Available online at www.sciencedirect.com
ScienceDirect
Cognition 106 (2008) 87–129
COGNITION
www.elsevier.com/locate/COGNIT

The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors ☆,☆☆

Ben Ambridge *, Julian M. Pine, Caroline F. Rowland, Chris R. Young

The University of Liverpool, School of Psychology, Eleanor Rathbone Building, Bedford Street South, Liverpool L697ZA, UK

Received 23 February 2006; revised 18 December 2006; accepted 22 December 2006

Abstract

Participants (aged 5–6 yrs, 9–10 yrs and adults) rated (using a five-point scale) grammatical (intransitive) and overgeneralized (transitive causative)[1] uses of a high frequency, low frequency and novel intransitive verb from each of three semantic classes [Pinker, S. (1989a). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press]: "directed motion" (*fall, tumble*), "going out of existence" (*disappear, vanish*) and "semivoluntary expression of emotion" (*laugh*, *giggle*). In support of Pinker's semantic verb class hypothesis, participants' preference for grammatical over overgeneralized uses of novel (and English)

☆ This manuscript was accepted under the editorship of Jacques Mehler.
☆☆ This research was funded by an ESRC Postdoctoral Fellowship award to Ben Ambridge (PTA-026-27-0705) and an ESRC Grant to Ben Ambridge, Julian Pine and Caroline Rowland (RES-000-22-1540).
* Corresponding author. Tel.: +44 151 794 1111.
E-mail address: Ben.Ambridge@Liverpool.ac.uk (B. Ambridge).

[1] Throughout this paper, the term "intransitive" (whether referring to a verb or construction) refers only to non-causative intransitives – sometimes termed "inchoative intransitives" – (e.g., *The man laughed*) and not to intransitives with unspecified or unexpressed objects (e.g., *The man ate*). The term "transitive causative" (e.g., *The sun melted the snow*) is used to contrast sentences of this type with both transitive non-causatives (e.g., *John saw Bill*) and periphrastic causatives (e.g., *The sun made the snow melt*).

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2006.12.015

verbs increased between 5–6 yrs and 9–10 yrs, and was greatest for the latter class, which is associated with the lowest degree of direct external causation (the prototypical meaning of the transitive causative construction). In support of Braine and Brooks's [Braine, M.D.S., & Brooks, P.J. (1995). Verb argument strucure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 352–376). Hillsdale, NJ: Erlbaum] entrenchment hypothesis, all participants showed the greatest preference for grammatical over ungrammatical uses of high frequency verbs, with this preference smaller for low frequency verbs, and smaller again for novel verbs. We conclude that both the formation of semantic verb classes and entrenchment play a role in children's retreat from argument-structure overgeneralization errors.

Keywords: No-negative-evidence problem; Overgeneralization errors; Entrenchment; Semantic verb classes; Grammaticality judgements; Verb argument-structure; Unaccusativity; Causativity; Transitivity; Syntax

1. Introduction

The *no-negative-evidence* problem has long been recognized as a central issue in language acquisition research (Bowerman, 1988). In order to produce novel utterances, children must use particular lexical items (such as nouns and verbs) in sentence structures in which these items have not appeared in the input data. However, children sometimes overgeneralize, using lexical items in structures in which they are not permitted in the adult grammar (e.g., **Do you want to see us disappear our heads?*, from Bowerman, 1988). Given that children do not appear to receive feedback about which of their utterances are ungrammatical (*negative evidence*), it is unclear how children learn not to produce such errors.

Many overgeneralization errors, for example past-tense overregularization (e.g., *runned*) are relatively non-problematic. As the child acquires the adult form (e.g., *ran*) she will cease to use the overregularized form, via a process such as blocking (Marcus, 1993; Marcus et al., 1992) competition (Bates & MacWhinney, 1987; Macwhinney, 1987) or pre-emption (Braine & Brooks, 1995; Clark, 1987). More problematic for such accounts are *argument-structure overgeneralizations*. These occur when the child uses a particular verb (e.g., the intransitive verb *disappear*) in an argument-structure construction (e.g., the transitive causative construction [SUBJECT][VERB][OBJECT]) in which it is not licensed in the adult grammar (e.g., **The magician disappeared the rabbit*). Such overgeneralizations are problematic, because there is no direct competitor in the input which could block the incorrect use. Many authors (e.g., Clark, 1987; Goldberg, 1995; Macwhinney, 2004) have argued that somewhat indirect competitors (e.g., in this case, a periphrastic causative such as *make disappear*) are sufficient to block the overgeneralized form. If this is the case, however, it is unclear how the child could learn that certain periphrastic causatives [e.g., *John made the baby stand up* (e.g., through giving an order)] do not block the corresponding transitive causative sentence [e.g., *John stood the baby up* (e.g., by propping it up against a wall); examples from Bowerman, 1988].

Some authors have proposed that children do receive corrective feedback from parents, in the form of recasts, requests for clarification and misunderstandings.

Whilst it seems likely that such feedback aids the recovery from overgeneralization errors (Farrar, 1992; Saxton, 2000), it is far from clear that such feedback is necessary or sufficient for the restriction process to occur (Bowerman, 1988; Marcus, 1993; Pinker, 1989a). Many parents, particularly those in non-Western cultures, do not provide such feedback (Lieven, 1994), and it seems unlikely that adult speakers would accept as grammatical any generalization that they had not produced, with a subsequent parental correction, during childhood.

The goal of the present study was to investigate two proposals (which are not necessarily mutually exclusive) for how children learn to appropriately restrict their argument-structure generalizations: the semantic verb class and entrenchment hypotheses. Although both proposals are potentially applicable to all argument-structure overgeneralization errors, we restricted our initial investigation to errors where an intransitive verb (e.g., *disappear*) is incorrectly used in a transitive causative construction (e.g., **The magician disappeared the rabbit*). Such errors are termed "causative overgeneralization errors" (or "causativization errors") as they involve the over-general application of a pattern, rule or alternation (sometimes termed the "causative alternation") that allows some verbs that appear in the intransitive construction (e.g., *The stick broke*) to appear also in the transitive construction (e.g., *The man broke the stick*). We do not consider here the converse error, where a transitive verb (e.g., *lose*) is incorrectly used in an intransitive construction (e.g., **I better put it down there so it won't lose*; Lord, 1979); an overgeneralization of what is sometimes called the "anticausative alternation".

2. The semantic verb class hypothesis

Under the *semantic verb class* hypothesis (Pinker, 1989a: 223) the causative (and anticausative) alternation is governed by a "broad-range lexical rule" which "allows a verb that specifies an event involving a thing to be embedded as an effect of an agent acting on that thing". A diagrammatic representation of this rule is reproduced in Fig. 1. The top part of the figure represents the non-causative event (e.g., *snow melting*), with the node labelled <+dynamic> representing the event itself (e.g., *melting*), and the node labelled THING representing the entity involved in that event (e.g., *snow*). The bottom part of the figure represents the causative event (e.g., *the sun causing snow to melt*). The first EVENT consists of an agentive THING [ ] (e.g., *the sun*) that ACTs upon the THING[Y] (e.g., *the snow*), causing the second EVENT: for this example *the snow* (THING Y) *melting* (<+dynamic>).

This broad-range rule includes the specification that "the predicate of the effect event can be either GO or ACT [i.e., dynamic, but not non-dynamic BE or HAVE]" (p. 223). Note that this rule transforms one lexicosemantic (NOT syntactic) structure into another (Pinker, 1989a: 63; c.f., Pinker, 1984). Innately specified linking rules spell out the syntactic structure for a given semantic structure. For example, consider the case where a transitive sentence (e.g., *The sun melted the snow*) is produced using the causative lexicosemantic structure shown in the second part of Fig. 1. The linking rules specify that the first ("agent") argument of the predicate be syntactically

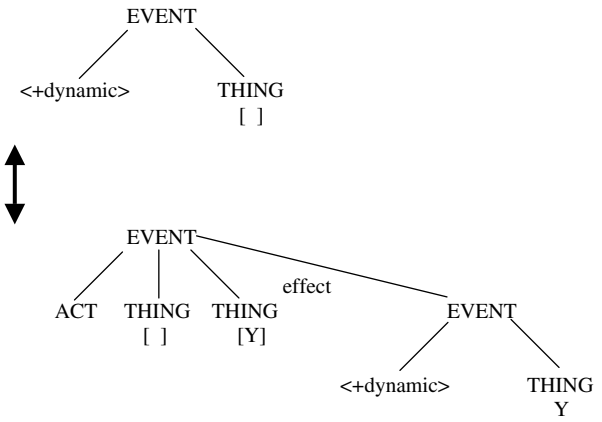

Fig. 1. Pinker's broad-range rule for the causative alternation (reproduced from Pinker, 1989a Learnability and cognition: The acquisition of argument structure, MIT Press, 223 Figs. 5.54). Note that this rule operates on lexicosemantic (as opposed to syntactic) structures.

realised as the subject (*the sun*), the predicate ACT itself as the verb (*melt*) and its second (''patient'') argument as the object (*the snow*).

Broad-range lexical rules are acquired via a two-stage process. First, the child applies these same ''universal linking rules in a backward direction'' (p. 258) to acquire lexicosemantic structures. For example, once a sentence has been recognized as having the syntax SUBJECT VERB OBJECT, these rules will spell out the lexicosemantic structure shown in 1b. A sentence of the form SUBJECT VERB will output the structure shown in 1a. Second, the child notices that some verbs occur in more than one lexicosemantic structure (and associated argument-structure) and so sets up a productive broad-range rule that derives one from another (e.g., Fig. 1).

At the same time as they are using this broad-range rule to generate novel causatives, children are forming semantically-based ''narrow conflation classes'' (p. 297) of verbs that undergo the ''possible regularity stated in [the] broad-range rule'' (p. 274). The child forms a narrow-range semantic class of alternating verbs by noticing that a particular verb alternates and extending this property to other verbs with the same ''grammatically relevant semantic structure'' (p. 274). As these narrow-range classes are formed, children stop applying the broad-range rule to verbs that are not members of such a class. For the causative alternation, some alternating classes are verbs denoting ''extrinsic change of physical state'' (e.g., *open*, *melt*, *shrink*) and ''contained motion taking place in a particular manner'' (e.g., *roll*, *bounce*, *skid*; p. 130).[2]

[2] Pinker (1989a, 131–133) also discusses some non-causativizable verb classes such as verbs of ''volitional or internally-caused actions'' (e.g., *eat*, *drink*, *run*), ''coming into or going out of existence'' (e.g., *disappear*, *vanish*, *disintegrate*), ''emotional expression'' (e.g., *laugh*, *cry*, *frown*), ''emission'' (e.g., *glow*, *buzz*, *leak*) and ''motion in a specified direction'' (*come*, *go*, *ascend*; Pinker, 1989b, 50). Presumably, however, Pinker does not wish to imply that the child actually forms these classes of nonalternating verbs as the very motivation for the account is the assumption that the child does not receive the evidence (i.e., evidence that some verbs do not alternate: ''negative evidence'') that would be necessary for these classes to be formed.

Under Pinker's (1989a: 350–351) account, children's overgeneralization errors are either "one-shot innovations based on broad-range rules" (which "have the same status as adult innovations" and thus "requir[e] no specific unlearning") or are caused by incorrect semantic representations of particular verbs, which cause them to be incorrectly assigned to a causative class. Errors of this latter type will "disappear as an automatic consequence of the fine-tuning of the verbs' semantic representations".

3. Semantic verb class hypothesis. Prediction 1: Broad-range rule vs narrow classes

Pinker's (1989a) account, therefore, predicts that young children who have yet to form accurate narrow-range semantic verb classes will accept as grammatical in a judgement task (and indeed will produce) causative sentences with verbs whose semantics are consistent with the broad-range rule (i.e., that denote dynamic events), even when they are not consistent with adult narrow-range semantic verb classes. Older children and adults, who have formed accurate narrow-range semantic classes, are predicted not to accept (or produce) causative sentences with verbs that are not members of an alternating narrow-range semantic class (or that, descriptively speaking, are members of a non-alternating class – see Footnote 2). One way to test this prediction is to obtain ratings for causative overgeneralizations with real English verbs that fit this profile such as *fall, disappear* and *laugh*. Although this is one approach adopted in the present study, it has two shortcomings: First, for real verbs, semantics are confounded with other factors such as frequency of occurrence and phonology. Second, this prediction does not distinguish the semantic verb class account from other accounts that also predict an age effect on ratings of acceptability for overgeneralization errors with real verbs (such as the entrenchment hypothesis, discussed in the next section).

A better approach is to create novel verbs with novel actions whose semantics are such that they are potentially causativizable under the broad-range rule (i.e., they are dynamic events) but that are not consistent with the semantics of any English causativizable class (but are consistent with the semantics of a non-causativizable class). Younger children would be predicted to accept (and produce) transitive causative uses of such verbs, whilst older children and adults, who are argued to have formed narrow-range semantic classes of causativizable verbs, would not.

An experiment conducted by Brooks and Tomasello (1999) provides some support for this prediction. Children were taught two novel verbs, presented in intransitive sentences only (e.g., *The ball meeked*). One verb was assigned a meaning consistent with the English narrow-range semantic class of verbs of "contained motion taking place in a particular manner" (e.g., *roll*, *bounce*; Pinker, 1989a: 130), which are causativizable under Pinker's (1989a) account. The other was assigned a meaning consistent with the (descriptive) non-alternating class of "verbs of motion in a specified direction (*come*, *go*, *ascend*; Pinker, 1989b: 50). Meaning was assigned by having the children watch the experimenter act out the sentences with puppets. In a test phase, where their task was to describe the experimenter's actions,

children aged 4;6 and above (though not a younger group aged 2;6) produced more utterances in which the novel verb was used in a transitive causative construction (e.g., *The man meeked the sock*) for the semantically-causativizable novel "manner" verb than the non-causativizable "direction" verb.

Similar effects have been found in comprehension studies (Lidz, Gleitman, & Gleitman, 2004; Naigles, Fowler, & Helm, 1992; Naigles, Gleitman, & Gleitman, 1993). When asked to act out ungrammatical sentences such as *The zebra goes the lion*, children aged 4 and under generally act out a causative ("frame compliant") reading (for this example, the zebra causing the lion to go), consistent with the application of Pinker's (1989a) broad-range rule for the causative alternation. Older children and adults were more likely to interpret the sentence as an omission error, acting out, for this example, the zebra going TO the lion (a "verb compliant" reading). This is consistent with the view that older children and adults had formed narrow range classes of potentially causativizable verbs, none of which included the verb *go*.

In summary, then, the present study tests the prediction of Pinker's (1989a) semantic verb class hypothesis that participants will respect the experimentally-assigned semantic class of a novel (here, intransitive) verb (and the semantic class of English verbs) when rating intransitive and transitive causative verb-uses, this effect increasing with age as precise classes are formed.

4. Semantic verb class hypothesis. Prediction 2: The role of direct causation

Pinker's theory (1989a) also makes a second prediction. In common with many other authors (e.g., Croft, 1991; Haspelmath, 1993; Levin & Rappaport Hovav, 1995; Marcotte, 2005; Smith, 1970; Wolff, 2003) Pinker (1989a) notes that, in English, a transitive causative sentence can be used only to denote events of direct external causation; that is, events where the subject directly causes the object to perform the action described by the verb. Prototypically, a transitive causative sentence describes an event where an animate agent deliberately causes an inanimate patient to perform an action that could not have occurred spontaneously, by means of direct physical contact. A sentence such as *The man washed the clothes* denotes an event that requires the direct physical contact of a causer-agent external to the patient, and as such requires the use of the transitive causative construction: Neither a periphrastic causative (**The man /made the clothes wash*) nor an intransitive sentence (**The clothes washed*) can be used to describe such an event.[3]

[3] It is important to note that it is the semantics of the *event*, not of the verb, that determine whether or not a transitive causative construction may be used. For example, when the verb *wash* describes an event of direct (unmediated) physical causation, the transitive construction is required. However, when the same verb is used to describe an event involving indirect (mediated) causation, this construction cannot be used. For example, *The man washed the clothes* is not a felicitous descriptions of an event where the man paid a drycleaner to wash the clothes (c.f., *The man had the clothes washed*).

On the other hand, events that are not caused by an external causer cannot be described using a transitive causative construction. The prototype of an event that is *not* caused by an external causer is an event that is caused entirely by forces internal to the entity performing the event. For example, events such as *swimming* cannot be described by transitive sentences (**The woman swam her child*) as it is not possible for an external entity to cause this event directly by means of physical contact (a periphrastic causative or intransitive is required; e.g., *The woman made her child swim*; *The child swam*). Thus it seems that the use of the transitive causative construction is prohibited for such events because the semantics of the event are too inconsistent with the semantics associated with this construction.

Events of inherent direct external causation (e.g., *wash*), which require the use of a transitive causative construction, and events for which no direct external causation is possible (e.g., *swim*), and so prohibit the use of this construction, can be thought of as forming two ends of a continuum. Events that can be construed as caused partly due to the action of an external causer and partly due to actions or properties internal to the entity undergoing the event (e.g., *break*, *melt*, *cheer up*, *roll*) can, in English, be used in either the transitive causative or the intransitive construction.[4] Indeed, a speaker may specifically choose to use one of other of these constructions in order to emphasise either the external or internal nature of the causation. For example, a speaker might choose to say *The photocopier broke* instead of *I broke the photocopier* in order to imply that the breakage was caused by the malfunction of mechanisms internal to the photocopier, and not by the actions of an external causer (i.e., himself).

This continuum corresponds to a cross-linguistic hierarchy of lexicalization (Croft, 1991; Haspelmath, 1993; Nedjalkov, 1969). For example, in a survey of 60 languages, Nedjalkov (1969) found no language in which the causative form of *laugh* (a semivoluntary, internally-caused action) was unmarked, with the non-causative form derived from it. At the other end of the hierarchy, the non-causative form of *split* (typically seen as requiring an external cause) was typologically unmarked in only one of 31 languages surveyed by Haspelmath (1993) (English is one of only three languages in which the alternation is non-directed). Verbs such as *burn* and

[4] The way in which speakers represent the knowledge that certain verbs can appear in both constructions is a matter of some debate. Under Pinker's (1989a) account, speakers store different semantic representations of a particular verb stem (e.g., causative and non-causative *melt*). In some cases the two representations will have been independently acquired from input sentences, using "universal linking rules in a backward direction" (p. 258). In others, one will have been acquired from the input in this manner, and the second generated from the first using a lexical rule (e.g., Fig. 1). Universal linking rules are then used to spell-out the syntax for a given semantic structure (e.g., causative *melt* requires the transitive, non-causative *melt* the intransitive). However, Goldberg (1995, 9) rejects this approach, arguing that it requires positing implausible verb senses. For example, for Pinker's (1989a) linking rules to spell out a the syntax of a sentence such as *She baked him a cake*, one would have to claim that speakers have a semantic representation of *bake* "something like 'X INTENDS to CAUSE Y to HAVE Z'". Under Goldberg's approach, this meaning instead would be attributed to the construction (here the double-object dative). Verbs take on part of the meaning of the construction in which they are used (e.g., *melt* has a causative meaning when it is used in a transitive causative construction).

dissolve, which can be construed as either internally or externally caused (and can appear in both transitive and intransitive constructions in English) are approximately equally split between having a basic non-causative and a basic causative form cross-linguistically.

There are also some parallels to be found between such a continuum and claims made with respect to the *unaccusative hypothesis* (Perlmutter, 1978). This hypothesis claims that "there are two classes of intransitive verbs, the *unaccusative* verbs [also called *ergative* verbs] and the *unergative* verbs" (Levin & Rappaport Hovav, 1995: 3). Strictly speaking, unaccusativity is a syntactic notion whereby "the single argument of an unaccusative verb is...equivalent to the direct object of a transitive verb, whereas the single argument of an unergative verb is...equivalent to the subject of a transitive verb" (Sorace, 2000: 879).[5] Although intransitive sentences containing unergative and unaccusative verbs have identical surface forms (they are said to have identical "S-structure", e.g., Chomsky, 1981), the unaccusative hypothesis claims that they have a different underlying structure "D-structure". Unaccusative verbs (e.g., *die*) have an object at D-structure which moves into subject position at S-structure, because in English, all verbs require a subject. For unergative verbs (e.g., *dance*) no such movement is required, as a subject is already present at D-structure, as illustrated below (adapted from an example given in Randall, van Hout, Weissenborn & Baayen, 2004: 333)

	D-structure		S-structure
Unaccusative:	[ ] [$_{VP}$ died [$_{NP}$ she]]	→	[$_{NP}$ she]$_i$ [$_{VP}$ died [t]$_i$]
Unergative:	[$_{NP}$ she] [$_{VP}$ danced]	→	[$_{NP}$ she] [$_{VP}$ danced]

Under this account one might expect that unaccusative verbs might be coerced into the transitive causative construction more easily than unergative verbs because the former, but not the latter, have an empty subject position into which the causer could be inserted. This account would not predict such a difference between two unaccusative (or two unergative) verbs, however.

[5] In fact, the classification of intransitive verbs into unaccusatives and unergatives is problematic, with many diagnostics giving different definitions (e.g., Legendre, 1989). The original diagnostic proposed was that languages (e.g., Dutch) that form impersonal passives may do so only with unergatives.

Dutch examples from Perlmutter (1978: 162) Unergative: *Er wordt hier door de jonge lui veel gedanst* [There is a lot of dancing by the young people (going on) here] Unaccusative: **Er wort in dit ziekenhuis door de patienten dikwijls gestorven* [*There is a lot of dying by the patients (going on) in this hospital].

However, this diagnostic has been criticised for providing classifications that are at odds with most other diagnostics (e.g., Zaenen, 1993). Another proposed diagnostic is whether, in Romance and Germanic languages, verbs form their past tense with BE or HAVE (Sorace, 2000). In French, for example, the verbs *disappear, fall, arrive* and *die* all form their past tense with BE (unaccusative), whilst *laugh, run, talk* and *dance* (unergative) select HAVE. However, as an anonymous reviewer noted, the link between auxiliary selection and the syntactic notion of unaccusativity has never been satisfactorily explained. Furthermore, it may not even be possible to tie together the various phenomena proposed as diagnostics into a single notion of unaccusativity. Fortunately, virtually all diagnostics agree on the classification of the verbs used in the present study.

More relevant to are purposes here are the semantic correlates of unaccusativity. Broadly speaking, unaccusative verbs (e.g., English *disappear*, *fall*, *arrive*, *die*) are seen as more externally caused than unergative verb*s* (e.g., English *laugh*, *run*, *talk*, *dance*). Intuitively, this is consistent with the syntactic claim that the single argument of a verb such as *die* is equivalent to an object, whilst the single argument of a verb such as *dance* is equivalent to a subject (i.e., dancing is something you do, dying is something that happens to you).

This suggests that causative overgeneralization errors might be expected to be more likely to be accepted as grammatical (or to be produced) when they occur with verbs that denote a higher degree of direct external causation (e.g., *fall*, *tumble*) than with verbs denoting more internal causation (e.g., *laugh*, *giggle*). Pinker's (1989a) broad-range rule for the causative (which specifies that only dynamic events can be causativized) is an attempt to capture this notion of direct causation. Indeed, Pinker (1989a) specifically makes the above prediction (see pp. 301–304) and finds support for it from an analysis of 106 causative overgeneralization errors reported by Bowerman (1982):

> Verbs that are uncausativizable by virtue of the subtle linguistic criteria that delineate narrow subclasses in English should be causativized more often than verbs that are uncausativizable by virtue of being cognitively incomparable with the notion of direct unmediated causation (p. 302).

For example, Pinker (1989a) argues that few causative overgeneralization errors are observed with verbs of voluntary action (e.g., *swim*, *climb*, etc... – a total of 14 errors) because there is no causative narrow-range class to which these verbs could erroneously be assigned. Conversely, such errors are frequent for verbs of directed motion (e.g., *fall*, *come*, etc... – a total of 30 errors) which are cases of "a legitimate semantic conflation class [i.e., the class that includes verbs such as "*bring, take, put*"] with an illegitimately assigned stem" (1989a: 303).[6] Pinker (1989a: 303) presents the errors from Bowerman (1982) ordered according to the following semantic verb classes:

Directed motion (*fall*, *come*),
Going out of existence (*disappear*, *vanish*),
Being/staying (*stay*, *wait*),
Possession (*have*, *take*),

[6] In fact verbs of being/staying and of possession together account for 29 errors, even though these violate the broad-range rule for the causative. In a footnote, Pinker argues that "Lexical abstraction would eventually trigger the slight narrowing that would be needed to bring the child's broad-range rule into conformity with the adults" p. 402. But, in fact, Pinker's (1989a) theory contains no mechanism for narrowing broad-range rules – only for forming narrow range classes of verbs that DO undergo the alternation. Clearly Pinker (1989a) does not mean to imply that adults abandon broad-range rules as they continue to play a "psychological role...in adults" (p. 153), for example, licensing "Haigspeak" errors (e.g., "We're gonna grin ya" p. 153). Another internal inconsistency is that *drop* is included in both a list of "non-causativizable" (p. 302) "directed motion" verbs (p. 303) and (on the same page) a class of "verbs conflating causation with the kind of event expressed by these verbs [i.e., verbs of directed motion].

Psychological (*remember*, *wish*),
Involuntary emission (*sweat*, *bleed*),
Internally-caused state change (*bloom*),
Semivoluntary expression of emotion (*laugh*, *giggle*),
Voluntary action (*swim*, *climb*).

Although this list clearly represents some kind of hierarchy - from greatest to smallest degree of direct external causation (with the predicted rate of causativization errors increasing) – it is unclear whether Pinker (1989a) intends to imply that each verb class corresponds to a distinct step along the hierarchy, or whether some classes are intended to be grouped at the same level.

In the text, Pinker (1989a: 303) seems to imply that the first five subclasses should be grouped together, as should the following three, with the voluntary-action verbs forming a class of their own. In the present study, to avoid confusion, we investigate only the relatively conservative prediction that causative overgeneralization errors with verbs of semivoluntary expression (penultimate in the list) will be rated as less acceptable than such errors with verbs of directed motion or going out of existence (first and second). Again, although we test this prediction on both English and novel verbs (with appropriate actions), the latter allow for a more stringent test of the hypothesis, as for English verbs semantics are confounded with factors such as frequency and phonology.

This second prediction of Pinker's (1989a) theory is shared by Marcotte's (2006) proposal that children make causativization errors because they "are prone to misacquire non-causative verbs that prototypically denote causative events as causative verbs, and to lump these verbs with causative alternation verbs on the basis of lexical semantic similarity" (p. 19). This occurs because "when speaking to children younger than 4, adults tend to use...intransitive verbs...as well as...transitive verbs...to talk about events in which a change of state or location occurs as a result of the action denoted by the verb" (p. 13; see Marcotte's (2005: 200), analysis of the CHILDES archive). Thus this account also predicts that causativization errors will be rated as more acceptable for verbs that describe events of "directed motion" and "going out of existence" – events which denote "a change of state or location" – than for verbs describing events of "semivoluntary expression of emotion", which do not. In Marcotte's (2005) analysis of the CHILDES archive, causativization errors were almost five times more frequent for intransitive verbs that are often used to describe events that are "necessarily or prototypically causative" (p. 206) (e.g., *fall*, *disappear*) than for those that are not (e.g., *laugh*, *giggle*; p. 338).

Experimental evidence suggests that the knowledge that the transitive construction can be used to describe only events of direct, unmediated external causation is acquired relatively late in development. Wolff (1999: 97–99) showed children an animation where a (green) marble knocked into a second marble, causing it to knock a third marble, causing it to move. When asked *What did the green marble dax?* adults and four-year-olds consistently chose the second marble (a direct causal chain) whilst three-year-olds chose the third marble (an indirect causal chain) on around 50% of trials. Participants' preference for the direct causal chain was even

greater when the green marble was replaced by a sentient causer, except for the youngest group. However, Ammon (1980) found that even three-year-olds preferentially selected a picture depicting an unmediated causal chain as the best illustration of a transitive causative sentence (but would accept either mediated or unmediated chains as illustrations of a periphrastic causative).

In summary, then, the present study tests the following predictions derived from Pinker's (1989a) semantic verb class hypothesis: (1) participants will respect the experimentally-assigned semantic class of a novel (here, intransitive) verb (and the semantic class of English verbs) when rating intransitive and transitive causative uses, and this effect will increase with age. (2) Causative overgeneralization errors will be rated as less acceptable for verbs denoting events with a low degree of external causation (*laugh*, *giggle*) than for verbs denoting events with a higher degree of external causation (*fall*, *tumble*, *disappear*, *vanish*) as the former are less compatible than the latter with the notion of direct external causation inherent in the transitive causative construction (and Pinker's broad-range rule).

5. The entrenchment hypothesis

The second proposal – which we term the *entrenchment hypothesis* (Braine & Brooks, 1995) – is that repeated presentation of a verb (e.g., *disappear*) in one (or more) attested construction (such as the intransitive construction; e.g., *The rabbit disappeared*) causes the learner to gradually form a probabilistic inference that adult speakers do not use that particular verb in non-attested constructions. The prediction of this hypothesis is that the more often children hear a particular verb used in particular constructions in the input (a) the less likely they will be to produce an overgeneralization error for this verb (i.e., to use this verb in a non-attested sentence construction) and (b) the more unacceptable they will consider such errors in a grammaticality-judgement task. Cross-sectionally, one would expect to find more overgeneralization errors, and higher ratings of acceptability for such errors, for a low than a high frequency verb, as children have had less opportunity to acquire the correct usage for the former. Looking longitudinally at a particular verb, as development proceeds towards the adult state, one would expect to see overgeneralization errors (and ratings of acceptability for such errors) fall as the verb becomes ever more *entrenched* in its correct usage (though, as we have seen, the semantic verb class hypothesis would also predict this effect, as children acquire a more precise representation of the verb's semantics over time). To date, only two studies have directly investigated the crucial cross-sectional prediction of the entrenchment hypothesis.

Brooks, Tomasello, Dodson, and Lewis (1999), using an elicited production methodology, found that children (aged between 3;4 and 9;0) produced fewer utterances that violated the verb's transitivity status for high frequency verbs (e.g., **The man disappeared the ball*, **The ball hit*) than for semantically-matched verbs of lower frequency (e.g., **The man vanished the ball*, **The ball struck*).

Theakston (2004) conducted a similar study using a grammaticality-judgement paradigm. Adults and children (aged 5;1–6;2 and 7;1–9;2) were asked to rate sentences containing argument-structure overgeneralization errors with high and semantically-matched low frequency verbs (e.g., **Somebody fell/tumbled it off*, **I'm gonna disappear/vanish it*, **That joke was so funny it really laughed/giggled me*). Children were asked to make binary judgements, whilst adults used a seven-point scale. In support of the entrenchment hypothesis, adults rated overgeneralization errors with low frequency verbs as more acceptable than corresponding errors with higher frequency verbs, and more children of both ages rated the former than the latter errors as acceptable in the binary-judgement task.

Also relevant here are the grammaticality-judgement studies of Hochberg (1986) and Mazurkewich and White (1984). Hochberg asked children aged 3;4–5;5 to select the correct form from sentence pairs containing an overgeneralization and a corresponding adult form (e.g., *I'm gonna dance the donkey* vs *I'm gonna make the donkey dance; My dinner's gonna eat* vs *I'm gonna eat my dinner*). Mazurkewich and White (1984) obtained binary grammaticality judgements for sentences constituting overgeneralizations of the double object construction (e.g., **Nancy drove Ted the car*) from children aged 9, 12 and 16 yrs. Both studies (like Theakston, 2004) found that the proportion of overgeneralizations chosen as the correct form or rated as acceptable decreased as the age of the participants increased (a finding equally compatible with both the semantic verb class and entrenchment hypotheses) though neither specifically tested the latter by manipulating verb frequency.

In summary, then, the present study tests the following prediction derived from Braine and Brooks (1995) entrenchment hypothesis: participants' preference for grammatical (here, intransitive) over ungrammatical (here, transitive causative) verb uses should be highest for high frequency verbs and lowest for novel verbs, with low frequency verbs in-between.

6. The relative contributions of the formation of semantic verb classes and entrenchment to the retreat from overgeneralization errors

Previous work certainly provides some support for both the semantic verb class and entrenchment hypotheses. It is clear from this work, however, that both factors are needed to explain the pattern of data observed.

It cannot be the case that the frequency of attested uses (i.e., entrenchment) is the only relevant factor in determining the (un)grammaticality of putative argument-structure generalizations. The elicited production study of Brooks and Tomasello (1999), as discussed above, found an effect of semantic verb class independent of verb frequency (as only novel verbs were used). For Theakston's (2004) adult grammaticality-judgement data, no significant correlation between simple verb frequency and the acceptability of overgeneralization errors was observed, because, as Theakston (2004: 31) acknowledges "the semantic class to which individual verbs belong also play[s] a role in constraining argument-structure errors". Viau (2006) also argues for effects of verb semantics that are independent of frequency on the basis of an

analysis of naturalistic data from the CHILDES corpus.[7] Considering briefly the domain of inflectional morphology, Yang (2002) argues convincingly that the pattern of past tense over-regularization errors (e.g., *runned*) observed in naturalistic data cannot be explained by simple verb frequency alone, but can be explained by the frequency of particular verbs relative to others in the same class (though here the classes are defined phonologically as opposed to semantically).

Conversely, it cannot be the case that the semantic class to which a particular verb belongs is the only relevant factor when determining the grammaticality of the use of that verb in a particular construction. If this were the case, then it is difficult to explain why even Theakston's (2004) adult participants – who presumably had accurate, fully-formed semantic representations of verbs such as *giggle* and *laugh* – still displayed a frequency-driven entrenchment effect with semantic verb class held constant (e.g., between *giggle* and *laugh*). Effects of verb frequency with semantic verb class held constant were also found in the elicited production study of Brooks et al. (1999) and the comprehension study of Naigles et al. (1992, both previously discussed). In this latter study significantly more children acted out a frame-compliant response for causative sentences with the lower frequency verb *fall* than the higher frequency verbs *come* and *go* (all are members of Pinker's (1989a: 303) class of verbs of "directed motion").

Given that both factors clearly play a role in the retreat from overgeneralization, it is important to ask how the formation of semantic verb classes and entrenchment interact. One extreme characterization of the relationship between the two factors is that semantic class determines the grammaticality or otherwise of argument-structure generalizations and that verb-frequency effects, although sometimes statistically significant, represent only slight (dis)preferences, that do not bear on the issue of whether or not a putative generalization is regarded as grammatical. At the other extreme is the view that, whilst it is generally true that verbs with similar semantics behave similarly, this is merely a descriptive fact about language that plays little role in learning, and that entrenchment is the primary mechanism by which children learn that some of their putative generalizations are disallowed (e.g., Bowerman & Croft, 2007). One viewpoint somewhere in between these extremes is Tomasello's (2003:

[7] Viau (2006) argues that the acquisition of the double-object dative construction ([x CAUSE [z HAVE y]]) depends on the acquisition of the semantic primitive HAVE (as instantiated by verbs such as *have* and *get*), and the acquisition of the prepositional dative (x CAUSE [y GO z]]) on the primitive GO (as instantiated by the preposition *to*). Viau (2006) claims that his finding of no significant difference between the age of acquisition of *have*/*get* and the double-object construction, and between *to* and the prepositional construction supports his decompositional analysis. In order to provide strong support for this analysis however, one would have to show that *have*/*get* is acquired BEFORE the double-object dative, and *to* BEFORE the prepositional dative. Viau (2006) also points to a significant correlation across children between the age of acquisition of *have*/*get* and the double object construction, and between *to* and the dative construction. However, to rule out the possibility that both correlations are due to the effect of some nuisance variable (e.g., general linguistic development) one would have to show that the age of acquisition of *have*/*get* is not also correlated with the age of acquisition for the prepositional dative (or *to* with the double-object dative), which Viau (2006) does not do. Having said all this, we do not wish to dispute the more general claim that there exist effects of verb semantics on language acquisition that are independent of frequency (a major finding of the present study).

180) claim that "entrenchment works early...and semantic subclasses [of verbs] begin to work later, perhaps not until about 4;6 or so".

7. The present study

The goal of the present study was to systematically investigate both the semantic verb class and entrenchment hypotheses using a methodology that allowed us to dissociate the effects of verb semantics and verb frequency and hence to investigate how best to characterize the relationship between the two factors.

A second aim was to address a number of methodological shortcomings of previous work in this area. The use of production data to investigate either the semantic verb class (Brooks & Tomasello, 1999) or entrenchment hypothesis (Brooks et al., 1999) is problematic in principle, as a child's apparent reluctance to produce overgeneralizations with particular verbs cannot be taken as firm evidence that such utterances are ruled out by her grammar. For example, the finding that a child *prefers* to say *The ball vanished* to **The man vanished the ball* does not necessarily imply that she regards the latter as unacceptable.

The grammaticality judgement studies of Hochberg (1986), Mazurkewich and White (1984) and Theakston (2004) do not suffer from this shortcoming. However, since these authors did not systematically manipulate semantic verb class independently of frequency (for example by introducing a novel verb, as in the studies of Brooks, Tomasello and colleagues), it is impossible to determine whether semantic class had any effect over and above frequency on participants' grammaticality judgements.

Furthermore, since these studies used only English (as opposed to novel) verbs, one cannot rule out the possibility that features of the experimental verbs other than frequency could have contributed to the observed effect. One such feature is verb-length: for all but one of Theakston's (2004) six high/low frequency verb pairs (*disappear/vanish*) the high frequency verb was monosyllabic, whilst the low frequency verb was bi-syllabic. Thus it is not possible, in principle at least, to rule out a prosodic explanation for the fact that overgeneralization errors containing high frequency verbs (presented aurally), sounded less acceptable than corresponding errors with lower frequency verbs. A second possible confound is that of semantic specificity: High frequency verbs are – almost by definition – less semantically-specific (and thus applicable in a wider range of situations) than low frequency verbs. It could be the case that Theakston's (2004) participants, particularly the youngest, considered the overgeneralization errors with the low frequency verbs to be more acceptable because they had yet to acquire the precise semantics associated with these more semantically-specific verbs.

In the present study, ratings of grammatical acceptability were collected for sentences using novel verbs, as well as high and low frequency English verbs. This allows us to control for (a) verb length in syllables (the novel verb is lowest in frequency, but is matched in length to the high frequency verb, except for *disappear/vanish*) and (b) semantic specificity (the novel verb is lower in frequency than the low frequency

English verb, but is matched for semantic specificity). The use of a novel verb also, in effect, adds an extra level of frequency to the design (high, low and very low [i.e., novel]), meaning that any entrenchment effect can show up as a stepwise progression, as opposed to a single comparison. Most importantly, the use of a novel verb allows us to investigate the effect of (experimentally assigned) verb semantics on participants' ratings of grammaticality for uses of that verb in particular argument-structure constructions, relatively uncontaminated by effects of frequency or phonology.[8]

Another potential limitation of the grammaticality judgement studies of Hochberg (1986), Mazurkewich and White (1984) and Theakston (2004) is that, with the exception of Theakston's (2004) adult participants, only binary judgements (i.e., grammatical or ungrammatical) were collected. It is not possible, therefore, to determine whether particular children who rated the overgeneralizations containing low-frequency verbs as unacceptable considered the corresponding overgeneralizations with high-frequency verbs to be even less acceptable. Additionally, since Theakston's (2004) adult and child participants used different rating scales it is impossible to determine whether entrenchment has a different effect on ratings of grammaticality at different points in development. The present study addresses these limitations by collecting graded grammaticality judgements from adults and children aged 5–6 yrs and 9–10 yrs, using a novel five-point scale designed to be easily understood by young children.

Whilst Theakston's (2004) study used grammatical sentences as fillers, judgements were not collected for uses of the experimental verbs in *grammatical* sentences (Mazurkewich & White, 1984; obtained ratings for both sentence types, but did not conduct any statistical analysis of their data). This is potentially problematic as, in the absence of grammatical sentences to serve as a baseline control, it is impossible to tell whether participants regarded sentences containing high frequency verbs as more grammatically acceptable than corresponding sentences containing lower frequency verbs per se (i.e., regardless of whether or not the sentence contained an overgeneralization error). Many authors have noted that, even when rating sentences that are perfectly grammatical, subjects give higher ratings of acceptability to sentences that contain high frequency items (Featherston, 2005; Keller, 2000; Kempen & Harbusch, 2003, 2004; Schuetze, 1996). Conversely, following the same logic, one might expect participants to display a general dispreference for sentences containing novel verbs; even those in which the novel verb is used grammatically (that is, in accordance with both the semantic features of the event which it denotes, and its attested usage).

To address this potential confound, in the present study we obtained ratings of grammatical acceptability for both grammatical (intransitive) and ungrammatical (transitive causative) sentences for each experimental verb. This allows us to control for the effect of simple verb frequency on ratings of grammatical acceptability by using participants' preference for the grammatical (intransitive) over the

[8] A reviewer pointed out that a novel verb might not be uncontaminated by the effect of phonology if it happened to have a relevant phonological shape. To guard against this possibility, each novel verb had two different phonological shapes (e.g., *blick* vs *tamm*), counterbalanced across participants.

ungrammatical (transitive causative) version of a sentence containing a particular verb – as opposed to simply the rating for the ungrammatical version – as a dependent measure.

To summarise, in the present study, we collected five-point ratings of grammatical acceptability for both grammatical (intransitive) and ungrammatical (transitive causative) uses of a high frequency verb, a low frequency verb, and a novel verb, from each of three non-causativizable semantic classes from participants aged 5–6 yrs, 9–10 yrs and adults.

The semantic verb class hypothesis predicts that participants will respect the experimentally-assigned semantic class of a novel (here, intransitive) verb (and the semantic class of English verbs) when rating intransitive and transitive causative uses, and that this effect will increase with age. This hypothesis also predicts that causative overgeneralization errors will be rated as less acceptable for verbs denoting events with a low degree of external causation (*laugh*, *giggle*) than for verbs denoting events with a higher degree of external causation (*fall*, *tumble*, *disappear*, *vanish*) as the former are less compatible than the latter with the notion of direct external causation inherent in the transitive causative construction (and Pinker's broad-range rule). The entrenchment hypothesis predicts that participants' preference for grammatical (here, intransitive) over ungrammatical (here, transitive causative) verb uses should be highest for high frequency verbs and lowest for novel verbs, with low frequency verbs in-between.

Crucially, our design allows us to dissociate the effects of semantic verb class and entrenchment and hence to investigate the contributions of each mechanism to the retreat from argument-structure overgeneralization errors. If only semantic verb class has an effect on this process, then grammatical (intransitive) verb uses should be preferred over ungrammatical (transitive) uses to approximately the same extent for all verbs (high frequency, low frequency and novel) within a given semantic class; this effect increasing with age. It is particularly important that this preference should be observed for novel verbs, as such an effect can realistically be attributed only to the effect of verb semantics. Furthermore, the predicted effect for causative overgeneralizations to be rated as less acceptable for verbs denoting more internally than more externally-caused events should be observed with approximately equal effect size for high and low frequency and novel verbs.

Conversely, if entrenchment is the only factor at work, participants' preference for grammatical (intransitive) over ungrammatical (transitive) uses should be highest for high frequency verbs, lower for low frequency verbs and lowest for novel verbs, regardless of the actions which they denote. Importantly, if semantic verb class is not a factor, then one would expect participants' preference for grammatical over ungrammatical uses of novel verbs to be approximately zero, as no entrenchment can have occurred for these verbs.

If both factors operate at all ages tested, then one would expect to observe significant main effects of semantic verb class and entrenchment (as outlined above) and perhaps, although not necessarily, an interaction between these effects. Since the youngest children studied were aged 5–6 yrs, our study does not allow us to investigate Tomasello's (2003: 180) specific claim that "entrenchment works early...and

semantic subclasses [of verbs] begin to work later, perhaps not until about 4;6 or so''. However, our study does allow us to test one element of the prediction; specifically that semantic verb classes will have begun to be operational for five-year-olds, in which case one would expect to see even the youngest children display a significant preference for grammatical (intransitive) over ungrammatical (transitive) uses of – crucially – the novel verbs (as well as the English verbs). If semantic verb classes are not operational for five-year-olds, then our study will allow for a test of a more general version of Tomasello's (2003: 180) claim; specifically that ''entrenchment works early...and semantic subclasses [of verbs]...later''. If this is the case then one would expect to find that the youngest children show virtually no preference for grammatical (intransitive) over ungrammatical (transitive) verb uses (i.e., no effect of semantic class) but a greater preference for grammatical over ungrammatical uses for high than low frequency verbs (i.e., an entrenchment effect). Older children and adults would then be expected to display both semantic verb class and frequency effects, as outlined above.

8. Methods

8.1. Participants

Participants were 27 children aged 5;3–6;4 ($M = 5;10$), 24 children aged 9;3–10;3 ($M = 9;9$) and 42 adults (all undergraduate students, mostly aged 18–19) recruited from schools or university in the North of England. All participants were monolingual English speakers.

8.2. Design

The experiment used a $3 \times 2 \times 2 \times 3 \times 3$ mixed design. The between-subjects variables were age (5–6 yrs, 9–10 yrs, adult) and counterbalance version (A or B, corresponding to the pairing of the novel verb forms with novel actions). For each counterbalance version, eighteen test sentences (and corresponding computer animations illustrating the event described), were constructed; one for each combination of the three within-subjects variables of sentence type (grammatical-intransitive, ungrammatical-transitive causative), semantic verb class (three classes; see following section) and verb frequency (high, low, novel).

8.3. Test sentences

Semantic verb classes were chosen such that all verbs in each class were non-alternating intransitive-only verbs. Thus all overgeneralization errors presented constituted causative overgeneralization errors. Verbs were selected from three (descriptive) intransitive-only verb classes set out in Pinker's (1989a: 303) hierarchy (see Section 4): verbs of ''directed motion'', ''going out of existence'' and ''semivoluntary expression of emotion''. In order to maximise the semantic similarity of the

high frequency, low frequency and novel verb in each class we selected from each class a pair of English verbs with very similar semantics but varying in frequency (*fall*/*tumble*, *disappear*/*vanish* and *laugh*/*giggle*). For each of the three novel verbs we created a novel action semantically similar to that denoted by the two English verbs in the target semantic class (i.e., a novel *falling* action, a novel *disappearing* action, and a novel *laughing* action).

Within each semantic class, each sentence of a particular type (i.e., intransitive or transitive causative) was of identical structure, and varied only with respect to the verb and NP (in each case, one of three proper names). For each intransitive/transitive causative sentence pair, the subject of the intransitive sentence and the object of the corresponding transitive causative sentence were different characters. For example, participants assigned to the first counterbalance condition rated the following sentences:

Verbs of falling
Intransitive: <*Bart fell*/*Lisa tumbled*/*Homer meeked*> *into a hole*.
Transitive causative: **The man* <*fell Homer*/*tumbled Bart*/*meeked Lisa*> *into a hole*.

Verbs of disappearing
Intransitive: <*Bart disappeared*/*Lisa vanished*/*Homer blicked*>.
Transitive causative: **The magician*< *disappeared Homer*/*vanished Bart*/*blicked Lisa*>.

Verbs of laughing
Intransitive: <*Homer laughed*/*Bart giggled*/*Lisa tammed*>.
Transitive causative: **The funny clown* <*laughed Bart*/*giggled Lisa*/*tammed Homer*>.

The sentences rated by participants in the second counterbalance condition were identical, except for the meaning assigned to each of the novel verb forms (*blick* = *verb of falling*, *tam* = *verb of disappearing*, *meek* = *verb of laughing*).

For each test sentence, participants viewed a computer animation depicting the event described. This was to allow for the creation of novel actions of *falling*/*disappearing*/*laughing in a particular manner*, corresponding to the three novel verbs. The use of animations was also designed to ensure that participants construed the meaning of each sentence as intended. For example, some of the adult participants in Theakston (2004) seemed to interpret the sentences *Don't laugh*/*giggle me* as *Don't laugh*/*giggle <u>at</u> me* rather than, as intended, *Don't make me laugh*/*giggle*. The animations also served to establish that the veracity of the sentence was not in doubt, thus encouraging (particularly the younger) participants to rate sentences according only to grammatical acceptability, and not truth-value (McDaniel & Cairns, 1996). Details of these animations can be found in Appendix A.

8.4. Novel-verb-training sentences and animations

In order to further train participants in the correct (non-causative) meaning and (intransitive) use of each of the novel verbs, two training trials involving different

characters were created for each of the novel verbs. For example, participants assigned to the first counterbalance condition heard the following training sentences, and saw corresponding animations:

<Marge/Maggie> meeked off the table.
<Marge/Maggie> blicked out of sight.
<Marge/Maggie> tammed when she heard a funny joke.

8.5. Rating scale

All participants rated sentences using a novel five-point rating scale, designed to be easily understood by young children. The scale consisted of five cartoon faces (see Fig. 2), and had a midpoint, denoted by a neutral face, and two levels each to represent differing levels of grammatical acceptability (smiling faces) and unacceptability (frowning faces). It was felt that young children would be unable to provide sufficiently fine-grained judgements to necessitate the use of a seven-point scale. In order to make the scale easier for young children to understand (and to allow for the collection of binary judgements in the event that they were unable to provide graded judgements) the scale was split at the midpoint into green faces, denoting "acceptable" and red faces, denoting "unacceptable". The face at the middle of the scale was divided into two different-coloured halves. The scale had no numerical or text key.

8.6. Grammaticality-judgement-training sentences

Seven grammaticality-judgement-training trials, each consisting of a sentence and corresponding animation were used to introduce participants to the procedure of rating grammatical acceptability (and not truth-value) on the five-point scale. Sentences 1–4 were designed to illustrate maximally acceptable and unacceptable sentences, with the remaining sentences – based around overgeneralizations of the dative alternation – illustrating more marginal cases. Details of these sentences, and mean ratings of acceptability given by each age group, can be found in Appendix B.

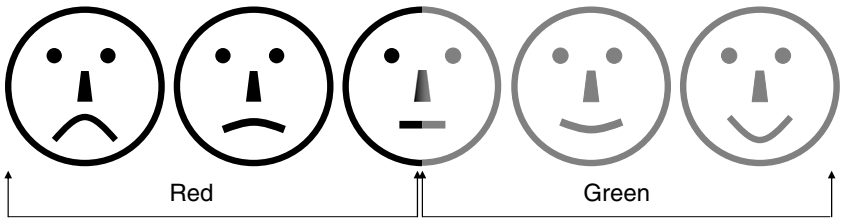

Fig. 2. The five-point rating scale used by all participants to rate sentences for grammatical acceptability.

8.7. Materials

Animations were created using *Flash MX Professional* and presented using *Swf.MAX*, running on a laptop PC computer. Stereo audio recordings were constructed such that one speaker, positioned next to the computer played the "sound effects", whilst the other (hidden inside a stuffed "talking dog" toy, except for the adults) played the sentences (spoken in a comical "dog's voice"). This was to enable the experimenter to present the task to (particularly the younger) children as a game in which the aim was to help a dog learning to speak English.

8.8. Procedure

8.8.1. 5–6-Year-olds

Children were tested individually, with the task divided into two sessions, given on consecutive days. The experimenter explained that the first task would be to learn the meanings of some new "made up" words, and then presented the six novel-verb-training trials. After each trial, the experimenter asked for a definition of the novel word. In most cases, the child simply gave a high frequency English gloss (e.g., *fall*, *disappear* or *laugh*), in which case the experimenter explained that the novel verb denoted, for example, not just falling, but falling in a particular way, and explained the distinctive manner of the novel action. This novel verb training procedure was given at the start of each of the two sessions.

The experimenter then introduced the grammaticality-judgement scale, using a procedure designed to enable even the youngest children to provide graded grammaticality judgments or, in the event that they were unable to do so, binary judgements. The experimenter initially explained that the dog was learning to speak English, but sometimes "gets it wrong, and says things a bit funny". The experimenter then told the child "we're going to help him, by telling him when he gets it right, and when he says it wrong. When he says it right we're going to choose the green counter and put it here [*indicates happiest face*]. When he says it wrong we're going to choose the red counter and put it here [*indicates saddest face*]. Don't worry about these other faces for now".

The experimenter then completed the first and second training trials, and invited the child to complete the third and fourth trials. All children correctly selected the happiest and saddest faces (using the appropriately coloured counters) for the third and fourth trials, respectively. Taking the green counter, the experimenter then explained that "Sometimes he [*indicates dog*] says it right but it's not perfect. If it's good but not perfect, you can put the counter here [*indicates second happiest face*]. If it's a little bit right and a little bit wrong, or somewhere in between you can put it here [*indicates middle face*]". Taking the red counter, the experimenter then continued, "Sometimes he says it wrong but it's not really terrible. If it's wrong but not terrible, you can put the counter here [*indicates second saddest face*]. If it's a little bit wrong and a little bit right, or somewhere in between you can put it here [*indicates middle face*]". The child then completed the three remaining training trials. For each trial the experimenter recorded the face that the child selected and (for the middle

face only) whether the green or red counter had been used (no child placed the green counter on a red face or vice versa).

The intention had been for the experimenter to give feedback for these practice trials, and to exclude children who did not appear to be using the scale appropriately. However, all children gave a good range of responses to the remaining three practice trials, and appeared to be using the scale as intended. Thus no feedback was given, and no children were excluded. Following the grammaticality-judgement training, the experimenter moved directly onto the test trials, which were presented in random order.

8.8.2. 9–10-Year-olds

The older children followed the same procedure as the younger children, except that no counters were used, as we were confident that these children would be able to provide graded judgements. Children marked their responses directly on the scale by circling the appropriate face.

8.8.3. Adults

Adults completed the same training and test trials as children, and rated sentences using the same scale. However, adults completed all trials in a single session, and were given only brief verbal definitions of the novel verbs, and brief verbal instructions in the use of the rating scale (see Appendix C).

9. Results and discussion

The mean ratings for all sentences are shown in Table A1 (Appendix D).

9.1. Preliminary analysis

A preliminary analysis was conducted to investigate the effect of counterbalance version, and to determine that participants of all ages were capable of using the rating scale as intended (i.e., that all participants rated the grammatical [intransitive] uses of the English verbs as significantly more acceptable than the ungrammatical, [transitive causative] uses). The data were entered into a $3 \times 2 \times 2 \times 3 \times 3$ mixed ANOVA with between-subjects variables of age and counterbalance-version and within-subjects variables of sentence type (grammatical – intransitive vs ungrammatical – transitive causative), semantic verb class (verbs of *falling*, *disappearing* and *laughing*[9]) and verb frequency (high, low, novel). The variable of counterbalance version was not associated with any significant main effects or interactions. Thus all subsequent analyses collapsed across the two versions.

[9] Recall that although, in effect, we used three very narrow semantic verb classes, these three classes are subsets of, and therefore correspond to, classes proposed in Pinker (1989a).

A $3 \times 2 \times 3 \times 3$ ANOVA collapsing across counterbalance-version revealed a significant main effect of sentence type ($F_{(1,90)} = 651.66$, $p < 0.001$, $\eta_p^2 = 0.88$), demonstrating that participants rated the grammatical, intransitive sentences ($M = 4.61$ [out of 5], $SE = 0.05$) as significantly more acceptable than their ungrammatical, transitive causative counterparts ($M = 2.91$, $SE = 0.06$).[10]

This ANOVA also yielded significant interactions of sentence type by age ($F_{(2,90)} = 30.69$, $p < 0.001$, $\eta_p^2 = 0.41$), sentence type by semantic class ($F_{(2,180)} = 47.124$, $p < 0.001$, $\eta_p^2 = 0.34$), and sentence type by verb frequency ($F_{(2,180)} = 85.43$, $p < 0.001$, $\eta_p^2 = 0.49$). Post hoc tests (all significant at $p < 0.001$) revealed that intransitive sentences were rated as significantly more acceptable than transitive causative sentences by all age groups (collapsing across frequency and semantic verb class), and across all semantic verb classes (collapsing across age and verb frequency), and across all levels of verb frequency (collapsing across age and semantic verb class). Thus, it would appear that our method for eliciting graded grammaticality judgements from both younger and older children and adults was extremely successful.

This finding notwithstanding, these interactions (and a three way interaction of sentence type by semantic verb class by age; $F_{(4,180)} = 3.76$, $p = 0.006$, $\eta_p^2 = 0.08$) demonstrate that the magnitude of participants' preference for grammatical over ungrammatical verb uses differed according to the variables of age, semantic verb class and verb frequency. We now consider these effects in the light of the predictions of the theories under investigation.

9.2. Testing the semantic verb class hypothesis

Pinker's (1989a) hypothesis makes two predictions: (1) since none of the verbs/actions (English or novel) were semantically consistent with any causativizing narrow-range semantic verb class, intransitive uses of all verbs (English and novel) should be preferred to transitive causative uses; with this effect increasing with age, as the narrow-range semantic verb classes are refined. (2) Causative overgeneralization errors will be rated as less acceptable for verbs denoting events with a low degree of external causation (*laugh*, *giggle*) than for verbs denoting events with a higher degree of external causation (*fall*, *tumble*, *disappear*, *vanish*) as the former are less compatible than the latter with the notion of direct external causation inherent in the transitive causative construction (and Pinker's broad-range rule).

Considering this first prediction, the preliminary analysis reported above demonstrated that participants did indeed display a significant preference for intransitive over transitive causative verb uses, collapsing across all ages and all levels of verb frequency. As we have argued previously, however, novel verbs allow for the stron-

[10] This ANOVA also revealed significant main effects of semantic verb class and frequency, and significant interactions of frequency by age, frequency by semantic verb class, and semantic verb class by frequency by age. However, because these effects collapse across grammatical (intransitive) and ungrammatical (transitive) sentence types, they are not interpretable.

gest test of this hypothesis as participants may be rejecting transitive causative uses of English verbs on the basis of attested usage, rather than verb semantics. The present set of analyses therefore investigated whether participants in each age group displayed a statistically significant preference for intransitive over transitive causative uses of the *novel verbs only* from each of the three semantic verb classes.

A 3×2 within-subjects ANOVA with independent variables of semantic class (verbs of *falling*, *disappearing*, *laughing*) and sentence type (intransitive – grammatical, transitive causative – ungrammatical) was conducted separately for each age group. In each case, the dependent measure was participants' ratings of grammatical acceptability for sentences including novel verbs only. These data are illustrated in Figs. 3–5.

For the 5–6-year-old children (see Fig. 3), a significant interaction of semantic verb class by sentence type ($F_{(2,52)} = 3.76$, $p = 0.03$, $\eta_p^2 = 0.13$) was observed, as were marginally significant main effects of semantic class ($F_{(2,52)} = 3.13$, $p = 0.052$, $\eta_p^2 = 0.11$) and sentence type ($F_{(1,26)} = 3.22$, $p = 0.08$, $\eta_p^2 = 0.11$). Analysis of this interaction revealed that, for the novel *verb of laughing*, children displayed a strong preference for grammatical, intransitive ($M = 3.70$, $SE = 0.25$) over ungrammatical, transitive causative ($M = 2.74$, $SE = 0.31$) uses ($p = 0.001$). For the novel *verb of falling* and *verb of disappearing*, however, children displayed no preference for grammatical uses over ungrammatical uses ($M = 3.98$, $SE = 0.24$ vs $M = 3.85$, $SE = 0.24$, $p = 0.9$, n.s.; $M = 3.48$, $SE = 0.27$ vs $M = 3.48$, $SE = 0.30$, $p = 1.00$, n.s.).

For the 9–10-year-old children (see Fig. 4), significant main effects of semantic verb class ($F_{(2,46)} = 4.06$, $p = 0.02$, $\eta_p^2 = 0.15$) and sentence type ($F_{(1,23)} = 50.34$,

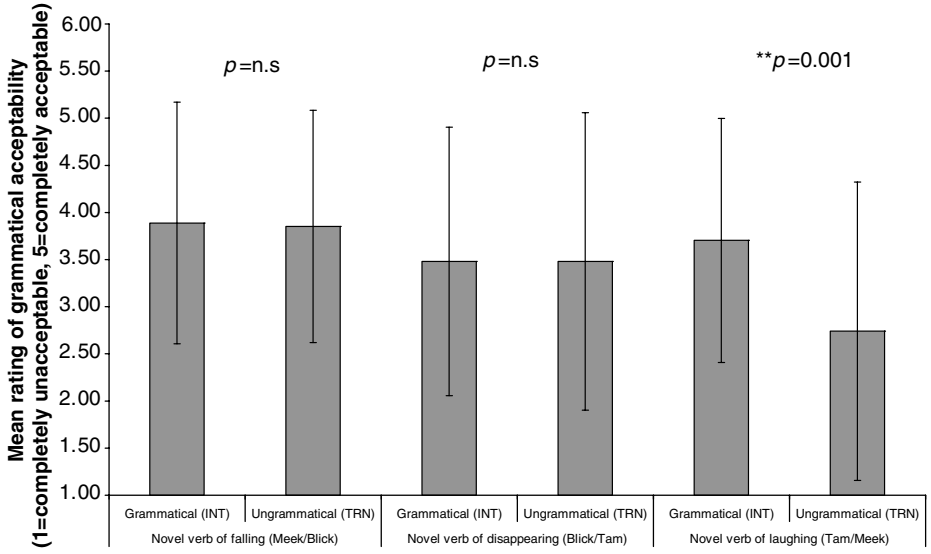

Fig. 3. Testing the semantic verb class hypothesis; 5–6-year-olds. Participants' mean ratings for grammatical (intransitive) and ungrammatical (transitive) uses of novel verbs by semantic verb class (error bars represent standard deviation).

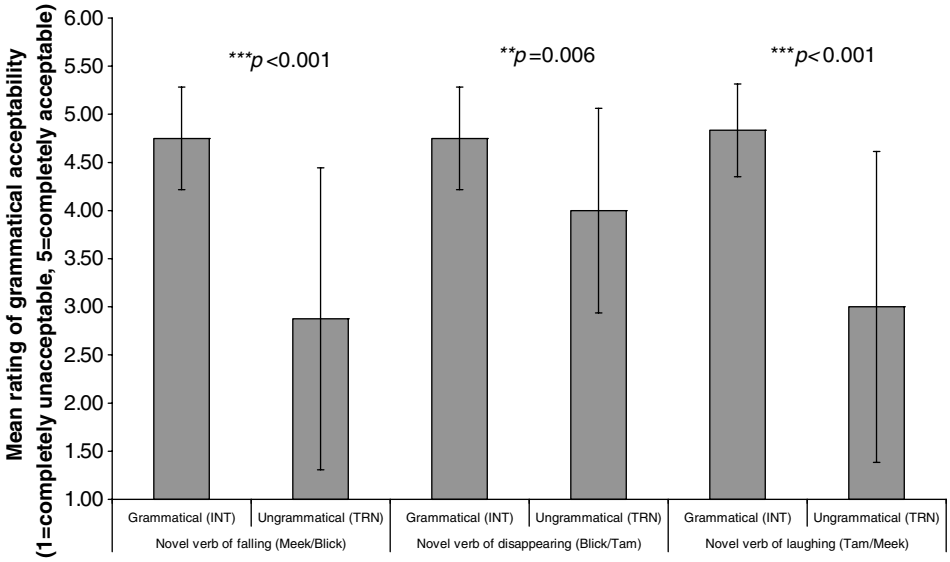

Fig. 4. Testing the semantic verb class hypothesis; 9–10-year-olds. Participants' mean ratings for grammatical (intransitive) and ungrammatical (transitive) uses of novel verbs by semantic verb class (error bars represent standard deviation).

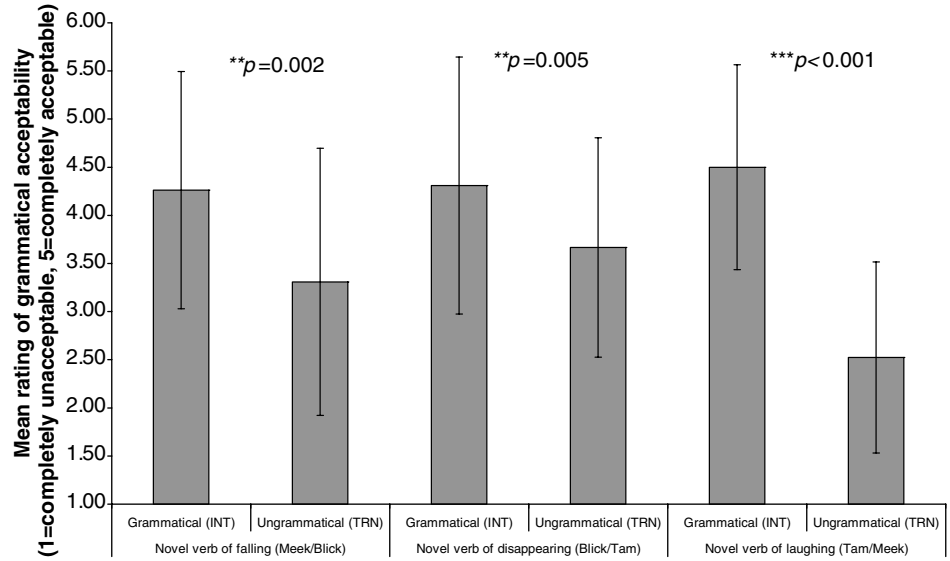

Fig. 5. Testing the semantic verb class hypothesis; adults. Participants' mean ratings for grammatical (intransitive) and ungrammatical (transitive) uses of novel verbs by semantic verb class (error bars represent standard deviation).

$p < 0.001$, $\eta_p^2 = 0.69$) were observed. The interaction of these two variables was also significant ($F_{(2,46)} = 5.49$, $p = 0.007$, $\eta_p^2 = 0.19$). Analysis of this interaction revealed that whilst these children strongly preferred grammatical to ungrammatical uses of novel verbs from all three semantic classes, this effect was more robust for the novel *verb of laughing* ($M = 4.83$, $SE = 0.10$ vs $M = 3.00$, $SE = 0.33$, $p < 0.001$) and *verb of falling* ($M = 4.75$, $SE = 0.11$ vs $M = 2.88$, $SE = 0.32$, $p < 0.001$) than for the novel *verb of disappearing* ($M = 4.75$, $SE = 0.11$ vs $M = 4.00$, $SE = 0.22$, $p = 0.006$).

The results for adults (see Fig. 5) were almost identical to those for the older children, with significant main effects of semantic verb class ($F_{(7,82)} = 7.51$, $p = 0.001$, $\eta_p^2 = 0.16$) and sentence type ($F_{(1,41)} = 45.92$, $p < 0.001$, $\eta_p^2 = 0.53$) and a significant interaction ($F_{(2,82)} = 13.20$, $p < 0.001$, $\eta_p^2 = 0.24$). Again, grammatical uses of novel verbs of each class were preferred to ungrammatical uses, and again this effect was least robust for the novel *verb of disappearing* ($M = 4.31$, $SE = 0.21$ vs $M = 3.67$, $SE = 0.18$, $p = 0.006$). As for the younger children, the largest mean difference was observed for the novel *verb of laughing* ($M = 4.50$, $SE = 0.16$ vs $M = 2.52$, $SE = 0.15$, $p < 0.001$), with the novel *verb of falling* in-between ($M = 4.36$, $SE = 0.19$ vs $M = 3.31$, $SE = 0.21$, $p < 0.001$).

This set of analyses therefore demonstrates that, as predicted by Pinker (1989a) all participants (except the youngest for two of the three verb classes) displayed a significant preference for intransitive over transitive causative uses of novel verbs whose semantics are not consistent with those of any causativizing narrow-range semantic verb class. A second analysis was conducted to investigate whether, as further predicted by Pinker (1989a), this effect would (1) increase with age and (2) be greater for verbs denoting more internally-caused events (*laugh*, *giggle*, *novel laughing verb*) than more externally-caused events (*fall*, *tumble*, *novel falling verb*, *disappear*, *vanish*, *novel disappearing verb*) and whether these effects would be observed for both English and novel verbs.

This analysis took the form of a mixed ANOVA with the dependent variable of preference for grammatical over ungrammatical use (i.e., rating for intransitive sentence minus rating for transitive causative sentence), and with independent variables of age and semantic class (verbs of *falling*, *disappearing*, *laughing*). Verb frequency (*high*, *low*, *novel*) was also included as an independent variable simply to allow us to investigate whether Pinker's (1989a) predictions are borne out with both high and low frequency verbs and – most importantly – with novel verbs, which, as we have argued above, constitute the strongest test of Pinker's (1989a) predictions. We reserve specific discussion of frequency effects until the next section.

Leaving aside, for the moment, frequency effects, this ANOVA yielded significant main effects of age ($F_{(2,90)} = 30.69$, $p < 0.001$, $\eta_p^2 = 0.41$) and semantic verb class ($F_{(2,180)} = 47.12$, $p < 0.001$, $\eta_p^2 = 0.34$), and a significant interaction of age by semantic verb class ($F_{(4,180)} = 3.76$, $p < 0.006$, $\eta_p^2 = 0.08$). This interaction is shown in Fig. 6.

Post hoc tests conducted to investigate the significant main effect of age revealed that the 5–6-year-olds showed a significantly smaller preference for the grammatical (intransitive) over the ungrammatical (transitive causative) sentences ($M = 0.96$ points on the five point scale, $SE = 0.12$) than either the 9–10-year-olds

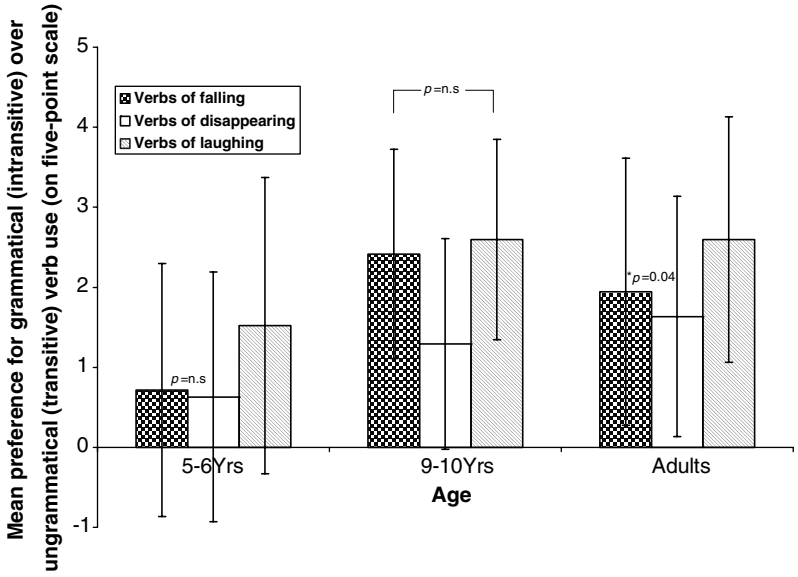

Fig. 6. Testing the semantic verb class hypothesis. Participants' mean preference for grammatical (intransitive) over ungrammatical (transitive) verb uses by age by semantic verb class, collapsing across verb frequency (error bars represent standard deviation). All differences illustrated are statistically significant at $p < 0.001$ unless otherwise indicated.

($M = 2.10$, $SE = 0.13$, $p < 0.001$) or the adults ($M = 2.06$, $SE = 0.10$, $p < 0.001$), who did not differ significantly.

Post hoc tests conducted to investigate the significant main effect of semantic verb class revealed the largest preference for grammatical over ungrammatical uses of *verbs of laughing* ($M = 2.24$, $SE = 0.09$), with this preference smaller for *verbs of falling* ($M = 1.69$, $SE = 0.10$) and smaller again for *verbs of disappearing* ($M = 1.19$, $SE = 0.09$; all comparisons significant at $p < 0.001$).

Further post hoc tests were conducted to investigate the interaction of age by semantic verb class (for means see Table A1: Appendix D; Fig. 6). Although the results pattern in the same direction for all age groups, some between-class comparisons did not reach significance for all ages. The advantage for *verbs of falling* over *verbs of disappearing* did not reach significance for the youngest children, and only narrowly reached significance for the adults ($p = 0.04$). The advantage for *verbs of laughing* over *verbs of falling* did not reach significance for the older children. All other comparisons were significant at $p < 0.001$.

This analysis provides support for Pinker's (1989a) predictions that participants' preference for grammatical (intransitive) over ungrammatical (transitive causative) verb uses should (1) increase with age and (2) be greater for verbs denoting more internally-caused events (*laughing* events) than more externally-caused events (*falling* or *disappearing* events). However, the data presented so far have included results for both novel verbs and real English verbs. It is therefore possible at this stage that

these findings do not reflect verb semantics at all but are an artefact of some other differences between the English verbs (such as frequency or phonology).

In fact, the age by semantic verb class by frequency ANOVA reported above found that verb frequency was not involved in any significant interactions with age or semantic verb class (though the latter was marginal at $p = 0.07$). This suggests that the findings presented so far, which are supportive of Pinker's (1989a) hypothesis are equally true for the high frequency verbs, the low frequency verbs and the novel verbs – the latter of which are relatively uncontaminated by effects of frequency or phonology.

Nevertheless, since, for this reason, novel verbs provide the most stringent test of Pinker's (1989a) hypothesis, post hoc tests were conducted to investigate these effects for novel verbs only. The findings were virtually identical. Again, participants' preference for grammatical over ungrammatical verb uses was smaller for the younger children ($M = 0.33$, $SE = 0.20$) than the older children ($M = 1.49$, $SE = 0.22$, $p < 0.001$) or adults ($M = 1.19$, $SE = 0.16$, $p = 0.002$), who did not differ significantly, and again, this preference was largest for *verbs of laughing* ($M = 1.59$, $SE = 0.15$), smaller for *verbs of falling* ($M = 0.96$, $SE = 0.18$) and smaller again for *verbs of disappearing* ($M = 0.46$, $SE = 0.09$; all comparisons $p = 0.02$ or better). The advantage for the novel *verb of falling* over the novel *verb of disappearing* did not reach significance for the youngest children or the adults and the novel *verb of laughing* and *verb of falling* did not differ for the older children. All other comparisons were significant at $p < 0.001$ (in the same direction as for the previous analysis with *all* verbs).

In summary, as predicted by Pinker (1989a), intransitive uses of the experimental verbs (which are consistent with their semantics) were generally preferred over transitive causative uses (which violate those semantics); this effect increasing between younger and older children. As further predicted by Pinker (1989a) participants were less tolerant of causative overgeneralizations of verbs that denote more internally-caused events (both English and novel *laughing* verbs) than of verbs that denote more externally-caused events (English and novel *falling* and *disappearing* verbs).

9.3. Testing the entrenchment hypothesis

The entrenchment hypothesis predicts that children's and adults' preference for grammatical (here, intransitive) over ungrammatical (here, transitive causative) verb uses will be greatest for high frequency verbs, for which considerable evidence regarding correct use has built up, and smallest for novel verbs, for which such evidence is minimal, with low frequency verbs in-between.

In the previous section, we reported a mixed $3 \times 3 \times 3$ age by semantic verb class by frequency ANOVA conducted on preference-for-grammatical use scores. In addition to the significant main effects of age and verb class (and their interaction) discussed above, this ANOVA also yielded a significant main effect of frequency ($F_{(2,180)} = 85.43$, $p < 0.001$, $\eta_p^2 = 0.49$). However, frequency did not significantly interact with either age ($F_{(4,180)} = 1.29$, $p = 0.27$, n.s.), or semantic verb class ($F_{(4,360)} = 2.21$, $p = 0.07$, n.s.) or both ($F_{(8,360)} = 0.97$, $p = 0.46$, n.s).

Planned comparisons conducted to investigate the main effect of frequency revealed that, as predicted by the entrenchment hypothesis, participants' preference for grammatical (intransitive) over ungrammatical (transitive) verb uses was greatest for high frequency verbs ($M = 2.43$, $SE = 0.08$), smaller for low frequency verbs ($M = 1.68$, $SE = 0.08$), and smaller again for novel verbs ($M = 1.00$, $SE = 0.11$), with all comparisons significant at $p < 0.001$.

The lack of a significant interaction of verb frequency by age ($F_{(4,180)} = 1.29$, $p = 0.27$, n.s.) demonstrates that verb frequency had a similar effect on judgements of grammaticality throughout development. Nevertheless, to be sure that the main effect of verb frequency held for each age group when analysed individually, we conducted planned comparisons to investigate the (non-significant) interaction of age by verb frequency (see Fig. 7). These revealed that, for all age groups, participants' preference for grammatical (intransitive) over ungrammatical (transitive) verb uses was significantly greater for high frequency than low frequency or novel verbs ($p < 0.001$ for all ages) and for low frequency than novel verbs ($p = 0.04$, $p = 0.02$ and $p < 0.001$ for the 5–6-year-olds, 9–10-year-olds and adults, respectively; for means see Fig. 7 and Table A1, Appendix D).

The marginally significant ($p = 0.07$) interaction of verb frequency by semantic class reported in the previous section suggests that the observed frequency effect might not hold across all the semantic verb classes studied. In fact, analysis of this

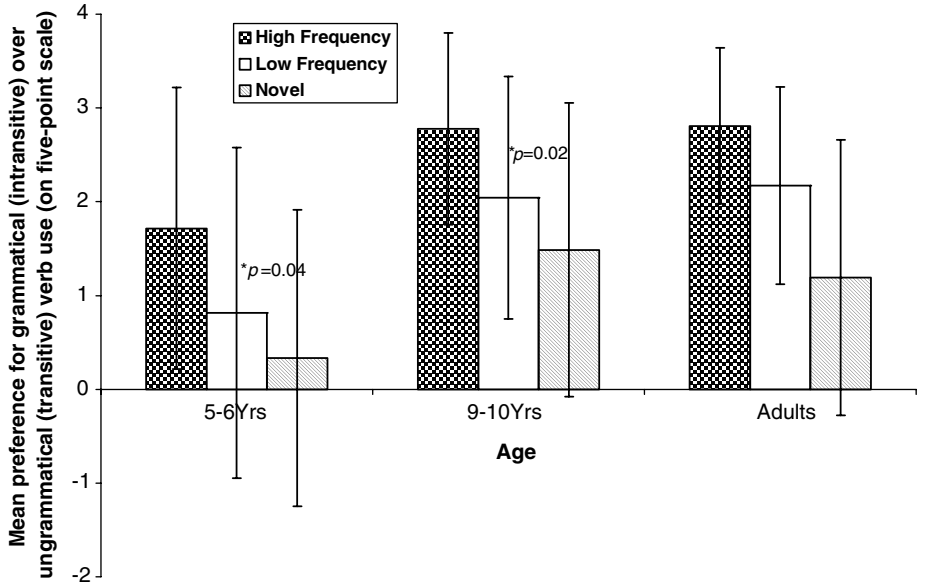

Fig. 7. Testing the entrenchment hypothesis. Participants' mean preference for grammatical over ungrammatical uses of high frequency, low frequency and novel verbs by age, collapsing across semantic verb class (error bars represent standard deviation). All differences illustrated are statistically significant at $p < 0.001$ unless otherwise indicated.

marginally significant interaction revealed that the frequency effect held across all semantic classes at $p = 0.01$ or better, with the exception of the (still significant) comparison between *laugh* and *giggle* ($p = 0.03$).

In summary, as predicted by the entrenchment hypothesis, participants' preference for grammatical over ungrammatical verb uses was highest for high frequency English verbs, lower for low frequency English verbs and lower still for novel verbs (indeed, non-existent for two of the three novel verbs for the youngest children). This effect was observed across all semantic verb classes and all age groups studied. Perhaps surprisingly, the magnitude of this frequency effect did not significantly increase with age. That is, although participants' *absolute* preference for grammatical over ungrammatical verb uses increased between the younger and older children, the extent to which this preference was greatest for high frequency verbs, smaller for low frequency verbs, and smallest for novel verbs did not significantly increase with age (though there was certainly a trend in this direction, particularly between the two child groups).

10. General discussion

In the present study, participants aged 5–6, 9–10 and adults rated grammatical (intransitive) and ungrammatical, overgeneralized (transitive causative) uses of a high frequency (*fall*, *disappear*, *laugh*), low frequency (*tumble*, *vanish*, *giggle*) and novel intransitive verb from each of three semantic classes, using a five-point scale. The aim of the study was to test the predictions of the semantic verb class hypothesis (Pinker, 1989a) and the entrenchment hypothesis (Braine & Brooks, 1995), and to dissociate the effects of the two factors.

Consistent with the first prediction derived from Pinker's (1989a) hypothesis, participants' preference for intransitive over transitive causative uses of (both English and novel) verbs whose semantics are not consistent with any causativizable verb class was greater for older children (9–10 yrs) and adults than younger children (5–6 yrs). For novel verbs, this effect can be due only to verb semantics. The lack of a difference between older children and adults is consistent with Pinker's (1989a: 349) "speculation...that an adult's narrow-class rules correspond to the verbs that happen to alternate in his...lexicon at a maturationally determined critical point, presumably around puberty".

Consistent with the second prediction made by Pinker (1989a) all participants were less tolerant of causative overgeneralization errors with English and novel verbs denoting more internally caused (verbs of *laughing*) than more externally caused (verbs of *falling* and *disappearing*) events, as the latter are more compatible with the notion of direct external causation inherent in the causative construction. However, Pinker's (1989a) account does not explain why participants were more tolerant of overgeneralizations of *verbs of disappearing* than *verbs of falling* (if Pinker's, 1989a, 1989b: 303 hierarchy is interpreted literally then any predicted effect would be in the opposite direction). Indeed, intuitively a *disappearing* event might seem to be more internally caused than a *falling* event, as a person might choose to disap-

pear (e.g., by leaving a party) but is unlikely to choose to fall: falling will often be caused – at least in part – by the actions or properties of external causers. The answer probably lies with the actual events to which the participants were exposed (i.e., the animations). The *disappearing* event did not involve figurative disappearing (e.g., leaving a party), but literal disappearing, as caused by a magician. Whilst a figurative *disappearing* event may be more under the control of the entity involved than is a *falling* event, a literal *disappearing* event *requires* the action of an external causer (e.g., a magician) whereas a *falling* event does not (non-magicians can make themselves fall, but not disappear). This would explain why a literal *disappearing* event is more apt to be described with a transitive causative construction than a (literal) *falling* event.[11]

Of course, this is only speculation as the difference in acceptability for causativization errors with *falling* and *disappearing* verbs was not predicted. Future work could investigate this account explicitly by obtaining ratings for causativization errors with disappearing (and other) verbs in figurative and literal contexts. However, if our speculation is accurate, then this reinforces the point made earlier that it is not verbs per se that are semantically (in)consistent with particular constructions, but the events that they describe (in particular contexts). Interestingly, Levin & Rappaport Hovav (1995: 126) argue that "verbs of existence and appearance [including verbs of *ceasing to exist* and *disappearing*] belong to a class for which the notions of internal and external causation are apparently not relevant". For these authors, such verbs are anomalous, because whilst "they would be most likely be classified as unergative by the linking rules" (p. 126) – as, in common with unergatives such as *laugh*, no external cause is present in the lexical semantic representation – they are "among the prototypical unaccusative verbs of many languages" (p. 126), including English. If our speculation is correct, this would suggest that there is nothing anomalous about such verbs. As for other verbs, the extent to which *(dis)appearing* verbs are appropriate in the transitive causative construction reflects the extent to which the event described involves direct external causation.

As we noted in Section 1, a purely syntactic version of the unaccusative hypothesis would predict that unaccusative verbs might be coerced into the transitive causative construction more easily than unergative verbs because the former, but not the latter, have an empty subject position into which the causer could be inserted. As we further noted, such an account could not explain a difference in the acceptability of causativization errors with two unaccusative verbs (e.g., *fall* and *disappear*), or indeed two unergative verbs (hence one cannot save the account by positing that *disappear* is in fact unergative, as the differential acceptability of causativization errors with *disappear* and *laugh* is then unexplained).

Consistent with the prediction of the entrenchment hypothesis, participants' preference for grammatical over ungrammatical verb uses was highest for high frequency English verbs, lower for low frequency English verbs and lower still

[11] An anonymous reviewer offered a similar explanation: a salient faller can have some control over falling (i.e., he can try to keep it from happening), whilst a disappearer has no control over disappearing (presumably this reviewer is referring only to literal disappearing).

for novel verbs (indeed, non-existent for two of the three novel verbs for the youngest children).

In Section 1, we argued that neither the semantic verb class nor the entrenchment hypothesis alone explain the pattern of data observed. The findings of the present study strongly support this assertion. The semantic verb class hypothesis, in its current form, does not explain the effect of verb frequency observed in the present study. "Indeed, Pinker's (1989a: 349) speculation about the role of maturation" would appear to predict that no verb frequency effect should be observed for adults. Neither do Pinker's (1989a) verb classes seem to capture the precise pattern observed in the present and previous studies. They do not explain why participants were more tolerant of causative overgeneralizations of *verbs of disappearing* than *verbs of falling*. More seriously they cannot account for the fact that overgeneralizations of verbs of "being/staying" and "possession" (Pinker's, 1989a: 303) account for around one third of the errors reported in Bowerman (1982) when such errors violate Pinker's (1989a) broad-range rule for the causative, and, as such, should not be produced at all (as Pinker acknowledges in Footnote 6, p. 402; see also Bowerman & Croft, 2007; Marcotte, 2005, 2006).

The entrenchment hypothesis alone cannot explain why participants showed a significant preference for intransitive over transitive causative uses of NOVEL *verbs of laughing* (all participants), *falling* and *disappearing* (older children and adults);[12] less still why this effect should increase between younger and older children. The finding that, for English verbs, preference for grammatical over ungrammatical verb uses does not increase between the older children and adults is also potentially problematic for the entrenchment hypothesis. According to the hypothesis, ratings of acceptability for sentences such as *The funny clown laughed Homer* should decrease as development proceeds, and tokens of attested (intransitive or periphrastic causative) uses build up in memory. This was not the case when comparing the older children and adults. It is possible that floor effects might be playing some role here. Although the mean rating for a causative overgeneralization error with a high frequency verb was around 2 (on the five-point scale) for both older children and adults (3 for younger children), the training procedure might have encouraged participants to reserve ratings of 1 (i.e., the lowest point on the scale) for sentences that do not preserve the constituent structure of the sentence (e.g., **His teeth man the brushed*).

[12] In principle, one might object that the process of entrenchment could have contributed to this effect, as each novel verb was presented in solely intransitive sentences during training. In fact, participants heard only two such sentences (each repeated twice for children) during training. Clearly entrenchment cannot operate on the basis of so few exposures, as children produce overgeneralization errors for verbs such as *go* and *fall* (Bowerman, 1988), which they are certain to have heard on more than four occasions. Furthermore, even if entrenchment were playing some small role here, it cannot explain the finding that this effect increased with age, as the younger and older children heard the novel verbs used in intransitive non-causative sentences an equal number of times.

So it would appear that neither of the extreme positions that we sketched in the input are accurate. Entrenchment does appear to be doing some work that the formation of semantic verb classes is not. However, it is not the case that the rough correlation between verb semantics and argument-structure privileges is simply a descriptive fact that plays little role in learning: Adults and children do seem to be making use of verb semantics to determine the argument-structure privileges of novel verbs.

A third possible characterization we discussed is that "entrenchment works early...and semantic subclasses [of verbs] begin to work later, perhaps not until about 4;6 or so" (Tomasello, 2003: 180). Certainly, the youngest children we studied showed clear evidence of an entrenchment effect (although even our youngest children were older than 4;6). Assessing the second part of Tomasello's claim is less straightforward. In one sense, only one of the three semantic subclasses of verbs (*verbs of laughing*) "worked" in that the 5–6-year-olds could use the semantics of the action to determine the argument-structure privileges of the novel verb (suggesting that some classes do not "work" until much later than 4;6). However, the fact that the youngest children did not reject transitive causative uses of the novel *verbs of disappearing* and *falling* does not necessarily constitute a "failure" to use semantic information. Presumably these children were more tolerant of overgeneralizations of the novel *disappearing* and *falling verb* than the novel *laughing* verb for the same reason that adults were (i.e., because the latter is less compatible with the notion of direct external causation inherent in the construction). Thus although the youngest children did not use verb semantics to reject causativizations of the novel verbs of *disappearing* and *falling*, they displayed, in their reluctance to reject them, some sensitivity to the semantic constraints on the causative construction that eventually come to render such sentences ungrammatical in the adult grammar.

This finding is compatible with the suggestion that semantic verb classes are not classical discrete categories that exist in an "all or nothing" fashion, but are more probabilistic categories (e.g., Lakoff, 1987) that build up gradually and increase in representational strength throughout the lifespan. Perhaps the gradient of direct unmediated external causation as discussed by Pinker (1989a) is most usefully thought of not as a hierarchy along which discrete semantic verb classes can be arranged, but as a continuous gradient along which individual verbs (or more properly the actions they denote in particular sentences) can be arranged. The closer they are to the prototype of a direct, unmediated externally-caused event, the more apt such verbs or actions are to be used in a causative construction, which itself bears this meaning (e.g., Goldberg, 1995). Verbs or actions that are semantically similar will naturally cluster together along the hierarchy, giving rise to emergent, descriptive "semantic classes" with more or less prototypical members. Children acquire, in a probabilistic, incremental manner, the "semantic core" of each particular argument-structure construction, and the precise semantics of each verb or action.

This notion of a gradient explains, far better than a class-based account, the status of borderline-acceptable utterances such as "Water will not deteriorate the sockets"

(Pinker, 1989a: 153), where the extent to which the action is conceived of as externally caused is unclear (and perhaps different for different speakers).[13] A gradient approach might also explain some anomalies associated with Pinker's (1989a) account. For example, as we have noted previously, this account predicts no causativization errors with *verbs of being/staying* and *possession* (as they are inconsistent with the broad-range rule for the construction), yet such errors constitute around one third of the errors reported by Bowerman (1982). A gradient account would allow for such errors when they describe acts of external causation (e.g., *I'm staying it in the water*) without having to predict errors for verbs like *be* (in the same class as *stay* under Pinker's account), to which it would seem more difficult to attribute a causal reading.

If the acquisition of the "semantic core" of a construction (and of the semantics of particular actions) and entrenchment are both probabilistic, incremental processes, then it is easier to imagine how the two processes could work together. For example, perhaps the frequency effect observed in the present study is mediated by the effect of verb semantics. Each time a verb is heard in a new context, the learner is given a further opportunity to increase her knowledge of the precise semantics of the verb, and hence the constructions that the verb can appear in. Each time the learner encounters a causative construction (and infers the speaker's intended meaning), this constitutes evidence that the construction encodes direct external causation.

A computer model that makes use of this statistical correlation between semantics and syntax (Alishahi & Stevenson, 2005) is able to replicate some of the present findings. The model takes as its input frames that are representations of scene-utterance pairings (i.e., of the utterance heard by the child, and the scene it describes), that specify the verb (e.g., *take*), its semantic primitives (*cause*, *move*), argument roles (*agent*, *theme goal*) and categories (*human*, *concrete*, *destination-predicate*) and syntactic structure (*arg1 verb arg2 arg3*). Whenever, the model encounters a new frame, it stores the semantic and syntactic information in the lexical entry for the verb. It then groups this frame together with the existing class of frames whose members share the greatest number of syntactic and semantic features, or, if none are sufficiently similar, creates a new class (since the classes are not pre-specified, this will be the case for the majority of frames in the early stages of learning). Whenever the model is presented with a previously encountered frame, it increases the representational strength of that frame and of its parent class.

[13] Bresnan and Nikitina (2003) propose an analogous analysis of the prepositional dative *NP V NP to NP* construction (e.g., *John sent the package to Chicago*). In this case the relevant gradient is the extent to which the first NP can be construed as causing the second NP to go to/be at the location denoted by the third. The notion of a gradient explains why phrases such as *give a headache/punch to someone* (though attested, and not wholly unacceptable) sound odd by comparison with more prototypical events of causing to go.

To simulate generalization, every five training pairs, the model is presented with a frame from which the syntactic representation has been deleted, but with all other information intact. The task of the model is to choose the most appropriate syntactic structure (this can be considered analogous to the child attempting to produce an utterance using a known verb to describe a scene which she is observing). In order to choose the most appropriate pattern, the simulation can make use of two sources of information: the syntactic structures stored in the lexical entry for that particular verb (*item-based* knowledge), and those stored in the lexical entry of the other verbs in the same class (i.e., the other verbs to which this verb is semantically and syntactically similar; *class-based* knowledge).

To see how the simulation models entrenchment, consider an example where its task is to choose the correct argument-structure for a frame containing the verb *fall*, the semantic primitives *cause, move*, argument roles *agent*, *theme*, *goal* and categories *human*, *human*, *destination-predicate*. Early in learning, when the model has encountered *fall* on only a few occasions, the influence of the syntactic structures stored in the lexical entry for *fall* (e.g., *arg1 verb*) will be relatively weak. However, the influence of the syntactic structures stored in the lexical entry for other verbs which the model has grouped into the same class (including those annotated with the syntactic structure *arg1 verb arg2 arg3*) will be relatively strong. Thus early in development the model may incorrectly select a structure such as *arg1 verb arg2 arg3* as its best prediction for our example scenario with the verb *fall*. This can be thought of analogous to the child producing (or rating as less unacceptable) a sentence such as *The man fell Homer into a hole*.

Later in development, when the model has encountered *fall* on many occasions, the influence of the syntactic structures stored in the lexical entry for *fall* (e.g., *arg1 verb*) will be relatively strong, compared to the influence of the model's class-based knowledge. Thus after a certain developmental point has been reached, the model will unerringly select *arg1 verb* as its best prediction of a syntactic structure for the verb *fall*, even if this results in some argument roles/categories remaining unspecified. This can be thought of as analogous to the child producing a sentence such as *Homer fell into a hole*, to describe a scene where a man is seen to cause this action (or rating a sentence such as *The man fell Homer into a hole* as more unacceptable).

Although Alishahi and Stevenson (2005) describe how their simulation models entrenchment longitudinally, it is clear that, cross-sectionally, the model will produce more incorrect syntactic predictions (''overgeneralization errors'') for verbs that have low than high frequency in the input (e.g., *fall* vs *tumble*). Additionally, the model also sometimes predicts (just as our human participants accept) non-attested syntactic structures for novel verbs (e.g., *meek*).

Although these authors do not specifically investigate the predictions of the semantic verb class hypothesis, it is clear that the simulation achieves its generalization-then-retreat effect via the formation of probabilistic verb classes that are largely based on semantic (but also syntactic) commonalities. One might therefore expect the model, like our (older) participants, to be capable of predicting the appropriate

argument-structure for novel verbs, based purely on their semantic class. It will be interesting to see whether or not this, in fact, turns out to be the case.

Returning to the present study, our findings have provided compelling evidence in support of both the semantic verb class hypothesis (Pinker, 1989a) and the entrenchment hypothesis (Braine & Brooks, 1995). However, our findings have also demonstrated that neither hypothesis on its own is sufficient to account for the pattern of data observed, and that the answer to the question of how children learn to restrict their argument-structure overgeneralization errors will necessarily include elements of both proposals. We conclude our discussion with some suggestions for future work that would be instructive in this area.

Although the present study allowed us to examine effects of semantic verb class independently of frequency (by comparing novel verbs) ideally one would also compare participants' ratings for overgeneralization errors for English verbs – matched for frequency – from different semantic classes.[14] (Though note that, in the present study, an effect of verb class was still observed for the English verbs, despite the potentially confounding effect of frequency).

Our study focussed on three effectively very narrow semantic verb classes (*verbs of falling*, *laughing* and *disappearing*). However, one important element of Pinker's (1989a) proposal is that the child forms classes of verbs which may be – on the surface – extremely heterogeneous (e.g., *directed motion*: *come*, *go*, *fall*, *rise*, *drop*; p. 303), but which share "grammatically relevant" (p. 273) semantic properties. Pinker's (1989a) account in its original form would therefore predict that overgeneralization errors involving members of the same class (e.g., *swim* and *eat* or *be* and *stay*) should receive similar ratings of (un)acceptability. This prediction would not be shared by a "fuzzy-class" account, under which the relevant dimension would be the extent to which the event denoted by a particular verb can be seen as involving direct unmediated external causation (this could be determined by asking participants to rate events along this dimension). Further studies using the current paradigm should investigate these competing models. It would also be interesting to investigate more systematically (and perhaps in a range of languages) whether causativization errors are rated as more ungrammatical for verbs that virtually always have a unmarked non-causative form cross-linguistically, than for verbs that are more equally split between having a unmarked non-causative form and an unmarked causative form across languages.

[14] Such a comparison is not possible with the current dataset, as the English verbs from different semantic classes are not matched for frequency, as the following frequency data from the *British National Corpus* (all texts, VERB uses only) demonstrate: *fall*, 26,367; *disappear*, 5317; *laugh*, 9115; *tumble*, 830; *vanish*, 1518; *giggle*, 686. However, these figures demonstrate that the semantic verb class effect holds even when the frequency effect works against it. For example, overgeneralizations with the unergative verbs *giggle* and *laugh* were rated as less acceptable than overgeneralizations with the other, unaccusative verbs, despite the fact that *giggle* is LESS frequent than either *tumble* or *vanish*, and that *laugh*, whilst more frequent than *disappear*, is far less frequent than *fall*.

Future work should also investigate whether the observed entrenchment effect operates on whole semantic classes of verbs, as opposed to just – as the present study has shown – on individual verbs themselves. It might be expected that a verb that belongs to a semantic class that includes a large number of high frequency verbs might be more resistant to overgeneralization than a verb of equal frequency belonging to a lower-frequency class. In the domain of inflectional morphology, Yang (2002) demonstrates that some low frequency verbs are nevertheless resistant to past-tense over-regularization errors (+*ed*) because they are members of a phonological class with a high overall frequency. Of course, conducting such a study with semantic verb classes would require a firm commitment to a particular set of classes (such as those proposed by Pinker, 1989a or Levin, 1993) and thus would necessarily constitute a strong test of one particular formulation of the semantic verb class hypothesis, rather than of the general principle.

For simplicity, our study focused exclusively on one particular type of argument-structure overgeneralization error (the causative). Future work should explore whether and how entrenchment and verb semantics work together to constrain other overgeneralization errors such as the converse error (the anti-causative) and errors pertaining to the dative and locative alternations (which are also discussed by Pinker, 1989a).

Although primarily conceived as explanations for the retreat from argument-structure overgeneralization errors, both entrenchment and the formation of classes of semantically-similar (or otherwise similar) items are also applicable, in principle, to other types of overgeneralization error such as morphological errors (e.g., reversative *un*-prefixation and past-tense/plural over-regularization errors). Future work in this direction could exploit the new experimental paradigm developed here, which allows the researcher to obtain graded grammaticality judgements from children.

Even the youngest children studied here were, at 5–6 years, relatively old in terms of syntax acquisition. A future goal of the field should be to obtain similar grammaticality judgements from younger children (perhaps even using paradigms such as preferential looking) to investigate the relative contributions of entrenchment and the formation of semantic verb classes early in development. Finally, although at least three experimental studies have now provided support for the entrenchment and semantic verb class hypotheses, we know of no study in which the frequency and semantics of the verbs presented to participants during training have been systematically manipulated (e.g., by presenting novel verbs with different levels of frequency, to describe novel actions). Future work using the novel paradigm for grammaticality judgements outlined here should address this shortcoming.

Acknowledgements

We thank the children, parents and teachers of St. Catherine's Roman Catholic Primary School, who took part in the research reported here. Our thanks are also due to three anonymous reviewers for helpful comments on an earlier version of this manuscript.

Appendix A

Actions and animations for experimental verbs.

Fall: character falls straight down (i.e., feet first) into a hole, without rotating.
Tumble: character falls into a hole, performing three complete 360° rotations.
Meek(/blick): character falls forward, then plunges headfirst into a hole, having rotated 180°.
Disappear: character gradually fades into background, accompanied by a ''boing'' sound effect.
Vanish: character suddenly vanishes from the screen, accompanied by an ''orchestral hit'' sound effect.
Blick(/tam): character gradually fades into the background, whilst spinning, accompanied by a ''magic'' sound effect.
Laugh: character laughs (standard ''laughing'' sound effect).
Giggle: character giggles (''giggling'' sound effect, i.e., higher pitch and greater pitch modulation than ''laughing'' effect).
Tam(/meek): character laughs very fast at an extremely high pitch (i.e., somewhat similar to a hyena's laugh).

For each intransitive sentence, the corresponding animation depicted the named character, alone, against a plain background, performing the action described by the verb. The actions described by the high frequency, low frequency and novel verbs in each semantic verb class were subtly different. For each semantic class, the novel action consisted of the performing of the action described by the high frequency verb in a particular manner.

For each transitive causative sentence, the corresponding animation depicted the second-named character performing the action denoted by the relevant verb as described above. For these animations, the first-named character (*the man/the magician/the funny clown*) was seen to INDIRECTLY cause the second-named character to perform the action by (a) pulling away the table on which the character was standing (*verbs of falling*), (b) waving a magic wand (*verbs of disappearing*) or (c) performing a juggling routine (*verbs of laughing*). It was important that the first-named character was seen to cause only INDIRECTLY the second-named character to perform the appropriate action in order to ensure that participants interpreted the novel verbs as non-causativizable intransitive verbs. Had the first character DIRECTLY caused the second to, for example, fall into the hole, the novel verb might have been interpreted as – taking the novel *falling* verb as an example – similar in meaning to *push* or *shove* (i.e., to a verb that *can* appear in transitive sentences, sometimes with a causative meaning). This is because, as discussed at length in Section 1, events of direct (but not indirect) external causation are associated with the transitive construction (not because all transitives are causative – many are not – but because all causatives [at least all direct causatives] are transitive).

An anonymous reviewer suggested that the fact that the novel verbs denoted *falling*, *disappearing* or *laughing in a particular manner* may be potentially problematic as causative verbs do not usually encode manner. However, the low frequency English verbs used in the present study – *tumble* (=''To fall; especially to fall in a helpless way''), *vanish* (''To disappear from sight, to become invisible, especially in a rapid and mysterious manner'') and *giggle* (''To laugh

continuously in a manner not uproarious, but suggestive either of foolish levity or uncontrollable amusement") also encode manner (at least according to these definitions from the *Oxford English Dictionary*). Thus whilst it may be true that causative verbs do not usually encode manner (though *manner of motion verbs* such as *roll* or *bounce* are an exception) – and this may be part of the reason why the verb classes we chose are non-causativizable – the meanings of the novel verbs would seem to be consistent with their classes, as was our primary aim.

This reviewer also pointed out that, in some languages, non-causative transitive sentences with *laugh/giggle* are permitted, with the meaning *laugh/giggle AT*, and was concerned that the existence of a similar but grammatical form might constitute a confound. However, if subjects were interpreting our *X laughed/giggled Y* sentences as *X laughed/giggled AT Y*, then we would presumably have replicated Theakston's (2004) finding of surprisingly high ratings of acceptability for these sentences. In fact, all age groups gave these sentences the lowest ratings of acceptability.

Finally, this reviewer was also concerned that, since children tended to gloss novel words with familiar ones (and adults were explicitly given glosses containing familiar words, see Appendix C), they might not have been treating the novel verbs as novel. However, if participants were not treating these verbs as novel, then one would presumably expect them to rate sentences with these novel verbs in a similar manner to sentences containing semantically-similar English verbs. In fact, as we outline in the Results section, children and adults treated the English and novel verbs differently when rating both grammatical (intransitive) and ungrammatical (transitive causative) sentences.

Appendix B

Grammaticality judgement training sentences and mean ratings on the five-point scale (1 = completely unacceptable, 5 = completely acceptable)

	5–6-Year-olds		9–10-Year-olds		Adults	
	M	*SD*	*M*	*SD*	*M*	*SD*
1. The cat drank the milk	n/a[a]	n/a	n/a	n/a	5.00	0.00
2. *The dog the ball played with	n/a	n/a	n/a	n/a	1.00	0.00
3. The frog caught the fly	5.00	0.00	5.00	0.00	5.00	0.00
4. *His teeth man the brushed	1.00	0.00	1.00	0.00	1.00	0.00
5. *The woman said the man a funny story	3.04	1.09	1.96	0.20	2.45	0.83
6. *?The girl telephoned her friend the news	2.70	1.17	3.21	0.83	3.33	0.90
7. ?The man whispered his friend the joke	3.96	1.13	3.58	1.02	3.36	0.82

[a] For child participants, the experimenter completed the first two practice trials to illustrate the procedure.

Sentences 1 and 3 and 2 and 4 were designed to illustrate maximally grammatically acceptable and unacceptable sentences, respectively. Sentences 5–7 were (in some cases marginally acceptable) overgeneralizations involving the double-object dative construction. Sentence 5 was designed to be the least acceptable of these three sentences but (arguably) more acceptable than the bizarre word-order errors illustrated in Sentences 2 and 4. Sentences 6 and 7 were designed to be less than completely acceptable but (again arguably) sufficiently acceptable as to be rated above the midpoint of the scale. In general, participants' ratings confirmed these intuitions.

Appendix C

Novel verb definitions and instructions (adults). Note that these definitions were designed to ensure that each novel verb was construed as a member of the same semantic verb class as the English verb mentioned in the definition.

[*VERB1*] *means something similar to fall, but it doesn't just mean fall. It means to fall in a particular head-first way, like you've just seen on the clip – and as you'll see on the next one. So this is [VERB1]ing.*
[*VERB2*] *means something similar to disappear, but it doesn't just mean disappear. It means to disappear in a particular spinning and fading way, like you just saw – and like this. So this is [VERB2]ing.*
[*VERB3*] *means to laugh, but to laugh in a particular high-pitched, crazy sounding way like you just heard. Here it is again. So this is [VERB3]ing.*

Your task for the remainder of the study is to rate sentences for how acceptable you find them using the scale on your sheet. Obviously the sad face on the left means "completely unacceptable" and the smiling face on the right means "completely acceptable". The others mean "bad but not terrible" (the other red face), "good but not perfect" (the other green face), and "in-between" (the red and green face). First we are going to do seven practice sentences. There are not any right or wrong answers. We just want to know how acceptable each sentence sounds to you, in your opinion.

Appendix D

Mean ratings for each test sentence.

Table A1
Raw data

Semantic verb class	Verb frequency	5–6-Year-olds ($N = 27$)			9–10-Year-olds ($N = 24$)			Adults ($N = 42$)			Total ($N = 93$)		
		Intrans (Gram.)	Trans (Ungr.)	Pref. for Gram.	Intrans (Gram.)	Trans (Ungr.)	Pref. for Gram.	Intrans (Gram.)	Trans (Ungr.)	Pref. for Gram.	Intrans (Gram.)	Trans (Ungr.)	Pref. for Gram.
Verbs of falling	High (*Fall*)	4.48 (0.20)	2.89 (0.25)	**1.60 (0.26)**	4.88 (0.07)	1.71 (0.13)	**3.17 (0.14)**	4.95 (0.05)	1.86 (0.11)	**3.10 (0.12)**	4.77 (0.07)	2.15 (0.10)	**2.62 (0.10)**
	Low (*Tumble*)	4.26 (0.17)	3.74 (0.27)	**0.52 (0.28)**	4.75 (0.09)	2.54 (0.25)	**2.20 (0.24)**	4.98 (0.02)	3.19 (0.14)	**1.79 (0.14)**	4.66 (0.06)	3.16 (0.12)	**1.50 (0.12)**
	Novel (*Meek*/*Blick*)	3.89 (0.24)	3.85 (0.24)	**0.04 (0.30)**	4.75 (0.11)	2.88 (0.32)	**1.88 (0.32)**	4.26 (0.19)	3.31 (0.21)	**0.95 (0.29)**	4.30 (0.12)	3.35 (0.15)	**0.96 (0.18)**
	All *verbs of falling*	4.21 (0.15)	3.49 (0.18)	**0.72 (0.20)**	4.79 (0.06)	2.38 (0.17)	**2.42 (0.17)**	4.73 (0.07)	2.79 (0.09)	**1.94 (0.13)**	4.58 (0.06)	2.89 (0.08)	**1.69 (0.10)**
Verbs of disappearing	High (*Disappear*)	4.63 (0.14)	3.26 (0.26)	**1.37 (0.26)**	4.92 (0.06)	2.92 (0.23)	**2.00 (0.24)**	5.00 (0.00)	2.60 (0.14)	**2.41 (0.14)**	4.85 (0.05)	2.92 (0.12)	**1.93 (0.12)**
	Low (*Vanish*)	4.70 (0.12)	4.19 (0.24)	**0.52 (0.25)**	4.92 (0.06)	3.78 (0.23)	**1.13 (0.26)**	4.95 (0.03)	3.10 (0.15)	**1.86 (0.15)**	4.86 (0.04)	3.69 (0.12)	**1.17 (0.12)**
	Novel (*Blick*/*Tamm*)	3.48 (0.27)	3.48 (0.30)	**0.00 (0.33)**	4.75 (0.11)	4.00 (0.22)	**0.75 (0.25)**	4.31 (0.21)	3.67 (0.18)	**0.64 (0.22)**	4.18 (0.13)	3.72 (0.14)	**0.46 (0.16)**
	All *verbs of disappearing*	4.27 (0.13)	3.64 (0.17)	**0.63 (0.20)**	4.86 (0.05)	3.57 (0.15)	**1.29 (0.17)**	4.75 (0.07)	3.12 (0.10)	**1.64 (0.11)**	4.63 (0.05)	3.44 (0.08)	**1.19 (0.09)**
Verbs of laughing	High (*Laugh*)	4.70 (0.14)	2.52 (0.24)	**2.19 (0.33)**	4.83 (0.08)	1.67 (0.13)	**3.17 (0.13)**	4.95 (0.03)	2.02 (0.13)	**2.92 (0.13)**	4.83 (0.05)	2.07 (0.10)	**2.76 (0.12)**
	Low (*Giggle*)	4.56 (0.17)	3.15 (0.38)	**1.41 (0.44)**	4.79 (0.09)	2.00 (0.17)	**2.79 (0.17)**	4.91 (0.06)	2.02 (0.13)	**2.88 (0.14)**	4.75 (0.06)	2.39 (0.14)	**2.36 (0.15)**
	Novel (*Tamm*/*Meek*)	3.70 (0.25)	2.74 (0.31)	**0.96 (0.25)**	4.83 (0.10)	3.00 (0.33)	**1.83 (0.34)**	4.50 (0.16)	2.52 (0.15)	**1.97 (0.19)**	4.35 (0.11)	2.76 (0.15)	**1.59 (0.15)**
	All *verbs of laughing*	4.32 (0.13)	2.80 (0.17)	**1.52 (0.22)**	4.82 (0.06)	2.22 (0.14)	**2.60 (0.15)**	4.79 (0.06)	2.19 (0.09)	**2.60 (0.11)**	4.64 (0.05)	2.41 (0.08)	**2.24 (0.09)**
All verb classes	High	4.61 (0.12)	2.89 (0.19)	**1.72 (0.21)**	4.88 (0.04)	2.10 (0.10)	**2.78 (0.11)**	4.97 (0.02)	2.16 (0.09)	**2.81 (0.09)**	4.82 (0.04)	2.38 (0.08)	**2.43 (0.08)**
	Low	4.51 (0.12)	3.69 (0.16)	**0.81 (0.18)**	4.82 (0.06)	2.78 (0.16)	**2.04 (0.15)**	4.94 (0.03)	2.77 (0.09)	**2.17 (0.09)**	4.76 (0.04)	3.08 (0.08)	**1.68 (0.08)**
	Novel	3.69 (0.18)	3.35 (0.21)	**0.33 (0.19)**	4.78 (0.07)	3.29 (0.19)	**1.49 (0.21)**	4.36 (0.17)	3.17 (0.13)	**1.19 (0.18)**	4.28 (0.10)	3.27 (0.10)	**1.00 (0.11)**

Participants' ratings for grammatical (intransitive) and ungrammatical (transitive) uses of high frequency, low frequency and novel verbs of falling, laughing and disappearing by age. Also shown are participants' preference-for-grammatical-use scores and (in parentheses), the standard error of the mean.

References

Alishahi, A., & Stevenson, S. A probabilistic model of early argument structure acquisition. In *Proceedings of the 27th annual conference of the cognitive science society*, Stresa, Italy, July 2005.

Ammon, M. S. H. (1980). *Development in the linguistic expression of causal relations: Comprehension of features of lexical and periphrastic causatives*. Unpublished PhD thesis. University of California, Berkeley, CA.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–193). New Jersey: Erlbaum.

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 overly general grammar? In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 73–101). Oxford: Blackwell.

Bowerman, M., & Croft, W. (2007). The acquisition of the English causative alternation. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability*. Hillsdale, NJ: Lawrence Erlbuam.

Braine, M. D. S., & Brooks, P. J. (1995). Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 352–376). Hillsdale, NJ: Erlbaum.

Bresnan, J., & Nikitina, T. (2003). On the gradience of the dative alternation. <http://lfg-stanford.edu/bresnan/download.html>.

Brooks, P. J., & Tomasello, M. (1999). How children constrain their argument-structure constructions. *Language, 75*(4), 720–738.

Brooks, P. J., Tomasello, M., Dodson, K., & Lewis, L. B. (1999). Young children's overgeneralizations with fixed transitivity verbs. *Child Development, 70*(6), 1325–1337.

Chomsky, N. (1981). *Lectures on Government and Binding*, Foris, Dordrecht.

Clark, E. V. (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. (1991). Possible verbs and the structure of events. In S. Tsohatzidis (Ed.), *Meaning and prototypes* (pp. 48–73). London: Routledge.

Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28*(1), 90–98.

Featherston, S. (2005). Magnitude estimation and what it can do for your syntax: some wh-constraints in German. *Lingua, 115*, 1525–1550.

Goldberg, A. E. (1995). Constructions: A construction grammar approach to argument structure. Chicago: University of Chicago Press.

Haspelmath, M. (1993). More on the typology of inchoative/causative verb alternations. In B. Comrie & M. Polinsky (Eds.), *Causatives and transitivity*. Amsterdam: John Benjamins.

Hochberg, J. G. (1986). Children's judgements of transitivity errors. *Journal of Child Language, 13*(2), 317–334.

Keller, F. (2000). *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Unpublished PhD thesis, University of Edinburgh.

Kempen, G., & Harbusch, K. (2003). An artificial opposition between grammaticality and frequency: comment on Bornkessel, Schlesewsky, & Friederici (2002). *Cognition, 90*, 205–210.

Kempen, G., & Harbusch, K. (2004). The relationship between grammaticality ratings and corpus frequencies: a case study into word order variability in the midfield of German clauses. In S. Kepser & M. Reis (Eds.). *Linguistic evidence: Empirical, theoretical, and computational perspectives*. Berlin: Mouton De Gruyter.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Legendre, G. (1989). Unaccusativity in French. *Lingua, 79*, 95–164.

Levin, B. (1993). English verb-classes and alternations: A preliminary investigation. Chicago: University of Chicago Press.

Levin, B., & Rappaport Hovav, M. (1995). Unaccusativity: At the syntax-lexical semantics interface. Cambridge, MA: MIT Press.

Lidz, J., Gleitman, H., & Gleitman, L. R. (2004). Kidz in the 'hood': syntactic bootstrapping and the mental lexicon. In G. Hall & S. R. Waxman (Eds.), *Weaving a Lexicon*. Cambridge, MA: MIT Press.

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition* (pp. 55–63). Cambridge: Cambridge University Press.

Lord, C. (1979). *Don't you fall me down: children's generalizations regarding cause and transitivity*. *Papers and reports on child language development (PRCLD) 17*. Stanford, CA: Stanford University, Department of Linguistics.

Macwhinney, B. (1987). The competition model. In B. Macwhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Erlbaum.

Macwhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language*(31), 883–914.

Marcotte, J. P. (2005). *Causative alternation errors in child language acquisition*. Unpublished PhD thesis. Stanford University, Stanford, CA.

Marcotte, J. P. (2006). Causative alternation errors as event-driven construction paradigm completions. In E. V. Clark & B. F. Kelly (Eds.), *Constructions in acquisition*. Stanford, CA: CSLI.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition, 46*(1), 53–85.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57* (54, Serial No. 228).

Mazurkewich, I., & White, L. (1984). The acquisition of the dative alternation: unlearning overgeneralizations. *Cognition, 16*, 261–283.

McDaniel, D., & Cairns, H. (1996). Eliciting judgments of grammaticality and reference. In D. McDaniel, C. McKee, & H. Cairns (Eds.), *Methods for assessing children's syntax*. Cambridge, MA: MIT Press.

Naigles, L., Fowler, A., & Helm, A. (1992). Developmental changes in the construction of verb meanings Cognitive Development. *Cognitive Development, 7*, 403–427.

Naigles, L., Gleitman, H., & Gleitman, L. R. (1993). Syntactic bootstrapping and verb acquisition. In E. Dromi (Ed.), *Language and cognition: A developmental perspective*. NJ: Ablex.

Nedjalkov, V. P. (1969). Nekotorye verojatnostnye universalnii v glagol'nom slovoobrazovanii. In I. F. Vardul (Ed.), *Jazykovye universalii i lingvisticheskaja tipologija* (pp. 106–114). Moscow: Nauka.

Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. In *Proceedings of the fourth annual meeting of the Berkeley Linguistics Society* (pp. 157–189). Berkeley: University of California.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1989a). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

Pinker, S. (1989b). Resolving a learnabiliity paradox in the acquisition of the verb Lexicon. In M. L. Rice & R. L. Shiefelsbusch (Eds.), *The teachability of language*. Baltimore: Paul H. Brookes.

Randall, J. H., van Hout, A., Weissenborn, J., & Baayen, H. (2004). Acquiring unaccusativity: a cross-linguistic look. In E. Alexiadoe, E. Anagnostopoulou, & M. Everaert (Eds.), *The unaccusativity puzzle: Explorations of the syntax-lexicon interface* (pp. 332–352). Oxford: Oxford University Press.

Saxton, M. (2000). Negative evidence and negative feedback: immediate effects on the grammaticality of child speech. *First Language, 3*(60), 221–252.

Schuetze, C. (1996). *The empirical base of linguistics: grammaticality judgments and linguistic methology*. Chicago: University of Chicago Press.

Smith, C. (1970). Jespersen's move and change class and causative verbs in English. In *Linguistics and literary studies in honor of Archibald A. Hill*. In M. A. Jazayery (Ed.). *Descriptive linguistics* (Vol. 2, pp. 101–109). The Hague: Mouton.

Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language, 76*, 859–890.
Theakston, A. L. (2004). The role of entrenchment in children's and adults' performance on grammaticality-judgement tasks. *Cognitive Development, 19*(1), 15–34.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Viau, J. (2006). Give = CAUSE + HAVE/GO: Evidence for early semantic decomposition of dative verbs in English child corpora. In *Proceedings of the 30th Boston university conference on language development*. *November 2005* (pp. 665–676). Boston University, Boston, USA.
Wolff, P. (1999). *Event construal and the linguistic coding of causal chains*. Unpublished PhD thesis. Northwestern University, Evanston, IL.
Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition, 88*, 1–48.
Yang, C. D. (2002). *Knowledge and learning in natural language*. Oxford: Oxford University Press.
Zaenen, A. (1993). Unaccusativity in Dutch: integrating syntax and lexical semantics. In J. Pustejovsky (Ed.), *Semantics and the lexicon* (pp. 129–161). Dordrecht: Kluwer.