The retreat from overgeneralisation errors: A multiple-paradigm approach

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Amy Bidgood.

July 2016

Contents

Abstract

The retreat from overgeneralisation errors: A multiple-paradigm approach

Amy Bidgood

This thesis examines children's argument structure overgeneralisation errors (e.g. **Don't giggle me!*). Errors of this kind arise from children observing that certain verbs can appear in more than one argument structure (e.g. *The ball rolled/Homer rolled the ball*). This pattern can be usefully generalised to allow children who have heard a verb produced in only one of these structures (e.g. *The window opened*) to produce it in the other (e.g. *Marge opened the window*). The ability to generalise patterns to new items is key to children becoming productive language users. However, if they *over*generalise this pattern, errors will result: *Bart giggled* is grammatical, but **Lisa giggled Bart* (meaning *Lisa made Bart giggle*) is not.

This thesis tested three hypotheses designed to explain how children retreat from such overgeneralisation errors, or, indeed, avoid making them altogether: the semantic verb class hypothesis (Pinker, 1989); the entrenchment hypothesis (Braine & Brooks, 1995); and the preemption hypothesis (Goldberg, 1995). Chapter 3 uses a novel-verb grammaticality judgment paradigm to investigate overgeneralisation errors in the locative construction (e.g. **Marge filled tea into the cup*). Chapter 4 investigates overgeneralisation errors in the transitive and intransitive constructions, using a grammaticality judgment paradigm with known verbs, as well as a production priming paradigm designed to elicit errors from young children (e.g. **Homer swam the fish*). Finally, in order to investigate the role of semantics in language development more generally, Chapter 5 moves beyond overgeneralisation errors to investigate children's acquisition of the passive construction (e.g. *Bart was helped by Lisa*).

This thesis adds to a growing body of work demonstrating that none of the individual theories (semantics, entrenchment, preemption) alone is able to explain children's retreat from overgeneralisation, and that an integrated approach, such as that proposed by Ambridge and colleagues' FIT account, is required to account for the data. The thesis moves our understanding forward by demonstrating both that this account can explain error patterns in production, and that the role of verb-in-construction semantic compatibility (a key aspect of the FIT account) can explain children's acquisition of argument structure more widely.

Chapter 1: Generativist and constructivist approaches to syntax acquisition

1.0 Introduction to the thesis

This thesis investigates children's argument structure overgeneralisation errors (e.g. **Don't giggle me!*) and the mechanisms that have been proposed to explain how children 'retreat' from such errors, or, indeed, avoid making them altogether. Through a series of experiments, several accounts of mechanisms aiming to explain the phenomenon are tested. However, it is first important to specify the theoretical framework adopted throughout the thesis. This is not simply a matter of terminology, but affects the assumptions about the knowledge children start out with, or develop along the way, to help them solve the overgeneralisation problem.

Theoretical approaches to language acquisition may be broadly split into two opposing views: generativist and constructivist. According to the generativist account, children's knowledge of syntax is abstract from the start of the process. Under this approach, lexical items are the basic unit of language and children are able to use their knowledge of syntactic rules to combine words (e.g. verbs and nouns) into larger units (e.g. verb phrases). The constructivist account, in contrast, assumes that constructions themselves (e.g. the transitive construction, X VERBed Y) are basic units into which lexical items can be placed. Under this approach, children are born with no knowledge of these constructions but, rather, they acquire them from the input. Following naturally from this is the fact that many of children's early constructions appear to be lexically restricted, based on the lexical items the child has heard in each construction in the input. This thesis investigates the psychological reality of lexical effects in children's acquisition of verb argument structures through the examination of argument structure overgeneralisation errors (e.g. **Don't giggle me!*).

This first chapter sets out in more detail the assumptions and predictions of the generativist and constructivist approaches to syntax acquisition. Evidence for and against the two approaches is then presented. While support is found for both generativist and constructivist accounts, the balance of the evidence indicates that lexical effects are a reality, and one that generativist accounts struggle to explain. The constructivist approach therefore informs the studies presented in later chapters of this thesis.

Chapter 2 examines in detail the main phenomenon of interest in this thesis: argument structure overgeneralisation errors (e.g. **Don't giggle me!*). It begins by explaining the origins of these errors and why the retreat from them may be such a challenge for children (and for researchers attempting to explain how this might come about). Three proposed mechanisms for the retreat from overgeneralisation are then described: the semantic verb class hypothesis (Pinker, 1989), the entrenchment hypothesis (Braine & Brooks, 1995) and the preemption hypothesis (Goldberg, 1995). Evidence in support of each is discussed, and the chapter concludes by suggesting that elements of all three mechanisms may need to be combined in order to successfully explain children's retreat from argument structure overgeneralisation errors.

Chapter 3 presents the first paper in this thesis (Bidgood, Ambridge, Pine & Rowland, 2014). This paper reports a grammaticality judgment study investigating overgeneralisation errors of the locative construction (e.g. *Lisa sprayed the flowers with water*/*Lisa sprayed water onto the flowers*; c.f. **Homer poured the cup with water*/**Homer filled water into the cup*). It therefore extends previous work (e.g. Ambridge, Pine, Rowland & Young, 2008) to a new construction, with the locative providing a critical test of the hypotheses under investigation because of the true bi-directionality of the alternation and the fine-grained distinctions between the semantic subclasses defined by Pinker (1989). The use of novel verbs is also a particular strength of this paper (c.f. Ambridge, Pine & Rowland, 2012), as any effects of semantics could not be attributed to participant's previous experience of the verbs in question appearing in locative sentences (i.e. any effect of semantics must be independent of frequency effects). Results show effects of both verb frequency and verb semantics, lending support to the entrenchment/preemption hypotheses and the semantic verb class hypothesis, respectively.

Chapter 4 presents a multi-method paper investigating two of the most frequent constructions in English: the transitive-causative (e.g. *The man rolled the ball*) and the intransitive-inchoative (e.g. *The ball rolled*). This alternation is an important test of the hypotheses due to its frequency (both full locative and full dative sentences are rare) and the higher frequency of overgeneralisation errors reported in these sentences types (see Pinker, 1989, pp. 22-25). So, whereas Bidgood et al. (2014) provides a critical test of the semantics hypothesis in particular, this paper investigates the ability of semantic and statistical approaches to explain the

errors children make most frequently. Extending the findings of previous work (e.g. Ambridge, Pine, Rowland, Freudenthal & Chang, 2014) to the transitive-intransitive alternation, using a large number of alternating and fixed-transitivity verbs in a grammaticality judgment study, this paper took a different approach to semantics to that of previous investigations of this alternation (e.g. Ambridge et al., 2008) by viewing semantics on a continuum rather than in discrete classes. Taking a novel methodological approach, syntactic priming was used to elicit overgeneralisation errors from young children. Results from both methods show strong support for entrenchment and semantic approaches, with more limited evidence for the preemption hypothesis.

Chapters 3 and 4 provide evidence for frequency-based and semantic accounts of children's retreat from overgeneralisation errors. However, these accounts must also be able to explain findings for constructions that do not involve the production of errors if they are to be taken seriously as general mechanisms in language acquisition. Chapter 5 therefore uses the priming method from Chapter 4 to investigate the role of semantics in children's acquisition of a construction known to cause significant difficulties for young children in terms of comprehension: the passive. This paper extends the findings of Messenger, Branigan, McLean and Sorace (2012) to demonstrate that, while young children have abstract knowledge of the passive construction (in line with generativist approaches), that knowledge is, in fact, semantically constrained, in line with the theoretical approach developed throughout this thesis.

Chapter 6 concludes the thesis, discussing the findings in Chapters 3, 4 and 5 and their implications for the entrenchment and preemption hypotheses and the semantic account of the retreat from overgeneralisation errors. Overall, the findings of the studies presented in this thesis all suggest the need for an integrated account that can explain statistical and semantic effects, which will also be discussed in this chapter.

Chapters 3, 4 and 5 are presented in a paper format rather than conventional thesis chapter format. Chapter 3 has already been published (Bidgood et al., 2014) and Chapters 4 and 5 are currently being prepared for submission to peer-reviewed journals. The format of these chapters has been standardised so as to fit with the thesis as a whole. For example, no abstract is provided and references are provided at the end of the thesis. In order to be suitable for publication, the introduction for each

of the experimental chapters is relatively short in comparison with a conventional thesis. More extensive introductory material is therefore presented in this chapter and Chapter 2, with Chapter 6 providing discussion to draw the findings of the experimental chapters together. Each experimental chapter is also introduced with an explanation of how it fits within the thesis as a whole and how it links with the preceding and following chapters. The primary reason for submitting the thesis in paper format is so that the data can be more quickly and easily disseminated to the wider academic community. The peer-review process has also been beneficial in strengthening the paper presented in Chapter 3. Finally, I hope that, by publishing work before the thesis is submitted, this will be beneficial for my future career.

The remainder of this chapter sets out the generativist and constructivist approaches in detail. Evidence for and against each approach is then discussed, along with why, ultimately, the constructivist approach was chosen as the theoretical framework for the current thesis.

1.1 General assumptions of generativist (nativist) approaches

Generative approaches to adult grammar assume that sentences are 'generated' via a set of formal rules in the domains of both morphology and syntax. For example, combining a determiner (*the*) and a noun (*bike*) creates a noun phrase (NP; *the bike*) (sometimes determiner phrase, although I will use noun phrase for consistency), combing a verb (*ride*) with an NP creates a verb phrase (VP; *ride the bike*), etc. Combining elements in the correct order by following the rules generates grammatical sentences.

In principle, it is possible to have a generativist approach to language acquisition that is non-nativist. However, generative approaches often assume that at least some aspects of language must be innate. At the core of this argument is the complexity of the task facing young children in acquiring their native language(s): the input they receive from their environment could not possibly provide them with sufficient evidence to learn all of the rules, or restrict them appropriately. This argument is known as the 'poverty of the stimulus' (Chomsky, 1980): there are simply too many possible rules children might posit in trying to work out which ones apply to their language, despite hearing millions of words and sentences, if they did not start out with a set of basic assumptions.

1.1.1 Principles and parameters

Famously, Chomsky (e.g. 1981) argued for a Universal Grammar (UG). Unlike descriptive grammars of individual languages, UG was designed to highlight the features of grammar that relate to all languages. These universal features give children a starting point: there are features they know to be true ('principles'); others they know might be possible in their specific language ('parameters'); and things they know are not possible ('constraints'). All of these enable children to correctly interpret complex sentences, avoid grammatical errors, and become fluent speakers of their native language(s) within just a few years.

One universal principle relates to syntactic categories: all languages have certain syntactic categories, such as NOUN (although others, such as DETERMINER, are not universal). Baker (2003) suggests that children have innate knowledge that labels for objects are members of the NOUN category. Children are thus able to populate their NOUN category with the object labels they hear – *baby*, *teddy*, *light*, etc. Knowing which words belong to which categories helps children to parse speech and to set language-specific parameters. This, in turn, helps them to quickly assimilate new examples into the NOUN category, including abstract nouns, through the use of distributional regularities in the input (e.g. Valian, Solt & Stewart, 2009).

An example of a parameter, which varies across languages, is that of head direction. The 'head' of a phrase is usually what gives it its name, so a verb is the head of a VP, a preposition is the head of a prepositional phrase (PP), etc. English is a head-initial language: the verb comes at the start of the VP *[Monkeys]* ***eat*** *bananas*; the preposition comes at the start of the PP ***in*** *the garden*. In contrast, Hindi is a head-final language: the verb comes at the end of the VP *[Bandarom] kēlē* ***khānē*** *[monkeys bananas* ***eat****]*; the preposition (actually a postposition here) comes at the end of the PP *bagīcē* ***mēm*** *[garden* ***in****]*. As head direction varies across languages, children must use language input to work out whether to set this parameter to head-initial (e.g. for English) or head-final (e.g. for Hindi). For a summary of several other proposed principles and parameters, see Ambridge and Lieven (2011: 122-123).

Unlike principles and parameters, which set out what is certain or possible in a language's grammar, constraints tell a child what is *not* possible (Crain & Thornton, 2012). Constraints therefore help children to avoid producing grammatical

errors in their own speech and to avoid misinterpretation of complex speech in the input. An example of a constraint is that of structure dependence (e.g. Crain & Nakayama, 1987), which ensures that children posit rules based on the abstract internal structure of a sentence, rather than on the surface order of words. The operation of this constraint is often exemplified by the formation of complex *yes/no* questions in English. To change a simple declarative sentence containing the auxiliary *be* (1) into a *yes/no* question (2), the following rule is sufficient: invert the auxiliary verb and the subject NP.

(1) All the monkeys **are** eating bananas.
(2) **Are** all the monkeys ___ eating bananas?

However, to change a complex declarative sentence (3) containing a restrictive relative clause (italicised) into a grammatical *yes/no* question (4), a more specific rule is required. The declarative sentence contains two auxiliary verbs from which to choose, and moving the wrong one would result in an ungrammatical utterance (5).

(3) All the monkeys *who **are** playing* **are** eating bananas.
(4) **Are** all the monkeys *who **are** playing* ___ eating bananas?
(5) ****Are*** all the monkeys *who* ___ *playing* **are** eating bananas?

The error in (5) would result if a child posited what seems to be the simplest rule: invert the *first* auxiliary verb and the subject NP. However, thanks to the structure dependence constraint, children will never posit this incorrect rule as they will always take the abstract internal structure of the declarative sentence into account. The correct rule is therefore posited: invert the auxiliary verb *of the main clause* and the subject NP.

Under a generative approach, then, children acquire a set of rules, based on the innate principles and constraints of UG, and the parameters they have set through linguistic input. Because of this, children are able to rapidly acquire the grammar of their language and avoid errors in their speech.

1.1.2 Semantic bootstrapping

Pinker's (1989) version of semantic bootstrapping (see also Pinker, 1984) suggests that innate linking rules complement other forms of innate knowledge to help children begin the process of acquiring their language's grammar. Through UG, children are born knowing about syntactic categories (N, V), phrases (NP, VP) and syntactic roles (subject and object). In addition, children have innate knowledge of semantic (or thematic) roles, such as AGENT and PATIENT. The semantic bootstrapping hypothesis proposes that children also possess innate linking rules allowing them to map semantic roles onto syntactic roles, thus facilitating the acquisition of argument structure. For example, the agent of a causal action (e.g. *The dog* in *The dog chased the cat*) maps onto the subject role, while the patient (e.g. *the cat*) maps onto the object role. As it is possible to observe from the environment which is the agent and which is the patient and, assuming that they have acquired the lexical items *dog* and *cat*, children are able to link *the dog* to the syntactic role of subject and *the cat* to the syntactic role of object. This gives children the information they need to work out that English word order is subject-verb-object (SVO).

Semantic bootstrapping is proposed as a mechanism that enables children to break into UG and start putting in place the grammar of their language. Once children have worked out that English has SVO order, for example, they will begin to comprehend less concrete examples whose interpretation is not immediately obvious from the environment. Indeed, children are able to determine syntactic roles in sentences even when they have no idea what the sentences means, as in the ubiquitous example, *The situation justified the measures*. Semantic bootstrapping, then, allows children to break into their innate knowledge and use it to build the grammar of their language. Once this grammar is built, they are able to use distributional analysis to work that, for example, *situation* must be a noun.

1.2 General assumptions of constructivist approaches

In contrast to nativist approaches, constructivist accounts of language acquisition assume no innate, language-specific knowledge. Instead, these approaches suggest that general cognitive systems, such as categorisation, enable children to construct a grammar from the input they receive. Children are not a 'blank slate', but they do not

require language-specific mechanisms to be encoded in their genes in order to be able to learn language: the input they receive is argued to be perfectly sufficient for children to succeed in acquiring their native language (i.e. there is no 'poverty of the stimulus'). Features of the input are also key in explaining the patterns of errors in children's language production that prove troublesome for nativist approaches.

Rather than building sentences by following a series of rules, as in generative grammar, constructivist approaches posit a series of 'constructions' into which lexical items can be placed to form sentences. For example, the construction for a simple transitive-causative sentence might be something like N1 V N2, where N1 acts on N2, causing N2 to be affected in some way. So, in the sentence *Bob annoys Wendy*, *Bob* (N1) acts on *Wendy* (N2) causing her to become annoyed. The approaches outlined below illustrate different theories of how children might acquire a construction grammar. Importantly, though, none of these theories rely on underlying, innate grammatical knowledge to account for the data.

1.2.1 Early constructivist accounts

One of the earliest instantiations of a construction grammar is Braine's (1963) 'pivot grammar'. Through examination of children's earliest 2-word utterances (when they are at the very beginning of syntax development), Braine noticed that these utterances tend to be fairly limited for the first few months, before increasing exponentially from 5 or so months after the first combinations appeared. This 'first phase' of multi-word speech seemed to be characterised by a limited number of words that always occurred in first or second position, and a greater variety of words with which they combined, e.g. *see boy*, *see sock*, *see hot*; *byebye plane*, *byebye man*, *byebye hot*; *boot off*, *light off*, *water off*. Braine termed these frequently-used words 'pivots', onto which more flexible 'X-words' could attach. X-words were essentially all words in the child's vocabulary except for the pivots, and could thus appear (in principle) with any pivot word. Hence, children were provided with a simple way to construct a number of combinations of the types *pivot + X* or *X + pivot*.

However, Brown (1973) argued that the three children studied by Braine appear to constitute the only evidence for pivot grammars, with children in other studies not conforming to these patterns (although he did note that the other children discussed were at a more advanced stage of development). Firstly, what constitutes a

pivot varies widely from child to child, and even within children (verbs, pronouns, prepositions, adjectives…). Secondly, one of the defining characteristics of a pivot word is that it only occurs in one position (first or second), but this did not appear to be the case for all children (e.g. Bowerman's 1973 cross-linguistic study). Pivot words also appear on their own (i.e. not in combination with an X-word) and some combinations consist of two X-words or even two pivots. A pivot grammar strategy cannot, therefore, be a universal strategy used to acquire syntax. One final problem facing Braine's pivot grammar is that, even if children do start out with a series of *pivot + X* schemas, it is not clear how they would be able to move from this to more advanced stages of syntactic development. The idea of pivot grammar has therefore been built upon by Braine (1976), Bowerman (1976) and Maratsos and Chalkey (1980), amongst others, to create accounts of development.

Further developing the idea that children's initial syntactic knowledge develops from lexically specific schemas, Tomasello (1992) outlined the Verb Island Hypothesis. Tomasello argued that the sort of fixed word-order expressions characteristic of pivot grammar do not reflect syntactic knowledge at all (cf. Ninio, 2014). Rather, the words simply appear in the order that children have heard them in the input. Moving from this stage to abstract syntactic knowledge is a complex task. Children begin with lexically specific knowledge: "in English, when you say *eat*, you first say the person who's doing the eating, then *eat* itself, then the thing being eaten". This leads to a schema such as [eater *eat* eatee]. The child builds a variety of lexically-specific schemas: [chaser *chase* chasee], [kicker *kick* kickee], [jumper *jump*], [talker *talk*], etc. By extrapolating across these schemas, something more abstract results: the first person mentioned in all of these schemas is the one initiating the action, i.e. the *agent*. Similarly, when a second person/object is given, they are often affected by the action; noticing this allows children to create a *patient* category. Words that come between the agent and patient are the action words (i.e. the verbs). (Other information can also help to form categories: words ending in *–ing* can also help form a verb category, for example. Morphology is especially important for languages with case marking, and vital if word order is free.) Eventually, extrapolating from verb-specific schemas allows for the creation of completely abstract constructions of the N1 V N2 type.

1.2.2 More recent constructivist approaches

More recent constructivist approaches (e.g. Tomasello, 2003) are based on the theories outlined above, although they do not posit such a reliance on verbs. Nevertheless, at the core of these approaches is the theory that children's syntactic knowledge is built on lexically specific items (fixed phrases such as *I want it* or *I'm doing it*), which are generalised to schemas (such as *I want X* or *I'm X-ing it*). The importance of frequency in the input is also more firmly established, explaining why certain constructions are more likely to be learnt early and also why lexical effects are so often observed in children's early language development.

Evidence from corpus and diary studies supports the view that children's earliest multi-word utterances are lexically constrained (e.g. Dąbrowska & Lieven, 2005; Lieven, Pine & Baldwin, 1997; Pine & Lieven, 1993; see e.g. MacWhinney, 1975, for cross-linguistic evidence). For example, Pine, Lieven and Rowland (1998) studied the early combinations of 12 children, finding a lack of overlap in the main verbs used with different auxiliaries and the nouns used as subjects and objects in transitive sentences. These effects did not appear to reflect the input directly, however, with children producing a disproportionate number of sentences with *I* in subject position compared to child-directed speech. Lieven, Behrens, Speares and Tomasello (2003) analysed the multi-word speech of a single 2-year-old child. Of the utterances the child had not produced before (which accounted for only 37% of the data), the majority required only a single change from a previously-produced utterance. Many of these changes involved the substitution or addition of a noun. Thus, many of this child's utterances appeared to be based on fixed phrases or simple schemas of the *I want X* type.

Evidence from experimental studies also provides support for lexical effects in language development. In Childers and Tomasello's (2002) elicited production study, children aged 2;4-2;10 underwent training with real verbs in SVO sentences, either containing pronouns or full noun phrases (e.g. *He's pulling it*; *The cow's pulling the car*). At test, participants were required to use a novel verb to describe a new scene, but only those who had received training including pronouns were able to do so (e.g. *He's meeking it*). This finding demonstrates that, although young children have knowledge of the SVO construction, their ability to use it is affected by the

frequency of the frame in which the verb occurs (although see Fisher, 2002, for an alternative explanation).

Evidence from additional experimental studies suggests that fixed phrases and schema play a role in language development in other ways. Bannard and Matthews (2008) tested 2- and 3-year-old children's ability to repeat phrases, such as *sit in your chair*, which is heard frequently in the input as a four-word phrase, and *sit in your truck*, which is not. They found that children were significantly more accurate at repeating the frequently heard sequences than the less familiar ones. Following up on this, Matthews and Bannard (2010) used a similar task to test the reality of slot-and-frame schemas in children's language. They chose phrases such as *a piece of X*: frequently heard three-word phrases (the 'frame') that varied in terms of the final word (the 'slot'). Children were better able to repeat four-word phrases when the words that fill the slot are more variable in the input and, thus, led to children having a more abstract schema. Together, these studies support the view that children store phrases such as *sit in your chair* as whole units, but that they also store phrases such as *a piece of X* as slot-and-frame schemas.

Through the processes of analogy and distributional analysis, children build on lexically specific knowledge to create fully abstract constructions, such as the transitive N1 V N2 construction. Showing support for the psychological reality of analogy and distributional analysis, evidence from artificial language learning studies has demonstrated that infants are able to learn simple 'grammars' through distributional analysis, and that they are able to generalise these to new instances (see Gómez & Gerken, 2000, for a review; for similar findings with adults, see e.g. Altman, Dienes & Goode, 1995; Reber, 1969; 1989). Gómez and Gerken (1999) trained 11- and 12-month-olds on an artificial grammar, using the head-turn preference procedure. In the training phase, infants listened to strings of 'words' which conformed to the grammar. At test, infants listened to new strings that either did or did not conform to that same grammar. Infants listened significantly longer to the strings which conformed to the grammar they had just learned than to the 'ungrammatical' strings. This demonstrated that, with just a few minutes of exposure, children were able to learn the rules of a very simple grammar and generalise these rules to new examples.

Before reaching an adult-like stage of syntactic abstraction, children may create a number of different 'transitive' schemas. Ambridge and Lieven (2015) suggest the following examples (although they stress that these have not been tested):

Contact (non-causative)	[AGENT] [ACTION] [PATIENT] (*John hit Bill*)
Causative	[CAUSER] [ACTION] [CHANGE] (*John broke the plate*)
Experiencer-Theme	[EXPERIENCER] [EXPERIENCE] [THEME] (*John heard Bill*)
Theme-Experiencer	[THEME] [EXPERIENCE] [EXPERIENCER] (*John scared Bill*)
"Weigh" Construction	[THING] [MEASURE/COST/WEIGH] [AMOUNT] (*John weighed 100lbs*)
"Contain" Construction	[CONTAINER] [CONTAIN] [CONTENTS] (*The tent sleeps four people*)

The examples at the top of this list (contact and causative) are the most frequent, and therefore prototypical, examples of the transitive construction, whereas those at the bottom are the least prototypical (for discussion of prototypes in language, see Ibbotson & Tomasello, 2009).

1.3 Evidence for nativist and constructivist approaches

Innate knowledge of language might seem a reasonable way to explain children's rapid, and relatively error-free, acquisition of their native language. However, this approach struggles to explain some features of children's language. Logically, once a rule has been acquired, the child should never make a mistake with that grammatical structure again. However, the following evidence suggests that this is not the case: error rates tend to vary across different lexical items, and that this variation is not random. Note that some generativist theories (e.g. Head-Driven Phrase Structure Grammar, Pollard & Sag [1994], Lexical-Functional Grammar, Bresnan [2001]) are better able to deal with lexical effects than those which assume a full dissociation

between syntax and the lexicon (e.g. minimalism, Chomsky [1993], X-bar theory, Jackendoff [1977]).

In an experimental study, Kidd, Lieven and Tomasello (2006) asked children aged 2;10 to 5;9 to repeat sentences containing sentential complements (e.g. *I think she is riding away on a horse*). Half of the sentences were grammatical and half ungrammatical (e.g. **I think him running away from the dog*). The complement-taking verbs were either high-frequency (e.g. *think*) or low-frequency (e.g. *pretend*) in corpora of child-directed speech. Results showed that children were better able both to repeat grammatical sentences and to correct ungrammatical sentences containing high-frequency than low-frequency complement-taking verbs. These findings suggest that frequency information plays a role in language acquisition, something that is not predicted under nativist accounts of language acquisition. Similar lexical effects have been observed, in both corpus and experimental studies, with other modal and auxiliary verbs (Pine et al., 1998; Rowland & Theakston, 2009; Theakston, Lieven, Pine and Rowland, 2005), negation (Cameron-Faulkner, Lieven & Theakston, 2007), and inflection (Wilson, 2003).

Further evidence of lexical effects have been demonstrated at sentence level through the use of 'weird word-order' studies, with both novel and known verbs (Akhtar, 1999). English has SVO word order (cf. Hindi, above, which has SOV word order). In weird word-order studies (in English), the experimenter describes an event (e.g. a doll pushing a toy car) using an ungrammatical sentence in which the word order has been changed from SVO (*Dolly pushed the car*) to SOV (**Dolly the car pushed*). Children are then encouraged to describe a version of the same event. The phenomenon of interest is whether children will imitate the experimenter's ungrammatical word order, or 'correct' it, by reverting to SVO. Matthews, Lieven, Theakston and Tomasello (2005; see also Abbot-Smith, Lieven & Tomasello, 2001) tested children aged 2;9 and 3;9 with English verbs of high, medium and low frequency in the input (e.g. *push, shove* and *ram*, respectively). They found that 2-year-olds were more likely to imitate the ungrammatical word order with low-frequency verbs than medium- or high-frequency verbs. 3-year-olds, on the other hand, were more likely to revert to the grammatical SVO order than to imitate the ungrammatical word order with verbs of any frequency. Not only has this study demonstrated lexical frequency effects, but it provides evidence of much more gradual acquisition of syntax than an innate principles-and-parameters approach

would allow for (as the head direction parameter would either have been set or not, and could not be in an intermediate state).

The bootstrapping proposals were designed to counter another issue for the UG account, namely how children are able to use this knowledge to analyse the strings of sounds they hear in their input. However, these proposals themselves are not without their own problems. For example, semantic bootstrapping appears to work quite well for canonical word orders in English (and children rarely hear examples of the passive, for example, in which the order of agent and patient is reversed: *The dog chased the cat/The cat was chased by the dog*). However, so-called ergative languages do prove problematic for Pinker's (1989) proposal. Nominative-accusative languages like English treat the subjects of both transitive and intransitive sentences in the same way, with the objects of transitive sentences being treated differently. This can be seen in the case-marking system (note that case is only explicit on pronouns in English): the subjects of both transitive and intransitive sentences have nominative case (e.g. *He chased Lisa; She ran*), whereas the objects of transitive sentences have accusative case (e.g. *Bart chased her*). In contrast, in ergative-absolutive languages, such as Basque, the subjects of transitive sentences are treated in one way (they have ergative case), whereas the subjects of intransitive sentences are treated in the same way as the objects of transitive sentences (they have absolutive case). In the following example, note the different case markings on the first person singular pronoun and the determiner (adapted from Ezeizabarrena and Larrañaga, 1996: 959):

(6) Nik lagun bat ikusi dot
I$_{[erg]}$ friend one$_{[abs]}$ seen has
'I have seen a friend'

(7) Lagun bat etorri da
Friend one$_{[abs]}$ come has
'A friend has come'

(8) Lagun batek ni ikusi nau
Friend one$_{[erg]}$ I$_{[abs]}$ seen has
'A friend has seen me'

If children always map AGENT/ACTOR to the subject role, they might assume that all initiators of the action/event, i.e. those in the subject position, would have the same case marking. In this sense, ergative languages may be problematic for the linking rules proposed in Pinker's (1989) semantic bootstrapping account. The proposed solution to this is that, rather than marking participants with the same syntactic roles (as in accusative language), the morphology of ergative languages mark the similarity of semantic roles. For example, the object of the verb *see* is a theme (*Lisa saw him*), as is the subject of the verb *laugh* (*He laughed*); in ergative languages, the pronouns in both of these sentences would have the same (absolutive) case markings. However, the real problem for the semantic bootstrapping hypothesis comes from 'split-ergative' languages (for a brief summary, see Pye, 1990: 1294). These languages sometimes act like accusative languages and sometimes like ergative languages. One example of a split-ergative language is Dyirbal, an Australian Aboriginal language, which behaves like an ergative language with nouns and third person pronouns, but like an accusative language with first and second person pronouns (Dixon, 1979: 63). It seems impossible for innate linking rules to be able to successfully explain children's acquisition of these languages (or, indeed, how such languages would have evolved in the first place, if innate linking rules link syntactic and semantic roles consistently).

Constructivist approaches are, of course, not without their critics. Arguing against the assumption that children's first combinations primarily consist of rote-learned, fixed phrases, Ninio (2014) presents an analysis suggesting that early 2-word combinations are, in fact, syntactic phrases resulting from children's use of productive rules. Ninio analyses the telegraphic speech of young children (e.g. *want bottle*, *bring chair*), focussing on their ungrammatical use of bare nouns (i.e. nouns which, in adult speech, would require a determiner in order to be considered grammatical; *want [a] bottle*, *bring [the] chair*, etc.). Analysis of child directed speech shows that adults do not make this kind of error when addressing children (although they do produce a number of grammatical, 2-word verb + noun utterances, such as with proper and plural nouns), so the errors in children's speech cannot be due to imitation. At the same age, these children produce large number of determiner + noun 2-word utterances, so errors also cannot be explained by children simply not being able to use determiners. Ninio suggests that children are using a productive rule, of verb + single-word object to create these telegraphic utterances, which she

posits is derived from hearing grammatical verb + noun and verb + pronoun utterances in the input. Ninio argues that these findings are best explained by generativist accounts and, indeed, they seem to count against the fixed-phrase stage of acquisition suggested by constructivist approaches. She suggests that performance limitations also play a role. It should be noted, however, that these findings do not necessarily count against a more abstract schema, such as action + thing, which might be formed under a constructivist account. (The lack of specificity in constructivist accounts in terms of *which* constructions/schemas children form, is also problematic for current versions of this theoretical approach.)

Inherent to constructivist approaches to language acquisition is the assumption that young children's knowledge of syntax is lexically restricted and, therefore, not fully abstract. These approaches therefore face challenges from studies suggesting that, in fact, even very young children have abstract syntactic knowledge. Gertner, Fisher and Eisengart (2006) used an intermodal preferential looking paradigm (IPLP) to test if children were able to use word order to determine the visual scene to which a transitive sentence containing a novel verb was referring. Children aged 21 and 25 months old watched videos in which a bunny and a duck performed novel actions on each other (e.g. in picture one, the bunny was pulling the duck along by its legs while the duck lay in a wagon; in picture two, the duck tipped the duck, who was sitting in a rocking chair). The children then heard the sentence *The duck is gorping the bunny*. Even the youngest children looked significantly longer at the correct picture, indicating that they were able to use abstract syntactic knowledge to understand that *The duck is gorping the bunny* meant that the duck is doing something to the bunny, and not vice versa.

While Gertner et al.'s (2006) study demonstrated that very young children are indeed able to use some sort of abstract syntactic knowledge to interpret sentences containing novel verbs, other studies using similar paradigms have shown that this knowledge is nevertheless restricted. Chan, Meints, Lieven and Tomasello (2010) used a similar paradigm to Gertner et al. (2006), except that the two videos showed identical actions, just with the agent and patient roles reversed. Under these conditions, 24-month-old children were unable to select to the correct interpretation, and it was not until they were 33 months olds that children performed above chance on this task. Noble, Rowland and Pine (2011) replicated the findings of Gertner et al. (2006) with novel transitive verbs using a forced choice pointing paradigm with

children aged 27, 31, 40 and 51 months old (note that all of these children were older than those in Gertner et al.'s 2006 study). This paradigm is similar to IPLP except that children have to make a conscious choice to point at the scene that they think matches the sentence. However, children did not perform above chance in a conjoined agent intransitive condition (e.g. *The duck and the bunny are daxing*) until after the age of three years. Finally, Fisher (1996) tested children's ability to comprehend giver and receiver roles. Watching a single video of two female participants exchanging an object, the children heard a sentence such as *She's trasking the balloon to/from her*. They were then asked to *Point to the one who was trasking*. Children were able to do this in the canonical *to* condition, but not the more unusual *from* condition, even when they were as old as 5 years.

Taken together, the IPLP and pointing tasks described here indicate that, while children as young as 21 months old demonstrate some abstract knowledge of canonical transitive constructions, knowledge of verb argument structure continues to develop over the third year of life. With more complex, non-canonical ditransitive events, this knowledge is not yet adult-like at the age of 5 years. Continuing development of argument structures is compatible with a constructivist view of language acquisition, but not so easy to explain under a generativist-nativist view.

Syntactic priming studies have allowed very young children to demonstrate their abstract syntactic knowledge in production. Syntactic priming (or structural priming) refers to the phenomenon that children and adults tend to re-use syntactic constructions that they have just heard (for reviews, see Branigan, 2007, and Pickering & Ferreira, 2008). Of particular relevance to the question of abstractness of syntactic knowledge is the assumption that priming is only possible when the person being primed has an abstract syntactic representation of the sentence structure in question (Branigan, Pickering, Liversedge, Stewart & Urbach, 1995). Thus, a child who did not have an abstract representation of the passive structure could not be primed to produce a passive sentence (e.g. *The banana was eaten by the monkey*). In one priming study, Messenger et al. (2012) took turns describing pictures with children as young as 3 years old. Contrary to previous findings that children up to the age of 7 struggle to comprehend passive sentences with certain types of verb (Maratsos, Fox, Becker & Chalkey, 1985), Messenger et al. (2012) found that priming occurs in these young children for the passive structure, irrespective of the verb type used in the prime sentence. Similar findings have been demonstrated with

passive, transitive and dative constructions (respectively, Bencini & Valian, 2008; Huttenlocher, Vasilyeva & Shimpi, 2004; Thothathiri & Snedeker, 2008).

However, Savage, Lieven, Theakston and Tomasello (2003) demonstrated that, while priming was present in the 3-year-olds they tested, this was only the case when there was a large degree of lexical overlap between prime and target. That is, when these young children were primed with a sentence containing pronouns (e.g. *It got pushed by it*) they were able to produce a passive target (e.g. *It got cut by it*). They were far less likely to produce a passive target when primed with a sentence containing full noun phrases (e.g. *The bricks got pushed by the digger*), which could not be reused in their own sentence. These results are therefore compatible with a constructivist interpretation whereby the children are using a semi-abstract slot-and-frame construction (e.g. *It got Xed by it*) to complete the priming task in the high lexical overlap condition.

1.4 Summary

This chapter has discussed two opposing views of language acquisition: generativist-nativist accounts and constructivist accounts. The literature reviewed here demonstrate that both approaches appear to have some merit. On balance, however, the generativist view is simply not able to explain critical aspects of language development. Key amongst these are the apparent lexical effects that are readily observable in children's acquisition data. The ability of constructivist approaches to explain these effects means this is the approach that must be adopted to explain the by-verb differences in the pattern of results in the studies reported in this thesis.

Chapter 2 presents a review of the literature of particular interest: overgeneralisation errors. Children observe patterns in the input, such as adding *–ed* to the end of English verbs to create the past tense (*walk**ed***, *jump**ed***, *stopp**ed***, etc.). This pattern can then usefully be generalised to other, newly acquired verbs to create forms such as *lik**ed*** and *dropp**ed***. However, children often over-apply these generalisations and create ungrammatical, overgeneralised forms such as **break**ed*** and **sleep**ed***. This type of morphological overgeneralisation error is frequently found in children's early language. The type of overgeneralisation under investigation in this thesis is a little more complex: that of argument structure overgeneralisation errors, such as **I filled toys into the box* and **She giggled me*.

1.5 The roles of authors in papers

Chapter 3 is a published paper (Bidgood et al., 2014). The conception and design of the experiment were undertaken by all authors. Amy Bidgood collected the data, performed initial analyses and wrote the first draft of the manuscript. Additional analyses and interpretation of the data were undertaken in collaboration with the other authors, as was revision of the manuscript.

Chapters 4 is a paper in preparation for submission (Bidgood, Ambridge, Pine, Rowland & Freudenthal, in prep.). The conception and design of the experiments were undertaken by the first four authors. The final author calculated the frequency counts from the corpus. Amy Bidgood collected the experimental data, performed initial analyses and wrote the first draft of the manuscript. Additional analyses and interpretation of the data were undertaken in collaboration with Ambridge, Pine and Rowland, as was revision of the manuscript.

Chapters 5 is a paper in preparation for submission (Bidgood, Ambridge, Pine & Rowland, in prep.). The conception and design of the experiments were undertaken by all authors. Amy Bidgood collected the experimental data, performed initial analyses and wrote the first draft of the manuscript. Additional analyses and interpretation of the data were undertaken in collaboration with the other authors, as was revision of the manuscript.

Chapter 2: Argument structure overgeneralisation errors

2.1 Overgeneralisation errors

In English, certain pairs of constructions allow verbs to 'alternate' between them to express related meanings. (Note that I am not suggesting that either construction is derived from the other; see Goldberg, 2002). Good examples of alternating pairs are the dative and locative constructions (1 and 2, respectively):

(1a) Paul gave Mary the book.
(1b) Paul gave the book to Mary.
(2a) Christine sprayed water onto the flowers.
(2b) Christine sprayed the flowers with water.

Becoming aware of these patterns is useful in language development. Initially, children may observe that several verbs can alternate between a pair of constructions, as with the following examples of the locative alternation:

(3a) Toby splashed the floor with water.
(3b) Toby splashed water onto the floor.
(4a) Ali loaded the car with bags.
(4b) Ali loaded bags into the car.

As children become productive language users, not restricted to simply imitating utterances they have already heard, they are able to generalise the pattern to new locative verbs (e.g. Ambridge, Pine & Rowland, 2012). After hearing a sentence such as (5a), they may produce a novel sentence, such as (5b).

(5a) Frances packed the cupboard with food.
(5b) Frances packed food into the cupboard.

Becoming aware of patterns and generalising these patterns to new instances is a key characteristic of human language use.

However, this process is not always straightforward. The problem of interest in this thesis is how children know when to *stop* generalising a pattern. If they assume that any verb heard in a ground-locative sentence, such as (5a), may also be used in a figure-locative sentence, such as (5b), they are likely to produce erroneous sentences such as those in (6):

(6a) *Howard poured the cup with tea.
(6b) *Becca spilt the carpet with juice.
(6c) *Glen spread the toast with peanut butter.

Errors such as these result from a process of generalisation and overgeneralisation, and are therefore known as overgeneralisation errors.

Many diary studies have reported examples of argument structure overgeneralisation errors of various different types (see Pinker, Lebeaux & Frost, 1987, for a summary of those reported in several previous papers). The following examples are all taken from Bowerman (1996):

(7a) Dative: *I said her no. (c.f. examples 1a/b)
(7b) Locative: *Can I fill some salt into the bear [salt-shaker]? (c.f. examples 2a/b)
(7c) Causative: *I saw a witch and she disappeared them.

Examples such as these, although not particularly common, attest to the fact that at least some children go through a stage of producing argument structure overgeneralisation errors, before correctly restricting their generalisations and retreating from error. (N.B. Adults have also been reported to produce overgeneralisations, although the source of these is unlikely to be the same as that of children's errors; see Pinker, 1989, pp. 154-160.)

The remainder of this chapter lays out the problem in more detail. Theoretical approaches to children's retreat from overgeneralisation errors, along with empirical evidence, will then be presented, focusing on the three hypotheses that have most informed the current thesis: the semantic verb class hypothesis (Pinker, 1989), the entrenchment hypothesis (Braine & Brooks, 1995) and the preemption hypothesis (Goldberg, 1995). As will become apparent, while evidence exists for all three

accounts, the data cannot be explained by any one theory alone. The final section of the chapter therefore discusses research investigating semantic and statistical approaches simultaneously, and accounts that integrate the various approaches.

2.2 Baker's Paradox and the 'no negative evidence' problem

The difficulty in explaining children's retreat from overgeneralisation errors was termed *Baker's Paradox* by Pinker (1989; see also Baker, 1979): children cannot know that a certain verb cannot participate in an argument structure alternation simply because they have not heard the verb used in that way. In addition, children do not receive sufficient negative evidence to learn all of the exceptions to the rules of English grammar. The following paragraph explains the 'no negative evidence' problem in more detail.

Brown and Hanlon's (1970) classic study is the foundation upon which the 'no negative evidence' claim is based. They found that the parents of the children they studied were no less likely to accept their child's utterance (in terms of how they responded to it) if it was ungrammatical than if it was grammatical. The focus of parents' responses was on the semantics, not the syntax. However, several follow-up studies have questioned this original finding. Hirsh-Pasek, Treiman and Schneiderman (1984) replicated Brown and Hanlon's finding that parental acceptance did not vary with grammaticality of child utterance but carried out an additional analysis, illustrating that parents of 2-year-olds (although not older children) were more likely to repeat or recast their child's utterance if it was ungrammatical than if it was grammatical (see also Chouinard & Clark, 2003; Demetras, Nolan Post & Snow, 1986). Penner (1987) also found that parents were more likely to recast or expand on children's ungrammatical utterances (and, again, found that this was more frequent in parents of younger children), whereas they were more likely to move the topic on following a grammatical utterance. These findings show that parents are certainly sensitive to the grammaticality of their children's utterances.

In contrast to Brown and Hanlon's (1970) study, then, follow-up research has shown that some children do receive feedback on grammatical errors, albeit often in the form of indirect negative evidence. However, this cannot be the case for all utterances produced by all children (in the case of utterances produced only in the

company of other young children, for example). Perhaps more important is the question of whether children are able take on board the feedback they receive. Below is one of several conversations reported by Pinker (1989, p. 13; from Braine, 1971) illustrating that children are, at least in some cases, unable to take on board either implicit (lines 2 and 3) or explicit feedback:

(8) Child: Want other one spoon, Daddy.
Father: You mean, you want THE OTHER SPOON.
Child: Yes, I want other one spoon, please, Daddy.
Father: Can you say "the other spoon"?
Child: Other… one… spoon.
Father: Say… "other."
Child: Other
Father: "Spoon."
Child: Spoon.
Father: "Other… spoon."
Child: Other… spoon. Now give me other one spoon?

Examples such as (8) suggest that parental feedback alone cannot account for children's retreat from overgeneralisation.

However, Saxton and colleagues have found that some forms of feedback do seem to reduce children's production of ungrammatical utterances. Saxton, Kulcsar, Marshall and Rupra (1998) conducted an experimental study, testing children at two time points, five weeks apart. They found that children who had received negative evidence on the over-regularisation of irregular past tense verbs at the first time point (as opposed to positive input) were more accurate at the second time point. Saxton, Backley and Gallaway (2005) also found improvement over a 12-week period, but only for three of the 13 types of grammatical error they examined. In contrast to the positive effect suggested by Saxton's work, for at least some error types, Morgan, Bonamo and Travis (1995) found no evidence that recasts led to children producing more grammatical utterances and, in the long term, the number of recasts was actually a *negative* predictor of grammaticality (see also Morgan & Travis, 1989).

The above studies provide mixed support for the role of negative evidence in children's retreat from the production of various error types. It therefore follows that

negative evidence is unlikely to be the primary means via which children retreat from overgeneralisation errors, and therefore this unlikely to be a solution to Baker's Paradox. The remainder of Pinker's (1989) book (see also an earlier account in Pinker, 1984) is devoted to solving Baker's Paradox, laying out an account suggesting that children are able to use semantic information to retreat from, or avoid, argument structure overgeneralisation errors. This semantic verb class hypothesis is discussed in depth in section 2.3.1.

2.3 Theoretical approaches to explaining the retreat from overgeneralisation

This section discusses the three theoretical approaches which have been most influential to this thesis, and in the field more generally: the semantic verb class hypothesis (Pinker, 1989), the entrenchment hypothesis (Braine & Brooks, 1995) and the preemption hypothesis (Goldberg, 1995). Each subsection will give an overview of the theory itself as well as presenting empirical evidence in support of that theory. Section 2.4 discusses recent attempts to integrate these three theories into an approach that better explains all of the data.

2.3.1 The semantic verb class hypothesis

Pinker's (1989) semantic verb class hypothesis involves a mechanism whereby fairly broad semantic structures, known as thematic cores, are linked to particular verb argument structures via innate linking rules. Numerous verbs may have the same thematic core, and each verb has a separate lexical entry for each thematic core that relates to it. Using the locative constructions as an example, the thematic cores for the figure locative (e.g. *Christine sprayed water onto the flowers*) and ground locative (e.g. *Christine sprayed the flowers with water*), respectively, are as follows (p. 77):

(9a) X moves Y into/onto Z

(9b) X causes Y to change its state by means of moving Z to Y

Thus, verbs such as *spray*, which are able to alternate between the two locative constructions, have one lexical entry denoting that a substance is moved to a location

via the method of spraying and second lexical entry denoting that moving a substance to a container or location (the flowers, in the example above), via a spraying motion, has changed the state of that container/location (the flowers are completely sprayed with water). Relating the two locative constructions allows children to set up a broad-range rule linking them together (although the set of possible rules is constrained by the innate linking rules that Pinker proposes).

Some verbs have only one lexical entry for the locative construction. *Pour* is a figure-only verb (cf. *Howard poured tea into the cup/*poured the cup with tea*), so has a lexical entry related to (9a) but not (9b). In contrast, *fill* is a ground-only verb (cf. *Howard filled the cup with tea/*filled tea into the cup*) and this has a lexical entry related to (9b) but not (9a). If children posit that all verbs that can appear in one of these two constructions are able to appear in the other, as is the case for *spray* but not *pour* or *fill*, overgeneralisation errors will result.

To solve the problem of why only some verbs are able to alternate between two related constructions, like the ground and figure locative, Pinker proposed that there are, in fact, more specific, narrow semantic classes (sometimes referred to as subclasses) within each of the broad semantic classes. Some of the narrow classes allow alternation between the two locative constructions, via a narrow-range rule, whereas others do not. According to Pinker (1989, pp. 126-7) *spray* belongs to the alternating narrow class defined as "force is imparted to a mass, causing ballistic motion in a specified spatial distribution along a trajectory". As children learn other verbs with semantics fitting this definition, such as *splash*, *sprinkle* and *squirt*, they know that both locative constructions are licensed. *Pour* belongs to the figure-only narrow class defined as "a mass is enabled to move via the force of gravity". As children learn other verbs with semantics fitting this definition, such as *drip*, *shake* and *spill*, they know that only the figure locative is licensed. *Fill* belongs to the ground-only narrow class defined as "a layer completely covers a surface". As children learn other verbs with semantics fitting this definition, such as *cover*, *line* and *pad*, they know that only the ground locative is licensed.

In contrast to the broad semantic classes and broad range rules, then, once the semantics of a particular verb have been learnt sufficiently well to place it into one of the narrow semantic classes, a child will know whether or not the alternative construction is licensed. This allows for productivity in language use, since a child need only hear an alternating verb in one of the constructions (or, in fact, in neither

construction; see Chapter 3) to be confident in using it in the other. In order to retreat from error, the narrow semantic classes and narrow-range rules must be properly learnt; until this is the case, children's language will be error-prone.

The studies of Gropen, Pinker, Hollander and Goldberg (1991a, b) used the locative construction to test the predictions of Pinker's semantic verb class hypothesis. Gropen et al. (1991a) conducted experiments using both elicitation and comprehension methodologies to test the prediction of the semantic verb class hypothesis that lack of detailed knowledge of verb semantics is the root cause of overgeneralisation errors. According to the innate linking rule of object affectedness, the direct object of the verb is the one which is affected by an event. Thus, in the figure-locative sentence *Howard poured tea into the cup*, the tea is most affected, having been moved from its original location in the teapot to its new location in the cup. In contrast, in the ground-locative sentence *Mark filled the cup with tea*, the cup is most affected as it has changed from being empty to being full.

To test participants' knowledge of verb semantics, Gropen et al. (1991a) used a forced-choice pointing task with a series of line drawings. Participants were first introduced to the illustrations for each of the verbs in question using an ambiguous sequence of pictures. For example, a woman pouring water from a jug to a glass, which ends up full, could be accurately described as a *pouring* or a *filling* event. They were told that this was either *pouring* or *filling* (they received trials with each verb). The ambiguous picture was then replaced with two unambiguous ones: the *pouring* picture showed a woman pouring water but the water spilling, so the glass remained empty; the *filling* picture showed a woman dripping water from a tap into the glass, which ended up full. Participants were asked to point to *pouring* or *filling*, whichever was the target verb on that trial. While adults performed significantly above chance for all verbs, children (aged 2-5 years) were only able to do so consistently with figure-only *pour*-type verbs, often misinterpreting *fill* as meaning something more like *pour*.

To test participants' knowledge of verb argument structures, participants were asked what the woman was doing to the water (figure/contents-focussed question) or the glass (ground/container-focussed question) in order to elicit verbal descriptions using each target verb. Again, adults almost always produced grammatical utterances to describe the events. However, children were more likely to produce ungrammatical forms, particularly using *fill* in the figure locative

construction (e.g. **She's filling it into the glass* in response to the question, *What's the woman doing to the water?*). Older children (aged 4;6-5;11) also produced fewer ungrammatical responses than younger children (2;6-4;5). Gropen et al. (1991a) did not find evidence of contingencies between semantic knowledge and error rates in their first experiment (just described), but repeated the study focussing solely on the verb *fill*, since this seemed to cause the most problems in terms of interpretation. This time, for children aged 3;6-6;6, results showed that children who were biased to interpreting the meaning of *fill* as being related to a manner (e.g. pouring) as opposed to an end-state (being full) were significantly more likely to produce ungrammatical, figure-locative utterances containing *fill*, than those who were not biased to this manner interpretation. Thus, these results support the semantic verb class hypothesis, albeit on a very limited scale: better semantic knowledge about the verb *fill* led to fewer overgeneralisation errors involving that verb.

In a further test of the semantic verb class hypothesis, Gropen et al. (1991b) tested participants' ability to categorise new verbs in line with Pinker's (1989) broad semantic classes and whether children (aged 3-9 years) and adults would be able to use this knowledge to produce sentences containing the novel verbs in the expected locative construction. Participants were taught manner-of-motion verbs (e.g. *keating* = moving in a zig-zag motion), designed to be interpreted as a figure-only verb, like *pour*. If participants were able to interpret these verbs as intended, they should produce more figure-locative than ground-locative responses (e.g. *You're keating the marble to the cloth*, rather than *You're keating the cloth with the marble*). Participants were also taught end-state verbs (e.g. *mooping* = changing colour), designed to be interpreted as a ground-only verb, like *fill*. If participants were able to interpret these verbs as intended, they should produce more ground-locative than figure-locative responses (e.g. *You're mooping table with the cloth*, rather than *You're mooping the cloth onto the table*).

Results showed that both adults and children were able to use the new verbs as expected, based on manner-of-motion or end-state interpretations. This provides strong evidence for the semantic verb class hypothesis: children and adults were able to use verb semantics alone (participants only heard the verb in its gerund form in the teaching phase) to identify the correct locative construction with which to produce a sentence (based on other verbs with broadly similar semantics). Using semantic

knowledge to restrict argument structure generalisations when learning real verbs is therefore highly plausible.

2.3.2 The entrenchment hypothesis

The entrenchment hypothesis was first proposed by Braine and Brooks (1995) as a mechanism by which children may avoid or retreat from overgeneralisation errors. Entrenchment is an inference-from-absence mechanism. As children hear more and more instances of a particular verb being used, they are able to track the statistical distribution of that verb in the input. The more a particular verb is heard in a particular construction, the more that verb becomes entrenched in that construction. Of course, some verbs may be heard in several different constructions:

(10a) Ruby cut the cake.
(10b) Ruby cut the cake with a knife.
(10c) The cake was cut.
(10d) The cake was cut by Ruby.
(10e) This cake cuts easily.
(10f) Ruby cut me a slice of cake.
(10g) Ruby cut a slice of cake for me.
(10h) Cut the cake!

Children hear the verb *cut* used in different constructions, such as those in (10), but they are highly unlikely to hear the verb use in an ungrammatical sentence such as **The cake cut into pieces*. Having heard *cut* in numerous constructions, but never in this ungrammatical intransitive construction, children might infer that it is not possible for *cut* to be used this way. Realising that this is the case enables children to retreat from, or avoid, overgeneralisation errors, such as **The cake cut into pieces*.

Since different verbs are heard with different frequencies in the input, the entrenchment hypothesis predicts by-verb differences in error rates. According to this hypothesis, children are less likely to produce errors with a verb the higher its frequency in the input, as the inference-from-absence is strengthened each time a verb is heard. Similarly, the more frequently a verb has been heard in the input, the

less likely that verb is to be judged as being grammatically acceptable when it is heard in an overgeneralised construction.

Brooks, Tomasello, Dodson and Lewis (1999) examined the predictions of the entrenchment hypothesis using the causative alternation. Some verbs are able to alternate between the transitive-causative construction (e.g. *Robert rolled the ball*) and the intransitive-inchoative construction (in which no external agent is expressed, e.g. *The ball rolled*). Others are only grammatical in the transitive-causative (cf. *Robert hit the ball/*The ball hit*) while others still are only grammatical in the intransitive-inchoative (cf. *The girl laughed/*Robert laughed the girl*). Brooks et al. tested the prediction of the entrenchment hypothesis that children will make fewer errors with verbs that are more frequent in the input (i.e. those that are more entrenched in their grammatical constructions and for which a stronger inference-from-absence is therefore available). Children observed puppets performing 4 different actions. These actions were each described with two different verbs: one with early age of acquisition (AOA) and one with late AOA. (Verbs with an early AOA are likely to be more frequent in the input than those with a late AOA.) Two verb pairs were transitive-only (e.g. *hit* and *strike*; early and late AOA, respectively) and two were intransitive-only (e.g. *come* and *arrive*; early and late AOA, respectively).

Participants were asked a series of questions to elicit descriptions of the events containing the target verbs. Questions were either agent-focussed (e.g. *What is the [agent] doing?*; designed to encourage production of transitive sentences), patient-focussed (e.g. *What is happening with the [patient]?*; designed to encourage production of intransitive sentences) or neutral (e.g. *What is happening now?*). The results were in line with the predictions of the entrenchment hypothesis: children produced more than twice as many overgeneralised responses with late AOA than early AOA verbs. This study therefore provides strong support for the entrenchment hypothesis. It is also interesting to note that children produced, on average, three times as many erroneous transitive sentences with intransitive-only verbs as they did erroneous intransitive sentences with transitive-only verbs.

Theakston (2004) employed a grammaticality judgment methodology to test the extent to which adults and children (aged 5 and 8 years) would accept overgeneralisation errors with high- and low-frequency verbs as being grammatically acceptable. Children completed a forced choice task (acceptable or unacceptable),

whereas adults provided their judgments on a 7-point scale (1 = completely acceptable, 7 = completely unacceptable). Participants of all ages heard sentences from various different construction pairs, some of which used verbs in their generally accepted argument structure, whilst the other contained the same verb in an overgeneralised structure (e.g. *It fell off* vs. **Somebody fell it off*). For each high-frequency verb (e.g. *fall*), a low-frequency equivalent was chosen (e.g. *tumble*), and similar sentence pairs were constructed.

Results showed that children of both age groups were significantly more likely to accept overgeneralisation errors with low- than with high-frequency verbs. In addition, the 5-year-olds were more likely than the 8-year-olds (who had presumably had more exposure to the verbs in question in grammatical constructions) to accept overgeneralisation errors. Finally, the results for the adults were in line with those of the children, with adults judging errors with low-frequency verbs as being more acceptable than those with their high-frequency equivalents. Overall, then, the results from Theakston's (2004) study again provide strong support for the entrenchment hypothesis, this time employing a judgment methodology and investigating a wide range of constructions.

2.3.3 The preemption hypothesis

The preemption hypothesis (e.g. Goldberg, 1995) incorporates elements of both statistical and semantic information. In common with entrenchment, preemption involves hearing instances of a verb in its authorised constructions and using this information to infer that certain other constructions are not compatible with that verb. The critical difference is that one construction can only preempt the use of an alternative, ungrammatical construction if the two are roughly equivalent in meaning. In the following example, sentences such as (11a), but not (11b), preempt (and therefore help children avoid) the error in (11c):

(11a) Howard poured tea into the cup.
(11b) Water poured out of the burst pipe.
(11c) *Deborah poured the pan with oil.

In a similar way to the entrenchment hypothesis, hearing a verb like *pour* used many times in the figure-locative construction (such as 11a), a construction with similar semantics to the ungrammatical ground-locative construction (such as 11c), allows children to infer that it cannot be used in the latter construction because the relevant meaning is conventionally expressed using the former. Having a readily available, alternative construction that is semantically very similar thus allows children to avoid errors such as (11c).

Like the entrenchment hypothesis, by-verb differences are also predicted by the preemption hypothesis. In this case, the more frequently a verb has been heard in the grammatical construction that preempts the erroneous one, the less likely children are to produce that verb in the ungrammatical construction.

To test the predictions of the preemption hypothesis in a controlled manner, Brooks and Zizak (2002) taught children (aged 4 and 6-7 years) two novel verbs. The verbs each described a novel action similar in semantics to Pinker's (1989) manner-of-motion verbs (e.g. *roll*). Importantly for the purposes of this study, these verbs are able to alternate between the transitive and intransitive constructions (e.g. *The ball rolled*/*Robert rolled the ball*) and, thus, the novel verbs could be used 'grammatically' in either construction (e.g. *The tree is tamming*/*The mouse is tamming the tree*). Participants were assigned to one of three conditions: the Alternative Construction group; the English Suppletive group; or the No Preemption group. All participants heard one action described with a novel verb in the intransitive construction (e.g. *The tree is tamming*) and another with a novel verb in the transitive construction (e.g. *The rabbit is dacking the car*). Participants in the Alternative Construction group also heard the intransitive novel verb in the preempting periphrastic causative construction (e.g. *The mouse is making the tree tam*) and the transitive novel verb in the preempting passive construction (e.g. *The car is getting dacked*). In contrast, as well as hearing the novel verbs in the intransitive or transitive constructions, participants in the English Suppletive group heard the actions described with real (English) verbs (e.g. *swing*, *bounce*) with the opposite transitivity. Thus, participants in both the Alternative Construction group and the English Suppletive group were able to provide the agent of the intransitive action and omit the agent of the transitive action whilst keeping the transitivity of the novel verbs as those verbs had been taught. In contrast, participants in the No Preemption group heard no alternatives and thus, if the preemption hypothesis is

correct, would be more likely to alter the transitivity of the novel verbs in order to provide or omit the agent in the intransitive and transitive conditions, respectively.

As in Brooks et al. (1999), participants were encouraged to use the novel verbs through a series of agent-focussed, patient-focussed and neutral questions. Participants' responses were coded for transitivity. The 4-year-old children in the English Suppletive group were significantly less likely to violate the assigned transitivity of the novel verbs than those in the No Preemption group (there was no difference between those in the Alternative Cosntruction group and the No Preemption group for these younger children). The 6- to 7-year-old children in both the English Suppletive and Alternative Construction groups were significantly less likely to violate the assigned transitivity of the novel verbs than those in the No Preemption group. These results therefore show support for the preemption hypothesis, particularly for the older children, as those given an alternative means of providing the agent for an intransitive novel verb or omitting the agent for a transitive novel verb were less likely to generalise those verbs into the alternative construction, even though the verb's semantics might lead the participants to believe that this was possible.

Goldberg (2011) used corpus evidence to test whether the probability of hearing a verb in one construction, when discourse context might have lead the language learner to think a second construction would have been more appropriate, is sufficient to infer that that second construction is, in fact, ungrammatical. By expressing the preemption hypothesis in this way, Goldberg was able to test the evidence for an explicit mechanism of how preemption might operate, using evidence from the input (or, at least, using corpus evidence as a proxy for this). The alternation examined in this paper was the dative:

(12a) Double object [DO] dative: *Mel told Sue the news.*

(12b) Prepositional object [PO] dative: *Mel told the news to Sue*.

Not all verbs are grammatical in the DO dative: **Mel shouted Sue the news* (cf. *Mel shouted the news to Sue*). These are PO-only verbs. (Note that some verbs are grammatical in the DO but not the PO dative [*Mel bet Sue £5* vs. **Mel bet £5 to Sue*], but these were not examined by Goldberg.) To test her theory, Goldberg took the situation in which the recipient of an action (*Sue* in 12) was expressed by a pronoun,

but the theme (*the news* in 12) was expressed with a full NP. In this specific context, the DO dative is generally preferred (i.e. *Mel told her the news* > *Mel told the news to her*). This context therefore provides a good test case for Goldberg's theory: if a dative verb with a pronominal recipient and full NP theme is heard in the PO dative more than in the DO dative, it suggests to learners that the DO is not grammatical.

Goldberg's corpus evidence showed that, using this statistic (the probability of hearing a dative verb with a pronominal recipient and full NP theme in the PO dative, e.g. *Mel told the news to her*), the input alone provided enough evidence for language learners to be able to discriminate statistically between alternating and PO-only verbs. For alternating verbs, such as *tell*, the probability of hearing a verb with a pronominal recipient and full NP theme in the PO dative was, on average, 0.04. In comparison, for PO-only verbs, such as *explain*, the average probability was 0.83. For example, all else being equal, children might expect to hear *explain* in the DO dative (with pronominal recipient and full NP theme) around 96% of the time, like other dative verbs. However, they actually hear *explain* in this context in only around 1% of cases (Goldberg, 2011, p. 137). They can therefore conclude that the PO use must be ungrammatical: the DO uses in this unexpected context preempt the ungrammatical PO use of the verb *explain*. Thus, on the basis of discourse context and distributional information, this test case shows that children would be able to use the preemption mechanism to learn which dative verbs are able to alternate between the PO and DO dative and which are only grammatical in the PO dative.

2.4 Integrating semantic and statistical approaches

As demonstrated in section 2.3, the semantic verb class hypothesis (Pinker, 1989), the entrenchment hypothesis (Braine & Brooks, 1995) and the preemption hypothesis (Goldberg, 1995) all enjoy a certain amount of empirical support. However, none of these theories is capable of explaining all of the data: the semantic verb class hypothesis cannot explain by-verb differences related to input frequency, and neither the entrenchment hypothesis nor the preemption hypothesis is able to explain how children are apparently able to use verb semantics to select the correct construction in which to use a novel verb without having ever heard it in a sentence. This section first presents several studies that have simultaneously examined the predictions of two or more of the semantic verb class, entrenchment and preemption hypotheses. It

then describes one account, the FIT account, that attempts to integrate aspects of all three theories in order to provide a more complete picture of the mechanism(s) involved in children's retreat from overgeneralisation.

Brooks and Tomasello (1999) used a novel-verb elicitation task, similar to Brooks et al. (1999) and Brooks and Zizak (2002), to test the predictions of both the preemption and semantic verb class hypotheses with the causative alternation (*The ball rolled/Robert rolled the ball*). Children were taught two novel verbs. For half of the children, *meek* was taught as a transitive-only verb of causation of directed motion (e.g. *raise*). These children heard *meek* used only in transitive contexts (e.g. *The mouse is meeking the flower*). They were also taught *tam*, an alternating manner-of-motion verb. Children in this group heard *tam* only in intransitive contexts (e.g. *The car is tamming*). For the other half of the children, *meek* was taught as an intransitive-only verb of inherently directed motion (e.g. *rise*). These children heard meek used only in intransitive contexts (e.g. *The flower is meeking*). Again, they were also taught *tam*, an alternating manner-of-motion verb. However, children in this group heard *tam* only in transitive contexts (e.g. *The doll is tamming the car*). This enabled Brooks and Tomasello to test the semantic verb class hypothesis: *meek* always belonged to a non-alternating class, and the semantic verb class hypothesis predicts children's production of this verb would be limited to the construction in which it was modelled; *tam* always belonged to the alternating manner-of-motion verb class, and the semantic verb class hypothesis predicts children would produce this verb in both constructions, in appropriate contexts.

To test the preemption hypothesis, in a similar way to Brooks and Zizak (2002), half of the children (in the No Preemption group) heard each verb in only one context: transitive or intransitive. The other half of the children (in the Preemption group) heard each verb in two contexts: transitive and truncated passive (e.g. *The mouse is meeking the flower* and *The flower is getting meeked*) or intransitive and periphrastic causative (e.g. *The car is tamming* and *The doll is helping the car tam*). This gave children in the Preemption group a way to place the discourse focus of either verb on the agent or patient of the action without changing the verb's transitivity. Children in the No Preemption group were, of course, able to use alternative structures in their own production if they wished, but they had not heard the novel verbs modelled in this way.

Throughout the test sessions, children were asked agent-focussed, patient-focussed and neutral questions to elicit a range of responses from the children. Results showed that, overall, children tended to produce verbs with the transitivity in which they had been heard in training. In support of the semantic verb class hypothesis, children aged 4 and 6-7 years (although not children aged 2 years) were significantly more likely to produce sentences with the opposite transitivity for *tam* (from the alternating manner-of-motion verb class) than for *meek* (from the fixed-transitivity verb classes). In support of the preemption hypothesis, the oldest children (aged 6-7 years) in the No Preemption group produced significantly more responses of the opposite transitivity to that heard in training than did the children of the same age in the Preemption group. This effect was not found for the younger children. Brooks and Tomasello therefore conclude that both semantic verb class and preemption have an effect on children's overgeneralisation errors, and that this effect increases with age.

Again using the causative alternation (*The ball rolled/Robert rolled the ball*), Ambridge et al. (Ambridge et al., 2008; Ambridge, Pine, Rowland, Jones & Clark, 2009; Ambridge, Pine & Rowland, 2011) ran a series of experiments using a grammaticality judgement methodology to test semantic and statistical accounts simultaneously. In grammaticality judgment studies, participants typically watch an animation depicting an event. They then hear a sentence describing the event, which may, or may not, be generally considered to be grammatical. For example, they may see an event in which a boy tickles a girl and the girl laughs and then hear a description such as *The girl laughed* or **The boy laughed the girl*. Participants provide a grammaticality judgement rating for the sentence. Ambridge et al. typically asked children to provide their judgements on a 5-point 'smiley-face' scale (see Figure 2.1), whereas adults provided judgements either using the same scale or a simple numerical scale (1-5 or 1-7, where 1 is completely ungrammatical and 5 or 7 is completely grammatical). In some studies, adults did not watch animations but judged written sentences instead.

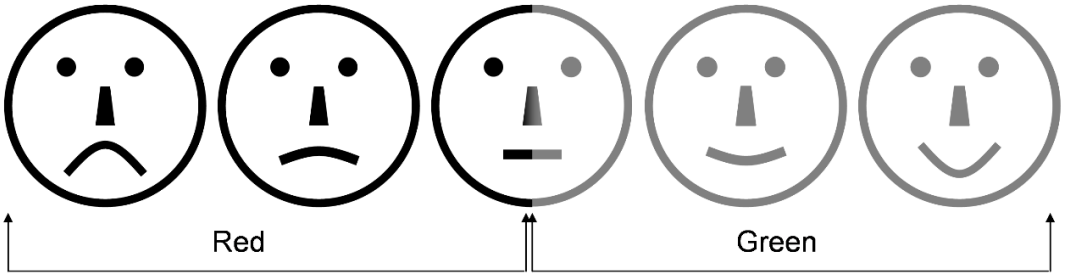

Figure 2.1. 5-point 'smiley-face' scale used in Ambridge et al.'s grammaticality judgement studies, where the face on left represents a completely ungrammatical sentence, the face on the right a completely grammatical sentence, and the remaining faces a rating between these two extremes.

Ambridge et al. (2008) collected grammaticality ratings from children (aged 5-6 and 9-10) and adults for grammatical intransitive sentences (e.g. *Bart fell into a hole*) and ungrammatical transitive sentences (e.g. **The man fell Bart into a hole*) containing either real or novel verbs. To test Pinker's semantic verb class hypothesis, the verbs were chosen from classes with a low degree of external causation (e.g. verbs of semi-voluntary expression of emotion; *laugh*) and with a higher degree of external causation (the prototypical meaning of the transitive-causative construction; e.g. verbs of going out of existence; *disappear*). In support of this hypothesis, participants judged transitive overgeneralisation errors with a novel verb from the class with a low degree of external causation (e.g. **The man laughed the boy*) to be less grammatical than such errors with a novel verb from a class with a higher degree of external causation (e.g. **The man disappeared the rabbit*). To test the entrenchment hypothesis, verbs of high and low frequency in the input, along with novel verbs, were included. In support of this hypothesis, participants judged transitive overgeneralisation errors with high-frequency verbs (e.g. **The man laughed the boy*) to be significantly worse, in comparison with their grammatical intransitive equivalents (e.g. *The boy laughed*), than such errors with low-frequency verbs (e.g. **The man giggled the boy*), in comparison with their grammatical intransitive counterparts (e.g. *The boy giggled*). Ambridge et al. (2008) concluded that both the formation of semantic verb classes and an entrenchment mechanism play a role in children's retreat from overgeneralisation errors.

Ambridge et al. (2011) used a similar method, although, this time, verbs were either intransitive-only (as in Ambridge et al. 2008) or alternating (e.g. *The ball*

rolled/The man rolled the ball). This had the advantage that participants could not use a strategy of consistently judging transitive sentences to be ungrammatical to give their responses. Children in this study judged sentences containing only novel verbs, whilst adults judged sentences containing both novel and real verbs. Consistent with earlier findings supporting the semantic verb class hypothesis, participants of all ages judged sentences containing novel verbs to be grammatical or ungrammatical in line with equivalent sentences containing known verbs from the same semantic class. Consistent with earlier findings supporting the entrenchment hypothesis, adults judged high-frequency real verbs to be significantly less grammatical in the ungrammatical transitive sentence than equivalent sentences containing their low-frequency counterparts. Ambridge et al. (2011) also manipulated the frequency of the novel verbs taught to adults, but no effect of entrenchment was observed with these verbs. The conclusion, again, was that both the formation of semantic verb classes and an entrenchment mechanism play a role in children's retreat from overgeneralisation errors.

Ambridge et al. (2011; see also Ambridge & Lieven, 2011) proposed the FIT account as an attempt to integrate semantic and statistical accounts. (Note that other accounts that integrate semantic and statistical elements have also been proposed, including Langacker [2000], MacWhinney [2004] and Tomasello [2003].) Under the FIT account, the sentence a speaker produces depends on various factors. These factors include the frequency of both the construction itself (higher-frequency constructions are more highly activated) and a particular verb in that construction, and the relevance of the construction in conveying the message. Frequency accounts, both entrenchment and preemption, are therefore accommodated under this account.

The key to the FIT account, though, and the factor that gives it its name, is the semantic compatibility (or 'fit') between the lexical items and the empty slots in the construction. For example, the slots in the transitive-causative construction are [AGENT] [ACTION] [PATIENT]. The prototypical semantics of these slots are gradually acquired through exposure to the input. Over time, children learn that the verbs that appear in the [ACTION] slot in transitive-causative construction are related to direct, external physical causation (e.g. *kick*). That means that new verbs children learn that also relate to direct, external physical causation are also compatible with the [ACTION] slot in the transitive-causative construction, and children will be able to deduce that they can use these verbs in this construction. In

contrast to the transitive-causative construction, the intransitive-inchoative construction contains the slots [PATIENT/ACTOR] [ACTION], where the prototypical semantics of the [ACTION] slot relate to *internal* causation. Again, the semantics of this slot are learnt over time, through exposure to the input (e.g. *The man laughed, The fish swam, The hot air balloon rose*). When children learn new verbs that also denote internal causation, they will be able to slot them into the [ACTION] slot in the intransitive-inchoative construction.

Some verbs, such as *roll* and *open*, have properties that relate to both internal and external causation. For example, it is only possible to roll a ball because of its round shape, and it is only possible to open a box because its lid allows us to do so. (Only unusual boxes would be able to be rolled, and only very unusual balls would be able to be opened.) Nevertheless, these actions usually require some external force to enable the actions to take place. Thus, verbs such as *roll* and *open* are semantically compatible with both the transitive-causative and intransitive-inchoative constructions and can be used in both.

The same argument applies to other constructions. For example, in the locative alternation (ground locative = *Frances filled the cupboard with food*; figure locative = *Frances placed food into the cupboard*), the ground locative construction has the following construction slots: [AGENT] [ACTION] [CONTAINER/LOCATION] with [CONTENTS]. In this construction, the semantics of [ACTION] slot relate to a change of state of the [CONTAINER/LOCATION] (e.g. the cupboard ends up completely full of food in the sentence *Frances filled the cupboard with food*); it is irrelevant how Frances moved the food to its new location. In contrast, the construction slots for the figure locative are [AGENT] [ACTION] [CONTENTS] into/onto [CONTAINER/LOCATION]. The semantics of the [ACTION] slot in this construction relate to the manner of motion of the contents when it is being transferred into the container. So, the manner in which the food enters the cupboard is different if it is placed/thrown/shoved into it; the cupboard may or may not end up in a full state.

As with the transitive-causative and intransitive-inchoative constructions, some verbs (e.g. *pack, spray*) have semantics that are compatible, to a certain extent, with the [ACTION] slots in both constructions and are, thus, able to be used grammatically in both. Nevertheless, the choice of construction is not irrelevant to the meaning of the resulting sentence – the ground and figure locative constructions

themselves carry meaning. For example, the ground locative sentence *Christine sprayed the wall with paint* implies that the entire wall ended up covered with paint, whereas the figure locative sentence *Christine sprayed paint onto the wall* does not. Thus, just because a verb is compatible with the [ACTION] slot in both constructions does not mean that both constructions are equally good choices to convey the message: if a change of state of the [CONTAINER/LOCATION] is important to the message, the ground locative is likely to be the best choice.

Importantly, the production of overgeneralisation errors can also be explained by the FIT account. If a child wants to express the causer of a *laughing* event, the grammatical choice would be the periphrastic causative (*The boy made the girl laugh*). However, this is a low-frequency construction and therefore has only low levels of activation. (Note that 'level of activation' here refers to the baseline level of accessibility of the construction in memory.) The intransitive-inchoative construction is high frequency (*The girl laughed*), but does not convey all aspects of the message. While the verb *laugh* is not semantically compatible with [ACTION] slot in the transitive-causative construction (no external causation is required by the verb *laugh*), the frequency of the construction itself and its ability to convey the entire message mean that this disadvantage may be overlooked, and the ungrammatical sentence **The boy laughed the girl* may be produced.

Ambridge and colleagues (Ambridge, Pine & Rowland, 2012; Ambridge, Pine, Rowland & Chang, 2012; Ambridge et al., 2014) have since used the grammaticality judgement methodology to test the predictions of semantic and statistical accounts, and the FIT account, with other constructions: the locative and the dative. By using mixed effect models to analyse the data, they were able to investigate the relative importance of the various semantic and statistical factors of interest. In addition, this allowed for the use of actual verb frequencies (both overall and in the relevant constructions) based on corpus counts, rather than relying on general high versus low frequency groups of verbs.

Ambridge, Pine and Rowland (2012) investigated the locative alternation (figure locative: *Lisa sprayed water onto the flowers*; ground locative: *Lisa sprayed the flowers with water*). While some locative verbs (e.g. *spray*) can appear in both the figure- and ground-locative constructions, others are figure-only (e.g. *pour*: *Lisa poured water into the cup*/**Lisa poured the cup with water*) or ground-only (e.g. *fill*: **Bart filled water into the cup*/*Bart filled the cup with water*). (Note that this study

used only known verbs and did not include novel verbs, which would have provided a stronger test of the semantics hypothesis.) Findings suggested that the data were best explained by a model that included both broad- and narrow-range semantic rules (see section 2.3.1, above) and overall verb frequency (entrenchment). Verb frequency in the preempting construction (in this case, the alternative locative construction) was found to have no dissociable effect from overall verb frequency, and therefore no support was found for a separate preemption mechanism. Although the paper reports support for Pinker's (1989) semantic classes, the semantic ratings collected in this study actually suggest that semantics should not be seen as categorical at all but, rather, should be viewed as a continuum. Variability in ratings also support this notion: sentences were not judged in a binary way as being grammatical or not, but on a scale, with some being slightly more or less grammatical than others. Thus, the probabilistic nature of the FIT account, outlined above, seems to capture the data better than Pinker's class-based proposal.

Ambridge et al. (2014) tested the predictions of the semantics, entrenchment and preemption hypotheses using the dative alternation (e.g. *Bill gave a present to Sue*/*Bill gave Sue a present*, c.f. *I said no to her*/**I said her no*). This study built on the new approach to semantics used in Ambridge, Pine and Rowland (2012). Instead of relying on the predetermined semantic classes proposed by Pinker (1989), Ambridge, Pine and Rowland collected semantic ratings from a group of adults (who did not participate in the grammaticality judgement task). Adults rated verbs based on the semantic features suggested by Pinker as being important for this alternation, but the researchers used Principal Components Analysis to produce a set of objective semantic criteria. This method provided a graded semantic measure, rather than the discrete classes of Pinker's original proposal. Findings showed that graded verb semantics predicted participants' grammaticality judgments. Thus, although Pinker's notions of the important semantic features for this alternation are likely to be correct (as the features against which verbs in this study were rated are those suggested by Pinker), the notion of discrete semantic classes is not the best way to capture these features. In terms of frequency measures, both entrenchment and preemption had dissociable effects on participants' grammaticality judgements. Unlike the case of the locative alternation, then, preemption *does* seem to play a role, over and above that of entrenchment, in the retreat from overgeneralisation errors with the dative construction. This finding is important in itself, as it indicates that different

mechanisms may be used to different extents in the retreat from error for different constructions. Overall, the model that best fits the data once again includes semantic and frequency information. Thus, an integrated account, such as the FIT account, is likely to be the way forward in terms of thinking about how children retreat from overgeneralisation errors.

Drawing together the findings of the grammaticality judgement studies of Ambridge and colleagues, a successful account must include roles for both semantic and statistical information, but must also be flexible enough to account for differences between different constructions. The FIT account fulfils all of these criteria.

2.4.1 Distinguishing between entrenchment and preemption

The discrepancies in the findings of Ambridge, Pine and Rowland (2012) and Ambridge et al. (2014) in terms of the roles of the entrenchment and preemption mechanisms deserves further comment. Entrenchment is an inference-from-absence mechanism: when a verb has been heard many times but has not been heard in a particular construction, a language learner may infer that that particular verb-construction pairing is not possible. Preemption takes into account both semantics and frequency: when a verb has been heard in a particular construction, but not in another related construction *with similar meaning* (e.g. the two locative constructions), a language learner may infer that the verb cannon be used in that second construction.

Stefanowitsch (2008) used adult grammaticality judgements of verbs in the two dative constructions to test the entrenchment and preemption hypotheses. He took corpus counts of 20 dative verbs that either appeared only in the double-object dative construction (e.g. *earn*: *Her books have earnt her a fortune/*Her books have earnt a fortune to her*) or the prepositional-object dative (e.g. *explain*: *He explained the procedure to them/*He explained them the procedure*). The corpus counts were used to calculate the degree of preemption and the degree of entrenchment (Stefanowitsch referred to this as *negative entrenchment*) for each verb. Findings showed a significant correlation between grammaticality judgements and degree of entrenchment, but no significant correlation between grammaticality judgements and degree of preemption. This finding contradicts that of Ambridge et al. (2014),

although the difference in methodology (correlation vs. mixed effects models) and the corpora from which frequency counts were obtained may go some way to explaining the different findings.

In an attempt to provide a more complete picture of the relative contributions of entrenchment and preemption to children's retreat from overgeneralisation errors, Ambridge, Bidgood, Twomey, Pine, Rowland and Freudenthal (2015) tested the predictions of the two hypotheses with several different alternations at the same time: intransitive-transitive (*The ball rolled/The man rolled the ball*); dative (*The boy gave a present to the girl/The boy gave the girl a present*); locative (*The boy sprayed paint onto the statue/The boy sprayed the statue with paint*); and active/passive (*The girl kicked the boy/The boy was kicked by the girl*). Children (aged 5-6 and 9-10 years) and adults provided grammaticality judgements for sentences from all four of these alternations. Results were in line with those of Stefanowitsch (2008): entrenchment was a significant predictor of grammaticality ratings, whereas preemption was not. However, Ambridge, Bidgood, Twomey, Pine, Rowland and Freudenthal (2015) acknowledge that preemption *does* appear to play a role "when the preempting construction is (a) particularly frequent relative to the error construction and (b) particularly closely synonymous with the error" (p. 17). Thus, a specific preemption effect is observed for the dative construction, for example (Ambridge et al., 2014). In conclusion, Ambridge, Bidgood, Twomey, Pine, Rowland and Freudenthal (2015) conclude that entrenchment and preemption should not be thought of as *mechanisms* so much as effects of a process of competition between constructions, such as that proposed under the FIT account.

When constructions are competing to convey the message of a boy causing a girl to laugh, the competing constructions might be the intransitive-inchoative (*The girl laughed*), the periphrastic causative (*The boy made the girl laugh*) and the (ungrammatical) transitive-causative (**The boy laughed the girl*). In avoiding producing the ungrammatical transitive-causative construction, an entrenchment effect may be observed as the speaker will have heard the verb used in the intransitive-inchoative and periphrastic causative constructions many more times than in the transitive-causative. This latter construction will therefore be activated less than the other two. However, a preemption effect may also be observed: the intransitive-inchoative construction, whilst much more frequent in the input than the periphrastic causative, may not be the most relevant, as it does not allow the speaker

to express the causer of the laughing event. Thus, in order to avoid the overgeneralisation error and still produce a sentence that expresses the entire message, the periphrastic causative would have to win out over both the transitive-causative and intransitive-inchoative alternatives. For verbs that appear only very rarely in the periphrastic causative construction, the activation level may not be high enough to block the production of the overgeneralisation error, **The boy laughed the girl*.

2.5 Summary

Evidence from the studies described in this chapter show some support for all three of the main hypotheses of interest: Pinker's (1989) semantic verb class hypothesis, Braine and Brooks's (1995) entrenchment hypothesis and, to a lesser extent, Goldberg's (1995) preemption hypothesis. However, what is clear is that no individual hypothesis is capable of explaining all of the data. What is needed is an approach that integrates both semantic and frequency information into account in its explanation of children's retreat from error, such as the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011).

The following three chapters present experimental evidence for an account that integrates semantic and statistical information, using various different methodological approaches. Chapter 3 examines the locative alternation (*Christine sprayed the flowers with water/sprayed water onto the flowers*), using a novel-verb grammaticality judgment study. Chapter 4 uses grammaticality judgments and error elicitation tasks to examine the transitive-intransitive alternation (*The ball rolled/Robert rolled the ball*). Finally, Chapter 5 goes beyond overgeneralisation errors to examine how verb-in-construction semantic compatibility influences children's production of a construction that is known to be problematic until a relatively advanced age: the passive. Chapter 6 concludes the thesis, summarising how the evidence gathered in the experimental chapters fits the theories in question, how this relates to nativist and constructivist approaches to language acquisition, and suggesting next steps for research in this field.

Chapter 3: The retreat from locative overgeneralisation errors: A novel verb grammaticality judgment study

3.0 Fit within the thesis

As discussed in Chapter 2, this thesis investigates three approaches to explaining how children retreat from argument structure overgeneralisation errors: semantics (Pinker, 1989), entrenchment (Braine & Brooks, 1995) and preemption (Goldberg, 1995). In addition, it will suggest how these three approaches might be integrated into a single account, such as the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011). This chapter investigates the locative alternation: *Lisa sprayed water onto the roses* (figure locative)/*Lisa sprayed the roses with water* (ground locative). One reason for choosing the locative construction as the first to investigate in this thesis is that children have been reported to make overgeneralisation errors in both directions, both ground-only verbs used in the figure locative (e.g. *I'm going to cover myself with a screen* → **I'm going to cover a screen over me*, Bowerman, 1982a) and figure-only verbs used in the ground locative construction (e.g. *I'm gonna pour water onto it* → **I'm gonna pour it with water*, Bowerman, 1981). For other often-studied construction pairs, such as the dative (*Bart gave Lisa the book/Bart gave the book to Lisa*), errors are usually only reported in one direction (e.g. *Don't say that to me* → **Don't say me that*, Bowerman, 1978); in contrast, attested errors in the locative alternation appear to be truly productive on both directions. A second reason for using the locative is that, in terms of testing Pinker's (1989) semantic verb class hypothesis, there are relatively subtle differences between the subclasses. This means that children might have difficulty realising which verbs are and are not able to alternate. If they prove unable to do so, this would be evidence against the semantic verb class hypothesis.

This chapter uses novel verbs to test the predictions of both the frequency and semantic approaches, although the use of novel verbs is a particularly appropriate, and stringent, way to test the importance of semantics. This is because, when participants judge the grammaticality of sentences containing novel verbs, they have no experience of hearing these verbs in sentences. They therefore have to judge the grammaticality of sentences based on semantics, which they have inferred from watching animations viewed during training. Novel verbs also allow us to use a

three-way contrast of verb frequency (high, low and novel) to test the statistical mechanisms in question.

The study in this chapter finds support for both semantic and statistical mechanisms in the retreat from overgeneralisation errors in the locative alternation. However, the locative is a relatively rare construction. Chapter 4 therefore investigates a much more frequent construction pair: the causative alternation (*John rolled the ball/The ball rolled*). This alternation also allows us to distinguish between the entrenchment and preemption hypotheses. Chapter 4 takes a different approach to semantics, moving away from discrete semantic classes towards continuum of semantic compatibility, more consistent with Ambridge and colleagues' FIT account. Alongside grammaticality judgments, Chapter 4 also moves on to investigate errors in production, and takes a different approach to investigating semantics.

This chapter has been published as a paper in *PLoS ONE* (Bidgood et al., 2014).

3.1 Introduction

As adults, we have the capacity for enormous creativity in language production: we often produce utterances that we have never heard. To reach this stage, children must acquire the grammar of the ambient language by forming generalisations about that language from the input. However, children must also learn to restrict these generalisations in order to avoid producing ungrammatical utterances (e.g. **I don't want it because I spilled it of orange juice* [= I spilled orange juice onto my toast]; Bowerman, 1981).

Pinker (1989) listed various grammatical constructions that have two alternating forms. The locative construction, for example, alternates between the ground- (or container-) locative, as in *The farmer loaded the wagon with hay*, and the figure- (or contents-) locative, as in *The farmer loaded hay into the wagon*. In the first sentence, the wagon is most affected, as it changes state from empty to full. In the second sentence, it is the hay that is most affected, as it is moved to a specific location; the wagon may or may not end up full. Pinker (1989, p. 79) described this change in how the event is construed as a "gestalt shift". (For earlier work on these constructions, see e.g. Hall, 1965; Fillmore, 1967; Anderson, 1971; Bowerman, 1982b; Levin & Rappaport, 1986).

When children hear verbs used in both the ground- and figure-locative constructions (*load*, *spray*, *stuff*, etc.), they may create a generalisation that any verb used in one of these constructions can also be used in the other, and this works well for some verbs. A child hearing *You splashed me with water*, a ground-locative construction, might generalise to the figure-locative construction to produce the grammatical utterance, *You splashed water onto me*. However, some English verbs, such as *fill* and *cover*, can only be used in the ground-locative construction (ground-only verbs) and generalising these verbs to the figure-locative construction would produce an ungrammatical utterance, such as **We filled toys into the box*. Conversely, some verbs, such as *pour* and *spill*, can only be used in the figure-locative construction (figure-only verbs). Generalising these verbs to the ground-locative construction would similarly produce overgeneralisation errors, such as **Daddy poured my cup with juice*.

One factor that could contribute to the retreat from overgeneralisation errors is parental feedback: so-called 'negative evidence'. It is undoubtedly the case that some parents provide feedback on errors that their children make, either through direct correction (e.g. C: **I filled mud into the hole*, M: *No, say "I filled the hole with mud"*) or implicitly, via rephrasing (e.g. M: *That's right, you filled the hole with mud*), facial expressions, misunderstandings or requests for clarification. Whilst evidence suggests that such feedback is helpful (Chouinard & Clark, 2003), children are unlikely to receive sufficient feedback of this type to account entirely for their retreat from overgeneralisation errors, particularly for low frequency verbs. Furthermore, some examples of parent-child interactions suggest that such feedback may have only a limited effect on children's language production (for reviews, see e.g. Pinker, 1989, pp. 9-14; Marcus, 1993).

The current paper investigates the extent to which two mechanisms constitute a solution to the 'no negative evidence' problem (Bowerman, 1988) and therefore explain the retreat from overgeneralisation with locative constructions. The first of these is Pinker's (1989) semantic verb class hypothesis: while evidence exists in support of this account, previous studies have primarily focussed on errors involving the transitive-causative and dative constructions, which, for reasons outlined in the following section, do not constitute as strong a test of the hypothesis. The second mechanism is statistical learning, in the form of entrenchment (Braine & Brooks,

1995) or preemption (Goldberg, 1995). Again, the locative alternation is a particularly good test of these hypotheses, as detailed below.

3.1.1 The semantic verb class hypothesis

Pinker's (1989) *semantic verb class hypothesis* attempts to explain how children's developing knowledge of verb semantics could explain the retreat from overgeneralisation errors. The proposed mechanism involves innate linking rules, which link generic semantic structures ('thematic cores') to verb argument structures: all verbs with the same thematic core are licensed in the same argument structure. These groups of verbs are known as *broad semantic classes*.

Some verbs, such as *spray* and *load*, can appear in more than one argument structure. Once children hear such examples, *broad-range rules* are formed (although the set of possible alternations is constrained by the innate linking rules). These allow verbs in related broad classes, such as figure and ground locative verbs, to alternate between the two structures. Until this point in development, learning is conservative and production is restricted to the use of verbs only in argument structures already heard by the child.

Of course, not all verbs that are grammatical in one locative construction are grammatical in the other, and this is due to idiosyncratic differences between verbs. Pinker (1989, pages 273-4) proposed that, by replacing "each idiosyncratic piece of information... with a parameter" and matching verbs on this more detailed level of semantics, *narrow semantic classes* (or 'subclasses') are formed. It is only membership in an alternating narrow class that enables a verb to be used grammatically in the other argument structure, via a *narrow-range rule*.

According to the semantic verb class hypothesis, the cause of children's overgeneralisation errors is that children do not initially have well-developed knowledge of verb semantics and do not necessarily know enough verbs in each narrow class for these classes to have been accurately formed. Thus, overgeneralisations occur as children occasionally apply the broad-range rule to some verbs to which a narrow-range rule would not apply. There is some evidence that children know that these productive forms are ungrammatical (1989, pp. 322-4). Children retreat from error as the operation of narrow-range rules gradually supersedes that of broad-range rules; the broad-range rules do remain in place,

however, and enable adults to produce 'Haigspeak' utterances (which the speaker/writer again does not necessarily consider to be grammatical, 1989, pp. 152-160).

Pinker (1989, pp. 126-7) specifies 15 narrow subclasses for locatives and allocates each of the 146 verbs to one of these subclasses (with two exceptions, *wrap* and *string*, which may each be the only members of their own respective subclasses). The defining semantics of each subclass specify whether the verbs contained within it can alternate between constructions, via a narrow-range rule, although even alternating classes have a bias towards one of the two constructions. Table 3.1 (adapted from Ambridge, Pine & Rowland, 2012, p. 262, based on Pinker, 1989, pp. 126-7) details the 15 subclasses.

Figure-(content-) oriented *(into/onto* verbs)	***Smear*-type, Alternating (*N*=10), designated reference category**. Simultaneous forceful contact and motion of a mass against a surface	*brush, dab, daub, plaster, rub, slather, smear, smudge, spread, streak*
	***Stack*-type, Alternating (*N*=3).** Vertical arrangement on a horizontal surface	*heap, pile, stack*
	***Spray*-type, Alternating (*N*=7).** Force is imparted to a mass, causing ballistic motion in a specified spatial distribution along a trajectory	*inject, spatter, splash, splatter, spray, sprinkle, squirt*
	***Scatter*-type, Alternating (*N*=4).** Mass is caused to move in a widespread or nondirected distribution	*bestrew, scatter, sow, strew*
	***Pour*-type, Content-only (*N*=10).** A mass is enabled to move via the force of gravity	*dribble, drip, drizzle, dump, ladle, pour, shake, slop, slosh, spill*
	***Coil*-type, Content-only (*N*=6).** Flexible object extended in one dimension is put around another object (preposition is *around*)	*coil, spin, twirl, twist, whirl, wind*
	***Spew*-type, Content-only (*N*=8).** Mass is expelled from inside an entity	*emit, excrete, expectorate, expel, exude, secrete, spew, vomit*
	***Glue*-type, Content-only (*N*=9).** Verbs of attachment	*attach, fasten, glue, nail, paste, pin, staple, stick, tape*
Ground-(container-)	***Stuff*-type, Alternating (*N*=6).** A mass is forced into a container against the limits of its capacity	*cram, crowd, jam, pack, stuff, wad*

oriented (*with* verbs)	***Load*-type, Alternating (*N*=3)**. A mass of a size, shape, or type defined by the intended use of a container is put into the container, enabling it to accomplish its function	*load, pack, stock*
	***Fill*-type, Container-only (*N*=21).** A layer completely covers a surface	*bandage, blanket, coat, cover, deluge, douse, edge, encrust, face, fill, flood, inlay, inundate, line, occupy, pad, pave, plate, shroud, smother, tile*
	***Pollute*-type, Container-only (*N*=22).** Addition of an object or mass to a location causes an aesthetic or qualitative, often evaluative, change in the location	*adorn, burden, clutter, deck, dirty, embellish, emblazon, endow, enrich, festoon, garnish, imbue, infect, litter, ornament, pollute, replenish, season, soil, stain, tint, trim*
	***Soak*-type, Container-only (*N*=15).** A mass is caused to be coextensive with a solid or layer-like medium	*drench, impregnate, infuse, interlace, interlard, interleave, intersperse, interweave, lard, ripple, saturate, soak, stain, suffuse, vein*
	***Clog*-type, Container-only (*N*=12).** An object or mass impedes the free movement of, from, or through the object in which it is put	*block, choke, clog, dam, plug, stop up, bind, chain, entangle, lash, lasso, rope*
	***Bombard*-type, Container-only (*N*=8).** A set of objects is distributed over a surface	*bombard, blot, dapple, riddle, speckle, splotch, spot, stud*
Alternating verbs with "unique geometry" that do not fit into the above classes (*N*=2)	Static of a linear object along a surface	*string*
	A flexible object conforms to part of the shape of an object along two or more orthogonal dimensions	*wrap*

Table 3.1. Pinker's (1989) narrow-range subclasses for locative verbs, adapted from Ambridge, Pine and Rowland (2012).

Further work has since been conducted aimed at defining the nature of the verb classes more precisely (e.g. Boas, 2008; Levin, 1993). However, this work does not change the basic prediction of the semantic verb class hypothesis to be tested

here, namely that children's production of, and retreat from, overgeneralisation errors will be predicted by their knowledge of the semantic class of the verb. In the present study, all of the verbs chosen were classified in the same way by both Pinker (1989) and Levin (1993), although it is worth noting that the organisation of verbs into classes of this kind is not universally accepted (e.g. Braine & Brooks, 1995; Goldberg, 1995; Ambridge, Pine & Rowland, 2012; Brinkmann, 1997; Fellbaum, 1990). It is also worth noting that the semantic verb class hypothesis cannot explain verb frequency effects, which are also pervasive in the literature (as reviewed below). Indeed, some authors (e.g. Stefanowitsch, 2008) have argued that apparent semantic verb class effects are epiphenomenal, with learners acquiring verbs' argument structure restrictions solely on the basis of surface-based statistical learning mechanisms such as entrenchment and preemption. It is to these mechanisms that we now turn.

3.1.2 The frequency hypothesis

Various accounts have attempted to explain how children are able to learn which verbs can be used in which constructions based on statistical properties of the input (e.g. Clark, 1988; Naigles & Hoff-Ginsberg, 1998). For example, the entrenchment hypothesis (e.g. Braine & Brooks, 1995; Theakston, 2004; Ambridge et al., 2008) proposes that, although children may be aware that it is possible to use certain verbs in two alternating constructions, such as the ground- and figure-locative constructions, they gradually learn that this is not the case for all verbs. While children hear figure-only verbs, such as *pour*, frequently in their input, they never hear them in the ground-locative construction. Eventually, this leads children to infer that, if it were possible to use *pour* in this construction, they "would have heard it by now", and hence that ground-locative uses of this verb are ungrammatical for adult speakers. An account that includes a related statistical mechanism (alongside a semantic element) is preemption (e.g. Goldberg, 1995; 2006; 2011; Boyd & Goldberg, 2011). This account proposes that only uses of the verb in a different grammatical pattern that nevertheless yields the same meaning will lead to the inference that the non-attested form is ungrammatical. For example, utterances such as *She poured water into the cup* would preempt **She poured the cup with water*, but

other semantically more distant uses (e.g., *It's pouring with rain*) would not (or, at least, would do so to a lesser degree).

Ambridge, Pine and Rowland (2012) attempted to distinguish between the effects of entrenchment and preemption on the retreat from overgeneralisation in the locative construction, suggesting that both may play a role. However, their entrenchment and preemption predictors were highly correlated, which made it difficult to distinguish effects of one from the other (see also Boyd, Ackerman & Kutas, 2012). For this reason, differentiating between entrenchment and preemption is beyond the scope of the present study (see also e.g. Wonnacott, 2011, p. 2; Perfors, Tenenbaum & Wonnacott, 2010, p. 612). For the remainder of this chapter, we will therefore simply refer to the *frequency hypothesis*. Our findings and conclusions could apply equally to the entrenchment and preemption hypotheses.

3.1.3 Existing evidence for the two accounts

Previous studies have provided evidence in support of both the semantic verb class hypothesis and statistical learning accounts. However, these have primarily been restricted to overgeneralisation errors relating to the causative alternation, such as *Homer broke the plate*/*The plate broke* (e.g. Boyd et al., 2012; Ambridge et al., 2009; 2011; Brooks & Tomasello, 1999; Brooks et al., 1999; Brooks & Zizak, 2002; Naigles, Fowler & Helm, 1992; Naigles & Lehrer, 2002). While these studies provide some support for both the semantic verb class hypothesis and the frequency hypothesis, any successful account must be able to deal with all of the alternations for which overgeneralisation errors are sometimes observed. Ambridge, Pine, Rowland and Chang (2012) tested the predictions of the semantic verb class and entrenchment hypotheses with the dative construction, finding support for both theories, but only in their adult participants (see also Gropen, Pinker, Hollander, Goldberg & Wilson, 1989, for support for broad and narrow verb classes in the dative construction).

So, while the results of studies involving the causative alternation appear to be consistent with both the semantic verb class and frequency hypotheses, both seem to struggle in the domain of the dative alternation. One possible explanation is that the dative is a special case, and that the semantic verb class and frequency hypotheses can explain the retreat from overgeneralisation across a range of different

constructions. Another is that it is the causative alternation that is the special case, with other constructions showing no semantic class or frequency effects. The aim of the present paper is, thus, to test the scope of the two hypotheses by testing their predictions against a third alternation: the locative.

3.1.4 The locative alternation

Like the dative, the locative alternation contains two relatively low frequency constructions with fine-grained distinctions between the relevant narrow semantic subclasses, and therefore constitutes a particularly good test case for both hypotheses. It provides a strong test of the semantic verb class hypothesis because of the sometimes very subtle differences between the narrow subclasses (see Table 3.1). For example, with alternating *spray*-type verbs, a mass is *caused* to move via a force imparted upon it, whereas, with ground-only *pour*-type verbs, a mass is simply *enabled* to move via the force of gravity. In contrast, differences between subclasses for the causative alternation seem more clear-cut: For example, verbs specifying the *manner* of motion, such as *bounce* (*The ball bounced / Bart bounced the ball*), alternate whereas verbs that specify the *direction* of motion, such as *fall* (*The ball fell / *Bart fell the ball*), do not (Pinker, 1989, pp. 130-4). In addition to the subtle subclass distinctions in the locative alternation, for children to form the appropriate subclasses, they would need to be able to observe the differences between them. Again, this seems far less plausible for locative verbs than for causative verbs since, in the locative example above, both the forces involved (e.g. gravity) and the subtle difference between causing and enabling motion are difficult to observe.

Like the dative, the locative alternation also provides a strong test of the frequency hypothesis due to the relatively low frequency of locative verbs, particularly in comparison with verbs involved in the causative alternation. A paucity of locative verbs (and, presumably, constructions) in the input could make it difficult for statistical learning mechanisms to operate.

A further advantage of studying the locative construction, in this case over both the causative and the dative constructions, is that it appears to be truly productive in both directions. With regard to the dative alternation, all known errors involve the overgeneralisation of prepositional-object (PO) verbs into the double-object (DO) dative construction (e.g. *Don't say that to me* → **Don't say me that*,

Bowerman, 1978). We are aware of no reported cases of DO verbs being overgeneralised into the PO construction (e.g. *Homer bet Marge $10* → **Homer bet $10 to Marge*). With regard to the causative alternation, the vast majority of errors involve the overgeneralisation of intransitive-only verbs into the transitive-causative construction (e.g. *She cried* → **You cried her*, Bowerman, 1982a). The converse error, whilst attested (e.g. *I didn't lose it* → **It won't lose*, Lord, 1979), is extremely rare. However, the locative is truly bidirectional, with many examples reported in the literature of ground-only verbs being used in the figure locative (e.g. *I'm going to cover myself with a screen* → **I'm going to cover a screen over me*, Bowerman, 1982b) and of figure-only verbs being used in the ground locative construction (e.g. *I'm gonna pour water onto it* → **I'm gonna pour it with water*, Bowerman, 1981).

This bidirectionality of errors is a useful feature of the locative, because it allows us to test for a possible confound: that children may be completing the judgment task using task-based strategies, especially for novel verbs. For example, in the causative study of Ambridge et al. (2008) and the dative study of Ambridge, Pine, Rowland and Chang (2012), a task-based strategy of always rating intransitives (in the former) or prepositional-object datives (in the latter) as acceptable would yield adult-like judgments for these sentence types, since all were, in fact, grammatical. Note that, in principle, children could quite easily establish such a strategy on the basis of the high frequency, familiar verbs in the studies (e.g. *Bart laughed; Homer gave a book to Marge*), and apply this strategy to lower frequency and novel verbs.

Thus, of the three argument structure alternations studied with respect to the problem of the retreat from overgeneralisation - in/transitive, dative and locative - the latter constitutes the strongest test case for both the semantic verb class and frequency hypotheses. It is therefore perhaps surprising that, of the three alternations, the locative has received by far the least experimental attention. We are aware of only three relevant studies: Gropen, Pinker, Hollander and Goldberg (1991a, b) and Ambridge, Pine and Rowland (2012). Both Gropen et al. studies showed support for Pinker's broad semantic classes, and Ambridge, Pine and Rowland found some support for both levels of semantic class, as well as frequency. However, Ambridge, Pine and Rowland investigated the semantic verb class hypothesis using known locative verbs; no novel verbs were included. Although the authors controlled for attested usage by using verb frequency as a predictor in the regression analysis, for

familiar verbs, the extent to which participants are basing their ratings on semantics alone, as opposed to attested usage, is difficult to ascertain.

3.1.5 The present study

The aim of the present study was to conduct a particularly strong test of the semantic verb class and frequency hypotheses by (a) focussing on the locative alternation, and (b) including both familiar and novel verbs. We obtained grammaticality judgment data from children (aged 5-6 and 9-10) and adults for uses of high frequency, low frequency and novel locative verbs (figure-only, ground-only and alternating) in both locative constructions. We tested whether participants would be able to use verb semantics to determine the grammaticality of sentences containing novel verbs, as predicted by the semantic verb class hypothesis. We also tested whether participants' tolerance of overgeneralisation errors when verbs are used in the inappropriate construction decreased with each increasing level of verb frequency (novel/low/high), as predicted by the frequency hypothesis.

A noteworthy aspect of this study is the fact that participants were taught novel verbs, each of which had semantics consistent with only one of Pinker's (1989) narrow subclasses: two novel verbs each from a ground-only subclass, a figure-only subclass and an alternating subclass. Participants' ability to use the semantics of each novel verb to make their grammaticality judgments is key to Pinker's (1989) proposal: without having the necessary subclasses in place, participants will be unable to judge which locative construction is (un)grammatical for each novel verb.

3.2 Method

3.2.1 Ethics Statement

This study was approved by the University of Liverpool Ethics Committee. Informed consent was obtained in writing both from adult participants and from the parents of the children who took part.

3.2.2 Participants

The participants were 20 children aged 5-6 years (5;6-6;5. *M*=5;11), 20 children aged 9-10 years (9;6-10;5, *M*=9;11) and 20 adults aged 20-25 years. The children were recruited from primary schools, and the adults from the University of Liverpool. All participants were monolingual speakers of English, and had no known language impairments.

3.2.3 Design and materials

3.2.3.1 Design

The experiment used a 3 x 2 x 3 x 3 x 2 mixed design. The between-subjects variables were age of participant (5-6 years, 9-10 years, adult) and counterbalance version (two groups based on which novel verb forms were paired with each meaning). The within-subjects variables were semantic verb subclass (*fill*-type, *spray*-type, *pour*-type; see below), verb frequency (high, low, novel) and sentence type (ground-locative, figure-locative).

3.2.3.2 Test sentences and animations

Table 3.2 shows all verbs and test sentences used. Locative verbs were chosen based on Pinker's (1989) narrow subclasses (subsequently referred to simply as 'classes'). The first of these is the ground-only (or container-only) *fill* class in which "a layer completely covers a surface", the second is the figure-only (or contents-only) *pour* class in which "a mass is enabled to move via the force of gravity", and the third is the alternating *spray* class in which "force is imparted to a mass, causing ballistic motion in a specified direction along a trajectory". For each class, two high frequency and two low frequency verbs with similar semantics were chosen. (Mean lemma frequency counts from the British National Corpus, 2007, are 5923 [range 750-18726] for high frequency verbs and 351 [range 111-658] for low frequency verbs; see Table 3.2 for details.) Participants were also taught novel verbs with similar meanings to the known verbs, two for each semantic class (see below for details of the training method). The form-meaning pairings for novel verbs differed

for each counterbalance group in order to control for any effect of phonological form.

Verb Class	Frequ-ency	Verb	Sentence Type	Sentence
Fill verbs	High (18726)	Cover	*Figure	*Bart covered mud onto Lisa
			Ground	Bart covered Lisa with mud
	Low (487)	Coat	*Figure	*Bart coated mud onto Lisa
			Ground	Bart coated Lisa with mud
	Novel	bredge/ blafe	*Figure	*Bart bredged/blafed mud onto Lisa
			Ground	Bart bredged/blafed Lisa with mud
	High (10546)	Fill	*Figure	*Lisa filled paper into the box
			Ground	Lisa filled the box with paper
	Low (111)	Line	*Figure	*Lisa lined paper into the box
			Ground	Lisa lined the box with paper
	Novel	chool/ tesh	*Figure	*Lisa chooled/teshed paper into the box
			Ground	Lisa chooled/teshed the box with paper
Spray verbs	High (750)	Spray	Figure	Lisa sprayed water onto the roses
			Ground	Lisa sprayed the roses with water
	Low (544)	Sprinkle	Figure	Lisa sprinkled water onto the roses
			Ground	Lisa sprinkled the roses with water
	Novel	tesh/ bredge	Figure	Lisa teshed/bredged water onto the roses
			Ground	Lisa teshed/bredged the roses with water
	High (750)	Splash	Figure	Homer splashed water onto Marge
			Ground	Homer splashed Marge with water
	Low (111)	Spatter	Figure	Homer spattered water onto Marge
			Ground	Homer spattered Marge with water
	Novel	dape/ nace	Figure	Homer daped/naced water onto Marge
			Ground	Homer daped/naced Marge with water
Pour verbs	High (3461)	Pour	Figure	Homer poured water into the cup
			*Ground	*Homer poured the cup with water
	Low (658)	Drip	Figure	Homer dripped water into the cup
			*Ground	*Homer dripped the cup with water
	Novel	nace/ dape	Figure	Homer naced/daped water into the cup
			*Ground	*Homer naced/daped the cup with water
	High (1306)	Spill	Figure	Marge spilt juice onto the rug
			*Ground	*Marge spilt the rug with juice
	Low (195)	Dribble	Figure	Marge dribbled juice onto the rug
			*Ground	*Marge dribbled the rug with juice
	Novel	blafe/ chool	Figure	Marge blafed/chooled juice onto the rug
			*Ground	*Marge blafed/chooled the rug with juice

Table 3.2. All verbs and test sentences used in test trials. Verb frequency counts (lemma counts from the British National Corpus, 2007) are provided in brackets.

For each of the verbs, a test sentence was created using each of the figure- and ground-locative constructions (see Table 3.2). Thus, for each verb in the ground-only *fill* class and the figure-only *pour* class, one sentence for each verb was grammatical and one ungrammatical (e.g. **Lisa filled paper into the box; Lisa filled the box with paper; Homer poured water into the cup; *Homer poured the cup with water*), whereas both sentences were grammatical for verbs in the alternating *spray* class (e.g. *Lisa sprayed the roses with water; Lisa sprayed water onto the roses*). Both sentences in each pair contained identical noun phrases.

For all test sentences, animations were created using Anime Studio Pro Version 5.5 (2006) and presented to participants using a laptop computer. Animations for both sentences in each test pair were identical, but each was presented with the relevant pre-recorded test sentence. Animations served to ensure that participants understood the intended meaning of the sentences, particularly those including novel verbs. They also established the veracity of each of the descriptions, thereby encouraging the participants, particularly the younger ones, to judge the sentences on the basis of their grammaticality rather than their truth value.

3.2.3.3 Novel verb training sentences and animations

Each novel verb was assigned a meaning similar to, but subtly different from, its semantic classmates in the study, whilst still being consistent with the class (e.g. filling *with a particular substance* or pouring *in a particular manner*; see Table 3.2). The English language includes verbs specifying both filling/coating *with a particular substance* (e.g. *to oil, to water, to paper*) and pouring *in a particular manner* (e.g. *to dribble, to drip, to ladle*). Thus, these novel verb meanings are neither non-language-like in general nor non-English-like in particular.

For each novel verb, three animations were created in order to convey the intended meanings to participants. For each of these animations, the novel verb was given three times, always as a gerund. The sentences were as follows:

1. (before clip) Look what CHARACTER's gonna do, it's called VERBing.
2. (during clip) Look what CHARACTER's doing, it's called VERBing.
3. (after clip) So VERBing is... [followed by a brief definition, see Table 3.3].

The definitions were intended to clarify the meanings of each verb and point out the important features of the action, which would enable learners to recognise each verb as being consistent with the intended narrow semantic class. Importantly, novel verbs were never presented in locative or transitive sentences during training (only as simple intransitives), to prevent participants basing their judgments of the novel-verb sentences on attested usage. Rather, according to the semantic verb class hypothesis, learners should determine the locative construction(s) in which each verb can be used on the basis of its semantics.

Novel verb	**Definition**
Novel cover/coat	like covering, except that it has to be with mud (like this)
Novel fill/line	like filling, except that it has to be with paper (like this)
Novel spray/sprinkle	like spraying, except that you have to press a button (like this)
Novel splash/spatter	like splashing, except that it has to be in big blobs (like this)
Novel pour/drip	like pouring, except that it has to be in one big lump (like this)
Novel spill/dribble	like spilling, except that it has to be straight down in tiny drops (like this)

Table 3.3. Novel verbs and definitions.

3.2.3.4 Grammaticality judgments

Participants rated sentences for grammatical acceptability using a five-point 'smiley face' scale (see Figure 3.1 and Ambridge et al., 2008). The scale was presented with no text or numbers. After viewing an animation and hearing the accompanying sentence, children were asked to first choose a coloured counter, with green indicating that the sentence 'sounded good' and red that it 'sounded silly'. They then placed the counter onto the scale to indicate how 'good' or 'silly' it sounded. The use of counters was intended to enable younger children to indicate that they found a sentence broadly acceptable or unacceptable, even if they were unable to provide a more graded judgment (although this did not turn out to be the case). The experimenter made a note of the judgment rating the child gave for each sentence. Adults and older children were asked simply to tick one of the faces to provide their judgment rating.

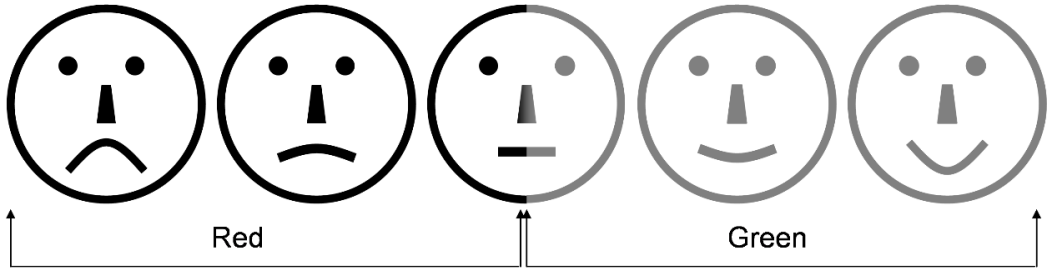

Figure 3.1. Five-point 'smiley face' scale for providing grammaticality judgments.

Participants were trained in the use of the judgment scale with a series of seven training animations. The first four of these were designed to be clearly acceptable or unacceptable, with the others designed to receive ratings somewhere in between. Sentences were chosen based on ratings given by participants in previous studies (see Table 3.4). Ratings for the first two sentences were given by the experimenter, to demonstrate the use of the scale, and participants were given feedback on their ratings for the five subsequent sentences. No feedback was given during the experiment proper. Detailed descriptions of the training procedure are given in Ambridge et al. (2008, pp. 106-7) and Ambridge (2011, pp. 122-3).

Sentence	**Typical score**
The frog caught the fly	5
His teeth man the brushed	1
The cat drank the milk	5
The dog the ball played with	1
The man tumbled Bart into a hole	2 or 3
The magician vanished Bart	2 or 3
The funny clown giggled Bart	1 or 2

Table 3.4. Grammaticality judgment training sentences. 'Sentences' used in the grammaticality judgment training trials, with their 'typical' scores (based on Ambridge et al., 2008). The experimenter completed the first two trials to demonstrate, with participants completing the remainder. Feedback was provided if judgments were thought to be inappropriate.

3.2.4 Procedure

Participants were first taught the novel verbs and then received training on the use of the grammaticality judgment scale (in both cases as described above). The main study consisted of 36 test trials: one ground-locative sentence and one figure-locative

sentence using each of the six high frequency verbs, six low frequency verbs and six novel verbs (see Table 3.2). Sentences were presented in a pseudo-random order, such that two sentences containing the same verb were never given in succession. In order to ensure that participants remembered the intended meaning of the novel verbs, one of the training trials was repeated immediately before each test trial containing a novel verb.

3.3 Results

Because the rating scale data are not true interval scale data, an empirical logit transformation (Agresti, 2002) was applied. First, the rating scale was converted to a proportion and were then transformed using the following formula: log (prop + 0.5) - log(1 – [prop + 0.5]). All means and SEs are reported for raw scores. All post hoc comparisons used Fisher's Least Significant Difference tests. Data are available to download from http://www.benambridge.com.

3.3.1 Preliminary analysis

A preliminary analysis, in the form of a 3x3x2 (age by verb class by sentence type) mixed ANOVA, was performed on known verbs in order to confirm that the verb type classifications (figure-only/ground-only/alternating) were correct for this group of adult participants and that children were rating the sentences as expected. Assuming that this is the case, the semantic verb class hypothesis predicts an interaction of sentence type by verb class such that ground-locative uses are preferred over figure-locative uses for verbs of the *fill* class with the reverse for verbs of the *pour* class, and no preference for the *spray* class. This analysis, and all subsequent analyses, were collapsed across the two counterbalance groups (which differed only with regard to the pairings of phonological stem forms and novel verb meanings), and across the two verbs in each cell of the design.

The ANOVA yielded several main effects. However, these will not be discussed as they collapse across grammatical and ungrammatical sentences, and so are not relevant to the hypotheses of the study. Importantly, as predicted, an interaction of verb class by sentence type was observed ($F_{(2, 114)}$=219.61, p<0.001, η_p^2=0.79). Analysis of this interaction revealed that, as predicted, for verbs in the *fill*

class, participants significantly preferred ground-locative uses (*M*=4.35, *SE*=0.05) over figure-locative uses (*M*=3.16, *SE*=0.07, $p<0.001$). Conversely, for verbs in the *pour* class, participants significantly preferred figure-locative uses (*M*=4.20, *SE*=0.09) over ground-locative uses (*M*=2.43, *SE*=0.10, $p<0.001$). Also as expected, for verbs in the alternating *spray* class, participants showed no preference for either sentence type (ground *M*=4.18, *SE*=0.06; figure *M*=4.09, *SE*=0.07; $p=0.12$, n.s.).

A significant 3-way interaction of verb class by sentence type by age ($F_{(4,114)}=9.05$, $p<0.001$, $\eta_p^2=0.24$; see Figure 3.2) indicated that the pattern of results outlined above differed according to age group. This interaction was driven by the fact that, whilst all age groups displayed the predicted pattern for the non-alternating *fill* and *pour* verb classes, the adults also displayed an unexpected preference for ground-locative uses of verbs from the alternating *spray* class, although a mean rating of 4 or above still indicates that both sentence types were rated as broadly acceptable. It is possible that this result reflects adults' sensitivity to the holism constraint: when an action has been wholly and successfully completed (as is the case for the animations using alternating verbs in the present study), the ground-locative construction is more felicitous that the figure-locative construction (cf. *Lisa taught the students French* vs. *Lisa taught French to the students*). This is an issue to which we will return in the Discussion.

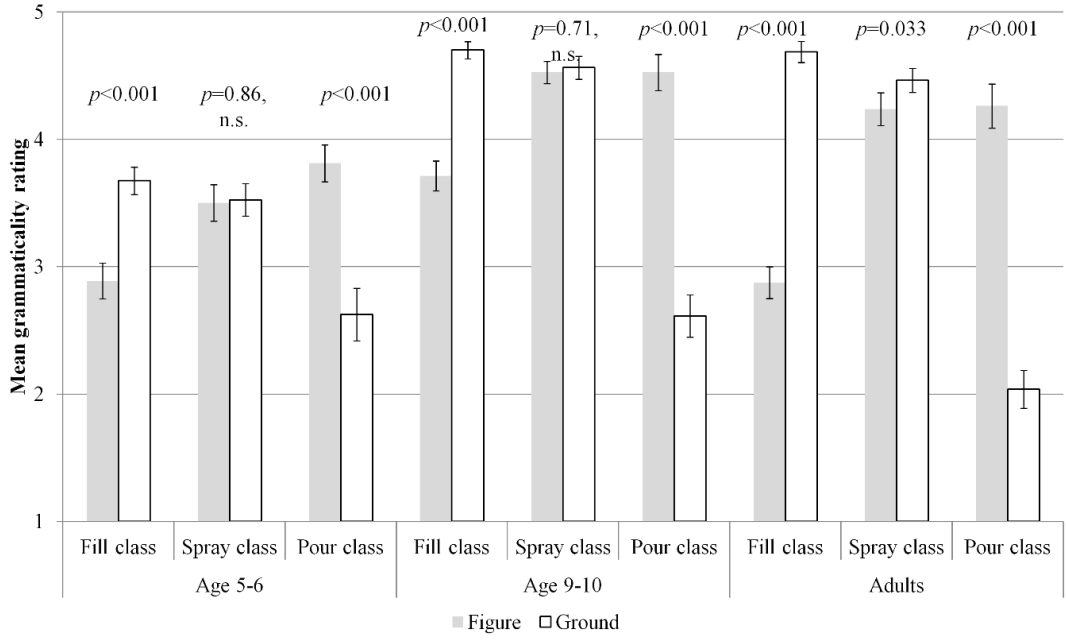

Figure 3.2. Three-way interaction of age by verb class by sentence type for familiar verbs.

3.3.2 Testing the semantic verb class hypothesis

In order to test the semantic verb class hypothesis, participants were taught six novel verbs, two of which were semantically consistent with the ground-only *fill* class, two with the alternating *spray* class and two with the figure-only *pour* class. Participants were then asked to judge sentences containing each of these novel verbs for their grammaticality. Each verb was presented in a figure-locative and a ground-locative construction. The semantic verb class hypothesis predicts that, as with known verbs of the same semantic classes, participants will judge figure-locative uses of the novel *fill* verbs to be less acceptable than ground-locative uses of these verbs, with the opposite pattern for the novel *pour* verbs, and no difference for the alternative uses of the novel *spray* verbs.

These predictions were again tested by means of a 3x3x2 (age by verb class by sentence type) mixed ANOVA, in this case conducted on the ratings for the novel verbs only. As before, this analysis yielded several main effects, which will not be discussed because they collapse across grammatical and ungrammatical sentences. Importantly, as predicted, and in line with the results for all verbs, an interaction of verb class by sentence type was observed ($F_{(2, 114)}$=42.45, p<0.001, η_p^2=0.43). Analysis of this interaction revealed that, as predicted, for novel verbs in the ground-only *fill* class, participants significantly preferred ground-locative uses (M=4.17, SE=0.07) over figure-locative uses (M=3.52, SE=0.09, p<0.001). Conversely, and again as predicted, for novel verbs in the figure-only *pour* class, participants significantly preferred figure-locative uses (M=4.19, SE=0.08) over ground-locative uses (M=3.18, SE=0.13, p<0.001). Unexpectedly, for novel verbs in the alternating *spray* class, participants also showed a small but significant preference for ground-locative uses (M=4.20, SE=0.10) over figure-locative uses (M=3.93, SE=0.10, p=0.031), although a mean rating of around 4 or above still indicates that both sentence types were rated as broadly acceptable. As previously noted, this may be due to the holism constraint.

A significant 3-way interaction of class by sentence type by age ($F_{(4,114)}$=4.27, p=0.003, η_p^2=0.13) indicated that the pattern of results outlined above differed according to age group. As outlined in more detail below, this interaction was driven by the fact that, whilst all groups displayed the predicted pattern for the novel verbs in the non-alternating *pour* class, only older children and adults showed the expected

preference for ground-locative uses of novel verbs in the non-alternating *fill* class, and only the adults displayed the unexpected preference for ground uses of novel verbs from the alternating *spray* class (see Figure 3.3).

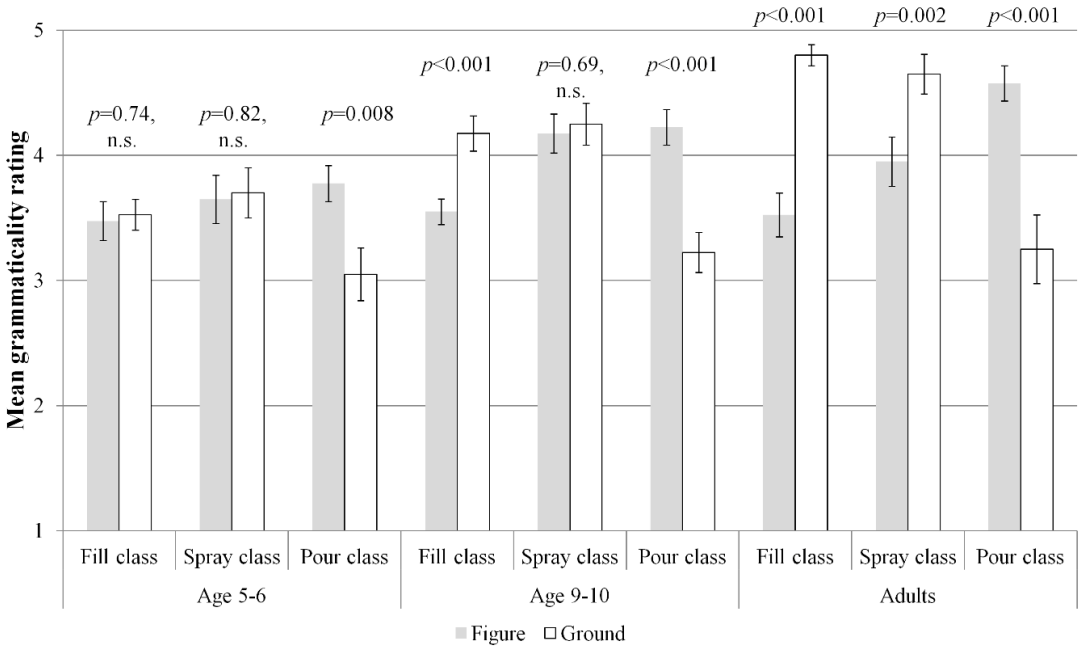

Figure 3.3. Three-way interaction of age by verb class by sentence type for novel verbs.

As predicted by the semantic verb class hypothesis, the 5-year-olds showed no significant preference for novel alternating *spray* class verbs in figure-locative uses (*M*=3.65, *SE*=0.19) or ground-locative uses (*M*=3.78, *SE*=0.20, *p*=0.82, n.s.). Also as predicted, they significantly preferred figure-only *pour* verbs in figure-locative uses (*M*=3.78, *SE*=0.14) over ground-locative uses (*M*=3.05, *SE*=0.21, *p*=0.008). These results suggest that they have identified the verb classes of these novel verbs correctly, and are using this information to judge the grammaticality of the verbs' use in the alternative locative constructions. Contrary to the prediction, however, the 5-year-olds displayed no significant preference for novel ground-only *fill* class verbs in ground-locative uses (*M*=3.53, *SE*=0.12) over figure-locative uses (*M*=3.48, *SE*=0.16, *p*=0.74, n.s.). It is possible that this youngest group of children had not fully grasped the complex semantics of *fill* class verbs, which may be more complex than those of the *pour* class (see Gropen et al., 1991b, and Introduction, above).

The results for the 9-year-olds are all as predicted by the semantic verb class hypothesis: no preference for novel alternating *spray* class verbs in either figure-locative uses (M=4.18, SE=0.16) or ground-locative uses (M=4.25, SE=0.17, p=0.69, n.s.), a significant preference for figure-only *pour* class verbs in figure-locative uses (M=4.23, SE=0.14) over ground-locative uses (M=3.23, SE=0.16, p<0.001), and a significant preference for ground-only *fill* class verbs in ground-locative uses (M=4.18, SE=0.14) over figure-locative uses (M=3.55, SE=0.10, p<0.001).

Adults also displayed the predicted preferences for the novel figure-only *pour* class verbs and the novel ground-only *fill* class verbs. They preferred figure-only *pour* class in figure-locative uses (M=4.58, SE=0.14) over ground-locative uses (M=3.25, SE=0.28, p<0.001) and they preferred novel ground-only *fill* class verbs in ground-locative uses (M=4.80, SE=0.08) over figure-locative uses (M=3.53, SE=0.18, p<0.001). Both of these results are in line with the predictions of the semantic verb class hypothesis. Unexpectedly, however, the adult participants also preferred the novel alternating *spray* class verbs in ground-locatives (M=4.65, SE=0.16) over figure-locatives (M=3.95, SE=0.20, p=0.002). This unexpected result parallels the findings observed for adults with familiar verbs, and may again be explained by the holism constraint (see Discussion). The fact that the 9-year-olds did not show this preference, whilst otherwise displaying an adult-like pattern of results, indicates that the holism constraint (as applied to the ground-locative construction, at least) may not be fully acquired until very late in development.

3.3.3 Testing the frequency hypothesis

To test the frequency hypothesis, we calculated difference scores for grammaticality judgment ratings for 'grammatical' sentences (ground-locative uses of *fill* class verbs; figure-locative uses of *pour* class verbs) minus 'ungrammatical' sentences (figure-locative uses of *fill* class verbs; ground-locative uses of *pour* class verbs) for high frequency, low frequency and novel verbs in both of these non-alternating classes. These difference scores represent the degree of preference for grammatical over ungrammatical verb uses (or, perhaps more importantly for our purposes, the degree of *dispreference* for *un*grammatical verb uses relative to matched grammatical alternatives). Alternating verbs were not included in this analysis since the frequency hypothesis only makes predictions regarding the degree of

unacceptability of ungrammatical verb uses (for alternating verbs, by definition, neither figure- nor ground-locative uses are ungrammatical).

The frequency hypothesis predicts that the largest difference scores will be observed for the high frequency verbs, smaller difference scores for the low frequency verbs and the smallest difference scores for the novel verbs. That is, increased exposure to a verb in grammatical sentences is predicted to increase the strength of the inference that non-attested uses are not permitted, and hence the extent to which participants will rate ungrammatical uses of that verb as unacceptable.

A 3x2x3 (age by verb class by verb frequency) ANOVA revealed that all three main effects were significant. The main effect of verb class ($F_{(1,57)}$=29.83, p<0.001, η_p^2=0.34) indicates that participants showed a larger dispreference for ungrammatical uses of *pour* class verbs (*M*=1.52, *SE*=0.10) than *fill* class verbs (*M*=1.01, *SE*=0.06). While the frequency hypothesis makes no predictions about verb class, this result is consistent with the results of the semantic verb class analysis, which found that participants were less tolerant of overgeneralisation errors with novel *fill*-type verbs than novel *pour*-type verbs.

The main effect of age ($F_{(2,57)}$=18.08, p<0.001, η_p^2=0.39) demonstrates that adults (*M*=1.78, *SE*=0.12) showed a greater degree of dispreference for ungrammatical sentences than both 9-year-olds (*M*=1.24, *SE*=0.08) and 5-year-olds (*M*=0.79, *SE*=0.14), and that 9-year-olds showed a greater degree of dispreference for such uses than 5-year-olds (all comparisons were significant at p<0.01 or better). This result could be interpreted as showing support for the frequency hypothesis, as adults will have had more exposure to grammatical uses of the relevant verbs than 9-year-olds who, in turn, will have had more exposure than 5-year-olds. For this interpretation to be correct, the important factor would have to be *absolute* frequency of exposure to the verbs in competing constructions (e.g. total number of ground-locative uses of *fill*), which obviously increases with age, as opposed to *relative* frequency (e.g. proportion of uses of *fill* in the ground-locative construction as opposed to other constructions), which presumably stays relatively constant across development. However, the lack of interaction between age and verb frequency (see below) suggests that this is not the case. That is, adults did not display a larger frequency effect (i.e. larger between-verb differences) than children, which one would expect if the relevant factor were absolute differences in verb frequency. It is

therefore likely that the main effect of age was simply due to older participants performing better on the task.

Importantly, as predicted by the frequency hypothesis, a main effect of verb frequency was observed ($F_{(2,114)}$=38.25, p<0.001, η_p^2=0.40; Figure 3.4) such that participants showed a greater dispreference for ungrammatical uses of the high frequency verbs (M=1.87, SE=0.11) than either the low frequency verbs (M=1.10, SE = 0.09, p<0.001) or the novel verbs (M=0.83, SE=0.10, p<0.001), which also differed significantly from each other in the predicted direction (p=0.050), although this last difference was much smaller.

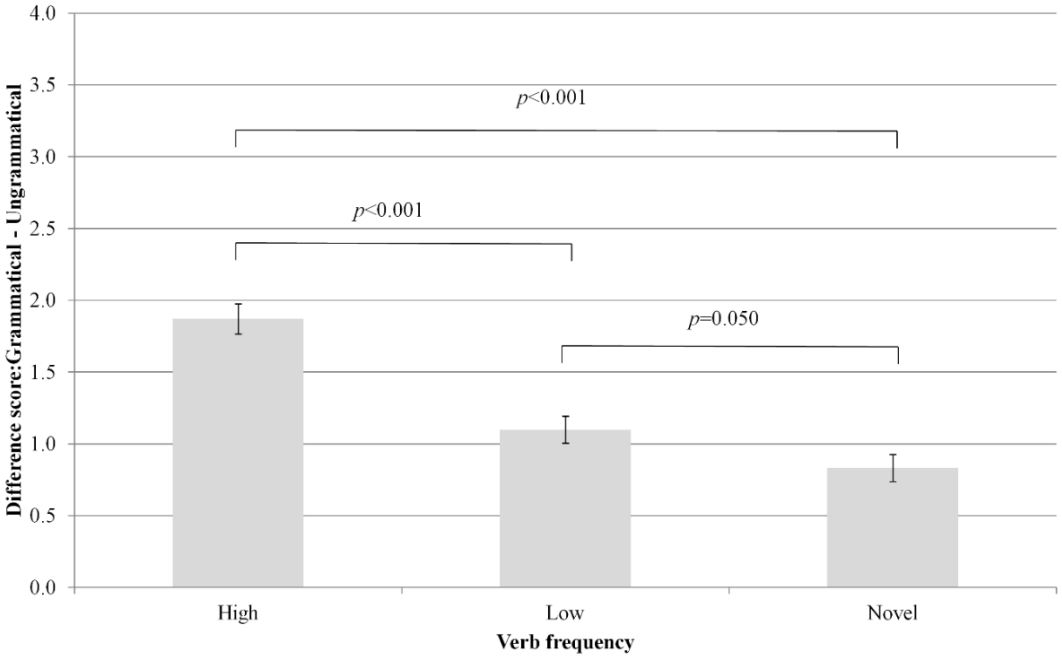

Figure 3.4. Main effect of verb frequency.

The analysis revealed no significant interactions of frequency by age ($F_{(4,114)}$=0.17, p=0.96, n.s., η_p^2=0.01), verb class by age,($F_{(2,57)}$=1.74, p=0.19, n.s., η_p^2=0.06), verb class by frequency ($F_{(2,114)}$=1.84, p=0.16, n.s.,, η_p^2=0.03) or frequency by verb class by age ($F_{(4,114)}$=0.94 , p=0.45, n.s., η_p^2=0.03).

3.4 Discussion

The aim of the present study was to conduct a particularly strong test of the semantic verb class hypothesis (Pinker, 1989) and the frequency hypothesis (e.g. Braine &

Brooks, 1995; Goldberg, 1995) by (a) focussing on the locative alternation, and (b) including both familiar and novel verbs. To this end, we obtained, from children (aged 5-6 and 9-10 years) and adults, judgments of figure- and ground-locative sentences containing high frequency, low frequency and novel verbs consistent with figure-only, ground-only and alternating narrow semantic classes.

The findings suggest that, in general, participants were able to use the semantics of each novel verb to align them with the ground-only *fill* class, the alternating *spray* class or the figure-only *pour* class, although the youngest group of children were unable to do so for novel *fill*-type verbs, and adults showed an unexpected preference for ground-locative uses of novel *spray*-type verbs. The findings of the present study also provide support for the frequency hypothesis: participants in all age groups displayed a greater dispreference for overgeneralisation errors with high frequency than with low frequency familiar verbs, and for errors with both of these groups than with novel verbs.

3.4.1 The role of semantics

According to Pinker's (1989) semantic verb class hypothesis, locative verbs fall into one of two broad semantic classes. A broad-range rule links entries for alternating verbs such as *spray*, which appear in both broad classes, allowing verbs attested in one locative construction to be used in the other (e.g. *Lisa sprayed the flowers with water* → *Lisa sprayed water onto the flowers*). Overgeneralisation errors occur when this rule is incorrectly applied to non-alternating verbs, such as *fill* and *pour*, and cease only when children acquire the more specific narrow semantic subclasses and narrow-range rules that allow the alternation to be restricted to verbs whose semantics are compatible with the core meanings of both locative constructions.

The main test of Pinker's hypothesis in the current study involved novel verbs. Participants were taught six novel verbs with semantics consistent with one of Pinker's narrow subclasses of locative verbs: two each were consistent with (a) the ground-only *fill* class, (b) the figure-only *pour* class, and (c) the alternating *spray* class. Participants provided grammaticality judgments for ground-locative and figure-locative uses of each of the novel verbs with results showing that, as predicted, participants judged ground-locative uses of novel *fill*-type verbs to be significantly more acceptable than figure-locative uses of these verbs, with the

opposite pattern observed for novel *pour*-type verbs. Since these verbs were never presented in locative constructions during training, participants must have been using verb semantics, as opposed to attested usage, to make these judgments. The subtle differences between subclasses of locative verbs, which are also not easily observable, make the locative alternation a particularly strong test of the semantic verb class hypothesis. In addition, the fact that both some figure-locative and some ground-locative sentences were ungrammatical allows us to rule out the possibility that participants were using a task-based strategy to identify the ungrammatical sentences (cf. Ambridge et al., 2008; Ambridge, Pine, Rowland & Chang, 2012). Thus, the results of this study clearly point to an important role for verb semantics in the retreat from overgeneralisation errors in the locative construction.

The semantic verb class hypothesis predicts no preference for either locative construction for alternating *spray*-type verbs. However, while both constructions were judged to be broadly grammatical, adult participants demonstrated an unexpected preference for ground-locative uses of both familiar verbs and novel verbs conforming to the semantics of this subclass. Therefore, one possibility is that adults simply have a general preference for the ground-locative construction for alternating verbs (although this is inconsistent with a strict interpretation of Pinker [1989, p. 127], who lists *spray*-type verbs as being "content-oriented", such that any preference involving these alternating verbs should have been for the figure-locative construction).

A possible explanation for the unexpected preference for ground-locative uses of alternating verbs can be found in the holism constraint. This constraint applies to ground-only locative verbs such as *fill* and *cover*, where the object must be completely filled or covered, respectively, in order for the sentence to be an accurate description of the event. The constraint also applies to the ground-locative construction itself: one semantic feature of this construction, but not the figure-locative construction, is that the 'ground' (e.g. the container) must be wholly affected. Indeed, it is the incompatibility of the semantics of the figure-locative construction and the semantics of verbs such as *fill* and *cover* that makes figure-locative sentences using these verbs ungrammatical.

It is possible that participants may have preferred the ground-locative uses of alternating verbs included in this study because, in the training for the novel verbs and all test animations, the 'location' or 'ground' was always completely affected

(e.g. water splashed onto all of it). It was necessary to create the animations in this way in order to keep the same methodology across all verbs and classes, since, without being completely splashed with water, the animation would have been inconsistent with the ground-locative construction. The animations could therefore be considered to be more consistent with the semantics of the ground-locative construction than with those of the figure-locative construction. The results also suggest a developing knowledge of the holism requirement, as applied to individual verbs, between the age of 5 and adulthood, which in turn provides further support for the semantic verb class hypothesis. Unlike the older children and adults, the 5- to 6-year-olds preferred figure-locative uses of novel *pour*-type verbs but showed no preference for either argument structure for novel *fill*-type verbs. This suggests that these children were unable to appreciate the holism requirement of the novel *fill*-type verbs they were taught based on the animations they viewed during training (see also Gropen et al., 1991a). The disparity between young children's judgment data with novel and familiar verbs may also indicate that these children are basing their grammaticality judgments with familiar verbs on attested usage as opposed to, or in addition to, verb semantics.

Additional support for the importance of a developing knowledge of the holism constraint, as applied to the ground-locative construction, is the fact that only the adult participants gave different judgment scores for the two locative uses of alternating *spray*-type verbs (for both known and novel verbs), although both constructions were judged to be broadly grammatical. This indicates knowledge of the importance of context to the semantics of the alternative locative constructions themselves, which may not yet have developed in the children we tested, leading adults to judge ground-locative uses of *spray*-type verbs as more acceptable than figure-locative uses of these verbs, based on the animations they viewed.

3.4.2 The role of frequency

The frequency hypothesis (e.g. Braine & Brooks, 1995; Goldberg, 1995) emphasises the importance of the statistical properties of the input in children's language acquisition. Under this hypothesis, children retreat from overgeneralisation errors by inferring, from their absence in the input, that certain argument structures cannot be used with certain verbs. The more a child hears, for example, the verb *fill* used in

different constructions with a similar meaning (preemption) or a different construction of any kind (entrenchment) without also hearing it in the figure-locative construction, the better able they are to determine that it is not possible to use *fill* in the latter. This hypothesis therefore predicts that participants will judge overgeneralisation errors with high frequency verbs to be less acceptable than equivalent overgeneralisation errors with low frequency verbs.

Results from the current study provide support for the frequency hypothesis. Participants of all ages showed the same patterns of dispreference for overgeneralisation errors, with higher dispreference scores for such errors with high frequency verbs, lower scores for low frequency verbs, and the lowest dispreference scores for novel verbs, which essentially have a frequency of zero in the input. This finding replicates that of Ambridge, Pine and Rowland (2012), who found a negative correlation between verb frequency and the acceptability of errors across a wider range of locative verbs. So, despite the low frequency of locative verbs and constructions in the input, the effects of this mechanism can clearly be seen in all age groups tested here.

The frequency hypothesis could be interpreted in two ways: either *absolute* frequency of a verb or the *relative* frequency of that verb in competing constructions could be taken as the important factor in the retreat from overgeneralisation. Initially, the finding that participants' dispreference for overgeneralisation errors increased with age appears to show support for the interpretation favouring absolute frequency, since the absolute frequencies of the relevant verbs in different constructions will increase with age, whilst the relative frequencies are likely to remain fairly constant throughout development. However, the fact that no interaction between age and verb frequency was observed counts against this interpretation. Provided that the ratio of high to low frequency verbs in the input remains relatively stable for all ages, an absolute frequency interpretation of the frequency hypothesis would have predicted an increasing difference in dispreference scores for overgeneralisation errors between verbs of different frequencies as the age of participants increased. The main effect of age observed here is therefore likely to be due to older participants simply performing better on the task. So, whilst the present study did not specifically investigate this aspect of the frequency hypothesis, findings suggest that the relative frequency of a verb in competing constructions might be the most important statistical factor in the retreat from overgeneralisation.

3.4.3 Explaining the retreat from overgeneralisation

The predictions of both the semantic verb class hypothesis and the frequency hypothesis have been supported by the findings of the current study: semantics and statistics clearly both have a role to play in the retreat from overgeneralisation. However, neither of these accounts in its current form can explain both the frequency effect and the fact that participants were able to provide grammaticality judgments for novel verbs in line with those of semantically-related familiar verbs. In order to explain the retreat from overgeneralisation errors more fully, an account must be posited that can explain both of these effects, such as Perfors et al.'s Bayesian account (2010) or Ambridge et al.'s FIT account (Ambridge et al., 2011; Ambridge & Lieven, 2011) (see also Alishahi & Stevenson, 2008; Chang, Dell & Bock, 2006; MacWhinney, 2004; Tomasello, 2003).

This study has shown that, as predicted by the semantic verb class hypothesis, children and adults are able to use the semantics of novel verbs to judge their grammaticality in locative sentences in line with verbs with similar semantics. As predicted by statistical learning accounts, children and adults judge errors with high frequency verbs to be worse (in comparison with their grammatical counterparts) than errors with low frequency verbs, which in turn are judged to be worse than errors with novel verbs. Thus, this paper adds to previous research indicating the importance of both semantics and statistics in children's retreat from overgeneralisation errors, and in language acquisition more widely. Future empirical and computational work should focus on testing accounts, such as those mentioned here, that integrate both of these mechanisms.

3.5 Acknowledgements

We would like to thank the schools, teachers, parents and children who made this research possible, as well as the undergraduate project students who assisted with data collection.

Chapter 4: How do children retreat from overgeneralisation error? Evidence from the causative alternation

4.0 Fit within the thesis

The study in Chapter 3 investigated the locative alternation. It found support for both semantic and statistical mechanisms. However, the locative is a relatively rare construction. This chapter investigates the causative alternation (*John rolled the ball/The ball rolled*). This alternation is an important test of the hypotheses in question due to its frequency and the higher frequency of overgeneralisation errors reported in these sentences types (see Pinker, 1989, pp. 22-25). So, whereas Chapter 3 provided a critical test of the semantics hypothesis in particular, this chapter investigates the ability of semantic and statistical approaches to explain the errors children make most frequently. Using a wider range of verbs, with a wider range of frequencies, this study also allows us to distinguish between the entrenchment and preemption hypotheses and to assess the relative contribution of each to the retreat from overgeneralisation. This chapter takes a different approach to semantics to that of Chapter 3. Rather than using discrete semantic classes, we created continuous semantic variables from semantic ratings provided by a group of adult participants. This approach allows more flexibility to investigate widely-observed lexical effects, as outlined in Chapter 1.

This chapter also takes a multi-method approach. As well as using the grammaticality judgment paradigm of Chapter 3, it uses production-priming to investigate the errors that children actually produce. Although it is, in principle, possible to investigate children's production of errors through the use of corpora and diary studies, the paucity of overgeneralisation errors in corpus studies and the potential for observer bias in noticing errors in diary studies means that elicitation in an experimental setting is preferable. The large number of errors that children made in the production study (Experiment 3, below) allowed us to use powerful statistical techniques to test the relative contribution to children's retreat from overgeneralisation errors of semantics, entrenchment and preemption in the same model.

As in Chapter 3, the experiments reported here find support for both semantic and statistical mechanisms. However, the way in which the semantics mechanism is

operating is not always clear, with some statistical predictions being in the opposite direction to our expectations. Possible reasons for this are discussed. While support for entrenchment is strong and fairly consistent, the picture for preemption is far less clear. In fact, preemption seems to operate as expected in the production study only. Taken together with the findings of Chapter 3, then, both statistical and semantic mechanisms appear to play a role in children's retreat from overgeneralisation errors, although further study is required to narrow down the relative contributions of these mechanisms in different alternations. Chapter 5 follows on from Chapters 3 and 4 by investigating the role of semantics in a wider context. It uses the priming methodology of the current chapter to examine children's acquisition of the passive. It also uses two approaches to semantics: the traditional, class-based approach taken in the previous chapter, and the fine-grained semantic continuum approach taken in the current chapter.

This chapter is currently being prepared for submission to a peer-reviewed journal.

4.1 Introduction

In order to attain adult-like levels of productivity in language, children must be able to generalise patterns they have observed in adults' speech to new forms. For example, a child might hear pairs of sentences such as *The plate broke* and *Homer broke the plate*, and *The window opened* and *Marge opened the window*. These pairs are examples of the *causative alternation*, in which verbs are used in both intransitive-inchoative sentences (with no external agent expressed) and transitive-causative sentences (with an explicit causative agent). Over time, the child will hear this pattern repeated with many verbs. On hearing an intransitive sentence containing a new verb, such as *The ball bounced*, they might generalise the pattern they have learned and produce the novel (for them) sentence, *Lisa bounced the ball*.

However, while generalisation is key to language development, *overgeneralisations* can also result. The same child who has created the sentence *Lisa bounced the ball* to enable them express the causer of the ball's bouncing might also want to express the causer of their own giggling, leading to an ungrammatical sentence like **You giggled me*. Such errors have been reported in various studies, perhaps most notably in Bowerman's diary studies, e.g. **You just cried me* (1981),

**Do you want to see us disappear our heads?* (1988). To avoid making transitivisation errors such as these, the child must learn to restrict the application of the causative alternation they have observed to only the appropriate verbs.

Three main approaches to solving this problem currently exist (explained in detail below), with varying levels of empirical support: the entrenchment hypothesis (Braine & Brooks, 1995); the preemption hypothesis (e.g. Goldberg, 1995); and the semantic verb class hypothesis (e.g. Pinker, 1989). Most of the empirical support for these hypotheses comes from grammaticality judgment studies, in which participants rate how (un)grammatical particular verbs sound in particular constructions. These studies are useful, but give us little information about children's *production* of overgeneralisation errors. While production data do exist, diary studies (e.g. Bowerman, 1982a; Lord, 1979) are, by their nature, limited in scope, and experimental production studies (e.g. Brooks & Tomasello, 1999; Boyd et al., 2012) have thus far tended to rely on a small number of novel verbs.

The current paper answers questions raised by the generalisability of current data by using a multi-methodological approach. Firstly, grammaticality judgments ensure that the current findings are in line with several recent studies of different alternating constructions. We then use production-priming to elicit overgeneralisation errors from young children, using a large number of real verbs, to investigate how these errors pattern as a function of distributional and semantic predictors. Our production data allow us to examine the validity of conclusions drawn on the basis of grammaticality judgment data alone, both in the current paper and in previous studies.

4.1.1 Entrenchment

Some proponents of statistical learning approaches have suggested that distributional information in the input can help children avoid and/or retreat from overgeneralisation errors. The entrenchment hypothesis (Braine & Brooks, 1995) proposes an inference-from-absence mechanism. For example, a child will hear the verb *giggle* used in various sentence structures (examples from CHILDES [MacWhinney, 2000], Thomas corpus [Lieven, Salomo & Tomasello, 2009]): *You always get hiccups when you giggle; Why are you giggling?; That just makes you giggle, doesn't it?; You're cheeky giggling away there;* etc. The more often a child

(or adult) hears a verb without hearing it in the alternative construction, the more information they have on which to make the inference that the unattested form must be ungrammatical. Therefore, the entrenchment hypothesis predicts that grammaticality judgments will vary with verb frequency: the higher the frequency of the verb, the worse the ungrammatical, overgeneralised sentence will be judged to be. Similarly, this hypothesis predicts that children will produce fewer errors with verbs of higher than lower frequency.

Theakston (2004) investigated the role of entrenchment in the retreat from overgeneralisation errors using a grammaticality judgment paradigm. Adults and children (aged 5 and 8 years) heard sentences containing argument structure overgeneralisation errors of different types (e.g. **She came me to school* [cf. *I came to school*], **I poured you with water* [cf. *I poured water onto you*]). Half of the verbs were high-frequency (e.g. *come, pour*) and half were low-frequency equivalents of the same semantic class (as classified by Levin, 1993) (e.g. *arrive, dribble*). Children indicated a binary grammatical/ungrammatical judgment, whereas adults' judgments were given on a 7-point scale. Results showed that participants of all ages judged the overgeneralised sentences containing high-frequency verbs to be significantly more ungrammatical than their equivalents containing low-frequency verbs. Thus, this study provided strong support for the entrenchment hypothesis.

In the current paper, we test the entrenchment hypothesis using corpus counts of verb frequency in all constructions. As verb frequency increases, we predict (a) decreasing acceptability of ungrammatical, overgeneralised sentences in our grammaticality judgment tasks, and (b) a decreasing error rate in our production-priming study.

4.1.2 Preemption

A related statistical-learning mechanism is preemption (e.g. Goldberg, 1995), although this hypothesis also involves an element of semantics. According to this approach, hearing a verb used only in constructions *with similar meanings* will lead to the inference that the unattested form is ungrammatical. In the case of the transitive-causative, the construction with closest meaning is the periphrastic causative: X made Y VERB. So, according to the preemption hypothesis, hearing a verb such as *giggle* in periphrastic causative sentences like *Bart made Maggie giggle*

(but never transitive-causative sentences like **Bart giggled Maggie*) should lead children to realise that this is the construction they must use if they want to express the causer of the giggling. The more often a child (or adult) hears a verb in the competing, preempting construction, the more information they have on which to make the inference that the unattested form must be ungrammatical. Therefore, the preemption hypothesis predicts that grammaticality judgments will vary with verb frequency in the preempting construction: the higher the frequency of the verb in that construction (e.g. the periphrastic causative), the worse the ungrammatical, overgeneralised sentence (e.g. the transitive-causative) will be judged to be. Similarly, this hypothesis predicts that children will produce fewer errors with verbs of higher frequency in the preempting construction than those with a lower frequency in that construction.

Brooks and Zizak (2002) tested the predictions of the preemption hypothesis in an elicited production study using novel verbs, with children aged 4 and 6-7 years. The use of novel verbs allows for strict control of the number of exposures participants receive to both preempting and non-preempting input sentences. All children were taught two novel verbs (*dack* and *tam*), one of which was heard in 36 transitive sentences (e.g. *The rabbit is dacking the car*), the other in 36 intransitive sentences (e.g. *The house is tamming*). Children were split into three groups. In the No Preemption group, only the transitive and intransitive sentences were heard, with no preempting alternatives. In the Alternative Construction group, prempting sentences were also presentented, with the periphrastic causative (e.g. *The rabbit made the house tam*) designed to preempt an 'ungrammatical' transitive sentence (e.g. **The rabbit tammed the house*), and the passive (e.g. *The car is getting dacked*) designed to preempt an 'ungrammatical' intransitive sentence (e.g.**The car dacked*). In the English Suppletive group, along with the transitive and intransitive training sentences, children heard sentences using known verbs (e.g. *The car is swinging, The rabbit bounced the house*), which provided an alternative, grammatical way to describe the action in the alternative transitivity. At test, the older children (but not the younger ones) in the Alternative Construction group were significantly less likely to produce a sentence violating the assigned transitivity of the novel verbs than the children in either of the other groups (which did not differ significantly from each other). Thus, support for the preemption hypothesis was found for these slightly older children.

In the current paper, we test the preemption hypothesis using corpus counts of verb frequency in preempting constructions. Following Brooks and Zizak (2002; see also Brooks & Tomasello, 1999), we use the periphrastic causative (e.g. *Homer made the fish swim*) as the preempting construction for transitivisation errors (e.g. *The fish swam* → **Homer swam the fish*) and the passive (e.g. *The ball was kicked*) as the preempting construction for intransitivisation errors (e.g. *Homer kicked the ball* → **The ball kicked*). As verb frequency in the appropriate preempting construction increases, we predict (a) decreasing acceptability of ungrammatical, overgeneralised sentences in our grammaticality judgment tasks, and (b) a decreasing error rate in our production-priming study.

4.1.3 Semantics

The final hypothesis under investigation in the current paper is the semantic verb class hypothesis (e.g. Pinker, 1989; Levin, 1993). Under this approach, the semantics of some verbs allow them to alternate between the transitive-causative construction and the intransitive-inchoative construction, whereas the semantics of other verbs mean that they are compatible with only one or other of these constructions. As children's knowledge of verb semantics is refined, they are able to avoid overgeneralisation errors. In Pinker's original proposal (1989), verbs are grouped into 'semantic classes'. Members of each class have related semantics and behave in the same way in terms of their permissible argument structures. So, for example, 'verbs of extrinsic change of physical state' (p. 130) such as *open, melt* and *shatter* can alternate between intransitive and transitive sentences: *The door opened/Marge opened the door*. In contrast, 'verbs of emotional expression' (pp. 130-1) such as *cry, smile* and *blink* are all intransitive-only verbs: *The girl smiled/*The joke smiled the girl*. More recent approaches (e.g. Boas, 2008; Fellbaum, 1990; Levin, 1993; van Valin, 2005) have offered slightly different explanations of the organisation of verbs in terms of their semantics.

Gropen et al. (1991a) tested Pinker's (1989) semantic verb class hypothesis using the case of the locative alternation. In this alternation, some verbs are able to alternate between the ground locative (e.g. *Lisa sprayed the flowers with water*) and the figure locative (e.g. *Lisa sprayed water onto the flowers*) constructions. Others are only grammatical in the ground locative (e.g. *Lisa filled the cup with water* vs.

Lisa filled water into the cup) or the figure locative (e.g. *Lisa poured water into the cup* vs. **Lisa poured the cup with water*). 'End-state verbs' such as *fill* describe a change of state of the container (here, the cup, which becomes full), whereas 'manner verbs' like *pour* describe the manner of motion of the contents (here, the water, which moves downwards in a steady stream). Gropen et al. taught two novel verbs (*keat* and *pilk*), one each from end-state and manner verb classes, to adult and children (aged 3, 5 and 7 years), then elicited sentences containing these verbs to test if participants were able to use verb semantics to produce 'grammatical' sentences, in line with known verbs of the same semantic class. At test, participants produced significantly more figure locative responses with novel manner verbs than novel end-state verbs, indicating that participants were able to use verb semantics to determine the appropriate, grammatical construction. Results therefore provided strong support for the role of verb semantics in the avoidance of overgeneralisation errors.

An alternative, but related, semantic approach is not to organise verbs into groups, but rather to take into account the consistency of the semantics of each verb with the semantics of the construction into which it is being placed. In the current chapter, we explore the influence of semantic verb-construction compatibility as a continuum, in line with the approach taken by Ambridge et al. (e.g. 2011, 2014; see also Bresnan, Cueni, Nikitina & Baayen, 2007). As verb compatibility with a construction decreases, we predict (a) decreasing acceptability of ungrammatical, overgeneralised sentences in our grammaticality judgment tasks, and (b) a decreasing error rate in our production-priming study.

4.1.4 Recent evidence for the three accounts

Various studies have shown support for all three of the accounts investigated here, with more recent studies tending to test more than one of these proposals in the same experiment. Some of these papers have focussed on the transitive-causative/intransitive-inchoative alternation (e.g. Braine, Brody, Fisch, Weisberger & Blum, 1990; Brooks & Tomasello, 1999; Maratsos, Gudeman, Gerard-Ngo & Dehart, 1987), although work has also been done on other construction pairs, such as the locative (e.g. Ambridge, Pine & Rowland, 2012; Bidgood et al., 2014; Gropen et al., 1991a, b) and dative alternations (e.g. Ambridge et al., 2014; Goldberg, 2011; Stefanowitsch, 2008).

In a combined grammaticality judgment and production study, Boyd et al. (2012) manipulated the frequency and construction type (intransitive only or intransitive and periphrastic causative) in participants' input for novel intransitive-only verbs (e.g. *The apple is yadding, The squirrel really made the apple yad*). They found evidence for independent effects of preemption and entrenchment. In support of the preemption account, participants were less likely to produce and accept transitive uses of the verbs when they heard them in both intransitive sentences and the preempting periphrastic causative than when they had heard them in only the intransitive (with the same total frequency). In support of the entrenchment account, participants in the intransitive-only condition were less likely to produce transitive sentences than a control group who were asked to produce sentences using the real (alternating) verb *bounce*.

Ambridge et al. (2008; see also Ambridge et al., 2011) also used grammaticality judgments with novel verbs, combined with known verbs. Support for statistical learning accounts was found, as adults were less accepting of transitive uses of high frequency than low frequency known intransitive-only verbs. For example, **The man fell the cup* (high frequency) was rated as less acceptable than **The man tumbled the cup* (low frequency). By manipulating semantics, the study supported the predictions of the semantic verb class hypothesis. Participants were taught novel verbs with semantics in line with one of Pinker's (1989) intransitive-only (intransitive-only or alternating in Ambridge et al., 2011) verb classes. For example, animations showed the novel verb *meek* with semantics in line with Pinker's *verbs of directed motion*, such as *fall* and *tumble*. Participants were able to use the semantic information to make grammaticality judgments for novel verbs in line with those of semantically related known verbs. For example, *The cup meeked* was judged as more acceptable than **The man meeked the cup*.

The current study uses a regression design, following Ambridge, Pine and Rowland's (2012) grammaticality judgment study of the locative alternation (e.g. ground locative *Lisa sprayed the flowers with water/*figure locative *Lisa sprayed water onto the flowers*). While some verbs, like *spray*, are able to alternate between the two locative constructions, others are only grammatical in the ground locative (e.g. *Lisa filled the cup with water* vs. **Lisa filled water into the cup*) or the figure locative (e.g. *Lisa poured water into the cup* vs. **Lisa poured the cup with water*). Ambridge, Pine and Rowland presented participants with a series of sentences of

both types, containing ground-only, figure-only or alternating verbs. Adults completed written questionnaires, whereas children (aged 5-6 and 9-10 years) heard sentences, for a smaller set of verbs, accompanied by an animation, and gave their judgments by placing a counter on a 5-point 'smiley-face scale'.

The predictor variables used in Ambridge, Pine and Rowland (2012) were: total verb frequency, calculated from a corpus, to test the entrenchment hypothesis; verb frequency in the preempting construction (here, the grammatical locative construction for non-alternating verbs), again calculated from a corpus, to test the preemption hypothesis; and judgments from a separate group of adults about each verb's semantic properties, based on Pinker's (1989) semantic verb class hypothesis, to test this approach. The results were analysed using mixed effects linear regression models. Regression allows each predictor variable to be entered as a continuous variable, rather than the discrete variables of previous studies (e.g. high- vs. low-frequency). It also allows the individual contribution of each variable to be assessed in relation to that of the other variables. Using mixed effects models allows each of the fixed effects (e.g. verb frequency, semantic score) to be analysed along with random effects (e.g. participant), which may explain additional variation in the results. Results showed significant effects of both overall verb frequency and semantics, but frequency in the preempting construction had no dissociable effect from overall verb frequency. Thus, support was found for the entrenchment and semantic verb class hypotheses, but not preemption.

Following a similar methodology, Ambridge et al.'s (2014) investigation of the dative construction found a significant effect of preemption, in all age groups tested, in addition to the effects of entrenchment and semantics. This suggests that the roles of entrenchment, preemption and semantic mechanisms may not play the same role in the retreat from overgeneralisation errors for all construction types.

In summary, previous findings suggest roles for entrenchment, preemption and semantics in the retreat from overgeneralisation errors resulting from the transitive-causative/intransitive-inchoative alternation. However, the majority of the studies reported here (a) use only a small number of verbs and/or (b) rely on evidence from novel verbs. Both of these issues limit the generalisability of their findings. Recent studies by Ambridge and colleagues (Ambridge, Pine & Rowland, 2012; Ambridge et al., 2014) have investigated the roles of all three mechanisms, whilst overcoming these generalisability issues. Using grammaticality judgment data

in a regression design, these papers found distinct roles for semantics and entrenchment, with preemption also playing a role in the retreat from dative, but not locative, overgeneralisation errors. These findings provide a template for further investigation of the transitive-causative/intransitive-inchoative alternation. However, the grammaticality judgment methodology used by these may be of limited use, since overgeneralisation is a phenomenon of production.

4.1.5 The current study

In this paper, we adopted a multi-method approach to the problem of the retreat from overgeneralisation in the transitive-causative/intransitive-inchoative alternation, using both grammaticality judgment and production tasks. We first sought to extend the findings of Ambridge et al. (2012, 2014), from the locative and dative constructions to the causative construction. To do this, in Experiment 1, we tested adults using a grammaticality judgment task with a large number of verbs. In Experiment 2, we used a scaled-down version of this task with both children and adults to investigate how the relative contributions of the three mechanisms to grammaticality ratings may change over the course of development. Having established the factors that influence grammaticality judgments from both adults and children, in Experiment 3, we used a production-priming methodology to investigate whether these findings hold for children's production. Specifically, we attempted to elicit overgeneralisation errors from 5- to 6-year-olds and examine whether the mechanisms under investigation (entrenchment, preemption and semantics) predicted errors in the same way that they predicted grammaticality judgments. If this is the case, further support will be added to the arguments presented in previous papers using grammaticality judgment data only.

4.2 Methods

The current study consists of three experiments. Experiment 1 was a grammaticality judgment study in which adults rated transitive and intransitive sentences containing each of 180 verbs, of which 60 were transitive-only, 60 intransitive-only and 60 alternating. For all non-alternating verbs, a difference score was calculated by subtracting each participant's rating for the ungrammatical sentence from their rating

for its grammatical equivalent. The bigger this difference, the less acceptable participants found the ungrammatical sentence to be, in comparison with its grammatical counterpart. Experiment 2 repeated the grammaticality judgment paradigm but with three age groups (adults, and children aged 5-6 and 9-10 years) and a reduced set of verbs (40 of each type), as well as using animations to ensure the meaning of each sentence was clear to the children. Experiment 3 used a production-priming methodology with 5- to 6-year-olds to elicit overgeneralisations with both transitive-only and intransitive-only verbs (e.g. **The ball kicked*, **Homer swam the fish*). The same set of verbs was used as in Experiment 2, with the alternating verbs used as fillers.

Before describing the individual experiments, we first outline the methods used to create the predictor variables to test the entrenchment, preemption and semantics hypotheses.

4.2.1 Frequency counts

In order to test the predictions of the entrenchment and preemption hypotheses, verb frequency counts were taken from the British National Corpus (2007). The entrenchment hypothesis (e.g. Braine & Brooks, 1995) posits an inference-from absence mechanism to explain children's retreat from overgeneralisation errors: the more a verb is heard in the input, without being heard in the ungrammatical construction, the stronger the inference that that verb-construction pairing must not be possible. Thus, the entrenchment hypothesis predicts that the more a verb has been heard *regardless of the construction*, the less acceptable it will be in ungrammatical sentences, and the less likely children will be to produce an error with that verb. Therefore, to test this account, we used counts of total verb frequency in the corpus.

The preemption hypothesis (e.g. Goldberg, 1995), while related to the entrenchment hypothesis, adds a semantic element: the more a verb is heard in constructions with a roughly equivalent meaning to the ungrammatical construction, the stronger the inference that the ungrammatical verb-construction pairing must not be possible. Thus, the preemption hypothesis predicts that the more a verb is heard *in a competing construction with similar meaning*, the less acceptable it will be in ungrammatical sentences, and the less likely children will be to produce this type of

error with that verb. To test this account, for the transitive-only verbs, we used counts of verb frequency in the passive (both full and truncated passives were counted), as this is the construction proposed by Brooks and Tomasello (1999) as the preempting construction for intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*). Like the intransitive construction, the passive construction puts the discourse focus on the patient by placing it first in the sentence (e.g. *The plate broke*; *The plate was broken [by Homer]*). The truncated passive also allows the sentence to exclude the agent altogether, as in the intransitive. In our corpus data, the majority of passive sentences were truncated (92.15% of all passive uses of transitive-only verbs), and thus the passive uses of these verbs are almost identical to the intransitive construction, except for the auxiliary *be*. Again following Brooks and Tomasello (1999), for intransitive-only verbs, we used counts of verb frequency in the periphrastic causative (e.g. *The man made the girl laugh*), since this construction expresses a similar meaning to the transitive-causative (e.g., **The man laughed the girl*) and overtly expresses both agent and patient. All verb frequency counts were log transformed.

4.2.2 Semantic ratings

Under the semantics hypothesis, verb semantics determine the permissible constructions for a particular verb, including the transitive and intransitive. Verbs with similar semantics tend to behave similarly in terms of the constructions in which they can appear. Pinker (1989) posited discrete classes of verbs, and identified the key semantic features of each class. In line with more recent work (e.g. Ambridge, Pine & Rowland, 2012; Ambridge et al., 2014), we treated verb semantics as a continuum and created a measure of verb semantics by conducting a rating task to determine the key characteristics of the transitive and intransitive constructions, as well as the individual verbs that can, and cannot, appear in these constructions.

4.2.2.1 Method

4.2.2.1.1 Participants

The participants were 10 adults aged 20-25, all undergraduate students at the University of Liverpool. They were each paid £50 for their participation. All participants were monolingual speakers of English, and had no known language impairments. They did not take part in the other experiments reported in this paper.

4.2.2.1.2 Test items

60 transitive-only verbs, 60 intransitive-only verbs and 60 verbs that can alternate between the two structures were chosen as test items, based on Pinker (1989) and Levin (1993). Table 4.1 lists the verbs used in the Experiments in this chapter.

Verb type	Also used in Experiments 2 and 3	Verbs
Transitive-only	Yes	cut, slice, chop, mash, hit, strike, bite, peck, touch, stroke, slash, saw, crush, squash, kick, tap, whack, punch, nudge, kiss, kill, destroy, demolish, take, bring, raise, hoist, lift, lower, leave, drop, tickle, amuse, feed, delight, give, madden, lend, pay, offer
	No	slice, slay, murder, assassinate, slaughter, execute, obliterate, poison, abandon, desert, gladden, sadden, sicken, donate, sell, nauseate, bribe, convince, enthuse, thrill, refund
Intransitive-only	Yes	go, come, rise, fall, tumble, ascend, descend, exit, enter, arrive, eat, jump, hop, run, drink, talk, swim, climb, sing, sleep, smile, cry ,laugh, frown, giggle, chortle, chuckle ,grin, groan, moan, glow ,glitter, leak, appear, disappear, vanish, die, stay, wait, live
	No	glisten, sweat, bleed, sparkle, twinkle, ooze, gush, decease, perish, expire, emerge, materialize, lapse, exist, cost, last, weigh, linger, sound, measure
Alternating	Yes	break, rip, shatter, smash, grow, change, bake, boil, cook, fry, burn, split, tear, melt, crack, improve, inflate, alter, shrink, freeze, crash, fold, crease, deflate, defrost, dissolve, enlarge, expand, open, close, snap, bend, slide, move, roll, bounce, turn, begin, start, stop
	No	transform, divide, explode, fill, flood, stretch, heal, evaporate, spin, rotate, float, hang, skid, cease, commence, continue, end, finish, proceed, resume

Table 4.1. Verbs used in the semantic ratings task and Experiment 1. Those used in Experiments 2 and 3 are also indicated.

4.2.2.1.3 Procedure

Each participant entered ratings on a spreadsheet. Both the verbs and the semantic criteria were randomised separately for each participant. The instructions were as follows:

> On the following sheet is a list of 180 verbs. Each verb denotes an event where one person or thing (A) causes another person or thing (B) to move or change. In what follows, we always use A to denote the "causer" and B to denote the thing that moves or changes. The spreadsheet also has a list of 26 different meanings, each of which applies to some verbs but not others. Your

task is to rate the extent to which each meaning applies to each verb, on a scale of 1-9.

The procedure was illustrated with verbs and features not used in the task proper. The 26 ‘meanings’ consisted of statements based on aspects of semantics thought to be related to the transitive-causative and/or intransitive-inchoative constructions (see Pinker, 1989; Levin, 1993). For example, *Another person or thing comes into physical contact with B* is related to the transitive-causative construction; *B is something that can move by itself* is related to the intransitive-inchoative construction. Table 4.2 for a full list of these semantic criteria. At no point were the verbs presented in sentences, nor were participants instructed to imagine them in sentences. Mean ratings were calculated for each semantic criterion for each verb.

Another person or thing comes into physical contact with B
B allows or enables the movement/change (as opposed to actually initiating it)
B causes its own movement/change
B comes into (or goes out of) existence
B emits light/a sound/a substance
B expresses a particular emotion
B initiates the movement/change (as opposed to merely allowing or enabling it)
B is a human or animal
B is something that can move by itself
B moves in a particular direction
B moves in a particular manner
B moves to/from a particular location
B must be willing for the movement/change to occur
B remains in the same place whilst its parts move
B requires an external cause for the movement/change to happen
B undergoes a change-of state (e.g., solid to liquid; whole to pieces)
For the movement/change to be possible, B must be a particular type of thing
For the movement/change to be possible, B must have some particular property
The movement/change can occur even if B resists
The movement/change is internally caused (i.e., caused by B)
The movement/change is something that B does (as opposed to something that "happens to" B)
The movement/change is voluntary as opposed to accidental on the part of B
The movement/change occurs to a particular degree/in a particular amount
The movement/change is something that "happens to" B (as opposed to something that B does)
The verb denotes a particular property that B possesses (rather than movement/change of B)
The verb denotes B being in a particular state (rather than movement/change of B)

Table 4.2. Semantic criteria used in the semantic rating task.

Principal Components Analysis (PCA) with oblique rotation was used to produce a number of composite factors, as many of the individual semantic criteria may be highly correlated. This analysis was conducted in RStudio (version 0.97.551; R version 3.1.1, R Core Team, 2014) following the procedure suggested by Field, Miles and Field (2012, pp. 772-793) and using the following packages and functions: *corpcor* (version 1.6.6; Schäfer, Opgen-Rhein, Zuber, Ahdesmäki, Duarte Silva & Strimmer, 2013), *GPArotation* (version 2012.3-1; Bernaards & Jennrich, 2005), *psych* (version 1.4.6; Revelle, 2014), *kmo()* (Kerns, 2007). The KMO statistic for the data is 0.76 ('good' according to Field et al.) with all values for individual items above the threshold of 0.5. Bartlett's test of sphericity ($\chi^2_{[df=190]}$=1970.70, p<0.001) indicated that items were sufficiently correlated. Five components were retained, based on parallel analysis (Horn, 1965). Details of how each of the original semantic features load on each of the five composite semantic components are provided in Table 4.3. Any loading of above 0.384 (or below -0.384) is statistically significant (Stevens, 2012, p. 332). In subsequent analyses, all five of the extracted components are entered into the models as predictor variables. Table 4.4 shows summary information for each component.

	Specific property/state	Animate/ volitional	External agent/cause	Manner/location of movement	Degree of change
Another person or thing comes into physical contact with B	-0.6	-0.01	0.76	-0.02	0.07
The movement / change is voluntary as opposed to accidental on the part of B	-0.6	0.61	0.03	0	0.11
B must be willing for the movement / change to occur	-0.2	0.72	-0.3	-0.1	0.09
B moves to/from a particular location	-0.1	0.06	-0.2	0.85	-0.2
The movement / change is internally caused (i.e., caused by B)	0.02	0.83	-0.1	0.14	-0.1
B requires an external cause for the movement / change to happen	0.03	-0.4	0.73	0.02	0.09
B is something that can move by itself	0.04	0.88	-0.03	0.15	-0.04
B moves in a particular manner	0.09	0.11	0.18	0.88	0.25
B remains in the same place whilst its parts move	0.12	0.23	0.39	-0.02	0.44
B comes into (or goes out of) existence	0.14	0.05	0.16	0.03	-0.9
B expresses a particular emotion	0.24	0.5	0.05	-0.4	0.16
B undergoes a change-of state (e.g., solid to liquid; whole to pieces)	0.25	-0.3	0.62	0.21	-0.2
The movement / change occurs to a particular degree/in a particular amount	0.28	-0.01	0.25	0.18	0.74
The movement / change can occur even if B resists	0.3	-0.1	0.58	-0.2	-0.2
B emits light/a sound/a substance	0.39	0.18	0.33	-0.2	0.22
For the movement / change to be possible, B must be a particular type of thing	0.43	0.56	0.21	-0.3	0
B allows or enables the movement / change (as opposed to actually initiating it)	0.67	-0.2	-0.2	0.07	0.11
The verb denotes B being in a particular state (rather than movement / change of B)	0.74	-0.1	0.11	0.02	-0.3
The verb denotes a particular property that B possesses (rather than movement / change of B)	0.74	-0.1	-0.2	-0.2	0.24
For the movement / change to be possible, B must have some particular property	0.8	0.21	0.16	0.03	0.03

Table 4.3. Semantic factor loadings.

Component label	Variance explained	Verb type related to (Pinker, 1989)	Direction of error-rate prediction: Transitive verbs	Direction of error-rate prediction: Intransitive verbs
Specific property/state	18%	Intransitive	+	-
Animate/ volitional patient	17%	Intransitive	+	-
External cause/ agent	13%	Transitive	-	+
Manner/ location of movement	10%	Intransitive	+	-
Degree of change	10%	Alternating	unclear	unclear

Table 4.4. Summary of composite semantic factors produced by PCA: component label (given by us); the percentage of variance it explains; whether it is related to transitive-only, intransitive-only or alternating verbs; and the direction in which it is expected to predict the rates of intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*) and transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*). Specifically, "Direction of error rate prediction" refers to whether we would expect more (+) or fewer (-) errors the higher a verb's rating is on each of the semantic components, in line with Pinker (1989).

Verbs with large values for s*pecific property/state* are those for which the patient must have a specific property (e.g. being able to glow; *glow*) or be in a particular state (e.g. solid; *melt*). Verbs with large values for *animate/volitional patient*, are those for which the patient is animate (e.g. *giggle*) and/or able to do things of its own accord (e.g. *run*). Verbs with large values for *external cause/agent* are those which cannot take place without an external cause (e.g. heat; *defrost*) or agent (e.g. an assassin; *assassinate*). Verbs with large values for *manner/location of movement* are those which specify the manner of a movement (e.g. *run*) or the location/direction of movement (e.g. *rise*). Verbs with large values for *degree of change* are those for which the patient changes in some way (e.g. *bend*), but for which that change does not involve coming into/going out of existence (verbs such as *die* and *materialise* have large negative values).

4.2.3 Experiment 1 (adults – grammaticality judgments)

From previous studies looking at the dative and locative alternations, we know that both corpus-derived entrenchment/preemption measures and experimentally derived semantic measures can predict grammaticality ratings (e.g. Ambridge, Pine and Rowland, 2012). Before examining children's error production, we wanted to ensure that these measures are also appropriate for investigating the transitive-causative/intransitive-inchoative alternation. Therefore, in Experiment 1, adults rated both transitive-causative and intransitive-inchoative sentences containing transitive-only, intransitive-only or alternating verbs for grammaticality.

4.2.3.1 Method

4.2.3.1.1 Participants

The participants were 44 adults aged 20-25, all undergraduate students at the University of Liverpool who had not taken part in the semantic ratings task. They each were paid £10 for their participation. All participants were monolingual speakers of English and had no known language impairments.

4.2.3.1.2 Test items

The same verbs were used as in the semantic ratings task. For each verb, transitive-causative and intransitive-inchoative sentences were created as follows:

> The man/woman/boy/girl [VERBed] the [object/person/animal] [modifying phrase]
> The [object/person/animal] [VERBed] [modifying phrase]

So, for example, the sentences for the alternating verb *grow* were *The man grew the flowers in the greenhouse* and *The flowers grew in the greenhouse*.

4.2.3.1.3 Procedure

Participants were asked to judge the grammatical acceptability of the test sentences described above. The sentences were presented, in a different random order for each participant, on a spreadsheet. Each sentence was presented with a semantic 'context' in which the sentences were to be judged (in lieu of the animations used in the subsequent studies), such as:

> Context: MAN CAUSE GIRL SING ALL NIGHT
> Sentence: The man sang the girl all night [or The girl sang all night]

Grammaticality ratings were given on a scale of 1-7, where 1 was completely unacceptable and 7 was completely acceptable. Example dative sentences with low, medium and high acceptability ratings (based on Ambridge, Pine, Rowland & Chang, 2012) were given to illustrate how grammaticality ratings should be distributed. Participants entered their rating for each sentence directly onto the spreadsheet. From these raw ratings, a difference score was calculated for the transitive-only and intransitive-only verbs by subtracting the rating for the ungrammatical sentence from the rating for the grammatical sentence (transitive minus intransitive for transitive-only verbs, intransitive minus transitive for intransitive-only verbs). Alternating verbs were not included in the analyses as, by definition, they are grammatical in both transitive and intransitive sentences. The larger the difference score, the less acceptable the ungrammatical sentence, in comparison with its grammatical counterpart. Previous studies (e.g. Ambridge et al. 2008) show that using a difference score as opposed to raw scores is important in order to avoid a possible confound resulting from participants simply disliking particular verbs, characters, events etc.

4.2.3.2 Results

Data were analysed separately for intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*) and transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*), allowing us investigate any possible differences in the contribution of statistical and semantic mechanisms to adults' judgments of errors of

each type. Analyses were carried out using linear mixed effects regression models in RStudio (version 0.97.551; R version 3.1.1, R Core Team, 2014) with the *lmer* function of the *lme4* package (version 1.1-7, Bates, Maechler, Bolker & Walker, 2014). Predictor variables were total verb frequency (the entrenchment measure), verb frequency in the preempting construction (passive for transitive-only verbs, periphrastic causatives for intransitive-only verbs) and all five semantic components extracted from the PCA (described above). Random intercepts for verb and participant were included in the model, with by-participant random slopes for as many predictor variables as possible while still allowing the models to converge (Barr, Levy, Scheepers & Tily, 2013). The predictor variables were initially standardised (transformed into *z* scores) before being entered into the model. Model summaries are shown in Table 4.5.

	Variable	Transitive-only verbs	Intransitive-only verbs
Fixed effects	Intercept	3.00 (0.16)	3.26 (0.20)
Estimate (SE)	Entrenchment	0.67 (0.22)	0.45 (0.14)
	Preemption	-0.45 (0.23)	0.09 (0.14)
	Specific property/state	-0.39 (0.12)	0.11 (0.14)
	Animate/volitional patient	0.21 (0.12)	0.37 (0.14)
	External agent/cause	0.21 (0.11)	-0.02 (0.13)
	Manner/location of movement	0.06 (0.12)	-0.31 (0.14)
	Degree of change	-0.14 (0.11)	0.10 (0.13)
Random effects	Verb (intercept)	0.49 (0.70)	0.62 (0.79)
Variance (SD)	Participant (intercept)	0.61 (0.78)	1.26 (1.12)
	Participant – Entrenchment	n/a	0.04 (0.19)
	Participant – Preemption	0.04 (0.19)	n/a
	Participant – Specific property/state	0.04 (0.21)	0.03 (0.17)
	Participant – Animate/volitional patient	n/a	0.06 (0.25)
	Participant – External agent/cause	0.05 (0.22)	0.02 (0.16)
	Participant – Manner/location of movement	0.02 (0.16)	n/a
	Participant – Degree of change	0.02 (0.14)	0.07 (0.27)
Model summary	AIC	10924	10986
	BIC	11106	11168
	Log Likelihood	-5430.8	-5461.8
	Deviance	10862	10942

Table 4.5. Model summary: Experiment 1, all verbs.

Model comparisons were computed using the *anova* function to determine significance levels for each of the predictor variables. Each variable was removed in reverse order, so the semantic component *Degree of change* was removed first and the *Entrenchment* measure last. Entrenchment is first in the model (and removed last in the model comparison process) as it is currently the mechanism with the most supporting evidence in the literature and, therefore, was deemed to be the most widely-accepted in the field. The other predictor variables would need to demonstrate that they could operate over and above entrenchment. Preemption was entered next, as it is also well-defined in the literature and the frequency counts are a subset of those used for the entrenchment predictor. Finally, the semantic predictors were entered in the order of the amount of variance they explained in the PCA (see Table 4.4).

The results of the model comparison analysis for transitive-only verbs are shown in Table 4.6 and for intransitive-only verbs in Table 4.7. As Table 4.6 shows, for intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*), significant predictor variables were total verb frequency, verb frequency in the preempting passive construction, and the semantic components *specific property/state* and *external agent/cause*. Both of these semantic components predict results in the expected direction: *specific property/state*, a feature of intransitive verbs, negatively predicts difference scores for intransitivisation errors with transitive-only verbs, whereas *external agent/cause*, a feature of transitive verbs, positively predicts these scores. These findings therefore provide support for the entrenchment and semantic hypotheses, although preemption was a negative predictor – the opposite of our expectation. We return to this issue in the Discussion.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	**χ^2**	**χ df**	***p* value**
Intercept	24	10953	11094	-5452.6	10905			
Entrenchment	25	10940	11087	-5444.9	10890	15.46	1	8.4^{e-5} ***
Preemption	26	10935	11088	-5441.6	10883	6.57	1	0.010*
Specific property/state	27	10927	11086	-5436.5	10873	10.15	1	0.0014 **
Animate/ volitional patient	28	10926	11091	-5435.2	10870	2.63	1	0.10
External agent/cause	29	10922	11092	-5431.9	10864	6.55	1	0.011*
Manner/ location of movement	30	10924	11100	-5431.8	10864	0.25	1	0.62
Degree of change	31	10924	11106	-5430.8	10862	1.97	1	0.16

Table 4.6. Model comparisons: Experiment 1, transitive-only verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	24	11009	11150	-5480.7	10961			
Entrenchment	25	10997	11144	-5473.7	10947	13.96	1	1.9^{e-4} ***
Preemption	26	10996	11149	-5472.0	10944	3.32	1	0.069
Specific property/state	27	10996	11155	-5471.0	10942	2.12	1	0.15
Animate/ volitional patient	28	10987	11152	-5465.5	10931	10.94	1	9.4^{e-4} ***
External agent/ cause	29	10989	11159	-5465.5	10931	0.053	1	0.82
Manner/ location of movement	30	10984	11161	-5462.1	10942	6.67	1	0.0098 ***
Degree of change	31	10986	11168	-5461.8	10942	0.78	1	0.38

Table 4.7. Model comparisons: Experiment 1, intransitive-only verbs.

As Table 4.7 shows, for transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*), significant predictor variables were total verb frequency, and semantic components *animate/volitional patient* and *manner/location of movement*. Both of these semantic components predict results in the expected direction: *animate/volitional patient*, a feature of intransitive verbs, positively predicts difference scores for transitivisation errors with intransitive-only verbs, whereas *manner/location of movement*, a feature of alternating verbs, negatively predicts these scores. Again, these results provide support for both the entrenchment and semantics hypotheses. Verb frequency in the periphrastic was only a marginally significant predictor (p=0.07).

In summary, the results from Experiment 1 show mixed support for the three theories under investigation. While there is strong support for both the entrenchment and semantics hypotheses, no support was found for the preemption hypothesis.

4.2.4 Experiment 2 (adults and children – grammaticality judgments)

Experiment 1 showed that total verb frequency (entrenchment) and verb semantics significantly predict difference scores for ungrammatical uses of both transitive-only and intransitive-only verbs, for adults in a grammaticality judgment task, with mixed

findings for the preemption hypothesis. However, our main goals in this paper were to explore the influence of these three mechanisms on overgeneralisation errors (a) developmentally and (b) in production. Thus before turning to production, we ran a second judgment study (Experiment 2) with both adults and children. This study had two aims. The first was to investigate the respective influences of entrenchment, preemption and semantics across development (which requires a judgment task, as older children and adults are unlikely to produce overgeneralisation errors). The second was to obtain judgment data that could be subsequently compared to production data obtained from children of the same age. This is important for generalising the findings from previous child judgment studies to the main phenomenon of interest: children's overgeneralisation errors in production.

4.2.4.1 Method

4.2.4.1.1 Participants

The participants were 96 children aged 5-6 (5;3-6;5, *M*=5;10), 96 children aged 9-10 (9;4-10;6, *M*=9;11), and 24 adults aged 18-25. The children were recruited from primary schools in the North West of England. The adults were all undergraduate students at the University of Liverpool and received course credit for their participation. They had not taken part in Experiment 1 or the semantic ratings task. All participants were monolingual speakers of English, and had no known language impairments.

4.2.4.1.2 Test items

The number of verbs was reduced from 180 in Experiment 1 (60 each of transitive-only, intransitive-only and alternating) to 120 (40 of each type). The verbs used were a subset of those previously used, chosen to be those most likely to be known to young children. The number of test items was reduced to ensure the task was not too long for the attention span of the younger group of children. The sentences used were similar to those in Experiment 1, with the exception that, in order to make the study more child-friendly, the generic terms describing the agents in transitive sentences (e.g. *the girl*) were replaced with names of familiar cartoon characters, such as *Lisa dropped the ball to the floor*.

Animations were created, using *Anime Studio Pro 5.5*, for each of the actions. The transitive and intransitive sentences for each verb were recorded by a native speaker of British English. The audio file played automatically after the animated action ended. Identical animations were used for the transitive and intransitive versions of each sentence. The use of animations ensured that the veracity of the sentences would not be in doubt and that participants' judgments would therefore be based on the grammaticality of the sentences, something that we have previously found to be important when testing young children (e.g. Ambridge et al., 2008).

4.2.4.1.3 Procedure

Test sentences and their accompanying animations were presented to participants using *VLC Media Player*. Grammaticality judgments were given on a 5-point 'smiley-face' judgment scale (see e.g., Ambridge et al., 2008), shown in Figure 4.1. Adults watched the full set of animations, in a pseudo-random order such that no two sentences containing the same verb were presented consecutively, in small groups of up to 10 participants. Adults marked their responses (individually) on an answer sheet containing one smiley-face scale for each sentence. Due to constraints on attention span, children were tested individually on one quarter of the sentences (60 in total), split over two days. Thus, 24 children rated each verb. Each child was tested on transitive and intransitive versions of sentences containing 10 each of transitive-only, intransitive-only and alternating verbs. Sentences were again presented in pseudo-random order. Children gave their responses by placing a green or red counter (indicating broadly grammatical or broadly ungrammatical, respectively) onto a single, larger smiley-face scale (see Figure 4.1). The experimenter noted down responses by hand. As in Experiment 1, a difference score was calculated for the transitive-only and intransitive-only verbs by subtracting the rating for the ungrammatical sentence from the rating for the grammatical sentence, resulting in a difference score for the ungrammatical sentence relative to its grammatical counterpart.

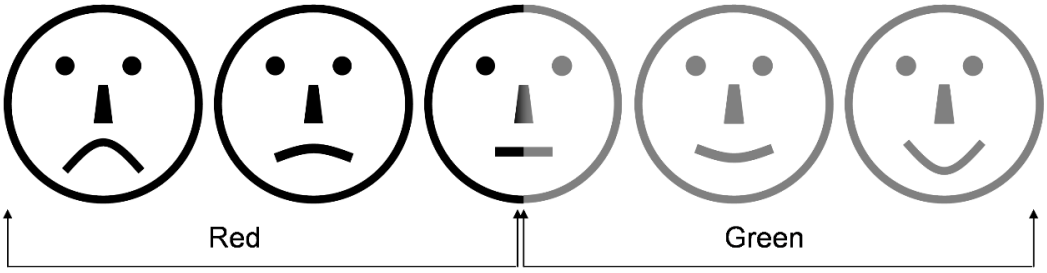

Figure 4.1. 5-point 'smiley-face'. The face on left represents a completely ungrammatical sentence, the face on the right a completely grammatical sentence and the remaining faces a rating between these two extremes.

4.2.4.2 Results

Results were analysed using the same method as in Experiment 1. Predictor variables in the initial analysis were age group, total verb frequency (the entrenchment measure), verb frequency in the periphrastic causative/passive (the preemption measure) and all five semantic components extracted from the PCA, as well as interaction terms for age group by each of the other predictor variables. As in Experiment 1, random intercepts were included for verb and participant, with by-participant random slopes for as many predictor variables as possible while still allowing the model to converge. Interaction terms, however, were not included as random slopes as the model could not converge when this was the case. A summary of the model is shown in Table 4.8. The results of model comparisons are shown in Table 4.9 for transitive-only verbs and Table 4.10 for intransitive-only verbs.

	Variable	**Transitive-only verbs**	**Intransitive-only verbs**
Fixed effects	Intercept	1.59 (0.10)	1.99 (0.10)
Estimate (SE)	Age group	0.07 (0.07)	0.97 (0.08)
	Entrenchment	0.20 (0.19)	0.55 (0.15)
	Preemption	-0.01 (0.18)	-0.15 (0.13)
	Specific property/state	-0.16 (0.10)	-0.14 (0.14)
	Animate/volitional patient	0.31 (0.10)	0.04 (0.13)
	External agent/cause	0.05 (0.10)	0.21 (0.12)
	Manner/location of movement	-0.12 (0.10)	-0.19 (0.12)
	Degree of change	0.05 (0.13)	0.17 (0.13)
	Age group:Entrenchment	0.15 (0.08)	-0.09 (0.06)
	Age group:Preemption	-0.01 (0.08)	0.04 (0.04)
	Age group:Specific property/state	0.07 (0.04)	0.08 (0.06)
	Age group:Animate/ volitional patient	0.13 (0.05)	-0.02 (0.06)
	Age group:External agent/cause	0.01 (0.05)	0.06 (0.05)
	Age group:Manner/ location of movement	-0.08 (0.04)	-0.15 (0.04)
	Age group:Degree of change	0.12 (0.06)	-0.10 (0.05)
Random effects	Verb (intercept)	0.24 (0.49)	0.28 (0.53)
Variance (SD)	Participant (intercept)	0.37 (0.61)	0.40 (0.64)
	Participant – Age group	n/a	0.07 (0.27)
	Participant – Entrenchment	0.08 (0.28)	0.09 (0.30)
	Participant – Preemption	0.06 (0.24)	n/a
	Participant – Specific property/state	n/a	0.07 (0.26)
	Participant – Animate/volitional patient	0.06 (0.24)	0.05 (.023)
	Participant – External agent/cause	0.04 (0.21)	0.11 (0.34)
	Participant – Manner/location of movement	0.02 (0.13)	n/a
	Participant – Degree of change	0.05 (0.21)	0.02 (0.14)
Model summary	AIC	10380	10477
	BIC	10654	10751
	Log Likelihood	-5143.7	-5192.5
	Deviance	10288	10385

Table 4.8. Model summary: Experiment 2, all verbs and age groups, with by-age interactions.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	χ^2	**χ df**	***p* value**
Intercept	31	10463	10648	-5200.6	10401			
Age group	32	10398	10589	-5167.2	10334	66.77	1	3.0^{e-16} ***
Entrenchment	33	10397	10594	-5165.6	10331	3.19	1	0.07
Preemption	34	10398	10600	-5165.0	10330	1.20	1	0.27
Specific property/state	35	10400	10608	5164.8	10380	0.41	1	0.52
Animate/ volitional patient	36	10390	10604	-5158.9	10318	11.70	1	6.3^{e-4} ***
External agent/cause	37	10391	10612	-5158.6	10317	0.61	1	0.43
Manner/ location of movement	38	10391	10618	-5157.6	10315	2.05	1	0.15
Degree of change	39	10393	10626	-5157.5	10315	0.11	1	0.74
Age group: Entrenchment	40	10394	10632	-5157.0	10314	1.16	1	0.28
Age group: Preemption	41	10394	10638	-5155.9	10312	2.18	1	0.14
Age group: Specific property/state	42	10393	10643	-5154.3	10309	3.08	1	0.08
Age group: Animate/ volitional patient	43	10384	10640	-5148.7	10298	11.17	1	8.3^{e-4} ***
Age group: External agent/cause	44	10385	10648	-5148.7	10297	0.08	1	0.77
Age group: Manner/ location of movement	45	10382	10650	-5145.7	10292	5.93	1	0.015 *
Age group: Degree of change	46	10380	10654	-5143.7	10288	3.99	1	0.046 *

Table 4.9. Model summary: Experiment 2, transitive-only verbs, all age groups, with by-age interactions.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	**χ^2**	**χ df**	***p* value**
Intercept	31	10588	10773	-5263.0	10526			
Age group	32	10516	10706	-5225.9	10452	74.39	1	2.2^{e-16} ***
Entrenchment	33	10508	10705	-5221.0	10442	9.64	1	0.0019 ***
Preemption	34	10510	10713	-5221.0	10442	0.032	1	0.86
Specific property/state	35	10512	10720	-5220.8	10442	0.38	1	0.54
Animate/ volitional patient	36	10509	10724	-5218.7	10437	4.28	1	0.039 *
External agent/cause	37	10508	10728	-5216.8	10434	3.85	1	0.0498 *
Manner/ location of movement	38	10506	10733	-5215.1	10430	3.34	1	0.068
Degree of change	39	10507	10740	-5214.6	10429	1.01	1	0.32
Age group: Entrenchment	40	10501	10740	-5210.7	10421	7.76	1	0.0054 **
Age group: Preemption	41	10502	10746	-5209.8	10420	1.75	1	0.19
Age group: Specific property/state	42	10485	10736	-5200.6	10401	18.42	1	1.8^{e-5} ***
Age group: Animate/ volitional patient	43	10488	10744	-5200.9	10402	0	1	1
Age group: External agent/cause	44	10488	10751	-5200.2	10400	1.30	1	0.25
Age group: Manner/ location of movement	45	10478	10747	-5194.2	10388	12.11	1	5.0^{e-4} ***
Age group: Degree of change	46	10477	10751	-5192.5	10385	3.26	1	0.071

Table 4.10. Model summary: Experiment 2, intransitive-only verbs, all age groups, with by-age interactions.

For intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*), model comparisons revealed main effects of age group and the semantic component *animate/volitional patient*, and significant interactions for age by three of the

semantic components (*animate/volitional patient*, *manner/location of movement* and *degree of change*). Neither the entrenchment nor the preemption predictors yielded significant main effects or interactions, although the main effect for entrenchment was marginal (p=0.07). For transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*), model comparisons revealed main effects of age, total verb frequency (entrenchment) and two of the semantic components (*animate/volitional patient* and *external agent/cause*). Significant interactions for age by total verb frequency and age by two of the semantic components (*specific property/state* and *manner/location of movement*) were also found.

The significant interactions indicate that different age groups were using semantic information differently when making grammaticality judgments for intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*), and both verb frequency and semantic information differently for transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*). We therefore carried out additional analyses by age group to investigate these differences further.

4.2.4.2.1 5- to 6-year-olds

Table 4.11 shows the model summaries for both verb types. Tables 4.12 and 4.13 show the results of the model comparisons for transitive-only and intransitive-only verbs, respectively. For transitive-only verbs, none of the variables significantly predicted difference scores, although the semantic component *specific property/state* was marginally significant in the expected direction. For intransitive-only verbs, total verb frequency (entrenchment) and the semantic component *animate/volitional patient* significantly predicted difference scores in the expected direction.

	Variable	Transitive-only verbs	Intransitive-only verbs
Fixed effects	Intercept	0.83 (0.11)	0.95 (0.13)
Estimate (SE)	Entrenchment	0.12 (0.19)	0.63 (0.19)
	Preemption	-0.10 (0.18)	-0.23 (0.15)
	Specific property/state	-0.19 (0.10)	-0.25 (0.18)
	Animate/volitional patient	0.09 (0.09)	0.14 (0.17)
	External agent/cause	0.07 (0.10)	0.15 (0.15)
	Manner/location of movement	-0.02 (0.10)	-0.05 (0.15)
	Degree of change	-0.10 (0.14)	0.28 (0.17)
Random effects	Verb (intercept)	0.12 (0.34)	0.34 (0.58)
Variance (SD)	Participant (intercept)	0.33 (0.57)	0.35 (0.59)
	Participant – Entrenchment	0.27 (0.52)	0.19 (0.44)
	Participant – Preemption	0.13 (0.36)	n/a
	Participant – Specific property/state	n/a	0.12 (0.34)
	Participant – Animate/volitional patient	n/a	0.08 (0.29)
	Participant – External agent/cause	0.03 (0.18)	0.19 (0.44)
	Participant – Manner/location of movement	0.07 (0.26)	n/a
	Participant – Degree of change	0.14 (0.38)	0.11 (0.33)
Model summary	AIC	3739.3	3947.3
	BIC	3889.5	4098.8
	Log Likelihood	-1838.7	-1942.7
	Deviance	3677.3	3885.3

Table 4.11. Model summary: Experiment 2, 5- to 6-year-olds, all verbs.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	**χ^2**	**χ df**	***p* value**
Intercept	24	3733.0	3849.3	-1842.5	3685.0			
Entrenchment	25	3734.0	3855.2	-1842.0	3684.0	0.94	1	0.33
Preemption	26	3735.5	3861.5	-1841.8	3683.5	0.49	1	0.48
Specific property/state	27	3734.5	3865.3	-1840.2	3680.5	3.08	1	0.08
Animate/ volitional patient	28	3735.1	3870.8	-1839.5	3679.1	1.37	1	0.24
External agent/cause	29	3735.9	3876.4	-1839.0	3677.9	1.18	1	0.28
Manner/ location of movement	30	3737.9	3883.9	-1839.0	3667.9	0.01	1	0.93
Degree of change	31	3739.3	3889.5	-1838.7	3677.3	0.57	1	0.45

Table 4.12. Model comparisons: Experiment 2, 5- to 6-year-olds, transitive-only verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	24	3995.6	4072.1	-1953.8	3907.6			
Entrenchment	25	3949.3	4070.7	-1949.7	3899.3	8.29	1	0.0040 **
Preemption	26	3950.7	4076.9	-1949.4	3898.7	0.60	1	0.44
Specific property/state	27	3951.3	4082.3	-1948.6	3897.3	1.48	1	0.22
Animate/ volitional patient	28	3946.6	4082.5	-1945.3	2890.6	6.67	1	0.0098 **
External agent/cause	29	3946.5	4087.3	-1944.3	3888.5	2.06	1	0.15
Manner/ location of movement	30	3948.5	4094.1	-1944.2	3888.5	0.057	1	0.81
Degree of change	31	3947.3	4097.8	-1942.7	3885.3	3.13	1	0.077

Table 4.13. Model comparisons: Experiment 2, 5- to 6-year-olds, intransitive-only verbs.

4.2.4.2.2 9- to 10-year-olds

Table 4.14 shows the model summaries and Tables 4.15 and 4.16 show the results of the model comparisons for transitive-only and intransitive-only verbs, respectively. For transitive-only verbs, total verb frequency (entrenchment) and the semantic component *animate/volitional patient* significantly predicted difference scores, while verb frequency in the preempting passive construction was not significant. However, the semantic prediction is not in the expected direction: *animate/volitional patient* is a factor related to intransitivity (Pinker, 1989) and, therefore, was expected to be a negative predictor of difference scores for intransitive uses of transitive-only verbs. (The more a transitive verb's semantics are related to having an animate or volitional patient, the more highly that verb was expected to be rated in the intransitive. The difference between judgment scores for the ['grammatical'] transitive and ['ungrammatical'] intransitive sentences was therefore expected to be smaller for verbs rated more highly for this semantic component. However, this was not the case here: the higher a verb's rating on the semantic component, the *larger* the difference score was.)

	Variable	Transitive-only verbs	Intransitive-only verbs
Fixed effects	Intercept	1.65 (0.16)	2.11 (0.14)
Estimate (SE)	Entrenchment	0.22 (0.30)	0.58 (0.19)
	Preemption	0.06 (0.29)	-0.11 (0.16)
	Specific property/state	-0.20 (0.17)	-0.15 (0.19)
	Animate/volitional patient	0.47 (0.16)	-0.06 (0.18)
	External agent/cause	0.01 (0.16)	0.24 (0.15)
	Manner/location of movement	-0.10 (0.16)	-0.17 (0.15)
	Degree of change	0.14 (0.21)	0.18 (0.17)
Random effects	Verb (intercept)	0.60 (0.78)	0.42 (0.65)
Variance (SD)	Participant (intercept)	0.42 (0.64)	0.49 (0.70)
	Participant – Entrenchment	0.02 (0.14)	0.08 (0.28)
	Participant – Preemption	0.09 (0.31)	0.11 (0.34)
	Participant – Specific property/state	n/a	0.08 (0.29)
	Participant – Animate/volitional patient	0.08 (0.28)	0.11 (0.34)
	Participant – External agent/cause	0.07 (0.26)	0.09 (0.29)
	Participant – Manner/location of movement	n/a	n/a
	Participant – Degree of change	0.06 (0.24)	0.14 (0.37)
Model summary	AIC	3591.3	3536.9
	BIC	3742.2	3721.9
	Log Likelihood	-1764.7	-1730.5
	Deviance	3529.3	3460.9

Table 4.14. Model summary: Experiment 2, 9- to 10-year-olds, all verbs.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	**χ^2**	**χ df**	***p* value**
Intercept	24	3596.5	3713.3	-1774.3	3548.5			
Entrenchment	25	3594.6	3716.3	-1772.3	3544.6	3.94	1	0.047 *
Preemption	26	3595.1	3721.6	-1771.5	3543.1	1.51	1	0.22
Specific property/state	27	3596.8	3728.2	-1771.4	2542.8	0.25	1	0.62
Animate/ volitional patient	28	3586.6	3722.9	-1765.3	3530.6	12.21	1	4.8^{e-4} ***
External agent/cause	29	3588.6	3729.8	-1765.3	3530.6	0.02	1	0.89
Manner/ location of movement	30	6589.8	3735.8	-1764.9	3529.8	0.78	1	0.38
Degree of change	31	3591.3	3742.2	-1764.7	3529.3	0.49	1	0.48

Table 4.15. Model comparisons: Experiment 2, 9- to 10-year-olds, transitive-only verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	31	3540.6	3691.5	-1739.3	3478.6			
Entrenchment	32	3533.9	3689.7	-1735.0	3469.9	8.65	1	0.0033 **
Preemption	33	3535.9	3696.5	-1735.0	3469.9	0.0036	1	0.95
Specific property/state	34	3537.2	3702.7	-1734.6	3469.2	0.70	1	0.40
Animate/ volitional patient	35	3537.7	3708.0	-1733.8	3467.7	1.54	1	0.21
External agent/cause	36	3535.6	3710.8	-1731.8	3463.6	4.09	1	0.043 *
Manner/ location of movement	37	3536.2	3716.3	-1731.1	3462.2	1.41	1	0.23
Degree of change	38	3536.9	3721.9	-1730.5	3460.9	1.26	1	0.26

Table 4.16. Model comparisons: Experiment 2, 9- to 10-year-olds, intransitive-only verbs.

For intransitive-only verbs, total verb frequency and the semantic component *external agent/cause* significantly predicted difference scores, while verb frequency in the preempting periphrastic causative construction was not significant. Again, the semantic component *external agent/cause* predicted results in the opposite direction to what was expected: the more a verb's semantics were in line with this factor, which relates to transitivity, the greater participants' difference scores for ungrammatical transitive uses of these verbs. Taken together with the findings for transitive-only verbs, it seems that 9- to 10-year-olds are not using semantic information as predicted, although this information does seem to affect their judgments in some way.

4.2.4.2.3 Adults

Table 4.17 shows the model summaries and Tables 4.18 and 4.19 show the results of the model comparisons for transitive-only and intransitive-only verbs, respectively. For transitive-only verbs, the semantic components *animate/volitional patient* and *manner/location of movement* were both significant predictors of difference scores, while total verb frequency (entrenchment) and verb frequency in the preempting

passive construction were marginally significant predictors. The semantic component *animate/volitional patient* relates to intransitivity but here is a positive predictor of difference scores for ungrammatical intransitive uses of transitive-only verbs, the opposite direction to our original expectation.

	Variable	**Transitive-only verbs**	**Intransitive-only verbs**
Fixed effects	Intercept	2.18 (0.13)	2.81 (0.13)
Estimate (SE)	Entrenchment	0.30 (0.19)	0.43 (0.15)
	Preemption	-0.03 (0.18)	-0.12 (0.13)
	Specific property/state	-0.05 (0.11)	-0.07 (0.15)
	Animate/volitional patient	0.29 (0.10)	0.04 (0.14)
	External agent/cause	0.12 (0.10)	0.22 (0.12)
	Manner/location of movement	-0.19 (0.10)	-0.35 (0.12)
	Degree of change	0.13 (0.13)	0.07 (0.13)
Random effects	Verb (intercept)	0.23 (0.48)	0.29 (0.45)
Variance (SD)	Participant (intercept)	0.19 (0.43)	0.18 (0.42)
	Participant – Entrenchment	0.06 (0.24)	0.02 (0.13)
	Participant – Preemption	0.03 (0.17)	n/a
	Participant – Specific property/state	0.02 (0.14)	n/a
	Participant – Animate/volitional patient	n/a	0.01 (0.08)
	Participant – External agent/cause	0.01 (0.08)	0.00 (0.06)
	Participant – Manner/location of movement	0.02 (0.12)	0.01 (0.09)
	Participant – Degree of change	n/a	n/a
Model summary	AIC	2874.4	2689.2
	BIC	3025.3	2810.9
	Log Likelihood	-1406.2	-1319.6
	Deviance	2812.4	2689.2

Table 4.17. Model summary: Experiment 2, adults, all verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	24	2887.0	3003.8	-1419.5	2839.0			
Entrenchment	25	2886.0	3007.7	-1418.0	2836.0	3.01	1	0.08
Preemption	26	2885.1	3011.6	-1416.5	2833.1	2.91	1	0.09
Specific property/state	27	2887.1	3018.5	-1416.5	2833.1	0.01	1	0.91
Animate/ volitional patient	28	2876.1	3012.4	-1410.0	2820.1	12.98	1	3.1^{e-4} ***
External agent/cause	29	2876.7	3017.8	-1409.3	2818.7	1.45	1	0.23
Manner/ location of movement	30	2873.5	3019.6	-1406.8	2813.5	5.11	1	0.024 *
Degree of change	31	2874.4	3025.3	-1406.2	2812.4	1.14	1	0.29

Table 4.18. Model comparisons: Experiment 2, adults, transitive-only verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	18	2701.8	2789.3	-1332.9	2665.8			
Entrenchment	19	2700.4	2792.9	-1331.2	2662.4	3.32	1	0.068
Preemption	20	2702.2	2799.6	-1331.1	2662.2	0.18	1	0.67
Specific property/state	21	2699.8	2802.0	-1328.9	2657.8	4.47	1	0.034 *
Animate/ volitional patient	22	2967.4	2804.4	-1326.7	2653.4	4.41	1	0.036 *
External agent/cause	23	2694.2	2806.1	-1324.1	2648.2	5.18	1	0.023*
Manner/ location of movement	24	2687.6	2804.4	-1319.8	2639.6	8.57	1	0.0034 **
Degree of change	25	2689.2	2810.9	-1319.6	2689.2	0.40	1	0.53

Table 4.19. Model comparisons: Experiment 2, adults, intransitive-only verbs.

For intransitive-only verbs, results show that semantic components *specific property/state*, *animate/volitional patient*, *external agent/cause* and *manner/location of movement* all significantly predicted difference scores, while total verb frequency (entrenchment) again did so only marginally, and preemption not at all. As with the transitive-only verbs, adults' difference scores were not always predicted in the expected direction by the semantic components. Predictions in the opposite direction to that expected were found with the components *specific property/state*, which

relates to intransitivity and was therefore expected to be a positive predictor, and *external agent/cause*, which related to transitivity and was therefore expected to be a negative predictor. Like the 9- to 10-year-olds, then, and contrary to findings in Experiment 1, adults in this study do not seem to be using semantic information as expected.

4.2.4.3 Discussion

Overall, the results from Experiment 2 show some support for the entrenchment hypothesis, but none for the preemption hypothesis. The entrenchment measure (total verb frequency) significantly predicted difference scores for transitive-only verbs for the 9- to 10-year-olds and marginally for adults. For intransitive-only verbs, entrenchment was a significant predictor for both groups of children, although not for adults (unlike Experiment 1). Therefore, it is likely that the influence of entrenchment on grammaticality judgments decreases with age, explaining the interaction of age by entrenchment for intransitive-only verbs in the original model in Table 4.8. In contrast, the preemption predictors (verb frequency in the passive for transitive-only verbs and in the periphrastic causative for intransitive-only verbs) did not significantly predict any results (although preemption was a marginally significant predictor for adults), contrary to findings with studies investigating other constructions.

Strong support for the importance of semantics was suggested in Experiment 2. For all age groups, at least one of the semantic measures significantly predicted difference scores for ungrammatical transitive uses of intransitive-only verbs (e.g. **Homer swam the fish*) and, for the two older age groups, for ungrammatical intransitive uses of transitive-only verbs (e.g. **The ball kicked*). However, the older children and adults did not always seem to be using semantic information as expected, with predictions for some of the semantic components going in the opposite direction to our expectations. Age by semantics interactions in the original model (Table 4.8) appear to be due to different semantic components playing different roles over the different age groups, with semantics in general apparently becoming more important as age increases (and as reliance on frequency information appears to decrease).

In addition, the finding that none of the variables was able to predict grammaticality judgments for intransitivisation errors with transitive-only verbs in the youngest age group indicates that these children might be unable to interpret errors of this type, therefore rendering children incapable of making systematic judgments on the acceptability of these sentences.

4.2.5 Experiment 3 (children – production-priming)

Experiments 1 and 2 confirmed that adults and older children use a combination of verb frequency and semantic information when making grammaticality judgments of intransitivisation and transitivisation overgeneralisations. For 5- to 6-year-olds, however, the results so far have suggested that both entrenchment and semantics, but not preemption, influence judgments for transitive overgeneralisations with intransitive-only verbs, with no significant findings for intransitive overgeneralisations with transitive-only verbs. While these results tell us something about children's metalinguistic knowledge of such errors, Experiment 3 allows us to investigate our main question of interest: whether children will actually *produce* overgeneralisation errors with transitive-only and intransitive-only verbs and, if so, what factors predict by-verb differences in error rates: entrenchment, preemption or semantics. Argument structure overgeneralisation errors in naturalistic diary studies are sparse and limited to those that the transcriber happens to notice. So, to test our three hypotheses, we used an elicited production-priming task aimed at encouraging the production of overgeneralisation errors in 5- to 6-year-old children.

4.2.5.1 Method

4.2.5.1.1 Participants

The participants were 64 children aged 5-6 (5;2-6;4, *M*=5;8) recruited from primary schools in the North West of England. All were monolingual speakers of English and had no known language impairments. None of these children had participated in Experiment 2.

4.2.5.1.2 Test items

Test items were the same as Experiment 2, with the addition of a single alternating verb (*float*, produced by the experimenter only), added for the purpose of the bingo game described below.

4.2.5.1.3 Materials

The 120 verbs (40 each of transitive-only, intransitive-only and alternating) were split into four sets, each containing 20 alternating verbs and 10 each of the transitive-only and intransitive-only verbs. Alternating verbs were therefore used twice as many times in total as fixed-transitivity verbs, since they were used in both priming conditions. Each child received a single verb set for their target verbs. The first experimenter used 10 transitive-only and 10 intransitive-only verbs for the prime sentences; there was therefore no overlap between the verbs received by the child and those produced by the experimenter.

4.2.5.1.4 Procedure

The aim of this experiment was to encourage children to produce both intransitivisation errors with transitive-only verbs (e.g. **The ball kicked; cf. Homer kicked the ball*) and transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish; cf. The fish swam*). In order to do this, we used a production-priming methodology in which an experimenter produced grammatical intransitive-inchoative sentences to encourage the child to use this construction with transitive-only verbs. On a separate day, children were primed with grammatical transitive-causative sentences to encourage the production of transitivisation errors with intransitive-only verbs. Examples of trials in each prime condition are given below, with the target error we were attempting to elicit.

(1) Intransitive prime condition (transitive-only target verbs)

a. Experimenter 2 (clue words): lightbulb, glow
b. Experimenter 1: The lightbulb glowed
c. Experimenter 2 (clue words): ball, hit
d. Child: *The ball hit

(2) Transitive prime condition (intransitive-only target verbs)

a. Experimenter 2 (clue words): lift, bag
b. Experimenter 1: Bart lifted the bag
c. Experimenter 2 (clue words): wait, boy
d. Child: *Lisa waited the boy

Each child participated on two occasions, on separate days. In each session, children took turns with an experimenter to describe a series of animations. These animations were presented using *Processing* (www.processing.org). Both experimenter and child were given 'clue words' by a second experimenter to encourage them to use the intended verb. The clue words consisted of the verb followed by the direct object, when transitive sentences were being primed, or the subject followed by the verb, when intransitive sentences were being primed. The second experimenter noted down children's responses, although all sessions were also audio-recorded using *Audacity* in order to check responses later if there was any doubt about what the child had said.

Half of the children received transitive primes on the first day and intransitive primes on the second, and vice versa for the other children. The first three pairs of animations were training trials containing only transitive-only or intransitive-only verbs for both experimenter and child, whichever the child was to be primed with on that day. These verbs were not in the child's verb set, nor were they used as primes by the experimenter in that child's test trials. Twenty test trials then followed, with the experimenter continuing to use transitive-only or intransitive-only verbs, depending on prime condition. The experimenter produced only grammatical sentences. In contrast, half of the target verbs given to the children were alternating verbs (and would therefore be grammatical whether the child produced a transitive or an intransitive sentence) and half were intransitive-only or transitive-only, whichever was the *opposite* of the prime condition. For these trials, if the child produced a

sentence using the same construction as that with which they had been primed, an overgeneralisation error would result.

In order to motivate the children to produce the sentences, a 'bingo game' was used (as in Rowland, Chang, Ambridge, Pine & Lieven, 2012). Each time Experimenter 1 or the child produced a sentence, Experimenter 2 (who could not see the computer screen) looked for a matching bingo card. In fact, Experimenter 2 had all of the bingo cards and whether or not the card was given to Experimenter 1/the child was predetermined: the games were fixed so that the child always won both games on the first day, lost the first game on the second day (in order to maintain tension) and the won the final game. This manipulation required an extra trial for Experimenter 1 only, on Day 2, always with the (alternating) verb *float*. Each 'game' lasted for ten trials each, in order to keep the child's attention and motivation.

4.2.5.2 Results

Children's responses were coded for sentence type: transitive (active), intransitive, passive (full or truncated), periphrastic causative, other use of the verb, irrelevant (target verb not included/no response). As we are investigating overgeneralisation errors, the constructions of interest were intransitive uses of transitive-only verbs and transitive uses of intransitive-only verbs. Sentences were only included in the analysis if the child used the target verb in his/her response, with error rate calculated as a proportion of errors from the total number of responses that included the target verb. Replacement of NPs with pronouns or generic terms was allowed (e.g. *the dad hit the ball* for *Homer hit the ball*; *it fell* for *the cup fell*), as were changes in tense/aspect (e.g. *Homer hit/hits/was hitting the ball*), morphological overgeneralisations (e.g. *The ball hitted*) and additional modifying phrases (e.g. *He kicked the ball in the goal*). The mean number of sentences of each type produced by each child is shown in Figure 4.2, of a possible maximum of 10.

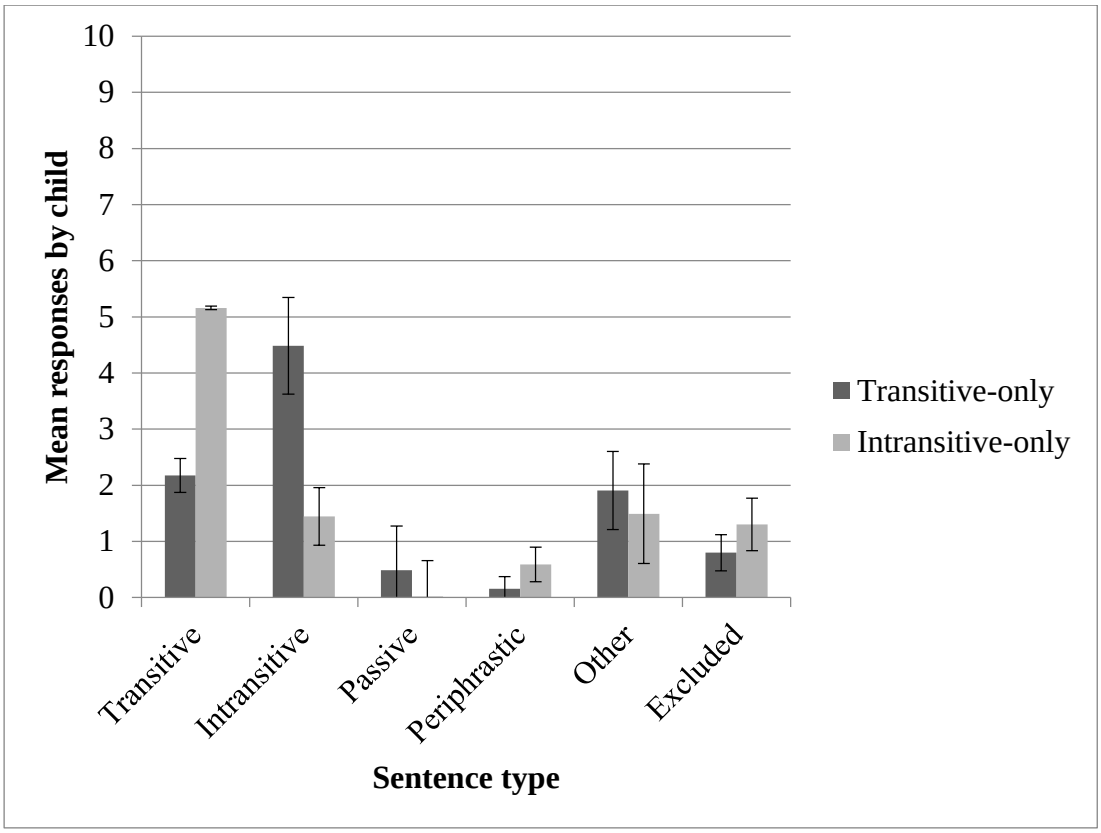

Figure 4.2. Mean number of sentences of each type produced by each child (*N*=64), with 95% CIs. Total number of sentences produced per child per verb type is 10. Black bars represent verbs that are considered grammatical only in the transitive construction (transitive-only verbs) and grey bars represent verbs that are considered grammatical only in the intransitive construction (intransitive-only verbs). The constructions given on X-axis are those produced by the child.

The binary dependent variable for this experiment is the child's response: overgeneralisation error (1) or other use of the target verb (0), with all responses in which the child did not use the target verb excluded from the analysis. As this is a binary variable, results were analysed using the *glmer* function of the *lme4* package (version 1.1-7, Bates et al., 2014), with family=binomial. Predictor variables were the same as Experiment 1. Random intercepts for verb and participant were included in the models, although all random slopes were removed in order to enable the models to converge.

As can be seen in Figure 4.2, our production-priming method succeeded in eliciting a large number of overgeneralisation errors, in both directions, from children age 5-6 years. The success of this task gives us a unique insight into the

mechanisms that influence children's error production for a large number of children with a large number of known verbs, a big advantage over corpus or diary studies, or elicitation tasks using a small set of novel verbs. The large number of errors children produced enabled us to examine by-verb differences, as well as testing if the mechanisms under investigation (entrenchment, preemption and semantics) pattern similarly in their predictions of error rates to their predictions of grammaticality judgments.

Table 4.20 shows the model summaries and Tables 4.21 and 4.22 show the results of the model comparisons for transitive-only and intransitive-only verbs, respectively. For intransitivisation errors with transitive-only verbs (e.g.*_The ball kicked_), semantic component *external agent/cause* was the only significant predictor of error rate. This prediction is in the expected direction. Neither the entrenchment measure nor the preemption measure predicted error rates.

	Variable	**Transitive-only verbs**	**Intransitive-only verbs**
Fixed effects	Intercept	1.24 (0.49)	1.98 (0.52)
Estimate (SE)	Entrenchment	-0.51 (0.49)	-1.00 (0.32)
	Preemption	-0.04 (0.48)	-0.24 (0.27)
	Specific property/state	0.20 (0.24)	0.20 (0.26)
	Animate/volitional patient	-0.18 (0.26)	-0.16 (0.27)
	External agent/cause	-0.78 (0.28)	-0.15 (0.24)
	Manner/location of movement	0.01 (0.27)	0.24 (0.23)
	Degree of change	-0.16 (0.29)	-0.72 (0.27)
Random effects Variance (SD)	Verb (intercept)	0.85 (0.92)	0.12 (0.35)
	Participant (intercept)	8.86 (2.98)	8.65 (2.94)
Model summary	AIC	461.09	391.07
	BIC	502.56	432.25
	Log Likelihood	-220.55	-185.53
	Deviance	441.09	371.07

Table 4.20. Model summary: Experiment 3, all verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	3	457.93	470.37	-225.97	451.93			
Entrenchment	4	459.39	475.97	-225.69	451.39	0.55	1	0.46
Preemption	5	461.38	482.11	-225.69	451.38	0.00	1	0.95
Specific property/state	6	461.68	486.56	-224.84	449.68	1.70	1	0.19
Animate/ volitional patient	7	463.36	492.39	-224.68	449.36	0.32	1	0.57
External agent/cause	8	457.44	490.61	-220.72	441.44	7.29	1	0.0049 **
Manner/ location of movement	9	459.40	496.71	-220.70	441.40	0.04	1	0.83
Degree of change	10	461.09	502.56	-220.55	441.09	0.30	1	0.58

Table 4.21. Model comparisons: Experiment 3, transitive-only verbs.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	3	412.77	425.12	-203.38	406.77			
Entrenchment	4	402.03	418.50	-197.01	394.03	12.74	1	3.6^{e-4} ***
Preemption	5	398.45	419.04	-194.22	388.45	5.58	1	0.018*
Specific property/state	6	400.45	425.16	-194.22	388.45	0.00	1	0.97
Animate/ volitional patient	7	395.34	424.16	190.67	381.34	7.11	1	0.0077 **
External agent/cause	8	395.43	428.38	-189.72	379.43	1.90	1	0.17
Manner/ location of movement	9	396.74	433.81	-189.37	378.74	0.69	1	0.41
Degree of change	10	391.07	432.25	-185.53	371.07	7.68	1	0.0056 **

Table 4.22. Model comparisons: Experiment 3, intransitive-only verbs.

For transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*), total verb frequency (entrenchment), verb frequency in the periphrastic causative (preemption) and the semantic components *animate/volitional patient* and *degree of change* all significantly predicted children's error rates in production. The semantic components *animate/volitional patient* and *degree of change* both relate to intransitivity and are negative predictors of error rates: the greater the extent to which

the semantics of an intransitive verb are in line with these components, the less likely children are to overgeneralise that verb into the transitive-causative construction. That is, children make fewer transitivisation errors with intransitive-only verbs the higher a verb is rated on these semantic components, usually associated with intransitive-only verbs (e.g. Pinker, 1989). This indicated that, the more typically intransitive a verb is, the less likely children are to overgeneralise it into the transitive construction (e.g. **Homer swam the fish*).

In summary, entrenchment and preemption measures predict error rates for transitivisation errors with intransitive-only verbs, but not the converse error. In terms of semantics, for both verb types, overgeneralisation errors were significantly predicted, in the expected direction, by individual semantic components. The results for this production study pattern in a similar way to the grammaticality judgment task in Experiment 2, thus validating the use of both methodologies here, as well as in previous studies.

4.3 General discussion

This paper used a multi-methodological approach to investigate three hypotheses that aim to explain children's retreat from overgeneralisation errors. The entrenchment hypothesis posits that children retreat from or avoid error through encountering verbs which are used multiple times in grammatical constructions (e.g. *Bart kicked the ball*, *The fish swam*). Children then infer from their absence in the input that ungrammatical verb-construction combinations (e.g. **The ball kicked*, **Homer swam the fish*) are not possible, and the more they hear the grammatical sentences, the greater the strength of this inference. The preemption hypothesis posits that hearing the verb used only in grammatical constructions *with a similar meaning* (e.g. *The ball was kicked*, *Homer made the fish swim*) will help children retreat from or avoid error, as they have a viable alternative available to express the same meaning. Again, the prediction is that the more children hear a given verb in these preempting constructions, the better able they will be to retreat from (or avoid) error. Finally, the semantics hypothesis predicts that, as children's knowledge of verb semantics and how these fit with the semantics of the construction increases, their errors will decrease because they will become aware that certain verb-construction pairings are not semantically compatible (e.g. it is not possible to say **The ball kicked* because an

external agent is required; it is not possible to say *_Homer swam the fish_ because the swimming motion is internally caused).

Previous studies (e.g. Ambridge et al., 2008; 2014; Bidgood et al., 2014; Brooks & Tomasello, 1999) have shown support for a mixture of the three hypotheses under investigation, with a number of different constructions. However, most of these have used only a small number of verbs and/or relied on the use of novel verbs. Studies using larger numbers of verbs have employed a grammaticality judgment methodology, whereas overgeneralisation is really a production phenomenon. The three experiments presented here investigated intransitivisation errors with transitive-only verbs (e.g. *_The ball kicked_) and transitivisation errors with intransitive-only verbs (e.g. *_Homer swam the fish_), using grammaticality judgment tasks with adults and children (Experiments 1 and 2) and a production-priming task with 5- to 6-year-olds (Experiment 3). Grammaticality judgments allow for a comparison with findings from previous studies with other constructions (e.g. Ambridge, Pine & Rowland, 2012), with the priming task giving us a crucial insight into children's production of errors. Overall, the results provided strong support for entrenchment and semantic mechanisms, suggesting both are likely to play a role in the retreat from overgeneralisation for errors involving the transitive-causative/intransitive-inchoative alternation. Evidence for a preemption mechanism is found, but only in the production task.

Echoing the findings of previous papers (e.g. Ambridge et al., 2008; Boyd et al., 2012; Naigles et al., 1992), clear support was found for a role played by entrenchment in the grammaticality judgments of adults and older children with both intransitivisation errors of transitive verbs (e.g. *_The ball kicked_) and transitivisation errors of intransitive verbs (e.g. *_Homer swam the fish_). For the younger group of children, the operation of this mechanism is clear with transitivisation errors of intransitive verbs in the judgment task and, importantly, in the production task. Therefore, the entrenchment mechanism clearly has a role to play in the retreat from overgeneralisation for errors involving the transitive-causative/intransitive-inchoative alternation.

In contrast to previous findings (e.g. Boyd et al., 2012; Brooks & Tomasello, 1999), the present study found only marginal support for the preemption hypothesis. In Experiment 1, verb frequency in the preempting construction predicted errors in an unexpected direction for intransitivisation errors with transitive-only verbs, and

only marginally predicted transitivisation errors with intransitive-only verbs. Preemption was unable to predict difference scores in the grammaticality judgment task for any age group in Experiment 2. However, in the production study, preemption did explain additional variance over and above that explained by entrenchment for transitivisation errors with intransitive-only verbs. This is likely to be due to a task effect (see also Blything, Ambridge & Lieven, 2014). In production, all possible constructions are competing to express the message the child wants to convey, thus forcing a choice between the alternative constructions – transitive vs. periphrastic in this case. In contrast, judgment tasks do not require participants to consider the alternative ways in which the message could have been conveyed. Other possibilities relate to the frequency of the preempting constructions: very low numbers of uses of the verbs in these constructions were found in the corpus. It could be that: (1) the verbs we chose for this study did not have a wide enough spread of uses in the preempting constructions for the mechanism's operation to be observed; (2) the corpus we used was too small to obtain accurate counts of preempting constructions; or, most likely, (3) preemption is genuinely unable to operate for the overgeneralisation errors under investigation due to the low frequency of the preempting constructions in the input. In any case, while the preemption mechanism is not well supported in our grammaticality judgment tasks, importantly, it does play a role in children's error production and, therefore, their retreat from error.

As in several previous studies (e.g. Ambridge et al., 2008; Brooks & Tomasello, 1999), evidence for a semantics mechanism is strong in all three experiments presented here, with semantic components predicting difference scores in both grammaticality judgment tasks and, importantly, in the production task. However, the direction of prediction was not always that which was expected for adults and older children in Experiment 2. Since results were as expected in Experiment 1, the smaller number of verbs in Experiment 2 could have caused problems with the analysis, possibly related to the fact that the PCA from which the semantic components were drawn involved the entire set of verbs used in Experiment 1. Another explanation lies in the process of labelling of the semantic components. While conducting PCA is essential in order to (a) make the number of components manageable, and (b) collapse individual features explaining overlapping variance, the labelling of components drawn from a PCA is inherently problematic due to the large

number of individual features loading onto each of these, thus making them difficult to interpret.

To investigate this possibility further, Table 4.23 shows the direction of prediction for each semantic component. The Table illustrates the fact that each semantic component either consistently predicts difference scores positively (and error rates negatively) or negatively (and error rates positively), irrespective of verb type. So, while these factors sometimes predicted results in the opposite direction to expectations, the direction of predictions are consistent across experiments and age groups. This could indicate that, rather than creating semantic components that indicate how 'transitive' or 'intransitive' a verb is, we have a set of semantic components that predict *whether or not a verb can alternate* between the two constructions under investigation. Regardless of whether this explanation is correct, the fact remains that verb semantics are able to predict both grammaticality judgments and error rates for both types of overgeneralisation investigated here, lending strong support to the semantics hypothesis. A mechanism involving verb semantics therefore clearly has a role to play in the retreat from overgeneralisation for errors involving the transitive-causative/intransitive-inchoative alternation.

	Exp. 1		**Exp. 2 (Age 5)**		**Exp. 2 (Age 9)**		**Exp. 2 (Adults)**		**Exp. 3**	
Semantic component	**TRN**	**INT**	**TRN**	**INT**	**TRN**	**INT**	**TRN**	**INT**	**TRN**	**INT**
Specific property/ state (INT)	-	n/a	n/a	n/a	n/a	n/a	n/a	-	n/a	n/a
Animate/volitional patient (INT)	n/a	+	n/a	+	+	n/a	+	+	n/a	-
External agent/ cause (TRN)	+	n/a	n/a	n/a	n/a	+	n/a	+	-	n/a
Manner/location of movement (ALT)	n/a	-	n/a	n/a	n/a	n/a	-	-	n/a	n/a
Degree of change (INT)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	-

Table 4.23. Direction of prediction of semantic components in all experiments. Expected relationship with (in)transtitivity is indicated in the first column: INT = intranstivity; TRN = transitivity; ALT = alternating.

4.3.1 Changes across development

The results of Experiment 2 suggest that the importance of entrenchment decreases with development, as verb frequency effects are observed for both age groups of children but not for adults. However, entrenchment does predict grammaticality judgments for the adults tested in Experiment 1, in line with previous findings (e.g. Ambridge et al., 2008; 2011). Experiment 1 included more verbs than Experiment 2, with the majority of these additional verbs being of low frequency. This suggests the existence of some sort of ceiling effect in the usefulness of statistical mechanisms: the more times a verb has been encountered in the input, the less of an influence hearing it yet more times will have on judgments of grammatical acceptability, perhaps with the effect levelling off for the most frequent verbs. The children in Experiment 2 may not have reached this threshold with the verbs included in that study, whereas adults have not reached the threshold with the lower frequency verbs in Experiment 1.

Results in Experiment 2 also suggest that, as the importance of entrenchment decreases over time, the role of semantics comes to play a more important role. This could imply simply that semantics explain a larger portion of variance once frequency effects are no longer important. However, it is likely that semantics are better-learnt in this oldest age group, therefore allowing the mechanism to operate fully in adults when it was unable to do so in children. This suggestion is consistent with Blything et al.'s (2014) investigation of *un-* prefixation errors (e.g. **unopen*, **unsqueeze*).

4.3.2 How do children retreat from overgeneralisation errors?

The results of the three experiments reported in this paper have demonstrated strong support for the entrenchment and semantics hypotheses in the retreat from overgeneralisation errors. Both factors were significant predictors for grammaticality judgment ratings in different age groups, and error rate in a production-priming task. The preemption hypothesis also received some support, notably in the critical production task.

These findings provide additional evidence for the claims of other recent work that suggest the need for an account of language acquisition, including the

retreat from overgeneralisation errors, that integrates semantic and statistical accounts (e.g. Ambridge et al., 2008; 2011). Taking evidence from the current paper along with previous papers allows us to take a wider view of the mechanisms involved in the retreat from overgeneralisation errors by looking across different constructions. Preemption is important when the erroneous construction and its preempting alternative have an almost identical meaning, as with the prepositional object and double object datives (e.g. *Homer gave the book to Marge* vs. *Homer gave Marge the book*; Ambridge et al., 2014) and especially the case of *un-* prefixation errors (e.g. *close* preempts **unopen*; Ambridge, 2013; Blything et al., 2014). Preemption also works best when the preempting construction is relatively high frequency, as in the dative (see above) and locative alternations (e.g. *Lisa sprayed the flowers with water* vs. *Lisa sprayed water onto the flowers*; Ambridge et al., 2011).

The current has demonstrated that the low frequencies of the passive and periphrastic causative, and a potential lack of close semantic compatibility with the constructions they could preempt, work against preemption in the case of the transitive-causative/intransitive-inchoative alternation. Entrenchment, on the other hand, has been shown to work well in the absence of a closely preempting alternative construction of high frequency. Semantics plays an important role in the retreat from overgeneralisation when there is a clear split between the two constructions involved in the alternation – this is the case in the current paper, but less so for the alternative dative constructions, for example. Finally, when children are forced to choose between possible alternative constructions, as in the production-priming task used in Experiment 3, but not in the grammaticality judgment task used in Experiment 2, the role of preemption plays a more important role. This echoes findings in Blything et al.'s (2014) paper on *un-* prefixation errors.

4.3.3 Conclusion

This paper tested the predictions of three accounts of the retreat from overgeneralisation in the transitive-causative construction: entrenchment, preemption and semantics. Using a combination of grammaticality judgment and production-priming methods, testing a large number of verbs and treating predictor variables as continuous has allowed for a particularly strong test of these hypotheses. Results

strongly support both the entrenchment and semantics hypotheses, and the preemption hypothesis to a lesser degree. We therefore suggest that a successful account of the retreat from overgeneralisation must therefore posit a role for both statistical and semantic mechanisms, such as the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011).

Chapter 5: Children's and adults' passive syntax is semantically constrained: Evidence from syntactic priming

5.0 Fit within the thesis

Chapters 3 and 4 investigated the roles of semantics, entrenchment and preemption in children's retreat from argument structure overgeneralisation errors, finding support for semantics and entrenchment in particular. However, the retreat from overgeneralisation is a relatively narrow focus for any theory. Finding support for these mechanisms in language acquisition more generally is therefore necessary if they are to be taken seriously. This chapter focuses on one of these mechanisms – semantics – investigating its role in children's acquisition of the passive construction. It presents a production-priming study with both adults and children, using the priming method of Chapter 4. It used two approaches to semantics: the more traditional semantic verb class approach taken Chapter 3, and the fine-grained continuum approach to semantics taken in Chapter 4. The idea of a continuum of semantic fit between the verb and the construction being a key factor in explaining by-verb differences is therefore developed here. If this approach can successfully explain differences in the acquisition of the passive construction, then semantic fit is not just relevant to the retreat from overgeneralisation, but to the acquisition of syntax more generally. The passive construction is particularly appropriate as a test of our semantics approach (verb-in-construction compatibility) as previous studies (e.g. Maratsos et al., 1985; Pinker et al., 1987) have found that children have difficulties with certain verbs in the passive until a relatively advanced age.

This chapter is currently being prepared for submission to a peer-reviewed journal. For the purposes of publication, it has been written as a response to Messenger et al. (2012). (Note that, unlike the previous chapters, verb frequency information has not been included in the analyses in the current chapter, as Messenger et al. did not include frequency information in their analyses. Nevertheless, see Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2015, for a similar analysis that included frequency information but, nevertheless, came to similar conclusions regarding the role of semantics.) Messenger et al.'s conclusion that children have an abstract representation of the passive by the age of 3 years is supported by our findings, but the details of their conclusion are questioned, as our

modified method yielded by-verb differences that could not be seen given the method used by Messenger et al. Although the initial focus of this chapter is on early semantic restriction versus early abstraction accounts, its focus is on finding an alternative approach to these, based on verb-in-construction compatibility. It therefore challenges generativist-nativist assumptions surrounding early abstraction, and illustrates how the approach developed throughout the Thesis so far can explain more than simply children's retreat from overgeneralisation errors.

5.1 Introduction

A key question in language acquisition is the extent to which children's earliest knowledge of syntax is abstract, with some researchers arguing that it is largely abstract from the beginning of multi-word speech (e.g. Wexler, 1998; Gertner et al., 2006) and others suggesting that this knowledge is initially lexically or semantically restricted (e.g. Schlesinger, 1988; Tomasello, 2003). A construction that has often been studied with regard to this debate is the English passive, which makes for a particularly useful test case as it is one of few constructions for which children (and even adults; e.g. Dąbrowska & Street, 2006) make errors in comprehension (e.g. Maratsos et al., 1985; Sudhalter & Braine, 1985; Gordon & Chafetz, 1990; Fox & Grodzinksky, 1998; Meints, 1999; Hirsch & Wexler, 2006).

Many studies of the passive have used syntactic priming (see Branigan, 2007, and Pickering & Ferreira, 2008, for reviews). This method is particularly suitable for investigating the abstractness of linguistic representations, as participants cannot be primed to produce a syntactic structure for which they do not have an abstract representation (Branigan et al., 1995). Priming has been used to investigate the passive in both adults (e.g. Bock, 1986; Bock & Loebell, 1990; Bock, Loebell & Morey, 1992) and children (e.g. Lempert, 1990; Savage, Lieven, Theakston & Tomasello, 2003; 2006; Huttenlocher et al., 2004; Bencini & Valian, 2008; Messenger, Branigan & McLean, 2011a; 2011b; Messenger et al., 2012). While a number of previous studies have found that performance differs on a verb-by-verb basis, supporting the idea of a semantic restriction on the passive, others have observed excellent performance across verbs, supporting the early abstraction approach. In the present article, we aim to reconcile these disparate findings by means of a new passive priming study with children (aged 4-6) and adults.

5.1.1 The early semantic restriction account

Previous evidence for the early semantic restriction account comes from studies showing that children appear to struggle more with comprehension and production of passive sentences with mental state verbs (e.g. *Bob was seen by Wendy*) than with actional verbs (e.g. *Bob was hit by Wendy*) (e.g. Maratsos et al., 1985; Pinker et al., 1987). Maratsos et al. (1985) used a sentence comprehension task to compare young children's (aged from 4 to 11 years) understanding of active and passive sentences using (a) prototypical actional verbs (agent-patient [AP] verbs, e.g. *hold*) and (b) mental verbs (experiencer-theme [ET] verbs, e.g. *like*). All children performed above chance with both verb types in active sentences (e.g. *Batman holds Donald [Duck]*, *Batman sees Donald*) and with AP verbs in passive sentences (e.g. *Donald is held by Batman*). In contrast, children as old as 7 years struggled with ET verbs in passive sentences (e.g. *Donald is seen by Batman*). The authors concluded that children begin with a prototypical passive construction, which involves verbs with a high degree of transitivity, such as AP verbs. As they get older, their representation of the passive is broadened to include other verbs, before reaching the adult-like state in which most (although not all) transitive verbs can be used in the passive (cf. e.g. *Donald was seen by Batman* but not **1lb was weighed by the package*). Additional evidence that early knowledge of the passive is semantically restricted comes from further studies of comprehension (e.g. Sudhalter & Braine, 1985; Gordon & Chaftez, 1990; Meints, 1999) and production (e.g. Tomasello, Brooks and Stern, 1998; Meints, 1999), naturalistic data (e.g. Israel, Johnson & Brooks, 2000) and syntactic priming (e.g. Savage et al., 2003).

An alternative interpretation of these data is that, although children's knowledge of syntax is, in general, abstract, the passive is subject to some additional construction-specific difficulty. For example, Borer and Wexler's explanation (1987) derives from how passive sentences are thought to be 'generated' via a generative grammar. As the object of an active sentence is 'raised' to become the subject of the passive sentence, an A-chain (Argument chain) is formed. This causes the correct form of the auxiliary verb *to be* to be inserted. Borer and Wexler (1987; see also Hirsch & Wexler, 2006) is that young children are unable to form A-chains and that their correct interpretation of truncated passives with some verbs (e.g. AP verbs) reflects the fact that they interpret these as adjectival uses (e.g. *Bob was hugged*).

Fox and Grodzinsky (1998) claim that young children *are* able to form A-chains, and show that they can correctly interpret truncated passives of all verb types, as well as full passives with actional verbs. In contrast, the young children in their study were unable to correctly interpret non-actional full passives (e.g. *Bob was seen by Wendy*). Fox and Grodzinsky propose that this is because children are not yet able to assign the correct thematic role to the NP in the *by*-phrase.

5.1.2 The early abstraction account

Previous evidence for the early abstraction account comes from syntactic priming studies showing that, even for children as young as 3;0, hearing a passive sentence increases the likelihood of subsequently producing a passive sentence with a different verb (e.g. Savage et al., 2003; 2006; Huttenlocher et al., 2004; Bencini & Valian, 2008; Messenger et al., 2011a). While these studies provide evidence that young children's knowledge of the passive is at least partly abstract, they do not provide direct evidence against the early semantic restriction account, as none of them directly compared children's performance with verbs of different semantic types.

Recently, Messenger et al. (2012) conducted production-priming and forced-choice comprehension tasks designed to investigate this question. Specifically, Messenger et al.'s aim was to investigate the possibility that children's early knowledge of the passive is indeed abstract, and that the by-verb differences observed in previous studies (e.g. Maratsos et al., 1985) could be explained by task effects. They argue that, because depicting non-actional events in experimental materials is more difficult than depicting events involving AP verbs, poor picture recognition may account for these results. Messenger et al.'s comprehension task replicated the findings of previous studies, with young children showing more accurate comprehension of passives with both agent-patient (AP; e.g. *hug*) and theme-experiencer (TE; e.g. *annoy*) verbs (both of which involve 'affectedness' of the passive subject) than with experiencer-theme (ET; e.g. *see*) verbs (in which the passive subject may not be affected). In contrast to Maratsos et al.'s (1985) findings (see also Hirsch & Wexler, 2006), Messenger et al. (2012) also found this pattern with *active* sentences, supporting their claim that it is ET *verbs* (or depictions of ET verbs) that are problematic for children, rather than ET *passives.*

To investigate adults' and children's abstract knowledge of the passive, Messenger et al. (2012) conducted a production-priming study. Participants took turns with an experimenter to describe pictures in which an animal was doing something to a human (e.g. a sheep hitting/shocking a girl). The experimenter produced an active or passive sentence containing either an AP or a TE verb (Experiment 1) or a TE or an ET verb (Experiment 2). Participants' pictures always depicted an actional event, designed to be described using an AP verb (e.g. *scratch, wash, hug*). Messenger et al. found no difference in the rate of passives produced following different prime types, and this finding was the same for both adults and children. They therefore concluded that children have adult-like abstract knowledge of passive syntax by 3-4 years of age.

5.1.3 A third possibility: The semantic construction prototype account

A third possibility, not investigated in previous studies, is that, while even young children have an abstract representation of the passive construction, this representation is nevertheless semantically constrained. A recent study suggests that adults' knowledge of the passive, while undisputedly abstract, may take the form of a semantic prototype construction. Ambridge, Bidgood, Pine, Rowland and Freudenthal (2015) conducted graded grammaticality judgment and timed forced-choice comprehension studies (both of which have the advantage of yielding a continuous, rather than binary, dependent measure). After controlling for both overall verb frequency and verb frequency in the passive construction, Ambridge et al. found that the verbs' *affectedness* ratings – designed to capture the semantics of the passive construction – significantly predicted both adults' judgments of grammatical acceptability and their reaction times for choosing the correct animation in the forced-choice comprehension task. (Note that almost all AP and TE verbs were rated highly for affectedness. This contrasts with ET verbs, the vast majority of which had low affectedness ratings.) Importantly, semantic effects for both grammaticality judgments and reaction times were significantly larger for the passive than the active construction. The implication is that the early semantic restriction account is right in highlighting by-verb semantic differences in the passive (which persist into adulthood), while the early abstraction account is right in highlighting the fact that, nevertheless, both adults and children have an abstract verb-general representation of

the construction. This suggests the need for an approach that integrates aspects of both the early semantic restriction account and the early abstraction account. One such approach is the semantic construction prototype account (e.g. Ibbotson & Tomasello, 2009): the passive construction is associated with the semantics of affectedness of the passive subject (e.g. Pinker et al., 1987), and the greater the extent to which a particular verb has compatible semantics, the greater the acceptability and ease of comprehension of the relevant passive sentence. For example, in the sentence *Bob was scared by Wendy*, Bob is definitely affected by the event; thus *scare* is highly compatible with the semantics of the passive. In contrast, in the sentence *Bob was seen by Wendy*, Bob may or may not be affected by the event and may even be unaware that it has occurred; thus *see* is less compatible with the semantics of the passive.

5.1.4 The present study

The aim of the current study is to test the predictions of (a) the early semantic restriction account (b) the early abstraction account and (c) the semantic construction prototype account. The study uses a production-priming task similar to that used by Messenger et al. (2012), but with one crucial difference: Messenger et al. varied the semantic type of the prime verb (AP/TE/ET) while holding constant the semantic type of the target verb (AP). We vary the semantic type of the *target* verb (AP/TE/ET), while holding constant the semantic type of the *prime* verb (AP). This small manipulation results in a more sensitive test of by-verb differences. From other studies, we know that the identity of the target verb can affect the size of the priming effect (see Coyle & Kaschak, 2008; Peter, Chang, Pine, Blything & Rowland, 2015). By using as targets only AP verbs – often suggested to be the prototypical verb type for passive sentences and, therefore, presumably the easiest for children to produce – Messenger et al. may have reduced the likelihood of observing by-verb differences. That is, even a prime verb whose semantics are less than fully compatible with the semantics of the passive construction (e.g. an ET verb) may still yield a "good enough" passive to prime production of an "easy" AP passive. By reversing the design of Messenger et al., the present study investigates whether, when primed with a prototypical passive, the extent to which children are able to *produce* a passive varies as a function of the semantics of the target verb (AP/TE/ET).

If the early abstraction account is correct, our results should essentially replicate the findings of Messenger et al. That is, they should yield a main effect of Prime Type (active vs. passive), with participants producing more passive sentences following passive primes, but no effect of Target Verb Type (AP/TE/ET) and no interaction; as Messenger et al. reported in their study, if the priming effect is underpinned by abstract knowledge of syntax, it should apply across the board. The size of the priming effect should also be equivalent for adults and children (i.e. no interaction of Prime Type by Age), as knowledge of the passive is considered to be fully adult-like even for young children. If the early semantic restriction account is correct, our results should show an interaction of Age by Target Verb Type such that children produce more passives with AP and TE than ET target verbs, while adults should produce an equal number of passives with all three verb types. This account also predicts a three-way interaction of Age by Prime Type by Target Verb type, such that passive priming occurs for all three target verb types for adults, but only AP and TE target verbs for children.

If the semantic construction prototype account is correct, our results should show main effects of both (a) Prime Type (i.e. more passives after passive primes) and (b) Target Verb Type (i.e. more passives for AP and TE than ET target verbs), but no interaction. This is because the account assumes that knowledge of the passive is (a) abstract from an early age (hence the main effect of Prime Type and absence of an interaction with Target Verb Type) but (b) takes the form of a semantic prototype construction, even for adults (hence the main effect of Target Verb Type). This account neither predicts nor rules out interactions of Prime Type by Age and Verb Type by Age, as it is agnostic as to whether or not children's representations are fully adult-like by age 4-6. The crucial prediction is that, any such interactions notwithstanding, the main effects of both (a) Prime Type and (b) Verb Type should hold for both children and adults when analysed separately. This is because, *for both children and adults*, knowledge of the passive (a) is abstract and (b) takes the form of a semantic prototype construction.

As an addition test of the semantic construction prototype account, and following Ambridge, Bidgood, Pine, Rowland and Freudenthal (2015), we use semantic ratings to create a continuous semantic variable, here termed *Affectedness*, which is hypothesised to be at the core of the passive semantic construction prototype. This variable was first used in Ambridge, Bidgood, Pine, Rowland and

Freudenthal (2015; labelled as *A affects B* in that paper). This allows us to move away from Pinker-style (1989) semantic classes and to test a more fine-grained approach to semantics. As with the analysis outlined above, if the semantic prototype account is correct, we predict that this additional analysis will show main effects of (a) Prime Type (more passives after passive primes) and (b) Semantics (more passives with verbs rated as having semantics more related to the semantic property of *affectedness*), but no interaction. Again, we make no prediction about interactions of Prime Type by Age and Semantics by Age, although each of the predicted main effects should hold for both adults and children when analysed separately.

5.2 Method

5.2.1 Participants

The participants were 60 children aged 4-6 years old (4;2-6;5, *M*=5;6) and 60 adults aged 19-24. The children were recruited from primary schools and nurseries in the North West of England and the adults were all undergraduate students at the University of Liverpool. All participants were monolingual speakers of English and had no known language impairments.

5.2.2 Test items

Prime verbs consisted of 24 basic agent-patient (AP) verbs (e.g. *hug*). Target verbs consisted of 12 AP verbs (e.g. *hug*), 12 theme-experience (TE) verbs (e.g. *annoy*) and 12 experiencer-theme (ET) verbs (e.g. *see*); all verbs used by Messenger et al. (2012) were included in our set. All prime and target verbs are given in Table 5.1. Eight different playlists were created, each of 36 trials. The eight original playlist orders were reversed to create a further 8 playlists in order to avoid order effects, as a pilot study suggested that rate of production of passive sentences increased over time. Prime sentences used 18 different AP verbs, each of which appeared in both an active and a passive sentence, on separate trials. The remaining 6 AP verbs were used as targets, along with 6 each of the TE and ET verbs. Participants in the action were one male and one female character from popular TV animation shows, chosen to be familiar to young children. Playlists were pseudo-randomised such that no more than 2 sentences of the same type (active/passive) or two verbs of the same type

(AP/TE/ET) appeared in a row. Active and passive sentences containing the same verb were never used in consecutive trials. The prime sentence always contained different participants to the target sentence, in order to minimise lexical overlap.

Prime verbs		**Target verbs**		
all AP		**AP**	**ET**	**TE**
avoid	hold	bite*	forget	amaze
bite*	hug	carry*	hate	annoy*
call	kick	chase	hear*	bother
carry*	kiss	dress	ignore*	frighten*
chase	lead	hit*	know	impress
cut	pat*	hug	like*	please
dress	pull*	kick	love*	scare*
drop	push	pat*	remember*	shock*
eat	shake	pull*	see*	surprise*
follow	squash*	push	smell	tease
help	teach	squash*	understand	upset*
hit*	wash	wash	watch	worry

Table 5.1. Verbs used as primes and targets. AP = agent-patient verb, ET = experiencer theme verb, TE = theme-experiencer verb. * denotes verbs also used in Messenger et al. (2012).

For each prime and target verb, animations were created, using *Anime Studio Pro 5.5*, to depict the action. The same animation was used for both the active and passive versions of prime and target sentences.

5.2.3 Procedure

Syntactic priming was used to encourage experimental participants to produce passive sentences. Experimenter 1 took turns with the participant to describe animations presented on a computer screen, using *Processing 2* (www.processing.org). A second experimenter, who was unable to see the screen, gave 'clue words' (the prime/target verbs) to Experimenter 1 and the participant. Experimenter 2 noted down participants' responses, although sessions were also audio-recorded, using *Audacity*, as a backup. Following Rowland et al. (2012), we used a 'bingo game' to motivate the participants to produce responses. For each sentence produced by Experimenter 1 or the participant, Experimenter 2 looked for a bingo card that matched the sentence. In fact, Experimenter 2 had all of the bingo

cards, but the game was fixed so that the participant always won the game. As the playlists were long, they were divided into 4 ‘games’ of 9 trials, with participants requiring 6 bingo cards to win. Before starting, the game was introduced to participants with three practice trials, using verbs that were not included in the experiment proper in active locative sentences (e.g. *Homer poured water into the cup*).

5.3 Results

Participants’ responses were coded for sentence type, irrespective of prime condition: correct active, correct (full) passive, incorrect active (with participants reversed), incorrect (full) passive (with participants reversed), other use of the verb, and excluded (target verb not used/no response). A response was coded as a correct active if it was an accurate description of the event, and contained both a subject and direct object bearing the appropriate role (agent/patient/theme/experiencer) and the target verb. A response was coded as a correct passive if it was an accurate description of the event, and contained both a subject and object bearing the appropriate role (agent/patient/theme/experiencer), an auxiliary verb (*get* or *be*), the target verb and the preposition *by*. These criteria are similar to those used by Messenger et al. (2012), with the exceptions that (i) participants in the current study were required to use the target verb, and (ii) the range of semantic roles was more varied, as we used AP, ET and TE verbs as targets, whereas Messenger et al. used only AP verbs as targets. Table 5.2 shows the frequency of each response type for both adults and children.

	Correct active	Correct passive	Incorrect (reversed) active	Incorrect (reversed) passive	Other use of verb	Excluded
Adults	1544	424	41	14	80	57
Children	1355	133	53	6	136	477

Table 5.2. Number of responses of each type by age group.

Again following Messenger et al., the dependent variable in our analysis was binary (correct active or correct passive response, with all other responses excluded). Results were analysed in *RStudio* (version 0.98.1103; R version 3.2.0, R Core Team, 2015). As the dependent variable was binary, results were analysed using the *glmer* function of the *lme4* package (version 1.1-7, Bates, Maechler, Bolker & Walker, 2014). Predictor variables were Age Group (adult/child), Prime Type (active/passive) and Target Verb Type (AP/TE/ET). Factors were centred prior to analysis and Helmert contrast coding was used for the 3-way factor of Target Verb Type. Random intercepts for Verb and Participant were included in the model, although no random slopes were included as their inclusion prevented the model from converging. Interactions for Age by Prime Type and Age by Verb Type were included in the model, although no other interactions were included as this also prevented the model from converging. Because we had to exclude the three-way interaction of Age by Prime Type by Target Verb Type, we ran separate models for adults and children in order to test the prediction that passive priming occurs for all three target verb types for adults, but only AP and TE target verbs for children. A plot of mean passive responses, by Age Group, Prime Type and Target Verb Type, is shown in Figure 5.1.

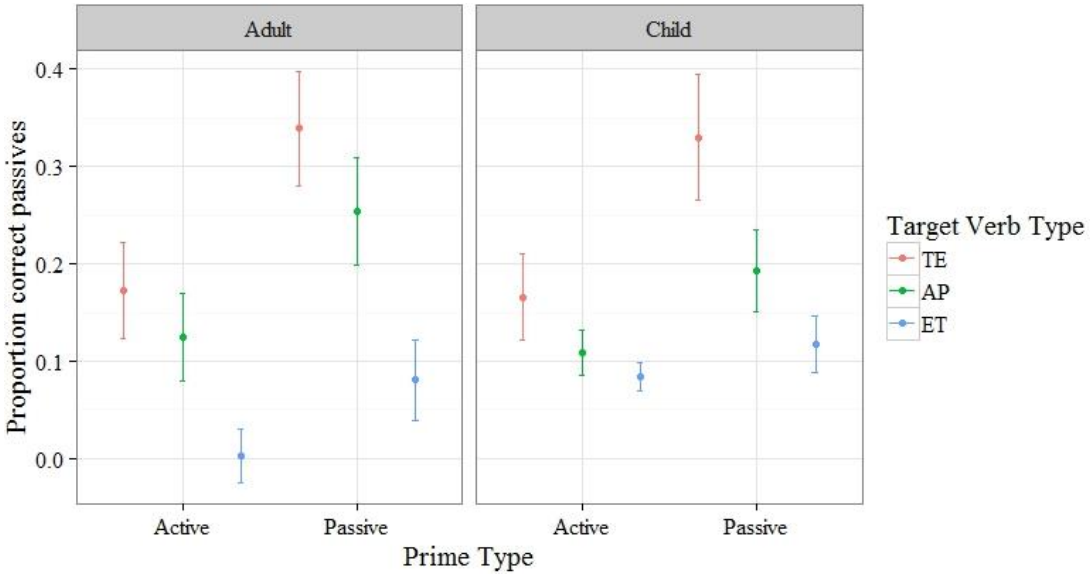

Figure 5.1. Proportion of correct passives produced, by age group (adult/child), prime type (active/passive) and target verb type (*theme-experiencer*, e.g. *scare* / *agent-patient*, e.g. *hug* / *experiencer-theme*, e.g. *see*).

Table 5.3 shows the model summary, with model comparison results shown in Table 5.4. Model comparisons revealed a main effect of Age Group ($\chi^2_{[df=1]}$=23.20, p<0.001), such that adults (M=0.22, SE=0.01) produced more passive sentences than children (M=0.09, SE=0.01). This effect was not specifically predicted by any of the three accounts under investigation, but is unsurprising given that passives are relatively difficult in general (i.e. they are longer than actives, and reverse active word order). Consistent with both the early abstraction and semantic construction prototype accounts, model comparisons also revealed a main effect of Prime Type ($\chi^2_{[df=1]}$=101.05, p<0.001), such that participants produced more passive sentences following passive (M=0.22, SE=0.01) than active primes (M=0.11, SE=0.01). The interaction between Age Group and Prime Type was non-significant, as predicted by the early abstraction account, but the p value was .061, so may have become significant with more power ($\chi^2_{[df=1]}$=3.51, p=0.061, n.s.). Therefore, on a conservative reading, the prediction of the early abstraction account here only receives tentative support.

Fixed effects Estimate (SE)	Intercept	-2.92 (0.19)
	Age Group	1.59 (0.33)
	Prime Type	1.23 (0.13)
	Target Verb Type (affected subject; AP+TE vs. ET)	1.67 (0.23)
	Target Verb Type (action; AP vs. TE)	-0.78 (0.22)
	Age Group x Prime Type	-0.53 (0.27)
	Age Group x Target Verb Type (affected subject; AP+TE vs. ET)	-0.17 (0.35)
	Age Group x Target Verb Type (action; AP vs. TE)	0.69 (0.27)
Random effects Variance (SD)	Participant (intercept)	1.72 (1.31)
	Target Verb (intercept)	0.19 (0.43)
Model summary	AIC	2433.3
	BIC	2494.8
	Log Likelihood	-1206.6
	Deviance	2413.3

Table 5.3. Model summary for all participants.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	χ^2	**χ df**	***p* value**
Intercept	3	2594.2	2612.6	-1294.1	2588.2			
Age Group	4	2573.0	2597.6	-1282.5	2565.0	23.20	1	1.46^{e-6} ***
Prime Type	5	2473.9	2504.7	-1232.0	2463.9	101.05	1	$<2.2^{e-16}$ ***
Target Verb Type (all)	7	2437.4	2480.4	-1211.7	2423.4	40.51	2	1.60^{e-9} ***
Target Verb Type (affected subject; AP+TE vs. ET)	6	2442.8	2479.6	-1215.4	2430.8	33.182	1	8.39^{e-9} ***
Target Verb Type (action; AP vs. TE)	7	2437.4	2480.4	-1211.7	2423.4	7.33	1	0.0068 **
Age Group x Prime Type	8	2435.9	2485.1	-1210.0	2419.9	3.51	1	0.061, n.s.
Age Group x Target Verb Type (all)	10	2433.3	2494.8	-1206.7	2413.3	6.62	2	0.036 *
Age Group x Target Verb Type (affected subject; AP+TE vs. ET)	9	2437.7	2493.0	01209.8	2149.7	0.25	1	0.62, n.s.
Age Group x Target Verb Type (action; AP vs. TE)	10	2433.3	2494.8	-1206.7	2413.3	6.37	1	0.012 *

Table 5.4. Model comparisons for all participants.

Importantly, and in support of the semantic construction prototype account, model comparisons revealed a main effect of Target Verb Type ($\chi^2_{[df=2]}$=40.51, p<0.001). This is also compatible with the early semantic restriction account, provided that the main effect is driven by differences in the children's responses, and not those of the adults (see separate analyses, below). Contrasts revealed that participants produced significantly more passive sentences with AP and TE verbs

(verbs involving passive subject affectedness) than with ET verbs (M=0.07, SE=0.01) ($\chi^2_{[df=1]}$=33.18, p<0.001) and significantly more passive sentences with TE (M=0.26, SE=0.01) than AP verbs (M=0.17, SE=0.01) ($\chi^2_{[df=1]}$=7.33, p=0.007).

The early semantic restriction account (see section 5.2.1) predicts an interaction of Age Group by Target Verb Type such that children will produce more passives with AP and TE than ET target verbs, while adults will produce an equal number of passives with all three types. Although a significant interaction was observed ($\chi^2_{[df=2]}$=6.62, p=0.037), inspection of Figure 5.1 reveals that it does not conform to the pattern predicted by this account. Rather, the pattern of by-verb differences was similar for the two age groups. In order to unpack this interaction, we conducted further analyses for each age group separately. These separate analyses also allow for the investigation of the prediction of the early semantic restriction account that passive priming will occur for all three target verb types for adults, but only AP and TE target verbs for children.

5.3.1 Adults

Results were analysed using linear mixed models, as above. A by-participant random slope for Prime Type was included in the model, although the model would not converge with additional by-participant random slopes (by-verb random slopes were not meaningful, given the design). A Prime Type by Target Verb Type interaction was also included. Table 5.5 shows the model summary, with model comparison results shown in Table 5.6. Model comparisons revealed a main effect of Prime Type ($\chi^2_{[df=1]}$=21.26, p<0.001), such that adults produced more passive sentences following passive (M=0.28, SE=0.01) than active primes (M=0.15, SE=0.01). Model comparisons also revealed a main effect of Target Verb Type ($\chi^2_{[df=2]}$=38.51, p<0.001). Contrasts revealed that adults produced significantly more passive sentences with AP and TE verbs (verbs involving passive-subject affectedness) than with ET verbs (M=0.09, SE=0.01) ($\chi^2_{[df=1]}$=34.83, p<0.001) as well as, unexpectedly, significantly more passive sentences with TE (M=0.31, SE=0.02) than AP verbs (M=0.24, SE=0,02) ($\chi^2_{[df=1]}$=3.92, p=0.048)[1]. No Prime Type by Target Verb Type

[1] This result may seem surprising, given that previous research has generally assumed prototypical passives to be actional AP verbs. However, it is worth noting that events denoted by TE verbs are, by definition, highly affecting for the passive subject (see Ambridge et al., 2015). In addition, Maratsos

interaction was observed ($\chi^2_{[df=2]}$=0.42, p=0.81, n.s.), suggesting that Prime Type affected adults' responses in the same way for all target verb types. Thus, in summary, the results for adults pattern as predicted by the semantic prototype account, but not the early semantic restriction or early abstraction accounts (neither of which predicts Target Verb effects in adults), with main effects of both (a) Prime Type (i.e. more passives after passive primes) and (b) Target Verb Type (i.e. more passives for AP and TE than ET target verbs), but no interaction.

Fixed effects Estimate (SE)	Intercept	-2.18 (0.20)
	Prime Type	1.10 (0.24)
	Target Verb Type (affected subject; AP+TE vs. ET)	1.64 (0.22)
	Target Verb Type (action; AP vs. TE)	-0.43 (0.21)
	Prime Type x Target Verb Type (affected subject; AP+TE vs. ET)	-0.17 (0.33)
	Prime Type x Target Verb Type (action; AP vs. TE)	-0.13 (0.28)
Random effects Variance (SD)	Participant (intercept)	1.46 (1.21)
	Participant – Prime Type	1.09 (1.04)
	Target Verb (intercept)	0.15 (0.39)
Model summary	AIC	1729.6
	BIC	1785.5
	Log Likelihood	-854.8
	Deviance	1709.6

Table 5.5. Model summary for adults.

et al. (1985) suggested that input frequency of the verb types in question may have an influence on passive production, with change-of-state verbs, such as TE verbs, being of higher frequency in the passive in child-directed speech than verbs describing temporary physical contact, such as *kick*, and many of the other AP verbs used in the current study.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	5	1780.1	1808.0	-885.03	1770.1			
Prime Type	6	1760.1	1794.3	-874.40	1748.8	21.26	1	1.02^{e-6} ***
Target Verb Type (all)	8	1726.1	1770.7	-855.03	1710.1	38.75	2	3.85^{e-9} ***
Target Verb Type (affected subject; AP+TE vs. ET)	7	1728.0	1767.1	-856.99	1714.0	34.83	1	3.60^{e-9} ***
Target Verb Type (action; AP vs. TE)	8	1726.1	1770.7	-855.03	1710.1	3.92	1	0.048 *
Prime Type x Target Verb Type (all)	10	1729.6	1785.5	-854.82	1709.6	0.42	2	0.81, n.s.
Prime Type x Target Verb Type (affected subject; AP+TE vs. ET)	9	1727.8	1778.1	-854.91	1709.8	0.24	1	0.63, n.s.
Prime Type x Target Verb Type (action; AP vs. TE)	10	1729.6	1785.5	-854.82	1709.6	0.19	1	0.67, n.s.

Table 5.6. Model comparisons for adults.

5.3.2 Children

Results were analysed using linear mixed models as for adults, above, except that no random slopes were included, as this prevented the model from converging. Table 5.7 shows the model summary, with model comparison results shown in Table 5.8. Model comparisons revealed a main effect of Prime Type ($\chi^2_{[df=1]}$=48.54, p<0.001), such that children produced more passive sentences following passive (M=0.14, SE=0.01) than active primes (M=0.04, SE=0.01). Model comparisons also revealed a main effect of Target Verb Type ($\chi^2_{[df=2]}$=22.19, p<0.001). Contrasts revealed that children produced significantly more passive sentences with AP and TE verbs (verbs involving passive-subject affectedness) than with ET verbs (M=0.03, SE=0.01)

($\chi^2_{[df=1]}$=14.19, p<0.001) and, again unexpectedly, significantly more passive sentences with TE (M=0.18, SE=0.02) than AP verbs (M=0.08, SE=0.02) ($\chi^2_{[df=1]}$=8.00, p=0.005). No Prime Type by Target Verb Type interaction was observed ($\chi^2_{[df=2]}$=0.15, p=0.93, n.s.), suggesting that Prime Type affected children's responses in the same way for all target verb types. Thus the results for children also pattern as predicted by the semantic prototype account, but not the early semantic restriction or early abstraction accounts, with main effects of both (a) Prime Type (i.e. more passives following passive primes) and (b) Target Verb Type (i.e. more passives for AP and TE than ET target verbs), but no interaction.

Fixed effects Estimate (SE)	Intercept	-3.95 (0.37)
	Prime Type	1.55 (0.37)
	Target Verb Type (affected subject; AP+TE vs. ET)	1.83 (0.46)
	Target Verb Type (action; AP vs. TE)	-1.16 (0.41)
	Prime Type x Target Verb Type (affected subject; AP+TE vs. ET)	0.04 (0.73)
	Prime Type x Target Verb Type (action; AP vs. TE)	-0.19 (0.52)
Random effects Variance (SD)	Participant (intercept)	2.43 (1.56)
	Target Verb (intercept)	0.55 (0.74)
Model summary	AIC	695.2
	BIC	737.6
	Log Likelihood	-339.6
	Deviance	679.2

Table 5.7. Model summary for children.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	3	756.05	771.96	-375.02	750.05			
Prime Type	4	709.51	730.73	-350.75	701.51	48.54	1	3.24^{e-12} ***
Target Verb Type (all)	6	691.32	723.15	-339.66	679.32	22.19	2	1.52^{e-5} ***
Target Verb Type (affected subject; AP+TE vs. ET)	5	697.32	723.85	-343.66	687.32	14.19	1	1.7^{e-5} ***
Target Verb Type (action; AP vs. TE)	6	691.32	723.15	-339.66	679.32	8.00	1	0.0047 **
Prime Type x Target Verb Type (all)	8	695.17	737.61	-339.59	679.17	0.15	2	0.93, n.s.
Prime Type x Target Verb Type (affected subject; AP+TE vs. ET)	7	693.31	730.45	-339.66	679.31	0.0057	1	0.94, n.s.
Prime Type x Target Verb Type (action; AP vs. TE)	8	695.17	737.61	-339.59	679.17	0.14	1	0.71, n.s.

Table 5.8. Model comparisons for children.

5.3.3 Fine-grained semantic analysis

In order to test a fine-grained approach to semantics, we re-ran the analyses above but replaced the discrete semantic classes with a continuous semantic variable, previously used in Ambridge, Bidgood, Pine, Rowland and Freudenthal (2015). The

variable was derived by collecting semantic ratings for each verb from ten adults (who did not take part in the priming study) for ten semantic features thought to be associated with the passive construction (e.g. the agent is doing something to the patient, the agent is responsible). Principal components analysis was then used to derive a single, composite variable. This variable is named *Affectedness*, as the biggest weightings are related to the extent to which the subject in an active sentence affects the object (irrespective of their semantic roles). As with the other predictor variables, this variable was centred before being entered into mixed effects models using the *glmer* function in R. The outcome variable was the proportion of correct passive responses out of the total of correct active and correct passives.

Predictor variables were Age Group (adult/child), Prime Type (active/passive) and Semantics (*Affectedness*). (Predictor variables were centred prior to analysis.) Random intercepts for Verb and Participant were included in the model, although no random slopes were included, as this prevented the model from converging. Interactions for Age by Prime Type and Age by Semantics were included in the model, although the three-way interaction was removed as this also prevented the model from converging. Table 5.9 shows the model summary, with model comparison results shown in Table 5.10.

Fixed effects Estimate (SE)	Intercept	-2.60 (0.20)
	Age Group	1.44 (0.30)
	Prime Type	1.25 (0.14)
	Semantics (*Affectedness*)	0.47 (0.11)
	Age Group x Prime Type	-0.48 (0.26)
	Age Group x Semantics	0.19 (0.12)
	Prime Type x Semantics	-0.10 (0.11)
Random effects Variance (SD)	Participant (intercept)	1.70 (1.30)
	Target Verb (intercept)	0.50 (0.70)
Model summary	AIC	2457.7
	BIC	2513.0
	Log Likelihood	-1219.8
	Deviance	2439.7

Table 5.9. Model summary for all participants.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	**χ^2**	**χ df**	***p* value**
Intercept	3	2594.2	2612.6	-1294.1	2588.2			
Age Group	4	2573.0	2597.6	-1282.5	2565.0	23.2	1	1.5^{e-6} ***
Prime Type	5	2473.9	2504.7	-1232.0	2463.9	101.05	1	$>2.2^{e-16}$ ***
Semantics (*Affectedness*)	6	2459.1	2496.0	-1223.5	2447.1	16.838	1	4.1^{e-5} ***
Age Group x Prime Type	7	2457.5	2500.6	-1221.8	2443.5	3.57	1	0.059, n.s.
Age Group x Semantics	8	2456.5	2505.7	-1220.2	2440.5	3.026	1	0.082, n.s.
Prime Type x Semantics	9	2457.7	2513.0	-1219.8	2439.7	0.826	1	0.36, n.s.

Table 5.10. Model comparisons for all participants.

In support of the semantic construction prototype account, and in line with the results of our initial analysis, model comparisons revealed a main effect of Semantics ($\chi^2_{[df=1]}$=16.84, p<0.001). Figure 5.2 shows that, the higher a verb is rated on the *Affectedness* scale, the more likely participants are to produce a correct passive sentence with that verb. The pattern is the same for both Age Groups, and for both Prime Types. As with the model presented in Table 5.3, main effects of Age Group ($\chi^2_{[df=1]}$=23.20, p<0.001) and Prime Type ($\chi^2_{[df=1]}$=101.05, p<0.001) were also observed. No significant interactions were found, although two marginal interactions were observed (Age Group by Prime Type: $\chi^2_{[df=1]}$=3.57, p=0.059; Age Group by semantics: $\chi^2_{[df=1]}$=3.03, p=0.082). Figure 5.2 indications that these marginal interactions may be driven by children's responses when primed with active sentences, which appear to be close to a floor effect.

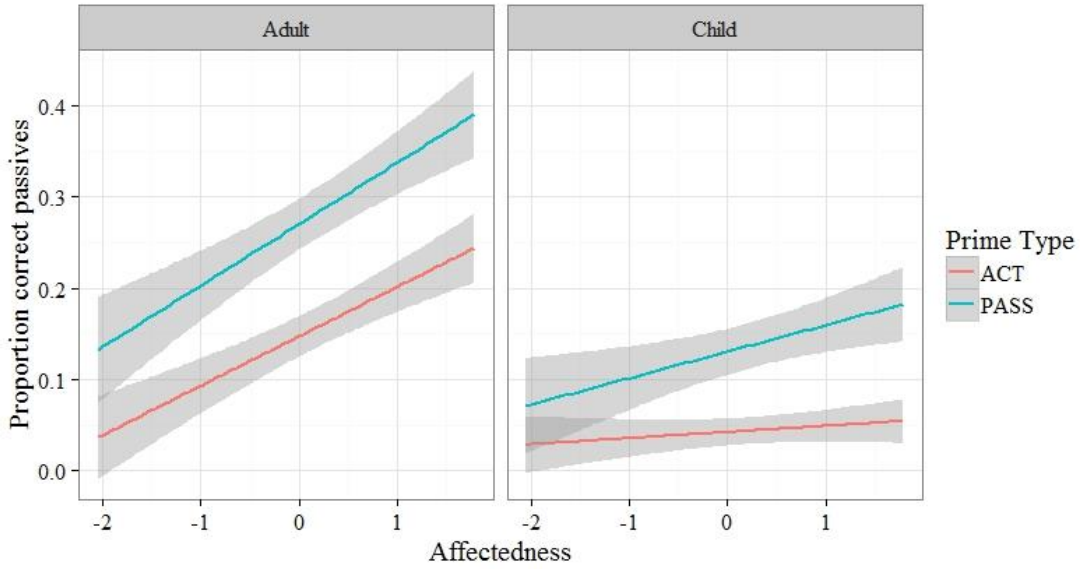

Figure 5.2. Proportion of correct passives plotted against semantic rating (*Affectedness*), split by Prime Type and Age Group.

As no Age Group by Semantics interaction was observed, the findings are incompatible with the predictions of the early semantic restriction account, as this account predicts that by-verb differences should only be observed for children. Nevertheless, as in the initial analysis, we ran analyses with each age group separately. Table 5.11 shows the model for summary for adults, with model comparisons for this age group shown in Table 5.12. Table 5.13 shows the model for summary for children, with model comparisons for this age group shown in Table 5.14. In both of these models, a random slope for Prime Type was included, although all other random slopes were removed to enable the model to converge. Main effects for Prime Type and Semantics (*Affectedness*) were observed for both Age Groups, but there was no interaction for either age group. These results again support the early abstraction account, with young children demonstrating abstract knowledge of the passive, and the semantic prototype account, with adults and children both increasingly likely to produce a passive sentence with verbs more compatible with the semantic feature of *Affectedness*. These results to not support the early semantic restriction account, as adults and children appear to be using semantic information in the same way.

Fixed effects Estimate (SE)	Intercept	-1.90 (0.21)
	Prime Type	1.11 (0.22)
	Semantics (*Affectedness*)	0.57 (0.11)
	Prime Type x Semantics	-0.17 (0.12)
Random effects Variance (SD)	Participant (intercept)	1.47 (1.21)
	Participant (Prime Type)	1.11 (0.06)
	Target Verb (intercept)	0.36 (0.60)
Model summary	AIC	1740.7
	BIC	1785.4
	Log Likelihood	-862.3
	Deviance	1724.7

Table 5.11. Model summary for adults.

Predictor	**df**	**AIC**	**BIC**	**logLik**	**deviance**	**χ^2**	**χ df**	***p* value**
Intercept	5	1780.1	1808.0	-885.03	1770.1			
Prime Type	6	1760.8	1794.3	-874.40	1748.8	21.255	1	4.0^{e-6} ***
Semantics (*Affectedness*)	7	1740.4	1779.5	-863.20	1726.4	22.416	1	2.2^{e-6} ***
Prime Type x Semantics	8	1740.7	1785.4	-862.35	1724.7	1.6922	1	0.19, n.s.

Table 5.12. Model comparisons for adults.

Fixed effects Estimate (SE)	Intercept	-3.87 (0.45)
	Prime Type	2.18 (0.62)
	Semantics (*Affectedness*)	0.38 (0.21)
	Prime Type x Semantics	0.09 (0.24)
Random effects Variance (SD)	Participant (intercept)	3.03 (1.74)
	Participant (Prime Type)	0.38 (0.61)
	Target Verb (intercept)	1.32 (1.15)
Model summary	AIC	711.73
	BIC	754.17
	Log Likelihood	-347.86
	Deviance	695.73

Table 5.13. Model summary for children.

Predictor	df	AIC	BIC	logLik	deviance	χ^2	χ df	*p* value
Intercept	5	732.25	758.78	-361.13	722.25			
Prime Type	6	711.78	743.62	-349.89	699.78	22.466	1	2.1^{e-6} ***
Semantics (*Affectedness*)	7	709.87	747.00	-347.93	695.73	3.9173	1	0.048 *
Prime Type x Semantics	8	711.73	754.17	-347.86	695.73	0.1411	1	0.71, n.s.

Table 5.14. Model comparisons for children.

5.3.4 Summary

Despite a significant Age Group by Target Verb Type interaction in the initial analysis, the pattern of results for adults and children is remarkably similar, as illustrated in Figure 5.1. This suggests that the pattern of results *per se* does not differ by age group but, rather, that the interaction reflects a difference in magnitude between the proportion of passives produced with the three different target verb types in the two age groups, particularly between TE and AP verbs. In addition, no Age by Semantics interaction was observed in the fine-grained semantic analysis. Thus, verb semantics appears to affect the production of passives for children and adults in the same way.

5.4 Discussion

Using production priming, the current study has provided additional support for Messenger et al.'s (2012) claim that both adults and young children have abstract syntactic knowledge of the passive, whilst our reversal of Messenger et al.'s design has highlighted the fact that, nevertheless, by-verb differences do exist. Our findings therefore support Ambridge, Bidgood, Pine, Rowland and Freudenthal's (2015) claim that adults' abstract knowledge of the passive takes the form of a semantic construction prototype, and add weight to this account by demonstrating that this holds for production and is also already the case for young children. The findings are not consistent with claims that children's acquisition of passive syntax is delayed (Borer & Wexler, 1987; Fox & Grodzinsky, 1998), nor that children's knowledge is

restricted to a core of actional verbs (e.g. Maratsos et al., 1985), at least not at the age of 4 years.

Thus, as we suggested in the introduction, the early semantic restriction account is right in highlighting by-verb semantic differences in the passive (although, in fact, these persist into adulthood) while the early abstraction account is right in highlighting the fact that, nevertheless, both adults and children have an abstract verb-general representation of the construction (though one that does not preclude by-verb semantic differences). Rather, we need an account that captures the insights of both of these accounts, explaining both early abstract knowledge and late by-verb semantic differences. The semantic construction prototype account is one such account, but this raises the question of exactly what it means to have an abstract construction that, nevertheless, constitutes a semantic prototype.

In fact, this scenario is one that is familiar in cognitive psychology (e.g. Rosch, 1975). For example, humans have an abstract concept of "bird" that they have formed (presumably) by generalising across concrete instances of this category that they have witnessed. Although this category is abstract (the 'bird' prototype may not be a real bird but an amalgamation of typical features in the birds people have encountered, c.f. exemplar theory, e.g. Nosofsky, 1986), it nevertheless has a prototype structure. People have no difficulty in recognising novel species of bird. However, prototypical instances (e.g. *robin*) enjoy an advantage over less prototypical instances (e.g. *ostrich*) in (a) judgment, (b) reaction time and (c) production priming tasks. What makes *robin* a more prototypical member of the category than *ostrich* is the fact that it shares greater overlap with other category members along the dimensions that are relevant to category membership (e.g. size, ability to fly). But, importantly, ostrich still shares *enough* similarities with other members (e.g. it has wings and a beak and lays eggs) to be included in the category.

In the same way, English speakers appear to have an abstract concept of the passive construction that they have formed by generalising across concrete instances to which they have been exposed. Although this construction is abstract (for example, people have no difficulty using novel verbs in this construction; Pinker et al., 1987), it nevertheless has a prototype structure (which we are capturing with the label "affectedness"). (See Ibbotson & Tomasello, 2009, for a discussion of how prototypes in language may lead to the formation of abstract schemas and constructions.) Prototypical instances (e.g. passives with *kick* [AP] and *frighten* [TE])

enjoy an advantage over less prototypical instances (e.g. passives with *see* [ET]) in (a) judgment (Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2015), (b) reaction time (Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2015) and (c) production priming tasks (the present study). What makes passives with AP and TE verbs more prototypical members of the category than passives with ET verbs is the fact that the former share greater overlap with other category members along the dimensions that are relevant to category membership (e.g. semantic affectedness). But, importantly, passives with ET verbs still share enough similarities with other members (i.e. they include a by-phrase and passive morphology, reverse canonical linking and discourse-focus patterns) to be included in the category.

In order to compare the semantic verb class account with the fine-grained semantic analysis we performed, we plotted the proportion of passives produced by adults and children against the semantic continuum of *Affectedness* (Figure 5.3). The colours indicate the three original verb classes (AP, ET and TE). The Figure shows that AP and TE verbs overlap, both in terms of their semantic ratings and the proportion of passive sentences the children and adults in the study produced with those verbs. Perhaps the traditional distinction between these classes is therefore not a particularly meaningful one. In contrast, the ET verbs seem to form a cluster in the low-affectedness/low-passive-production quartile of the Figure. Nevertheless, particularly for adults, the figure shows that verbs of this type vary noticeably in terms of the proportion of passive sentences produced based on the verb's *Affectedness* rating. Overall, then, meaningful by-verb differences in semantics seems to be best captured using a continuum, rather than discrete classes. This supports the idea of a prototypical structure, with prototypical passive sentences containing verbs that are highly rated in terms of *Affectedness*, such as *frighten* and *hit*.

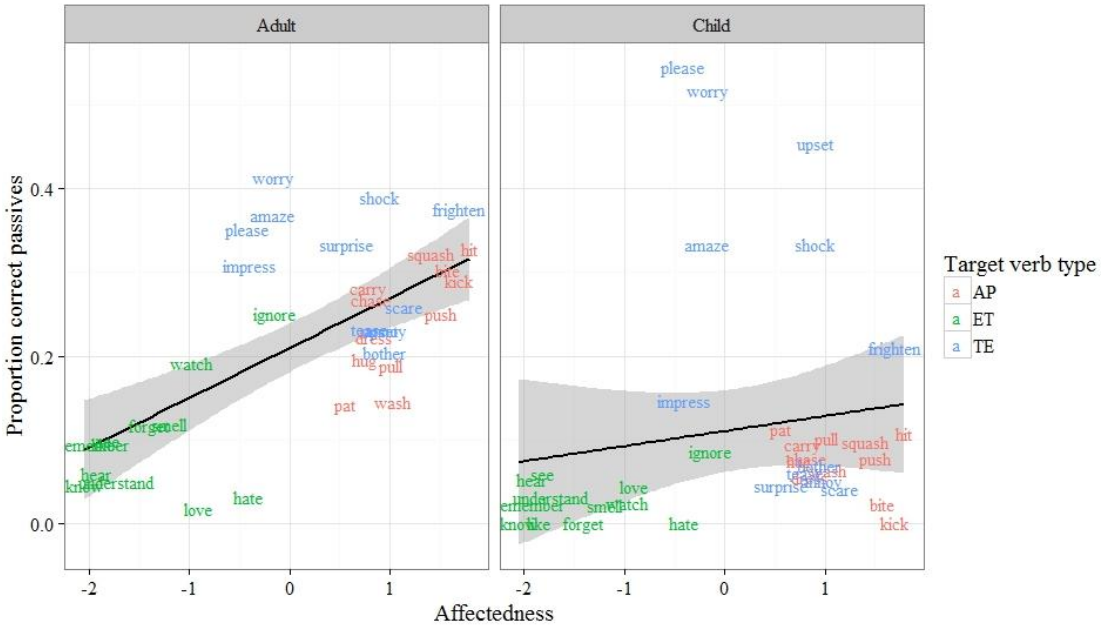

Figure 5.3. Proportion of correct passives out of correct actives and passives combined produced by adults and children, plotted against *Affectedness* rating. Colours indicate the traditional semantic classes.

Of course, the idea that linguistic categories have a semantic prototype structure is not a new one (e.g. Langacker, 1987; Goldberg, 2006; Dąbrowska, Rowland & Theakston, 2009). Although, until recently, there had been little supporting evidence for this view in the domain of verb argument structure constructions, the present study joins a growing body of research finding exactly such effects. For example, in a grammaticality judgment study, Ambridge, Pine and Rowland (2012) found evidence supporting Pinker's (1989) claim that the prototypical semantics of the ground locative construction (e.g. *Homer filled the cup with water*) involve *end-state*, whereas the prototypical semantics of the figure locative construction (e.g. *Homer poured water into the cup*) involve *manner of motion*. Adult participants' grammaticality judgments were significantly predicted by the extent to which verb semantics were consistent with the end-state manner-of-motion semantics for ground- and figure-locative sentences (as rated by independent participants), respectively. Children (aged 5 and 9 years) also showed this effect but for figure-locative sentences only, suggesting development of locative construction semantic knowledge continues beyond 9 years of age (see also Bidgood et al., 2014).

In a similar grammaticality judgment study, Ambridge, Pine, Rowland, Freudenthal and Chang (2014) demonstrated the psychological reality of semantic

constraints on the dative constructions, again based on Pinker (1989). The prototypical semantics of the double-object (DO) dative construction (e.g. *Bart threw Lisa the ball*) relate to *causing to have*, whereas the prototypical semantics of the prepositional-object (PO) dative construction (e.g. *Bart threw the ball to Lisa*) relate to *causing to go (in a particular manner*). Both adults and children rated verbs with semantics consistent with the prototypical semantics of the construction in which they appeared as more acceptable (see Ambridge, 2013; Blything et al., 2014, for similar findings for the un-VERB reversative construction).

More generally, the findings of the present study suggest that the familiar dichotomy between early-abstraction accounts and those that posit a stage characterised by lexical and semantic restrictions is too simplistic. Instead, these findings suggest the need for an account of acquisition that combines the insights of both approaches. Yes, children rapidly abstract across concrete exemplars to acquire abstract representations relatively early in development, but these abstract representations nevertheless retain the character of the original exemplars that gave rise to them. The semantic prototype theory constitutes one way (but not necessarily the only way) of explaining these findings.

Setting aside these broader considerations, in conclusion, the current paper suggests that, by the age of 4 years, children have an abstract underlying syntactic representation of the passive. This knowledge, however, has semantic structure that persists into adulthood.

5.5 Acknowledgements

This work was supported by the ESRC International Centre for Language and Communicative Development (LuCiD). The support of the Economic and Social Research Council [ES/L008955/1] is gratefully acknowledged. This work was also supported by Grant RPG-158 from the Leverhulme Trust. We would also like to thank the schools, teachers, parents and children who made this research possible, and the undergraduate students who assisted with data collection.

Chapter 6: Discussion

6.1 Introduction

This thesis set out to investigate children's argument structure overgeneralisation errors (e.g. **Don't giggle me!*) and the mechanisms that have been proposed to explain how children 'retreat' from such errors, or, indeed, avoid making them altogether. Due to the prevalence of lexical effects reported in the literature, the thesis was based on a constructivist framework. Through a series of experiments, three approaches to explaining children's retreat from argument structure overgeneralisation errors were examined: semantics (e.g. Pinker, 1989), entrenchment (Braine & Brookes, 1995) and preemption (e.g. Goldberg, 1995). Chapter 3 presented a critical test of the semantics and frequency accounts: a novel-verb grammaticality judgment study of the locative alternation. Chapter 4 examined the most frequent alternation – the causative – using both grammaticality judgment and error elicitation tasks. Chapter 5 extend the investigation of the role of semantics beyond argument structure overgeneralisation errors to the passive construction, with the aim of testing whether semantics might play a wider role in children's acquisition of argument structure.

This Discussion chapter will first summarise the findings of the three experimental chapters presented in this thesis. It will discuss how these results contribute to the growing body of evidence that, although individual mechanisms each enjoy some support, children's retreat from overgeneralisation errors is best explained by a theory that integrates elements of the entrenchment, preemption and semantic accounts. The chapter concludes by suggesting that future research should focus on testing integrated accounts of argument structure acquisition, such as Ambridge et al.'s FIT account (Ambrige & Lieven, 2011; Ambridge et al., 2011).

6.2 What do the studies in this thesis tell us about the retreat from overgeneralisation?

Overgeneralisation errors, such as **Don't giggle me!*, are thought to stem from children noticing that some verbs can appear in two argument structures with similar meaning, e.g. *Christine sprayed water onto the flowers* and *Christine sprayed the*

flowers with water. As various different verbs appear in the same distribution in the input (e.g. *splash, load*), children are able to usefully generalise this pattern to new verbs (e.g. *sprinkle, pack*). If children apply this pattern too widely, to verbs that can only appear in one or other of the constructions (e.g. *pour, fill*), overgeneralisation errors will result, such as **Howard poured the cup with tea/filled tea into the cup*. This thesis began by investigating how three proposed mechanisms were able to explain the retreat from such overgeneralisation errors: semantics, based initially on the semantic verb class hypothesis (Pinker, 1989); entrenchment (Braine & Brooks, 1995); and preemption (Goldberg, 1995).

According to Pinker's semantic verb class hypothesis, verbs with similar semantics are grouped into classes (e.g. 'locative verbs') and more specific subclasses (e.g. one in which "a mass is enabled to move via the force of gravity"). Each of the subclasses may be alternating or non-alternating, but children who are yet to construct these subclasses may erroneously treat all verbs as alternating. Not until children have correctly set up these subclasses will they be able to avoid, or retreat from, overgeneralisation errors.

According to the entrenchment hypothesis, children may make overgeneralisation errors when they have not received enough information in the input to rule out the possibility that certain verbs may not alternate between competing constructions. By frequently hearing verbs in grammatical sentences, children are eventually able to infer from their absence in the input that certain verb-construction pairings must not be acceptable.

Finally, like the entrenchment hypothesis, the preemption hypothesis proposes that lack of evidence in the input can lead to overgeneralisation errors. However, this hypothesis proposes that only hearing verbs used in constructions with very similar semantics to the ungrammatical construction will lead children to infer that the latter must be unacceptable, and thus lead to the retreat from overgeneralisation.

Chapter 3 used the locative alternation (*Christine sprayed water onto the flowers/sprayed the flowers with water*) to test the predictions of the semantic verb class hypothesis and the 'frequency' hypothesis (it was not possible to differentiate between the predictions of the entrenchment and preemption hypotheses in this study). Chapter 4 used the causative alternation (*John rolled the ball/The ball rolled*) to investigate all three hypotheses separately, although semantics was treated as a

continuum (e.g. Ambridge et al., 2014) rather than relying on the discrete classes proposed by Pinker. Chapter 5 moved beyond overgeneralisation errors to investigate the predictions of the semantic verb class hypothesis in more detail with a construction that children find notoriously difficult: the passive (e.g. *Bob was helped by Wendy*). Altogether, the experimental evidence presented here provides strong support for the roles of both verb semantics and verb frequency effects, although the role played by preemption is less clear than that of entrenchment. Overall, the conclusion is that none of these mechanisms alone can explain children's retreat from overgeneralisation, or their acquisition of argument structure more widely. Instead, an integrated account, such as the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011), is needed to explain the results presenting here and in previous papers.

6.2.1 Chapter 3: Locatives

Chapter 3 presented a novel-verb grammaticality judgment study of the locative alternation (e.g. *Christine sprayed water onto the flowers/sprayed the flowers with water*). Adults and children (aged 5-6 and 9-10) were taught novel verbs. These were presented with animations illustrating their meanings, but were never presented in locative sentences. Thus, grammaticality judgments of these novel verbs had to be based on their semantics. The semantics of the verbs were designed to match with one of three of Pinker's (1989) fine-grained semantic subclasses of locative verbs – one ground-only class (containing verbs like *fill*), one figure-only (containing verbs like *pour*) and one alternating class (containing verbs like *spray*). If participants judged sentences containing the novel verbs in line with sentences containing real verbs from the same semantic subclass, then they must be able to use verb semantics to determine the construction(s) in which those verbs can be used grammatically. To test the 'frequency' hypothesis (it was not possible to distinguish between the entrenchment and preemption hypotheses in this study), we included high-frequency and low-frequency real verbs from each of the three semantic subclasses, along with the novel verbs. If participants use verb frequency information to determine if a verb is grammatical in a certain construction, they should be less willing to accept ungrammatical sentences the more frequently that verb has previously been heard.

The results of the experiment were analysed using ANOVAs. When looking at only novel verbs – they key test of the semantic verb class hypothesis in this chapter – results for all participants followed the predicted pattern for novel non-alternating verbs, providing strong support for the semantic verb class hypothesis. Unexpectedly, participants showed an overall preference for the ground-locative uses of alternating novel *spray*-type verbs. We speculated that this might be related to the holism constraint: in order for animations to be compatible with the ground-locative construction, the whole of the 'ground' object needed to be affected, e.g. all of the roses needed to be sprayed with water for the sentence *Lisa sprayed the roses with water* to be felicitous. Therefore, participants could have judged the figure locative, *Lisa sprayed water onto the flowers*, to be less acceptable (although ratings were still broadly grammatical for both sentence types for the alternating verbs).

Strong support was also provided for the frequency account. Results revealed a greater dispreference for ungrammatical uses of high-frequency non-alternating verbs (e.g. *fill, pour*) than for low-frequency (e.g. *line, drip*) or novel verbs from the same subclasses. The dispreference for ungrammatical uses of low-frequency non-alternating verbs was also greater than the dispreference for novel verbs in the same class, although this difference was much smaller.

There were also differences between the age groups in this study. Unlike the adults and older children, the 5- to 6-year-olds showed no preference for either construction type for novel ground-only verbs (e.g. *fill*). We suggested that the fine-grained semantics of this verb class might have been too difficult for these young children to have fully grasped. Only the adults showed the unexpected preference for ground locative uses of novel alternating verbs. As mentioned above, this could have been due to the holism constraint: perhaps the complex semantics of this constraint have not yet been acquired by either group of children in the current study. Adult participants also showed a greater degree of dispreference for the ungrammatical sentences for non-alternating verbs than either group of children, and the older children showed a greater degree of dispreference for the ungrammatical sentences than did the younger children. This could be interpreted as providing further support for the frequency hypothesis, since adults have, presumably, had more exposure to all of the (real) verbs in question in competing constructions than the children have, and therefore would have been more confident in their judgments of the relative

(un)grammaticality of the sentences presented here. All of these by-age differences, then, could be taken as further support for the hypotheses under investigation.

Overall, the results from Chapter 3 cannot be explained by either the semantics or frequency accounts alone. An integrated account that can explain both of these effects seems necessary.

6.2.2 Chapter 4: Causatives

This chapter presented three experiments testing the semantics, entrenchment and preemption hypotheses, using the causative alternation (e.g. *John rolled the ball/The ball rolled*). Experiment 1 was a grammaticality judgment study with adults, using a large number of verbs. Experiment 2 was another grammaticality judgment study, but it used a smaller number of verbs in order to enable us to test children as well as adults (young children would have been unlikely to know all of the verbs used in Experiment 1 and not all of the verbs would have been suitable for them, e.g. *murder*). Experiment 3 employed a novel use of the production-priming method in a (very successful) attempt to elicit overgeneralisation errors from 5- to 6-year-old children. Unlike the locatives study in Chapter 3, this chapter used only real verbs. In addition, rather than testing the semantic verb class hypothesis specifically, we tested a more general semantic account by collecting semantic ratings from a group of adults (who did not take part in the main experiments) to create a continuum of causative semantics. This approach allows for more flexibility to investigate widely-observed lexical effects, and follows a method similar to that used by Ambridge and colleagues' investigations of the locative and dative constructions (Ambridge et al., 2014, and Ambridge, Pine, Rowland & Chang, 2012, respectively). Corpus counts were used to test both the entrenchment and preemption hypotheses, with overall verb frequency used as the predictor variable for the entrenchment hypothesis and verb frequency in a semantically-related competing construction used to test the preemption hypothesis (the periphrastic causative for intransitive-only verbs, e.g. *Homer made the fish swim*, and the passive for transitive-only verbs, e.g. *The ball was kicked*).

In both grammaticality judgment studies, we calculated difference scores for non-alternating verbs by taking the rating for the ungrammatical sentence away from the rating for the grammatical sentence. This was important as it allowed us to

control for any general preferences for certain animations, or judgments about the likelihood of certain situations occurring. These scores were calculated separately for intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*) and transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*). Experiment 1 provided strong support for the semantics hypothesis, with semantic factors significantly predicting difference scores for both transitive-only and intransitive-only verbs. Overall verb frequency also significantly predicted difference scores for both verb types, with difference scores increasing as verb frequency increased. This is the pattern of results expected according to the entrenchment hypothesis and thus provides strong support for this account. In contrast, the preemption hypothesis was not supported in Experiment 1, with the only significant predictive relationship being in the unexpected direction.

Results for Experiment 2 were analysed separately for each age group due to significant by-age interactions in the initial analyses. For the youngest children, nothing predicted difference scores for intransitivisation errors with transitive-only verbs, although scores for transitivisation errors with intransitive-only verbs were predicted by one semantic factor and total verb frequency. These predictions were both in the expected direction, therefore providing some support for the semantics and entrenchment hypotheses respectively. For the older children, total verb frequency predicted difference scores for both verb types, thus providing further strong support for the entrenchment hypothesis. Again, though, no support was found for the preemption hypothesis. The role of semantics here is less clear, as semantic factors were significant predictors of difference scores, but not in the expected directions. Finally, for adults, neither entrenchment nor preemption found any support. While various semantic factors significantly predicted difference scores for adults, the direction of these predictions was often in the unexpected direction, just like with the older children, meaning the role of semantics is unclear. The role of entrenchment and semantics also seems to change over the age groups in Experiment 2, with entrenchment becoming less important as age increases, but semantics potentially increasing in its influence.

Experiment 3 used production-priming in a novel way: as a method to elicit overgeneralisation errors from 5- to 6-year-old children. While grammaticality judgment studies tell us what participants consciously know about overgeneralisation errors, these errors are primarily a phenomenon of production in young children.

Testing the predictions of our hypotheses against the errors that children actually produce was therefore an important test, and something that is almost impossible to do using data from corpora or diary studies, primarily due to their sparsity in naturally-occurring data. We were successful in eliciting large numbers of both intransitivisation errors with transitive-only verbs and transitivisation errors with intransitive-only verbs. Semantic factors predicted error rates in both directions, whereas the entrenchment and preemption predictors predicted error rates only for transitivisation errors with intransitive-only verbs. In Experiment 3, then, support was found for all three hypotheses under investigation.

Taken together, the experiments reported in Chapter 4 provide strong support for the entrenchment hypothesis, with total verb frequency predicting difference scores and error rates in the expected direction for many parts of the experiments. Very little support for the preemption hypothesis was found, however, with the only positive evidence coming from the production of transitivisation errors in Experiment 3 (although this is perhaps the best test of this hypothesis). Evidence for the semantic account was somewhat mixed: although semantic factors predicted results in almost all parts of the experiments, some of the predictions were in the opposite direction to our expectations. Again, evidence from Chapter 4 points to the need for an integrated account that takes into consideration the observed effects of entrenchment, preemption and semantics.

6.2.3 Chapter 5: Passives

Chapters 3 and 4 focussed on children's retreat from argument structure overgeneralisation errors, with these studies finding evidence for the influence of both semantic and statistical factors. However, any mechanism involved in language acquisition should also be able to explain effects outside of the limited domain of overgeneralisations. This final experimental chapter, therefore, examined the influence of semantics in children's acquisition of the passive, a construction that is notoriously difficult to master.

Taking the study of Messenger et al. (2012) as a starting point, this chapter used structural priming to encourage children to produce passive sentences to describe a series of animations. Previous studies (e.g. Maratsos et al., 1985) found that children's difficulties in comprehension of passive sentences varied by verb

type, with agent-patient (AP) verbs (e.g. *hold*) proving relatively easy to comprehend in comparison with experiencer-theme (ET) verbs (e.g. *see*). While Messenger et al. replicated these findings in a comprehension task, a priming task found that children were primed equally by passive sentence containing verbs of three different semantic classes – AP, ET and TE (theme-experiencer verbs, e.g. *frighten*) – to produce passive sentences with AP verbs. This led the authors to conclude that children, in fact, already have an abstract representation of the passive construction at the age of three years, and that the differences between the semantic classes in comprehension tasks were likely due to the relative difficulty in depicting ET events in still pictures.

In this chapter, we reversed Messenger et al.'s method, using only AP primes but target verbs from all three semantic classes. We found priming effects for all verb types, thus replicating Messenger et al.'s finding that young children already have an abstract representation of the passive (as they would not be able to be primed without having this), but we also found significant differences in rate of production of passives between the verb classes. Both adults and children produced the greatest proportion of passive sentences with TE verbs, a smaller proportion with AP verbs and the smallest proportion with ET verbs. This result was replicated using a more fine-grained approach to semantics, similar to that taken in Chapter 4 (see also Ambridge, Pine, Rowland & Chang, 2012, and Ambridge et al., 2014). For this analysis, we used a composite semantic factor derived via Principal Components Analysis, based on adult judgments (see Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2015). This enabled us to treat semantics as a continuum, based on the core semantics of the passive construction – that of *Affectedness* of the passive subject. The results of this analysis showed that, the more a verb's semantics fitted with these semantics, the more likely both children and adults were to produce passive sentences with that verb.

These findings led us to conclude that a semantic prototype account is the best explanation for the acquisition of the passive: just as a robin is a more prototypical member of the 'bird' category than an ostrich, TE verbs such as *frighten*, which are high in *Affectedness*, fit better with the prototypical meaning of the passive construction than ET verbs such as *see*, which are relatively low in *Affectedness*. These findings are in line with the recent comprehension and grammaticality judgment studies of Ambridge, Bidgood, Pine, Rowland and Freudenthal (2015). The findings also add further support to the idea that the semantic fit between the verb

and the construction, as proposed by the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011), plays an important role in the selection of an appropriate construction when conveying a message.

6.3 Summary

The studies presented in this thesis investigated three hypotheses: the semantic verb class hypothesis (Pinker, 1989), the entrenchment hypothesis (Braine & Brooks, 1995) and the preemption hypothesis (Goldberg, 1995). Between the studies, support was found for all three hypotheses, although evidence for the preemption hypothesis was less convincing than for semantics and entrenchment. This section will summarise the new evidence for these theories presented in the thesis. Finally, a discussion of how this evidence supports the need for an integrated account, such as the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011), will be discussed.

6.3.1 Evidence for the semantics hypothesis

The role of semantics was tested in all three chapters in this thesis. The initial approach, in Chapter 3, was based on Pinker's (1989) semantic verb class hypothesis. This study was also a particularly stringent test of this hypothesis as it used novel verbs: participants could not have used previous experience with these verbs in context (as required for the operation of statistical-learning mechanisms) to provide their grammaticality judgments and must have made these judgments on the basis of semantic information. The fact that children and adults were able to able to provide grammaticality judgments for locative sentences containing the novel verbs in line with real verbs from the same semantic class, for at least two of the three verb classes tested in each age group, provides strong support for the semantic verb class hypothesis. These findings are in line with previous work (e.g. Ambridge et al., 2008; 2009; 2011; Brooks & Tomasello, 1999; Gropen et al., 1991a, b).

Following more recent work by Ambridge and colleagues (e.g. Ambridge, Pine & Rowland, 2012; Ambridge, Pine, Rowland & Chang, 2012; Ambridge et al., 2014), Chapter 4 took a slightly different approach to the influence of semantics. Rather than viewing semantics in terms of discrete verb classes, we created

continuous semantic variables, related to the causative alternation, by collecting semantic ratings from adults and using Principal Components Analysis (PCA) to reduce these to a smaller number of composite semantic factors. These were then used as the semantic predictors in the statistical models. This approach allows more flexibility to investigate widely-observed lexical effects. Results of the analysis suggested that semantics did indeed have an influence on the way in which both adults and children judge the grammaticality of overgeneralisation errors in the causative alternation, and on the rate at which young children actually produce overgeneralisation errors. However, the direction of prediction of each of the composite semantic factors was not always as expected, and sometimes appeared to be contradictory. One possible reason for this is that our labelling of the semantic factors produced by the PCA was not reflective of the complex make-up of each of these factors, since each was created from 26 original semantic features (see Chapter 4 for further discussion). Nevertheless, the fact remains that semantics did have an effect both on grammaticality judgments and error production.

So far, then, semantics has been shown to have an influence on children's retreat from overgeneralisation errors. Chapter 5 set out to test if semantics also has a role to play in a more general language acquisition mechanism. Since work with adults (Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2015) has demonstrated an influence of semantics on adults' comprehension and grammaticality judgments of (primarily grammatical) passive sentences, and given that the passive is a construction that children are known to struggle with, in comprehension at least, until a relatively advanced age (7 years in Maratsos et al., 1985), Chapter 5 tested whether children's and adults' production of the passive would similarly be affected by verb semantics. As this study was based on the methodology of Messenger et al. (2012), we tested semantics using the three semantic classes/types (AP, ET and TE) used in her study, as well as using a semantic continuum (borrowed from Ambridge, Bidgood, Pine, Rowland & Freudenthal, 2015). Although results support Messenger et al.'s finding that even young children have an abstract representation of the passive construction, a clear influence of semantics was also found, for both adults and children, using both the class-based and fine-grained semantic continuum approaches. The conclusion from this is that the abstract syntactic knowledge that is already in place at 4 years of age is, nonetheless, semantically constrained. Thus,

semantics has a wider influence on language development beyond its role in children's retreat from overgeneralisation errors.

Investigating both class-based and fine-grained continuum approaches to semantics in the same study allowed for comparison of these two approaches. Results suggest that the semantic continuum approach is the best fit for the data: Figure 5.3 showed that verbs in all three of the original classes (AP, TE and ET) vary in their level of *Affectedness*, and that the proportion of passive sentences produced by participants varied in line with this. In addition, the level of *Affectedness* (the prototypical meaning associated with the passive construction) for verbs in each of the classes overlapped with one other. In summary, while elements of the class-based semantics approach proposed by Pinker (1989) have merit (the fine-grained semantic factors used in both Chapters 4 and 5 were derived from these, after all), evidence points to the need to take a more fine-grained, graded approach to semantics, in line with that proposed by the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011). (Note that the use of a Likert scale, rather than categorical grammatical/ungrammatical judgments in Experiments 1 and 2 in Chapter 4 may have forced the participants into providing judgments that favour a continuous scale. Pinker's original [1989] semantic verb class hypothesis is therefore not fully countered by the findings of this thesis.)

6.3.2 Evidence for the entrenchment hypothesis

The entrenchment hypothesis was tested in Chapters 3 and 4 (although it was not distinguished from preemption in Chapter 3, and was labelled the *frequency hypothesis*). In Chapter 3's investigation of locative overgeneralisation errors, verb frequency was treated as a categorical variable, with verbs classed as high-frequency, low-frequency or novel (essentially zero-frequency), based on corpus data. Results showed strong support for the frequency hypothesis: participants showed a significantly larger dispreference for overgeneralised sentences (compared to their grammatical counterparts) containing high-frequency verbs than either low-frequency or novel verbs, as well as a larger dispreference for overgeneralised sentences containing low-frequency verbs than those containing novel verbs.

In Chapter 4, entrenchment was the most consistent of our predictors across the grammaticality judgment and production studies. In Experiment 1, adults'

grammaticality judgments were predicted by our measure of entrenchment: total verb frequency (determined by corpus counts). In Experiment 2, total verb frequency predicted grammaticality judgments for both groups of children tested, although this was no longer the case for adults. This discrepancy from Experiment 1 could have been related to the smaller set of verbs used, which were mostly the higher-frequency verbs used in Experiment 1. This result may indicate a developmental effect of entrenchment, and may also be indicative of a ceiling effect: perhaps it is not possible for verbs to become yet more entrenched once they have been experienced a certain (presumably large) number of times. Importantly, verb frequency also predicted the rate at which children produced overgeneralisation errors with individual verbs: the higher the frequency of the verb, the less likely children were to produce an overgeneralisation error with that verb.

Taken together, then, the studies in Chapters 3 and 4 provide strong support for the role of entrenchment in children's retreat from overgeneralisation errors, in line with previous work (e.g. Ambridge et al., 2008, 2011; Brooks et al., 1999; Theakston, 2004). Chapter 4 also raises interesting possibilities for future research related to developmental effects of entrenchment, and a possible ceiling effect.

6.3.3 Evidence for the preemption hypothesis

Like the entrenchment hypothesis, the preemption hypothesis was tested in Chapters 3 and 4. However, in Chapter 3, the high correlation between total verb frequency and verb frequency in the preempting construction (here, the alternative locative construction) made it infeasible to distinguish between the predictions of preemption and entrenchment (see also Ambridge, Pine & Rowland, 2012; Boyd et al., 2012; Wonnacott, 2011, p. 2; Perfors et al., 2010, p. 612). Thus, the findings in support of the entrenchment hypothesis in Chapter 3, described above, apply equally to the preemption hypothesis.

Chapter 4 tested the preemption hypothesis directly. Following Brooks and Tomasello (1999) and Brooks and Zizak (2002), we used the passive as the preemptive construction for avoiding intransitivisation errors with transitive-only verbs (e.g. *The ball was kicked* preempts **The ball kicked*) and the periphrastic causative as the preemptive construction for avoiding transitivisation errors with intransitive-only verbs (e.g. *Homer made the fish swim* preempts **Homer swam the*

fish). The results of the experiments provided somewhat mixed evidence for the preemption hypothesis. In Experiment 1, preemption predicted difference scores for adults' grammaticality judgments of intransitivisation errors with transitive-only verbs, but in the opposite direction to our expectations: the more a transitive-only verb had appeared in the passive, the *more likely* adults were to accept it in an overgeneralised intransitive sentence. This contradicts the preemption hypothesis. In Experiment 2, preemption did not predict grammaticality judgments at all. Thus, these studies found no evidence that preemption plays a role in grammatical acceptability judgments of intransitivisation errors with transitive-only verbs or transitivisation errors with intransitive-only verbs. This goes against the findings of a previous study using the same methodology with the dative alternation (Ambridge et al., 2014) and suggests that preemption might play a role for the retreat from overgeneralisation errors in some alternations but not others (see also Ambridge, Pine & Rowland's 2012 grammaticality judgment study of the locative alternation, in which no dissociable effect was found between entrenchment and preemption).

Chapter 4's Experiment 3, in the other hand, did find a significant effect of preemption. The rate at which children produced transitivisation errors with intransitive-only verbs was significantly predicted by verb frequency in the preempting periphrastic causative construction (although intransitivisation errors with transitive-only verbs were not predicted by our preemption measure). The task effect seen here (see also Blything et al., 2014) might be due to competition between constructions: when producing a sentence to convey the desired message, all possible constructions are competing to express the message, thus forcing a choice between the alternative constructions. In a grammaticality judgment task, however, this is not necessarily the case. The results of Experiment 3, then, do lend support to the preemption hypothesis, suggesting that it may play a role in the *production* of overgeneralisation errors which is, after all, the phenomenon it is designed to explain.

Taken together, these findings suggest that the role preemption plays in children's retreat from overgeneralisation errors is not as clear-cut as those played by semantics and entrenchment. That role may differ by construction and might depend on the demands of the task at hand, with production tasks perhaps providing the best evidence that a preemption mechanism does, indeed, have a role to play.

6.3.4 Comparing frequency effects across studies

Verb frequency measures were in included in the analyses in Chapters 3 and 4. In Chapter 3, a single frequency measure was included as a categorical variable: verbs were classified as high or low frequency (or novel). Frequency effects were observed as expected, with all participants judging ungrammatical sentences to be less acceptable with high-frequency than low-frequency verbs. In Chapter 4, total verb frequency counts were used as the entrenchment predictor with verb frequency in a competing construction with similar semantics as the preemption predictor. Results in this study were not straightforward. Firstly, neither entrenchment nor preemption were significant predictors in all studies. As discussed above, this could be due to these mechanisms operating differently with the different ages (Experiment 2 appeared to suggest a ceiling effect for entrenchment) and different tasks (preemption seemed to be a better predictor in the production task). Furthermore, the unexpected direction of the prediction of our preemption predictor in Experiment 1 could have been influenced by the large number of additional predictors in the model: a preemption-only model might have given different results. In addition, it is worth noting that the frequency measures seemed to be more successful predictors for transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*) than intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*). These apparent construction-specific effects are not entirely surprising, however. Previous studies by Ambridge and colleagues have suggested that frequency information may be used differently in locative and dative construction pairs, for example (Ambridge, Pine & Rowland [2012] and Ambridge et al. [2014], respectively). Nevertheless, it potentially dangerous to draw firm conclusions about the frequency effects in these studies (particularly the apparent null effects) due to the small numbers of participants in the studies (see Tversky & Kahneman, [1971] and Dienes [2014]).

Overall, frequency effects were found, to some extent, in all studies in which they were examined. In fact, even though Chapter 5 did not include a frequency measure, apparent frequency effects were suggested by the analysis presented in Figure 5.3. As this figure show, a group of theme-experiencer verbs (e.g. *please, worry, amaze*) were produced by children far more frequently in the passive construction than other verbs of that class. It is likely that these verbs are heard frequently in truncated passive/adjectival uses (e.g. *I was amazed*), which might have

increased the likelihood with which children were likely to produce them in full passives here. Frequency effects in the production of passives would be in line with the comprehension/reaction-time and grammaticality judgment studies (with adults) of Ambridge, Bidgood, Pine, Rowland and Freudenthal (2015) and Street and Dąbrowska (2014).

6.3.5 The FIT account: An integrated approach

The evidence presented above suggests that children's retreat from overgeneralisation errors is influenced by verb-in-construction semantics, verb frequency in the input and, perhaps to a lesser extent, the frequency of those verbs in particular (preempting) constructions. None of these factors can explain all of the variance in adults' and children's responses to grammaticality judgment tasks, nor their performance in production-priming tasks. Several accounts have been proposed that integrate aspects of semantics and statistics to explain aspects of children's language acquisition, including, including Langacker (2000), MacWhinney (2004) and Tomasello (2003). The FIT account, proposed and developed by Ambridge and colleagues (e.g. Ambridge & Lieven, 2011; Ambridge, Pine & Rowland, 2011), is a more recent version of these accounts, incorporating aspects of each of them. By doing so, it aims to provide a more complete picture of children's retreat from overgeneralisation errors, and of their language development more generally. It is worth noting that the FIT account is still being developed, with several aspects still underspecified. For example, the precise way in which the factors discussed below (verb frequency, construction frequency, etc.) interact with each other is not well-defined, nor, in fact, are some of the factors themselves. This lack of specificity means that the predictions of the FIT account are not necessarily clear. Nevertheless, to the extent that is currently possible, this section will discuss the FIT account in the context of the findings of the current thesis.

The FIT account proposes that the effects of semantics, entrenchment and preemption can all be explained by constructions competing to convey the speaker's message. If the message the speaker wants to convey is that a particularly amusing joke made a girl laugh, initially, all constructions will be available for the speaker to use. Constructions that appear more frequently in the language environment, such as the transitive-causative and intransitive-inchoative constructions, are initially likely

to be more highly activated. (Recently-heard constructions will also have increased levels of activation, as indicated by structural priming effects.) The relevance of the competing constructions to convey the entire message increases the activation of some constructions relative to others. In this case, the most relevant constructions are those with slots for the verb, *laugh*, and each of the participants, *joke* and *girl*, such as the transitive-causative and the periphrastic causative. Constructions that are completely irrelevant, particularly those of low frequency (e.g. the locative constructions), may be dismissed. The frequency of the verb in each of the constructions also plays a role – as *laugh* has been frequently heard in the intransitive-inchoative construction, its activation level increase relative to the transitive-causative, despite it having less relevance for conveying the message.

Finally, and importantly, the semantic compatibility (or 'fit') between the [ACTION] slot in the competing constructions and the semantics of the verb come into play. The [ACTION] slot in the transitive-causative construction implies direct, physical causation. While causation between the joke and the girl laughing does exist, it is neither direct nor physical – the joke amused the girl and the amusement caused the physical act of laughing. Thus, the semantics of *laugh* do not fit with the semantics of the transitive-causative construction. The semantics of the [ACTION] slot in the intransitive-inchoative construction, on the other hand, relate to internally-caused events. While this ignores the causal aspect of the message (and makes the intransitive-inchoative construction less relevant to conveying the entire message), the verb is still compatible with the construction – the girl's internal amusement at the joke caused the physical response. The [ACTION] slot in the periphrastic causative implies indirect causation. This [ACTION] slot is therefore an excellent semantic fit for the verb, as well as the construction itself being highly relevant for conveying the message. However, the periphrastic causative is disadvantaged in terms of both its overall frequency in the input and, in comparison with the intransitive-inchoative construction, the verb-in-construction frequency.

Thus, with relevance to the message, overall frequency, verb-in-construction frequency and semantic compatibility all playing roles in the choice of construction, the construction that will win out is not necessarily obvious. The construction with the best fit, and possibly the most relevance for conveying the message, is probably the periphrastic causative (*The joke made the girl laugh*). However, the verb's frequent appearance in the intransitive-inchoative construction, and its not-

incompatible semantics, may lead the intransitive-inchoative to be used (*The girl laughed*). A child might produce the overgeneralised transitive-causative sentence, **The joke laughed the girl*, if the overall frequency of the construction itself (much higher than the periphrastic causative) and its relevance to conveying the message (it is more relevant than the intransitive-inchoative) outweigh the fact that the verb has not been heard in the construction and that its semantic fit is not particularly good.

The FIT account clearly yields effects of semantics (through the compatibility between the verb and the construction) and entrenchment (via verb-in-construction frequency statistics). Preemption effects can also be seen, through a combination of verb-in-construction frequency and verb-in-construction compatibility. Importantly, the differences in the efficacy of preemption to explain the retreat from overgeneralisation in different construction pairs can also be explained via the competition model. The two dative constructions (e.g. *Lisa gave the book to Bart/Lisa gave Bart the book*) have high degrees of overlap in terms of their semantics, their relative frequency in the input and the messages that they are likely to be relevant for conveying. Thus, the frequency with which a verb appears in one or other construction in the input will have a large effect on how likely it is to be chosen over the other. A clear effect of preemption will therefore be seen, as in Ambridge et al. (2014). However, when the competing constructions are very rare in the input, as with the locative construction, the preemption effect may be too small to be seen over and above a more general entrenchment effect (see Ambridge, Pine & Rowland, 2012).

In contrast to the dative alternation, the preempting structures for overgeneralisations errors in the transitive-intransitive alternation, investigated in Chapter 4 of this thesis, are generally considered to be different constructions entirely (e.g. Brooks & Tomasello, 1999) – the passive construction (e.g. *The ball was kicked*) preempts intransitivisation errors with transitive-only verbs (e.g. **The ball kicked*), whereas the periphrastic causative construction (e.g. *Homer made the fish swim*) preempts transitivisation errors with intransitive-only verbs (e.g. **Homer swam the fish*). In addition, both of these preempting constructions are much lower in frequency than the transitive and the intransitive constructions themselves. This makes preemption effects much more difficult to observe, and goes some way to explaining the finding that no preemption effects were apparent in the judgment studies in Chapter 4.

The competition mechanism is clearly effective in explaining both children's initial production of overgeneralisation errors and how they might retreat from these. The mechanism has recently been instantiated as a connectionist model (Ambridge & Blything, in press) using the dative alternation (e.g. *Paul gave the book to Mary/Paul gave Mary the book*), with the model producing the pattern of overgeneralisation and retreat from overgeneralisation observed in children's acquisition data. This model was also able to correctly predict the construction(s) in which novel verbs would be grammatical using semantics alone (as adults and children were able to do in Chapter 3 of this thesis) and to reproduce the pattern of grammaticality judgments given by adults in Ambridge, Pine, Rowland and Chang (2012). These findings therefore offer strong support for the FIT account as a learning model that can truly account for the data observed in both production and grammaticality judgment studies.

This thesis has provided additional evidence for the FIT account, and thus enabled further development of thoughts about its operation, in a number of ways. The locatives study in Chapter 3 demonstrated that both semantic and frequency elements of the FIT account are in operation with this construction pair. While Ambridge, Pine and Rowland (2012) found similar results, the current study had the advantage that it also used novel verbs as a conclusive test of the role of verb semantics, showing that children and adults are both able to use the semantic compatibility between the new verb and the constructions alone (with no verb-in-construction information at all) to make judgments in line with real verbs that have similar semantics.

The causatives study in Chapter 4 showed that overall verb frequency information, as well as semantics, predicted grammaticality judgments from both adults and children. Verb frequency was used as the predictor variable to test the entrenchment hypothesis, but it is also equivalent to the sum of all verb-in-construction frequency counts, one of the four elements of the FIT account. Chapter 4 also presented some evidence for the use of semantic information (verb-in-construction compatibility) in grammaticality judgments with the causative alternation, although the nature of the continuous predictor variables created via Principal Components Analysis meant that the exact nature of the relationship between verb semantics and grammatical acceptability was unclear. In the production study, effects of verb frequency (equivalent to the sum of verb-in-construction frequencies), verb frequency in a competing construction (a combination of verb-in-

construction frequency and verb-in-construction compatibility) and semantics (verb-in-construction compatibility) were observed. The fact that all of these effects were significant in the same model adds weight to the argument that the various factors are all working alongside each other to influence the eventual choice of competing constructions.

Chapter 5 moved beyond overgeneralisation errors to look at the acquisition of a construction known to be problematic for young children: the passive. This chapter tested only the effect of verb-in-construction compatibility. The fact that both children and adults were primed to produce more passive sentences with a verb the more its semantics fit with those of the construction demonstrate that, not only does the FIT account provide a viable mechanism for explaining the retreat from overgeneralisation, but that this mechanism is able to apply more widely in the acquisition of argument structure.

6.4 Possible future research directions

The research presented in this thesis has demonstrated that both semantic and statistical accounts have a role to play in children's retreat from overgeneralisation errors and, in the case of semantics at least, in language acquisition more generally. However, the findings of both Chapter 3 and, in particular, Chapter 4 emphasise that further work is required to investigate how these mechanisms interact to enable children to retreat from overgeneralisation errors or, indeed, avoid them altogether. Since overgeneralisation errors are really a phenomenon of language production, the priming method used in Chapters 4 and 5 seems a promising way to move forward, rather than relying on metalinguistic knowledge using methodologies such as grammaticality judgments (although this method has proven instrumental in improving our understanding of the mechanisms in question up to this point). The priming method could easily be extended to study dative and locative overgeneralisation errors, for example.

Additionally, no mechanism for language acquisition can be language-specific: children learn the language of their environment, irrespective of the language of their biological parents. Nevertheless, the operation of this mechanism might be somewhat different in languages other than English. One way to test this would be to investigate how both statistical (entrenchment-like and preemption-like)

and semantic mechanisms might work together in languages that are typologically different to English, but in which argument structure overgeneralisation errors are still observed. For example, in some languages (e.g. K'iche' Mayan), no systematic variable in the morphological marker for the causative have been noted. Thus, effects of statistical variable may be observed, while no effect of semantics would be predicted. Hindi also has different morphological markers for causativisation, but semantic differences have been noted in this language. A semantic effect would therefore be expected. Conversely, while an effect of preemption would be predicted (both morphological markers are effectively synonymous), no entrenchment effect over and above this would be expected for Hindi due to the high frequency of the two main competing constructions and the comparatively very low frequency of any other constructions that can contain the verbs in question. By testing languages in which not all mechanisms would seem feasibly able to operate, language-specific and language-general aspects of the mechanism of argument structure acquisition can be investigated.

A further avenue that ought to be explored is how these mechanisms relate to language acquisition more generally, beyond the retreat from overgeneralisation. Chapter 5 has begun to explore this question by investigating the role of semantics in the production of passive sentences. This work could be extended to look at the roles of semantics and entrenchment in children's comprehension of the passive, following Ambridge, Bidgood, Pine, Rowland and Freudenthal's (2015) passive comprehension work with adults. Other constructions should also be investigated, partly because a successful mechanism involved in language acquisition cannot be specific to a single construction (this would be too specific) but also because the work presented here has suggested that the relative importance of different mechanisms may we weighted differently with different argument structures.

Importantly, future studies should focus on how to test integrated accounts, such as the FIT account, rather than continuing to test entrenchment, preemption and semantic mechanisms separately.

6.5 Conclusion

The three experimental chapters of this thesis have demonstrated clear effects of both frequency and semantics in children's retreat from overgeneralisation errors. They

have replicated and extended the results of previous papers in this area through the use of improved and innovative methodologies. While Chapter 4 provided clear evidence for the entrenchment hypothesis, however, no strong evidence for the preemption hypothesis was found. In addition, Chapter 5 demonstrated that semantic effects are to be found in other domains of language acquisition; here, in the acquisition of the passive construction. Despite the evidence for each of the hypotheses presented here, what is clear is that none of them can explain all of the lexical effects observed in the data on its own. What is needed is an approach that integrates both semantic and frequency information into account in its explanation of children's retreat from error, such as the FIT account (Ambridge & Lieven, 2011; Ambridge et al., 2011). While the current paper has provided important additional evidence for this account, it has also demonstrated that much more work is needed to clarify how the factors of construction frequency, verb-in-construction frequency, verb-in-construction compatibility and relevance interact with each other in the competition between constructions to convey a speaker's message.

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What preschool children do and do not do with ungrammatical word orders. *Cognitive Development, 16*(2), 679-692.

Agresti, A. (2002). *Categorical data analysis Vol. 359*. Hoboken, NJ: John Wiley & Sons.

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26*(2), 339-356.

Alishahi, A. & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science, 32*(5), 789-834.

Altman, G. T. M., Dienes, Z., & Goode, A. (1995). Modality Independence of Implicitly Learned Grammatical Knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(4), 899-912.

Ambridge, B. (2011). Paradigms for assessing children's knowledge of syntax and morphology. In Hoff, E. (Ed.). *Guide to Research Methods in Child Language*. London: Blackwell-Wiley, pp. 113-132.

Ambridge, B. (2013). How do children restrict their linguistic generalizations?: An (un-)grammaticality judgment study. *Cognitive Science, 37*(3), 508-543.

Ambridge, B., Bidgood, A., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2015). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science*.

Ambridge, B., Bidgood, A., Twomey, K. E., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2015). Preemption versus entrenchment: Towards a construction-general solution to the problem of the retreat from verb argument structure overgeneralisation. *PLoS ONE, 10*(4): e0123723.

Ambridge, B., & Blything, R. P. (in press). A connectionist model of the retreat from verb argument structure overgeneralization. *Journal of Child Language*.

Ambridge, B. & Lieven, E. V. M. (2011). *Child Language Acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Ambridge, B. & Lieven, E. V. M. (2015). A constructivist account of child language acquisition. In B. MacWhinney & W. O'Grady (Eds.), *Handbook of Language Emergence* (pp. 478-510). Hoboken, NJ: Wiley Blackwell.

Ambridge, B., Pine, J. M. & Rowland, C. F. (2011). Children use verb semantics to retreat from overgeneralization errors: A novel verb grammaticality judgment study. *Cognitive Linguistics, 22*(2), 303-323.

Ambridge, B., Pine, J. M. & Rowland, C. F. (2012). Semantics versus statistics in the retreat from locative overgeneralization errors. *Cognition, 123*, 260-279.

Ambridge, B., Pine, J. M., Rowland, C. F. & Chang, F. (2012). The roles of verb semantics, entrenchment, and morphophonology in the retreat from dative argument-structure overgeneralization errors. *Language, 88*(1), 45-81.

Ambridge, B., Pine, J. M, Rowland, C. F., Freudenthal, D., & Chang, F. (2014). Avoiding dative overgeneralization errors: Semantics, statistics or both? *Language, Cognition and Neuroscience, 29*(2), 218-243.

Ambridge, B., Pine, J. M., Rowland, C. F., Jones, R. L. & Clark, V. (2009). A semantics-based approach to the "no negative evidence" problem. *Cognitive Science, 33*(7), 1301-1316.

Ambridge, B., Pine, J. M., Rowland, C. F. & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition, 106*(1), 87-129.

Anderson, S. R. (1971). On the role of deep structure in semantic interpretation. *Foundations of Language, 7*, 387-396.

Anime Studio Pro Version 5.5 (2006). Scotts Valley, CA: e frontier America, Inc.

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Enquiry, 10*, 533-581.

Baker, M. C. (2003). *Lexical categories: Verbs, nouns and adjectives* (Cambridge Studies in Linguistics, vol. 102). Cambridge: Cambridge University Press.

Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning. *Psychological Science, 19*(3), 241-248.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language, 68*(3), 255-278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and s4. R package version 1.1-7. http://cran.r-project.org/package=lme4.

Bencini, G. M. L., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language, 59*(1), 97-113.

Bernaards, C. A. & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Education and Psychological Measurement: 65*, 676-696. http://www.stat.ucla.edu/research/gpa.

Bidgood, A., Ambridge, B., Pine, J. M., & Rowland, C. F. (2014). The retreat from locative overgeneralisation errors: A novel verb grammaticality judgment study. *PLoS ONE, 9*(5), e97634.

Bidgood, A., Ambridge, B., Pine, J. M. & Rowland, C. F. (in prep.). Children's and adults' passive syntax is semantically constrained: Evidence from syntactic priming.

Bidgood, A., Ambridge, B., Pine, J. M., Rowland, C. F., & Freudenthal, D. (in prep.). How do children retreat from overgeneralisation error? Evidence from the causative alternation.

Blything, R., Ambridge, B., & Lieven, E. V. M. (2014). Children use statistics and semantics in the retreat from overgeneralization. *PLoS ONE, 9*(10) e110009.

Boas, H. C. (2008). Towards a frame-constructional approach to verb classification. *Revista Canaria de Estudios Ingleses, 57*, 17-48.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355-387.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*(1), 1-39.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*(1), 150-171.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 123-172). Dordrecht: D. Reidel Publishing Company.

Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish*. Cambridge: Cambridge University Press.

Bowerman, M. (1976). Semantic factors in the acquisition of rules for word use and sentence construction. In *Directions in normal and deficient language development* (pp. 99-179). Baltimore: University Park Press.

Bowerman, M. (1978). Systematising semantic knowledge: changes overtime in the child's organisation of word meaning. *Child Development, 49*, 977-987.

Bowerman, M. (1981). The child's expression of meaning: Expanding relationships among lexicon, syntax and morphology. *Annuls of the New York Academy of Sciences 379*(1), 172-189.

Bowerman, M. (1982a). Evaluating competing linguistic models with language acquisition data: Implications of developmental errors with causative verbs. *Quaderni di Semantica, 3*(1), 5-66.

Bowerman, M. (1982b). Reorganizational processes in lexical and syntactic development. In Wanner, E. & Gleitman, L. R. (Eds.), *Language acquisition: the state of the art*. Cambridge: Cambridge University Press, pp. 319-346.

Bowerman, M. (1988). The "no negative evidence" problem: how do children avoid constructing an overly general grammar? In Hawkins, J. A. (Ed.), *Explaining language universals*. Oxford: Blackwell, pp. 73-101.

Bowerman, M. (1996). Argument structure and learnability: Is a solution in sight? *Proceedings of the 22nd Annual Meeting of the Berkley Linguistics Society: General Session and Parasession on The Role of Learnability in Grammatical Theory*, 454-468.

Boyd, J. K. & Goldberg, A. E. (2011). Learning what NOT to say: The role of statistical preemption and categorization in A-adjective production. *Language, 87*(1), 55-83.

Boyd, J. K., Ackerman, F. & Kutas, M. (2012). Adult learners use both entrenchment and preemption to inter grammatical constraints. In *2012 IEEE International Conference on Development and Learning and Epigenetic Robots*. Red Hook, NY: Curran Associates, pp. 1-2.

Braine, M. D. S. (1963). The ontogeny of English phrase structure: The first phase. *Language, 39*(1), 1-13.

Braine, M. D. S. (1971). On two types of models of the internalization of grammars. In D. I. Slobin (Ed.), *The Ontogenesis of Grammar* (pp. 153-186). New York: Academic Press.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development, 41*(164), 1-104.

Braine, M. D. S., Brody, R. E., Fisch, S. M., Weisberger, M. J., & Blum, M. (1990). Can children use a verb without exposure to its argument structure? *Journal of Child Language, 17*(2), 313-342.

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

Branigan, H. (2007). Syntactic priming. *Language and Linguistics Compass, 1*(1-2), 1-16.

Branigan, H., Pickering, M. J., Liversedge, S. P., Stewart, A. J., & Urbach, T. P. (1995). Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research, 24*(6), 489-506.

Bresnan, J. (2001). *Lexical-Functional Syntax.* Oxford: Blackwell.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. *Cognitive Foundations of Interpretation*, 69-94.

Brinkmann, U. (1997). *The locative alternation in German: Its structure and acquisition*. Amsterdam: John Benjamins Publishing Co.

British National Corpus (2007). BNC XML edition (3rd Ed..). Oxford: Oxford University Computing Services (distributor), on behalf of the BNC Consortium.

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.

Brooks, P. J. & Zizak, O. (2002). Does preemption help children learn verb transitivity? *Journal of Child Language, 29*, 759-781.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Brown R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), Cognition and the development of language (pp. 11-53). New York: Wiley.

Cameron-Faulkner, T., Lieven, E., & Theakston, A. (2007). What part of no do children not understand? A usage-based account of multiword negation. *Journal of Child Language, 34*(2), 251-282.

Chan, A., Meints, K., Lieven, E., & Tomasello, M. (2010). Young children's comprehension of English SVO word order revisited: Testing the same children in act-out and intermodal preferential looking tasks. *Cognitive Development, 25*(1), 30-45.

Chang, F., Dell, G. S. & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234-272.

Childers, J. B., & Tomasello, M. (2002). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology, 37*(6), 739-748.

Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1993). A minimalist program for linguistic theory. *MIT occasional papers in linguistics no. 1.* Cambridge, MA: MIT Working Papers in Linguistics.

Chouinard, M. M. & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language, 30*(3), 637-669.

Clark, E. (1988). On the logic of contrast. *Journal of Child Language, 15*(2), 317-335.

Coyle, J. M., & Kaschak, M. P. (2008). Patterns of experience with verbs affect long-term cumulative structural priming. *Psychonomic Bulletin & Review, 15*(5), 867–970.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language, 63*(3), 522-543.

Crain, S., & Thornton, R. (2012). Syntax acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(2), 185-203.

Dąbrowska, E., & Lieven, E. (2005). Towards a lexically specific grammar of children' question constructions. *Cognitive Linguistics, 16*(3), 437-474.

Dąbrowska, E., Rowland, C., & Theakston, A. (2009). The acquisition of questions with long distance dependencies. *Cognitive Linguistics, 20*(3), 571-197.

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences, 28*(6), 604-615.

Demetras, M. J., Nolan Post, K., & Snow, C. E. (1986). Feedback to first language learners: the role of repetitions and clarification questions. *Journal of Child Language, 13*(2), 275-292.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781.

Dixon, R. M. W. (1979). Ergativity. *Language, 55*(1), 59-138.

Ezeizabarrena, M. J, & Larrañaga, M. P. (1996). Ergativity in Basque: A problem for language acquisition? *Linguistics, 34*(5), 955-991.

Fellbaum, C. (1990). English verbs as a semantic net. *International Journal of Lexicography 3*, 278-301.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: SAGE.

Fillmore, C. J. (1967). The case for case. In Bach, E. & Harms, R. (Eds.), *Proceedings of the Texas Symposium on Language Universals, April 13-15 1967*. New York: Holt, Rinehart and Winston.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology, 31*(1), 41-81.

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition, 82*(3), 259-278.

Fox, D., & Grodzinsky, Y. (1998). Children's passive: A view from the by-phrase. *Linguistic Inquiry, 33*(2), 311-332.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684-691.

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

Goldberg, A. E. (2002). Surface generalizations: An alternative to alternations. *Cognitive Linguistics, 13*(4), 327-356.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, A. E. (2011). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics, 22*, 131-154.

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109-135.

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science, 4*(5), 178-186.

Gordon, P. & Chaftez, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition, 36*(3), 227-254.

Gropen, J., Pinker, S., Hollander, M. & Goldberg, R. (1991a). Affectedness and Direct Objects - the Role of Lexical Semantics in the Acquisition of Verb Argument Structure. *Cognition, 41*(1-3), 153-195.

Gropen, J., Pinker, S., Hollander, M. & Goldberg, R. (1991b). Syntax and Semantics in the Acquisition of Locative Verbs. *Journal of Child Language, 18*(1), 115-151.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R. & Wilson, R. (1989). Learnability and acquisition of the dative alternation in English. *Language, 62*(2), 203-257.

Hall, B. C. (1965). *Subject and object in modern English*. PhD Thesis: MIT.

Hirsch, C., & Wexler, K. (2006). Children's passives and their *resulting* interpretation. In K. Deen, J. Nomura, B. Schulz, & B. Schwartz (Eds.), *The proceedings of the inaugural conference on generative approaches to language acquisition – North America* (pp. 125-136). University of Connecticut occasional papers in linguistics 4.

Hirsh-Pasek, K., Treiman, R., & Schneiderman, M. (1984). Brown & Hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language, 11*(1), 81-88.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language, 50*(2), 185-195.

Ibbotson, P., & Tomasello, M. (2009). Prototype constructions in early language acquisition. *Language and Cognition, 1*(1), 29-85.

Israel, M., Johnson, C., & Brooks, P. J. (2000). From states to events: The acquisition of English passive participles. *Cognitive Linguistics, 11*(1-2), 103-129.

Jackendoff, Ray (1977). *X-bar-Syntax: A Study of Phrase Structure: Linguistic Inquiry Monograph 2.* Cambridge, MA: MIT Press.

Kerns, G. J. (2007). kmo() function for R. URL http://tolstoy.newcastle.edu.au/R/e2/help/07/08/22816.html

Kidd, E., Lieven, E., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development, 21*(2), 93-107.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford: Stanford University Press.

Langacker, R. W. (2003). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.). *Usage-based models of language acquisition*, pp.1-63. Stanford: CSLI.

Lempert, H. (1990). Acquisition of passives: The role of patient, animacy, salience, and lexical accessibility. *Journal of Child Language, 17*(3), 677-696.

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

Levin, B. & Rappaport, M. (1986). The formation of adjectival passives. *Linguistic Inquiry 17*(4), 623-661.

Lieven, E., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language, 30*(2), 333-370.

Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language, 24*(1), 187-219.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics, 20*(3), 481-508.

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

MacWhinney, B. (1975). Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language, 2*(1), 65-77.

MacWhinney, B. (2000) *The CHILDES project: Tools for analyzing talk* (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Maratsos, M., & Chalkey, M. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (Ed.), *Children's language*, (Vol. 2, pp. 127-151). New York: Gardner Press.

Maratsos, M., Fox, D., Becker, J., & Chalkey, M. (1985). Semantic restrictions on children's passives. *Cognition, 19*(2), 167-191.

Maratsos, M., Gudeman, R., Gerard-Ngo, P., & DeHart, G. (1987). A study in novel word learning: The productivity of the causative. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Matthews, D., & Bannard, C. (2010). Children's production of unfamiliar word sequences is predicted by positional variability and latent classes in a large sample of child-directed speech. *Cognitive Science, 34*(3), 465-488.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development, 20*(1), 121-136.

Meints, K. (1999). Prototypes and the acquisition of the English passive. In B. Kokinov (Ed.), *Perspectives on cognitive science*, Vol. 4, pp.67-77. Sofia: NBU Press.

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language, 66*(4), 568-587.

Messenger, K., Branigan, H. P., & McLean, J. F. (2011a). Evidence for (shared) abstract structure underlying children's short and full passives. *Cognition, 121*(2), 268-274.

Messenger, K., Branigan, H. P., & McLean, J. F. (2011b). Is children's acquisition of the passive a staged process? Evidence from six- and nine-year-olds' production of passives. *Journal of Child Language, 39*(5), 991-1016.

Morgan, J. L., Bonamo, K. M., & Travis, L. L. (1995). Negative evidence on negative evidence. *Developmental Psychology, 21*(2), 180-197.

Morgan, J. L., & Travis, L. L. (1989). Limits on negative information in language input. *Journal of Child Language, 16*(3), 531-552.

Naigles, L. R., Fowler. A. & Helm, A. (1992). Developmental Shifts in the Construction of Verb Meanings. *Cognitive Development, 7*, 403-427.

Naigles, L. R. & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25*: 95-120.

Naigles, L. R. & Lehrer, N. (2002). Language-general and language-specific influences on children's acquisition of argument structure: a comparison on French and English. *Journal of Child Language, 29*: 545-566.

Ninio, A. (2014). Variables and values in children's early word-combinations. *Psychology of Language and Communication, 18*(2), 106-125.

Noble, C. H., Rowland, C. F., & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: Evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive Science, 35*(5), 963-982.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39-57.

Penner, S. G. (1987). Parental responses to grammatical and ungrammatical child utterances. *Child Development, 58*(2), 376-384.

Perfors, A., Tenenbaum, J, B, & Wonnacott, E. (2010). Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language, 37*: 607-642.

Peter, M., Chang, F., Pine, J.M., Blything, R., & Rowland, C.F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language, 81*, 1-15.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin, 134*(3), 427-459.

Pine, J. M., & Lieven, E. V. M. (1993). Reanalysing rote-leaned phrases: Individual differences in the transition to multi-word speech. *Journal of Child Language, 20*(3), 551-571.

Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the development of the English verb category. *Linguistics, 36*(4), 807-830.

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

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

Pinker, S., Lebeaux, D. S., & Frost, L. A. (1987). Productivity and constraints in the acquisition of the passive. *Cognition, 26*(3), 195-267.

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar.* Chicago: University of Chicago Press.

Pye, C. (1990). The acquisition of ergative languages. *Linguistics, 28*(6), 1291-1330.

R Core Team (2014). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.r-project.org/.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Reber, A. S. (1969) Transfer of syntactic structure in synthetic languages. *Journals of Experimental Psychology, 81*(1), 115-119.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*(3), 219-235.

Revelle, W. (2014). psych: Procedures for personality and psychological research. Version 1.4.6. Evanston, IL: Northwestern University. http://cran.r-project.org/package=psych.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of experimental psychology: General*, *104*(3), 192-233.

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. V. M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition, 125*(1), 49–63.

Rowland, C. F., & Theakston, A. L. (2009). The acquisition of auxiliary syntax: A longitudinal elicitation study. Part 2: The modals and auxiliary DO. *Journal of Speech, Language and Hearing Research, 52*(6), 1471-1492.

Savage, C., Lieven, E. V. M., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science, 6*(5), 557-567.

Savage, C., Lieven, E. V. M., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development, 2*(1), 27-149.

Saxton, M., Backley, P. & Gallaway, C. (2005). Negative input for grammatical errors: Effects after a lag of 12 weeks. *Journal of Child Language, 32*(3), 643-672.

Saxton, M., Kulcsar, B., Marshall, G., & Rupra, M. (1998). Longer-term effects of corrective input: An experimental approach. *Journal of Child Language, 25*(3), 701-721.

Schäfer, J., Opgen-Rhein, R., Zuber, V., Ahdesmäki, M., Duarte Silva, A.P., & Strimmer, K. (2013). corpcor: Efficient estimation of covariance and (partial) correlation. Version 1.6.6. http://cran.r-project.org/package=corpcor.

Schlesinger, I. M. (1988). The origin of relational categories. In Y. Levy, I. M. Schlesinger & M. D. S. Braine (Eds.), *Categories and processes in language acquisition* (pp. 121-178). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stefanowitsch, A. (2008). Negative evidence and preemption: A constructional approach to ungrammaticality. *Cognitive Linguistics, 19*(3), 513-531.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences* (5th Ed.). New York: Routledge.

Street, J. A., & Dąbrowska, E. (2014). Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. *Applied Psycholonguistics, 31*(1), 97-118.

Sudhalter, V. & Braine, M. D. S. (1985). How does comprehension of passives develop? A comparison of actional and experiential verbs. *Journal of Child Language, 12*(2), 455-470.

Theakston, A. L. (2004). The role of entrenchment in children's and adults' performance on grammaticality judgement tasks. *Cognitive Development, 19*(1), 15-34.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2005) The acquisition of auxiliary syntax: BE and HAVE. *Cognitive Linguistics, 16*(1), 247-277.

Thothathiri, M. & Snedeker, J. (2008) Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language, 58*(2), 188-213.

Tomasello, M. (1992). *First verbs: A case study of grammatical development.* Cambridge: Cambridge University Press.

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

Tomasello, M., Brooks, P. J., & Stern, E. (1998). Learning to produce passive utterances through discourse. *First Language, 18*(53), 223-237.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*(2), 105-110.

Valian, V., Solt, S., & Stewert, J. (2009). Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language, 36*(4), 743-778.

van Valin, R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge, UK: Cambridge University Press.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*(1), 23-79.

Wilson, S. (2003). Lexically specific construction in the acquisition of inflection in English. *Journal of Child Language, 30*(1), 75-115.

Wonnacott, E. (2011). Balancing generalization and lexical conservatism: An artificial language study with child learners. *Journal of Memory and Language, 65*, 1-14.