
Mean	15.8		30.3

Table 2 shows the proportion of the children in each age group who accepted the sentences containing each verb as grammatical.

The data in Table 2 reveal that for the 5-year-olds, five of the six pairs of verbs show the expected trend with a higher proportion of children accepting the sentences containing low frequency verbs than those containing high frequency verbs. For the verb pair *pour/dribble* a similar proportion of the children accept the sentences containing both the high and low frequency verb. For the 8-year-olds, the observed pattern is even stronger with all six pairs of verbs showing the expected trend.¹

3.4. The role of input frequency

To determine whether the input frequency of individual verbs determined the likelihood that the children would judge argument structure errors with these verbs to be grammatical, a Spearman's rank order correlation was calculated comparing the input frequency of each verb (in the Manchester corpus) with the number of children who judged the sentence including that verb to be grammatical. The correlation was not significant for either the 5- or the 8-year-olds, 5-year-olds $\rho(10) = -0.38$, $P < .05$; 8-year-olds $\rho(10) = -0.42$, $P < .05$, revealing that there is not a straightforward relation between the overall input fre-

¹ One possibility is that as the low frequency verbs tend to be two syllables whereas the high frequency verbs tend to be one syllable, the observed preference for sentences containing low frequency verbs may reflect increased processing demands. This explanation seems unlikely as the preference for low frequency verbs holds for the verb pair *disappear/vanish* where the high frequency verb is the longer of the two verbs, and the 5-year-olds do not show a strong preference for the low frequency verb *dribble* over the high frequency verb *pour*. Nevertheless, further research is needed to investigate the possible impact of word length.

quency of an individual verb and the likelihood that children will judge argument structure errors containing that verb to be grammatical.

However, scatterplots plotting log frequency in the Manchester corpus against the number of children accepting sentences containing individual verbs show that for both the 5- and the 8-year-olds, the two verbs *whisper* and *shout* are clear outliers that do not follow the otherwise apparently linear relation between verb frequency and judgments of grammaticality. Sentences containing these two verbs are judged to be grammatical somewhat more often than would be predicted on the basis of their frequency in the input. When these two verbs are excluded from the analysis, there is a significant rank order correlation between verb frequency and the number of children judging individual sentences to be grammatical for both the 5-year-olds, $\rho(8) = -0.67$, $P < .05$, and the 8-year-olds, $\rho(8) = -0.74$, $P < .05$.

4. Summary

The data show that the children's judgments of grammaticality were influenced by the frequency of the verbs modeled in ungrammatical argument structures, with respect to both their overall judgments of low versus high frequency verbs and their judgments of individual verb pairs. Moreover, the relative frequencies of use of the verbs in the input to children was related to their relative judgments of grammaticality, with larger numbers of children accepting sentences containing lower frequency verbs than higher frequency verbs. A second study was then carried out to investigate whether similar frequency effects could be observed in adults' grammaticality judgments of verb argument structure errors.

5. Study 2: adults' grammaticality judgments

5.1. Method

5.1.1. Participants

Thirty-six adults recruited predominantly from the student population at the University of Manchester participated in the study.

5.1.2. Materials

Twenty-four sentences were selected for use with the adults as follows (see [Appendix B](#)), plus a further three practice sentences.

- *Group 1*: 12 sentences containing argument structure errors with high frequency verbs.
- *Group 2*: 12 sentences containing argument structure errors with low frequency verbs matched to the verbs appearing in Group 1 for semantic class.

Six of the sentence pairs were identical to those used in Groups 1 and 2 for the children. A further six pairs of sentences were created by selecting additional examples of argument structure errors recorded in the literature and pairing each verb with a verb matched for semantic class but differing in its frequency as determined by frequency in the Manchester corpus.

Additional sentences were included for two reasons. First, the number of test sentences used with the children was limited by their attention span and the need to include control sentences to assess their ability to make grammaticality judgments. These limitations were not present to the same degree with the adults which allowed data to be collected for a larger number of verbs than was possible with the children. Second, the likelihood that the adults would recognize a relation between the sentences containing high and low frequency verbs matched for semantic class was reduced by including a larger number of sentences, as this meant that a larger number of sentences was presented between the high and low frequency sentences in each matched pair.

To ensure that the frequency of verb use in child-directed-speech (the Manchester corpus and CHILDES frequency database) was also representative of verb use in adult language, the frequency of the selected verbs in the Brown corpus of spoken English (Brown, 1984), and the Kucera–Francis corpus of written English (Kucera & Francis, 1967) was calculated. Pearson's correlations were then calculated between the four corpora. The six inter-corpus correlations were all above $r(22) = .98$, $P < .001$ for the verbs used in the study, thus the frequency of each verb in the Manchester corpus was used as a measure of relative frequency in adult language.

Before beginning the study, the adult participants were given three practice sentences to encourage them to rate sentences on a relative scale of grammatical–ungrammatical, rather than in absolute terms as either grammatical or ungrammatical. This was because the kinds of argument structure errors found in children's early speech and used in the current study are generally recognized to be ungrammatical by adults, but the purpose of the current study was to ascertain whether there was an effect of frequency on the relative ratings of grammaticality across children and adults. The first two practice sentences were unambiguously grammatical (sentence one: *The car is clean and shiny*) and ungrammatical (sentence two: *The cat the dog jumped*). The third sentence was selected as one that was ambiguous to help participants to understand the possibility of rating sentences on a relative scale (*He spun the girl a fairytale*).

To determine what this sentence should be, 18 undergraduate students were asked to rate 10 sentences selected from Goldberg (1995) that are indicated to be marginally acceptable, or acceptable to some speakers of English, on a scale of one (fully grammatical) to seven (completely ungrammatical). The sentence above was given a mean rating closest to the midpoint on the scale (a score of four) and was selected for use in the study.

5.1.3. Procedure

The adult participants were told that they would be read a series of sentences and their task was to indicate on a scale of one to seven how grammatically acceptable they found each sentence, where a score of one was grammatically acceptable, and a score of seven was ungrammatical. Adults were told that people tend to differ in their judgments of how acceptable some sentences are, and therefore there were no right or wrong judgments.

Before hearing the test sentences, adults were presented with three practice sentences designed to encourage them to use the interim points on the scale. One sentence was grammatical, one was ungrammatical, and the third was ambiguous. The investigator checked that the participants were using the scale correctly for the first two sentences, i.e., a low score for the grammatical sentence and a high score for the ungrammatical sentence. If the

participant rated the ambiguous sentence on an interim point on the scale, the investigator began the test sentences. If the participant rated the ambiguous sentence using an endpoint on the scale, the investigator explained that although there are no right or wrong responses, she would have rated that sentence somewhere between the previous two sentences because although it sounded a bit odd, she thought that it sounded better than the previous (ungrammatical) sentence, but not as good as the first (grammatical) sentence. The investigator then began the test sentences.

Each test sentence was read to the participant by the investigator (no written information was available) and the participant rated each sentence on the scale. One sentence from each pair was presented forming the first 12 sentences, followed by the second sentence from each pair in the same order forming the second 12 sentences. This was to maximize the number of sentences appearing between the members of each pair to avoid the adults making any direct comparisons between sentences containing matched verbs. The order in which high and low frequency verbs were presented was randomized across pairs of sentences.

6. Results

6.1. Mean rating for high and low frequency verbs

The first analysis examined whether the adults judged the sentences containing low frequency verbs as a group to be more grammatical than those containing high frequency verbs. On a scale of one to seven, higher scores indicate sentences judged to be less grammatical. A Wilcoxon test comparing the mean rating for high and low frequency verbs for each participant revealed that there were significant differences in the adults' mean ratings for sentences containing high, $M = 5.43$, $S.D. = 0.98$, and low, $M = 4.92$, $S.D. = 1.07$, frequency verbs, $Z(35) = -4.42$, $P < .01$, showing that overall the adults judged sentences containing low frequency verbs to be more grammatical than those containing high frequency verbs.

6.2. Mean rating for individual verbs

This analysis compared the mean rating given for each sentence containing a high frequency verb with the mean rating given for the corresponding low frequency verb sentence across participants. A Wilcoxon test comparing the mean rating for each high and low frequency verb across participants revealed that there was a marginally significant difference in the mean rating for the high frequency verbs, $M = 5.43$, $S.D. = 0.79$, and the low frequency verbs, $M = 4.92$, $S.D. = 0.99$, in each pair, $Z(11) = -1.88$, $P = .06$. This shows that the adults judged the sentences containing the low frequency verbs in each matched pair to be more grammatical than those containing high frequency verbs.

6.3. Individual verbs

Overall, the data showed that argument structure errors involving lower frequency verbs were judged as more grammatical than argument structure errors involving higher frequency

Table 3

The mean grammaticality rating given by adults to sentences containing each high and low frequency verb (1: grammatical, 7: ungrammatical)

High frequency verbs	Mean rating	Low frequency verbs	Mean rating
Brush	5.08	Shampoo	5.08
Come	6.25	Arrive	6.56
Disappear	5.92	Vanish	4.33
Fall	6.69	Tumble	5.67
Fill	5.06	Line	3.94
Laugh	5.75	Giggle	6.22
Pour	5.67	Dribble	5.11
Shout	4.78	Whisper	3.22
Spill	5.56	Drip	5.19
Squash	3.89	Squeeze	3.81
Stay	6.03	Remain	4.56
Touch	4.53	Stroke	5.40
Mean	5.43	Mean	4.92

verbs. The adults' data were then examined to determine whether this pattern could be observed for individual verb pairs. Table 3 shows the mean rating for the sentences containing each high and low frequency verb. Of the 12 verb pairs, 8 showed the predicted trend with the high frequency verbs being rated as less grammatical than their low frequency counterparts. One verb pair (*brush/shampoo*) was rated equally, and three verb pairs (*come/arrive*, *laugh/giggle*, and *touch/stroke*), showed a trend in favor of the higher frequency verb.

Interestingly, two of these three verb pairs (*come/arrive*, *laugh/giggle*) show the opposite trend to that observed in the children's data.

- The adults failed to show the expected preference for the high frequency verb *laugh* over the low frequency verb *giggle*. One possible explanation for this finding is that the adults were judging the sentences as having the meaning *Don't laugh/giggle at me* rather than the intended *Don't make me laugh/giggle*, and consequently judged the higher frequency verb *laugh* to be more acceptable in this construction than *giggle* because they have heard it used in this construction more often.
- The adults rate the ungrammatical sentence containing the high frequency verb *come* as more grammatical than the sentence containing the low frequency verb *arrive*. Given the vast difference between the frequencies of these two verbs (*come* is 282 times more frequent than *arrive*), this is particularly surprising. Although the children appear to base their judgments of these sentences on the frequency of the verb, the adults may be showing a greater sensitivity to the lexical preferences of each of these verbs. The test sentences took the form *She came/arrived me to school/the park*. However, one normally comes to a location, whereas one normally arrives at a location. The adults may have paid more attention to the relation between the verb and the preposition, and therefore preferred the sentences where the verb and preposition were compatible, even though the verb was of higher frequency.

To test whether these factors played a role in the adults' judgments of sentence grammaticality, 50 additional adults were asked to judge five pairs of sentences presented using a

similar methodology to that outlined earlier. One pair of sentences was included to determine whether the meanings of the original test sentences with *laugh* and *giggle* were misinterpreted (*That joke was so funny it really laughed/giggled me*). A second pair of sentences was included to determine whether the preposition affected adult's judgments of grammaticality (*Her Mum came her to school/Her Dad arrived her at school*). The results show that with these modifications, the adults showed the predicted preference for the sentences containing the low frequency verbs *giggle*, M rating = 4.04, and *arrive*, M rating = 5.28, over the high frequency verbs *laugh*, M rating = 6.02, and *come*, M rating = 6.28.

6.4. The role of verb frequency

To examine whether the frequency of individual verbs was related to relative judgments of grammaticality, a Pearson's correlation comparing verb frequency with the mean grammaticality rating for each verb was carried out. The correlation was not significant, $r(22) = .32$, $P > .05$, showing that there is not a clear relation between verb frequency and judgments of grammaticality.

7. Summary

The results reveal that as for the children, the adults' judgments of sentences containing argument structure errors were affected by the frequency of the verb, both in terms of their overall rating of sentences containing low versus high frequency verbs, and for individual verb pairs. However, in contrast with the children's data, there was not a clear relation between the relative frequencies of individual verbs and relative ratings of grammaticality in the adult data. This could reflect a complex interaction between the overall frequency of use of individual verbs and their frequency of use in a range of different constructions that are known to the adult participants. These factors may have influenced grammaticality judgments but were not considered in the present study. Alternatively, the different methodologies used with the adults and children (forced choice vs. a 7-point scale) may have influenced the results. In any case, the overall results demonstrate that both the adults and the children show a greater acceptance of sentences containing verb argument structure errors if the verbs are of relatively low frequency in the language in comparison with errors involving high frequency verbs.

There are, however, a few instances where the pattern of results goes against the general trend. In addition to the differences between the children's and adults' judgments of the verb pairs *come/arrive* and *laugh/giggle* discussed above, judgments on the following sentence pairs may cast additional light on how adults and children process sentence structure:

- The trend for accepting argument structure errors with low frequency verbs more than with high frequency verbs is observed for the verb pair *shout/whisper*. However, all groups of participants showed a higher than expected tendency to accept ungrammatical sentences with these verbs, given their frequency in the input. The children's tendency to accept these sentences may reflect a weaker knowledge of more complex three-argument constructions in comparison with the two-argument SVO construction modeled in other

argument structure errors. However, a lack of familiarity with complex constructions cannot explain the similar pattern in the adults' responses. One possible explanation is that for both of these verbs, the ungrammatical double-object dative *She shouted/whispered her something* is very close to the grammatical prepositional dative *She shouted/whispered something to her* both in lexical content and in length. This is in contrast with argument structure errors where intransitive verbs are used in a simple transitive SVO construction, for example, *She disappeared it* where the correct form would require the use of the two-verb periphrastic causative construction, for example *She made it disappear*, or a change of main verb, for example, *She hid it*. Thus, the extent to which a potentially preempting construction is similar to the ungrammatical form may influence participants' judgments of grammaticality.

- The 5-year-olds fail to show the frequency distinction between the verbs *pour* and *dribble* observed in the data from the 8-year-olds and the adults. It is not clear why this is the case. These verbs take three arguments, but they appear in the locative rather than the dative construction. Thus, it is necessary to both rearrange the object and location, and replace the preposition *with* with an appropriate preposition (e.g., *I poured you with water/I poured water over you*) to produce a grammatical construction. The 5-year-olds may be showing a general lack of knowledge of the construction, a sensitivity to the fact that the preposition *with* is incorrect without recognizing the relation to its grammatical counterpart, or a lack of understanding of the specific sentences modeled. In any case, it is clear that understanding this developmental change requires a detailed understanding of how children build up knowledge about many different aspects of verb argument structure.

Although further research is needed to examine these possibilities, the similarity between the ungrammatical sentences and a grammatical alternative, the complex relation between the individual lexical items in a given sentence, and the perceived meaning of each sentence may interact with an entrenchment effect based on verb frequency. This complex interaction of factors may explain why there is not a direct relation between the frequency of individual verbs and adults' perception of those verbs as grammatical in an ungrammatical structure.²

8. Discussion

The study examined whether frequency-based entrenchment affected children's and adults' judgments of verb argument structure errors to determine whether the entrenchment effects previously observed in young children's production of verb argument structure errors might also be present in adult language processing. The results provide support for the role of entrenchment in constraining argument structure overgeneralizations in children and in adults. The 5-year-olds, 8-year-olds, and the adults were more likely to accept argument structure errors containing low frequency verbs than those containing high frequency

² It is also possible that the measures of verb frequency used in the current study may not accurately reflect the experience of individual speakers, or the frequency of some of the verbs in everyday spoken language for adults. Alternatively, it may be the case that for some verbs it is the frequency of that verb in a particular construction rather than its overall frequency that determines judgments of grammaticality. Further research is needed to disentangle the relative influence of these factors.

verbs matched to the low frequency verbs for semantic class. These findings provide further support for the experimental findings of [Brooks et al. \(1999\)](#), and extend the process of entrenchment to adult language.

The apparent influence of verb frequency on adults' judgments of sentence grammaticality is especially interesting given the verbs the adults were asked to judge. The verbs were selected on the basis that they were present in a corpus of language addressed to 2-year-olds, and therefore they should be very well known to adults. Thus, these verbs would not be considered marginal or ambiguous in the adult grammar with respect to their argument structure or indeed their semantic class ([Pinker, 1989](#)). Moreover, adults are not widely thought to produce these types of errors often in their own speech. Nevertheless, the adults continued to show a significant preference for argument structure errors with low frequency verbs in comparison with high frequency verbs when asked to rate sentences with respect to their grammaticality.

This suggests that first, knowing the semantic class to which a verb belongs, even for adult speakers, is not the only process involved in determining the argument structures in which that verb may appear. There is clearly an interaction between verb frequency and the degree to which a given verb's argument structure is governed by the semantic class to which it belongs. This suggests that narrowly defined semantic classes may be organized around prototypes based on the frequency of use of individual verbs rather than reflecting absolute membership criteria.

Second, [Brooks et al. \(1999\)](#) suggest that the processes of entrenchment and preemption may be confounded when testing young children with known verbs, because verbs that are early acquired are also likely to be verbs that children have learned in a number of different structures. They may therefore have preempting structures available for high frequency verbs, but not for low frequency verbs. However, the current study suggests that entrenchment operates independently of preemption to some extent. The adult participants in the current study would be expected to have preempting structures for both the high and low frequency verbs because all of the verbs are well known to adult speakers. However, the assumed availability of preempting structures did not remove the frequency effect observed in the participants judgments of sentence grammaticality.

Third, the use of a grammaticality judgment methodology may uncover differences in adult linguistic representations that are not immediately evident from their language use. The lack of argument structure errors in adults' speech with familiar verbs of differing frequencies could be taken as evidence that frequency does not affect the extent to which individual verbs are found acceptable in different argument structure constructions. However, adults are expected to have strongly entrenched preempting structures available, and are therefore unlikely to make errors in production. Nevertheless, it may be that although adults tend not to make the argument structure errors observed in children's speech, when asked to make fine-grained judgments of their acceptability, differences in their representations as a function of their frequency of use are uncovered.

The results of the current study provide further evidence for a developmental increase in the role of entrenchment in constraining argument structure overgeneralization errors. Although the 5-year-olds showed sensitivity to the frequency of individual verbs in the input, they were significantly more likely than the 8-year-olds to accept sentences containing argument structure errors with both high and low frequency verbs. This suggests that

although the 5-year-olds have begun to build up differential knowledge about the behavior of individual verbs that reflects their frequency in the input, their knowledge of the verb argument structures associated with even the high frequency verbs is somewhat less developed than it is for the 8-year-olds. This is assumed to reflect the fact that the 8-year-olds have had considerably more experience with the target language, and have consequently built up a greater knowledge of the behavior of individual verbs based on their frequency of use in the input.

However, the current study also reveals that it is not simply the frequency of use of individual verbs that affects children's and adults' judgments of sentence grammaticality. As previous studies have shown, preemption and the semantic class to which individual verbs belong also play a role in constraining argument structure errors. The current study suggests that in addition to these factors, researchers need to pay attention to two other aspects of verb argument structure.

First, the differing judgments between the adults and children of the sentences containing the verbs *come* and *arrive* suggest that even when semantic class is controlled, the verb's choice of preposition may have a stronger influence on adults' perceptions of grammaticality than the verb's frequency. Regardless of whether it is an understanding of the relation between verbs and their prepositions or some other aspect of the verbs' argument structure, there is apparently some development between 8 years of age and adulthood that affects judgments of grammaticality with these verbs. Similar developments in the understanding of the locative construction between the ages of 5 and 8 may explain the lack of a frequency effect in the 5-year-olds' data for the verbs *pour* and *dribble*.

Second, the children were more accepting of argument structure errors with the verbs *shout* and *whisper* than would be predicted on the basis of their frequency. These verbs were modeled in a double-object dative construction, where the grammatical counterpart is the prepositional dative, which suggests that it may take children longer to learn which verbs can be used in the ditransitive construction in contrast with the simple transitive construction modeled in some of the other test sentences. In addition, the prepositional and double-object dative constructions share more similarities in terms of length and lexical content than the transitive and periphrastic causative constructions. This suggests that the overall similarity in lexical content and length between a sentence containing an argument structure error and its grammatical counterpart may contribute to the children's judgments of grammaticality.

Thus, it may be important to examine not only whether there is a preempting structure available to children and adults, but also the extent to which that structure differs from the ungrammatical alternative. The verb argument structure errors modeled in the current study represent a variety of different types of errors. Most of these errors have direct counterparts with the same verb in the forms of the periphrastic causative, prepositional dative, or locative alternation, but others require a verb change (e.g., *She came me to school* requires the verb *brought* to render the utterance grammatical), or have no straightforward alternative (e.g., *How do you unsquash it?*). Further research is needed to tease apart the precise influence these factors have on children's ability to constrain their argument structure overgeneralizations.

In conclusion, the current study adds to a body of evidence that suggests that children's sensitivity to the distributional properties of the input may provide an important part of the explanation for their ability to generalize verb argument structures and also for their retreat

from verb argument structure overgeneralizations (Brooks & Tomasello, 1999; Brooks et al., 1999; Brooks & Zizak, 2002; Tomasello, 2000). Moreover, the same sensitivity to the relative frequencies of use of individual verbs in individual constructions appears to be present in adult language processing.

This provides support for a usage-based constructivist approach in which both children's early knowledge of grammar and the adult end-state grammar are thought to reflect the relative frequency of use of individual lexical items and constructions. From this perspective, it is the gradual building up of type and token frequencies that leads to abstraction beyond individual lexically-based constructions, but that also ultimately constrains productivity. Although the precise mechanisms involved in this process are as yet not well understood, further detailed investigation of the interaction between children's developing knowledge of verb argument structure and the availability of different types of preempting constructions in the input should provide a useful avenue for further research.

Acknowledgments

I would like to thank the schools who took part in the research reported here.

Appendix A

Test sentences: children

High frequency verbs	Low frequency verbs
She came me to school	She arrived her to the park
I'm gonna disappear it	I'm gonna vanish it
Somebody fell it off	Somebody tumbled it off
Don't laugh me	Don't giggle me
I poured you with water	I dribbled teddy with water
Shall I shout you something?	Shall I whisper you something?

Appendix B

Test sentences: adults (as for children plus)

High frequency verbs	Low frequency verbs
I'll brush him his hair	I'll shampoo him his hair
Can I fill some salt into there?	Can I line some paper into the box?
I spilled the carpet with juice	I dripped the table with milk
How do you unsquash it?	How do you unsqueeze it?
Can you stay this open?	Can you remain this open?
Touch your hand to that	Stroke your hand to that

References

- Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.
- Bock, J. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Bowerman, M. (1978). Systemizing semantic knowledge: Changes over time in the child's organization of word meaning. *Child Development*, 49, 977–987.
- Bowerman, M. (1982). Evaluating competing linguistic models with language acquisition data: Implications of developmental errors with causative verbs. *Quaderni di Semantica*, 3, 5–66.
- Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overgeneral grammar? In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 73–101). Oxford: Blackwells.
- Braine, M. (1971). On two types of models of the internalisation of grammars. In D. I. Slobin (Ed.), *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.
- Braine, M., & Brooks, P. (1995). Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs*. Hillsdale, NJ: Erlbaum.
- Brooks, P., & Tomasello, M. (1999). How children constrain their argument structure constructions. *Language*, 75, 720–738.
- Brooks, P., Tomasello, M., Dodson, K., & Lewis, L. (1999). Young children's overgeneralizations with fixed transitivity verbs. *Child Development*, 70, 1325–1337.
- Brooks, P., & Zizak, O. (2002). Does pre-emption help children learn verb transitivity? *Journal of Child Language*, 29, 759–781.
- Brown, G. (1984). A frequency count of 190,000 words in the London-Lund corpus of English conversation. *Behavioural Research Methods Instrumentation and Computers*, 16, 502–532.
- Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 1–33). Hillsdale, NJ: Erlbaum.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425–455.
- Bybee, J. (1998). The emergent lexicon. In M. C. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *Papers from the 34th annual meetings of the Chicago Linguistics Society, Part 2: The panels* (pp. 421–439). Chicago: Chicago Linguistics Society.
- Clark, E. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Ellis, N., & Schmidt, R. (1998). Rules of associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes*, 13, 307–336.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: The University of Chicago Press.
- Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, 16, 173–184.
- Hochberg, J. (1986). Children's judgments of transitivity errors. *Journal of Child Language*, 13, 317–334.
- Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence: Brown University Press.
- Kuiper, K. (1996). *Smooth talkers: The linguistic performance of auctioneers and sportscasters*. Hillsdale, NJ: Erlbaum.
- Levin, B. (1993). *English verb classes and alternations*. Chicago: University of Chicago Press.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: Vol. 2. The database*. London: Lawrence Erlbaum Associates.
- Marslen-Wilson, W. (1990). Activation, competition, and frequency in lexical access. In G. Altmann (Ed.), *Cognitive models of speech processing* (pp. 148–172). Cambridge, MA: MIT Press.
- Mazurkewich, I., & White, L. (1984). The acquisition of the dative alternation: Unlearning overgeneralizations. *Cognition*, 16, 261–283.
- Naigles, L., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language*, 25, 95–120.

- Pinker, S. (1989). *Learnability and cognition: The acquisition of verb-argument structure*. Cambridge, MA: Harvard University Press.
- Randall, J. (1990). Catapults and pendulums: The mechanisms of language acquisition. *Linguistics*, 28, 1381–1406.
- Rowland, C. F., & Pine, J. M. (2000). Subject-auxiliary inversion errors and wh-question acquisition: 'What children do know?'. *Journal of Child Language*, 27, 157–181.
- Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, 28, 127–152.
- Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2002). Going going gone: The acquisition of the verb 'Go'. *Journal of Child Language*, 29, 783–811.
- Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74, 209–253.
- Trueswell, J. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, 35, 566–585.
- Van Valin, R. (2002). The development of subject-auxiliary inversion in English wh-questions: An alternative analysis. *Journal of Child Language*, 29, 161–175.