Meaning and Context in Children's Understanding of Gradable Adjectives

Abstract

Children as young as three years of age and adults distinguish between three types of gradable adjective (GA) meanings – relative, absolute maximal, and absolute minimal – based on the role of the context in setting the standard of comparison. Results lend support to a typology of GAs determined by differences in scalar structure. While relative GAs such as *big* depend on the context for the standard of comparison, absolute GAs (e.g., *full*, *spotted*) do not. Evidence comes from a pragmatically-oriented task in which we exploit participants' awareness of the existence and uniqueness presuppositions of definite descriptions to highlight differences among GAs, seen in their willingness to accommodate when the presuppositions are violated. Reaction time analysis of children's non-adult-like acceptance of infelicitous requests with maximal standard GAs reflects an allowance of imprecision that we argue is pragmatic in nature, and therefore distinct from the vagueness encoded in the semantic representations of relative GAs.

Meaning and Context in Children's Understanding of Gradable Adjectives

1.Introduction

A significant part of becoming a competent language user requires understanding the relation between context and meaning. In this paper, we explore three ways that meaning and context interact in child and adult language: the interpretation of expressions whose meanings are partially determined by the context of utterance (semantic context dependence); context-based conditions on the felicitous use of an expression (presupposition accommodation); and informative uses of expressions in contexts in which they strictly speaking do not apply (imprecision). Our empirical focus involves uses of gradable adjectives in definite descriptions (noun phrases), in which each of these three types of context/meaning interaction is manifested. Specifically, we focus on definite descriptions in which a gradable adjective modifier takes on different interpretations in different contexts in order to ensure that the presuppositions of existence and uniqueness introduced by the definite determiner are satisfied. We show that by the age of 3, children are sensitive to all three varieties of context dependence and that their application of the different context/meaning relations is appropriately guided by the relevant aspects of their linguistic representations. Thus we demonstrate that by age 3, children's semantic representations for gradable adjectives and definite determiners are fully adult-like and moreover, that the consequences of these representations for relating meaning and context are in place.

1.1   Relative Gradable Adjectives and Context Dependence

Gradable adjectives (GAs) are adjectives whose core meaning involves reference to a scalar concept on the basis of which objects can be ordered (e.g., height, weight, cost, etc.). Distributionally, gradable adjectives are identified by the fact that they can appear felicitously in

comparative constructions (e.g., *taller than*, *as heavy as, less expensive than*) and with various types of degree morphemes (measure phrases, intensifiers, etc.) whose function is to specify where the argument of a gradable adjective is located on the corresponding scale. Our focus in this paper is the meaning and use of gradable adjectives in contexts that lack degree morphology of any kind: GAs in the so-called positive form, such as *tall, heavy* and *expensive*. From a semantic perspective, the positive form is interesting because it is context dependent: intuitively, what 'counts as' tall can vary from context to context, depending on the syntactic environment in which it appears, the topic of discussion, or the interests and expectations of the participants in the discourse.

For example, although there is clearly a shared meaning between the three uses of *tall* in (1a-c) (which we will characterize in more detail below), it is equally clear that on the most natural interpretations of these examples, the actual height that is sufficient to qualify as *tall* changes: tall snowmen are shorter than tall buildings, and tall buildings are shorter than tall mountains.

1.   a.   That is a tall snowman.

      b.   That is a tall building.

      c.   That is a tall mountain.

The examples in (2a-b) involve similar variation in what it takes to count as tall, but in this case it is not the modified noun that provides the basis for this shift, but rather extralinguistic knowledge about the heights of snowmen that is brought to bear in light of the meanings of the agents of the described events: third graders tend to build smaller snowmen than fraternity brothers (Kamp & Partee, 1995).

2.   a.   The third graders built a tall snowman.

b.   The fraternity brothers built a tall snowman.

Finally, we see context dependence at work even in a simple predication such as (3a): we might know that Anna is a woman, a gymnast, and taller than the average gymnast, but whether (3a) is judged to be true or false of Anna depends on how we understand the adjective. If we understand (3a) as spelled out in (3b), then we would probably judge the sentence to be true. If we understand (3a) as in (3c), however, we might very well judge it to be, since even a tall gymnast might be a short woman.

3.   a.   Anna is tall.

      b.   Anna is tall *for a gymnast.*

      c.   Anna is tall *for a woman.*

In what follows, we will refer to GAs that are context dependent in the positive form as RELATIVE GAs (to be distinguished later in section 1.3 from a second class of GAs which are not). A common analysis of the positive form of relative GAs is that they denote properties that are true of an object just in case it possesses a degree of the scalar concept encoded by the GA that exceeds a contextually determined STANDARD OF COMPARISON (see e.g., Bartsch & Vennemann, 1972; Cresswell, 1977; Kennedy, 1999, to appear; Klein, 1980, 1991; Ludlow, 1989; von Stechow, 1984; Wheeler, 1972; and many others). The standard of comparison is a degree that typically corresponds to an 'average' or 'norm' of the measured concept relative to some salient set (usually referred to as a 'comparison class'), which may be explicitly indicated by a constituent that the adjective is in construction with (as in (1a-c) and (3b-c)), may be inferred based on other information in the sentence (as in (2a-b)), or may be implicit (as in (3a)). (The standard of comparison can also be fixed on a more arbitrary basis in special contexts; see Fara (2000) for discussion.)

There are a number of ways in which this kind of analysis can be implemented in a compositional semantics, which differ primarily on the semantic type assigned to lexical adjectives and in more general assumptions about the relation between the morphologically unmarked positive form and the various forms that have explicit degree morphology. For the purposes of this paper, we will assume a simplified semantics that abstracts away from these issues, showing just those features of the positive form that a more fully articulated compositional account will have to end up with. Specifically, we will assume that relative GAs in the positive form have context-dependent denotations along the lines of the one given for *tall* in (4), where $[[\alpha]]^c$ means 'the denotation of α in context c', and $\lambda x.\phi$ means 'the property that is true of an object o just in case ϕ is true when o is substituted for all occurrences of x in ϕ.'

4. $[[\text{tall}_{\text{pos}}]]^c = \lambda x.x$'s height exceeds the standard of comparison for height in c

We will have nothing to say here about how the standard of comparison is actually computed, or differences that may exist between the various kinds of examples illustrated above based on the extent to which the comparison class is made explicit. (See Kennedy (to appear) for detailed discussion of these issues.) While these are all important issues, they are orthogonal to the point that is of central concern to us here, and which is made explicit in (4): the meaning of a relative GA such as *tall* varies from context to context, depending on where the standard of comparison is actually set. Specifically, the feature of meaning of *tall* that changes is its extension: the things that it is true and false of. Returning to the example in (3a), let us use *c(gymnast)* to represent a context in which we are talking about gymnasts, and intend (3a) to be understood as (3b), and *c(woman)* to represent a context in which we are talking about women and intend (3a) to be understood as (3c). Whenever the standards of comparison in *c(gymnast)* and *c(woman)* are distinct (which is normally the case), we will end up with the result that

$[[\text{tall}_{\text{pos}}]]^{c(\text{gymnast})} \neq [[\text{tall}_{\text{pos}}]]^{c(\text{woman})}$, because the set of things whose heights exceed the standard in the former context is a proper subset of the set of things whose heights exceed the standard in the latter context. In particular, if Anna's height falls in between the standard of comparison in *c(gymnast)* and *c(woman)*, as represented graphically in (5), then $[[\text{tall}_{\text{pos}}]]^{c(\text{gymnast})}$ is true of Anna but $[[\text{tall}_{\text{pos}}]]^{c(\text{woman})}$ is false of Anna.

5.   HEIGHT: 0 ---------- $\text{standard}_{c(\text{gymnast})}$ ----- $\text{height}_{\text{Anna}}$ ----- $\text{standard}_{c(\text{woman})}$ ------->

The denotation in (4) (and comparable denotations for other relative GAs) thus captures the fact that all uses of the positive form involve a shared meaning: a relation between the degree to which an object manifests some scalar concept (in this case, tallness) and a contextually determined standard of comparison for that concept. However, it is not until we actually fix the standard of comparison that that we get a property that can be true or false of an object – the kind of thing that can be used to make assertions or denials, ask questions, and so forth. It is in this sense that the meaning of the positive form is context dependent.

Before moving on, we should explicitly mention that in addition to being context dependent, the positive form is also vague, in the sense that there are contexts in which even with complete knowledge of the relevant facts, it may not be possible to judge the truth of a predication involving it. For example, in the case of (3a), we might know exactly what Anna's height is, that the intended meaning of (3a) is (3b), and that the average height of a gymnast is such-and-such, and yet still be unsure about whether it is actually true that Anna is tall (for a gymnast): she could be a borderline case. For example, if we know that the average height of a gymnast is 1.45 meters, and Anna is 1.6 meters, then (3b) is clearly true. However, if Anna is 1.5 meters tall, then (3b) does not necessarily follow.

Similarly, we find it difficult or impossible to judge adjacent objects along a scalar continuum differently relative to a positive form GA whose meaning makes reference to that continuum: the conditional in (6), for example, seems to be straightforwardly valid.

6.   For any individuals *a* and *b*, if *a* is tall and *b* is just a bit shorter than *a*, then *b* is also tall.

But if (6) were valid, it would support (through successive applications) the unacceptable conclusion that a 1-meter-tall man is tall given the initial (uncontroversial) premise that a 2-meter-tall man is tall; this is an example of the so-called Sorites Paradox (the paradox of the heap).

Vagueness and context dependence go hand in hand in the case of the positive form (of a relative GA), and a full account of the semantic properties of this construction must take both into account. Simply put, the mere assumption that the positive form is context dependent, even in the way out lined above, is not enough to account for vagueness, since denotations along the lines of (4) do not by themselves explain borderline cases or our (mistaken) intuitions about the inductive premise of a Sorites argument. However, most accounts of vagueness make crucial use of context dependence (e.g., Fara, 2000; Kamp, 1981; Kamp & Partee, 1995; Kennedy, to appear; Raffmann 1994, 1996; Soames, 1999), and we will assume that ultimately these two aspects of meaning are related features of the semantics of the positive form. To keep the exposition as clear as possible, however, we will continue to talk strictly in terms of the context dependence of the positive form rather than its vagueness, both because the crucial factor that will be important for our experimental investigation of the context/meaning relation is the way that different contexts result in different standards of comparison, and because the other features of vagueness outlined above do not play a role in our investigation.

1.2   Gradable Adjectives in Definite Descriptions:

Accommodating the Standard of Comparison

The context dependence of relative GAs provides a great deal of flexibility in their use: because they can convey different properties in different contexts, unified by their 'relation to a standard' semantic core, the same form can be used to convey different information about the degree to which an object manifests some scalar concept, as illustrated by the examples in (1)-(3) above. A particularly striking example of this flexibility comes from the use of relative GAs in definite descriptions, which is the empirical focus of this paper.

Definite descriptions have been the focus of a great deal of work in semantics, pragmatics and the philosophy of language, most of which converges in some form or another on two central semantic/pragmatic claims: a singular definite NP of the form *the ϕ* is associated with two presuppositions. First, it presupposes that there is an object that satisfies the property encoded by ϕ (the EXISTENCE PRESUPPOSITION). Second, it presupposes that the object uniquely satisfies ϕ (the UNIQUENESS PRESUPPOSITION). (See Abbott, 1999; Birner & Ward, 1994; Heim, 1990; Kadmon, 1990; Neale, 1990; Roberts, 2003; Russell, 1905; Strawson,1950; and many others for representative discussion.) There are important differences of opinion on the relation between these conditions, the extent to which they can be overridden, and the way that apparent violations should be handled, but for our purposes we can take them as reasonable approximations of what a speaker is committed to in order to felicitously use a definite description, and what a hearer takes to be the case if she accepts such a use.

The effect of these presuppositions on judgments of felicity can be illustrated by the following example. Consider a context in which two individuals A and B are sitting across from each other at a table, there are two blue rods of unequal lengths on the table in front of B, and A's goal is to get B to pass over one of the rods. In such a context, A cannot felicitously use (7)

to make this request, because the uniqueness presupposition of the definite description *the blue rod* is not met: there are two objects in the context that satisfy the property denoted by the nominal constituent *blue rod*.

7.   #Please give me the blue rod.[1]

By the same token, A's utterance of (8) would be just as infelicitous, in this case because the existence presupposition of the definite is not met: there is no object that satisfies the property *red rod* in the context.

8.   #Please give me the red rod.

        Speaker A *can*, however, felicitously use (9) to request the longer of the two rods.

9.   Please give me the long rod.

Importantly, (9) is felicitous when the two rods are independently judged to be long, when they are independently judged to be not long, and when one is considered long and the other not: all that matters is that there is a difference in length between them. On the surface, this appears to conflict with what we saw with (7), which fails uniqueness because both rods are blue rods, and (8), which fails existence because neither rod is a red rod. The crucial difference, of course, is that *long* is a positive form relative GA, and so denotes the context dependent property in (10).

10. $[[\text{long}_{\text{pos}}]]^c = \lambda x.x$'s length exceeds the standard of comparison for length in c

        Two features are particularly salient about the context in which (9) is used, and these bear on the choice of a standard of comparison: the presuppositions of the definite NP require there to be a unique long rod in the context, and the two rods are of unequal lengths. Given that there are two salient rods in the context *c(rod)*, the only way to satisfy these presuppositions is for one of

---

[1] In this paper, we will use the '#' symbol to indicate an utterance that is grammatical, but infelicitous.

them to count as long and the other one not, relative to the context. Because the rods have unequal lengths, this result can be achieved by 'shifting' the prevailing standard of comparison – the one on the basis of which the rods are independently judged to be both long or both not long – so that it differentiates between them, as represented in (11).

11.   LENGTH: 0 ------------ $length_{rod\_1}$----- $standard_c$ ----- $length_{rod\_2}$ ----------->

This move from the prevailing standard (of what is considered to be a long rod) to one that allows the presuppositions of the definite to be satisfied in that context is an instance of presupposition accommodation (Lewis, 1968) – treating the context as though it contained the information relevant to satisfying some presupposition. In this case, the crucial move is to treat the context as though the standard of comparison for *long* is one that makes it true of one rod and false of the other. Given the statement of the truth conditions of the positive form GAs, this amounts to making it true of the longer of the two rods (if one object has a degree of the scalar concept named by a gradable adjective α that exceeds the standard for α, and a second object does not, then the first is more α than the second).[2]

In sum, definite descriptions involving nouns modified by positive form relative GAs are particularly interesting for studies of the context/meaning interface because it is possible to construct scenarios (the differentiation tasks described above) in which they invoke two different

---

[2] There is a way of using a positive form GA in a definite description to ask for the lower-ranked object: use the polar-negative antonym, e.g., *the short rod*. This strategy does not always work due to markedness effects associated with negative members of antonym pairs, but for those adjectives that do not show such effects (or show them weakly, such as *short*), the basic account runs the same as for *long*, modulo the difference between the antonyms in the orderings they impose: one is the inverse of the other, so 'more shortness' corresponds to 'less longness'.

features of this interface: 1) contextually adjusting the meaning (extension) of the adjective by manipulating the standard of comparison, and 2) doing so in such a way that accommodates the presuppositions of the definite description.

1.3   Absolute Gradable Adjectives and Imprecision

As it turns out, not all GAs are context dependent (or vague) in the positive form, a fact that provides the basis for a set of control data in experimental tasks that use facts like those discussed in the previous section to investigate understanding of the context/meaning interface. In addition to the large set of relative GAs, there is a class of adjectives that are demonstrably gradable (in that they readily allow comparisons) but which do not have context-dependent standards of comparison in the positive form. Following Unger (1975), Kennedy and McNally (2005), and Kennedy (to appear), we will refer to these as ABSOLUTE GAs (see also Rotstein & Winter, 2004).[3] Absolute GAs come in two varieties: minimum standard absolute GAs, which require their arguments to merely possess some degree of a gradable property (e.g., *spotted* and *bent*), and maximum standard absolute GAs, which require their arguments to possess a maximal degree of a gradable property (e.g., *full* and *straight*). Rotstein and Winter (2004), Kennedy and McNally (2005), and Kennedy (to appear) provide extensive arguments on the basis of modifier selection and entailment patterns for the absolute/relative distinction; what is important for our purposes here is that neither type of absolute GA has the sort of 'differentiation use' in definite descriptions that we see with relative GAs.[4]

---

[3] Rips and Turnbull (1980) also made use of a relative/absolute distinction among adjectives, but meant something slightly different.

[4] A question that we will not address in this paper concerns the factors that determine whether a GA is relative or (minimum/maximum) absolute, as these are discussed in detail in other work

For example, maintaining the kind of two-object context discussed in the previous section, a definite description such as *the spotted disk* can be felicitously used to request one of the two disks just in case one has spots and the other does not, but not to pick out the more spotted of two disks that are both spotted (or, obviously, one of two spotless disks). Likewise, *the full jar* can be felicitously used to request one of two jars only if one jar is completely full and the other is not; it cannot normally be used to refer to the fuller of two partially-full jars, or one of two jars that are both full. These judgments are expected if the positive forms of *spotted* and *full* have denotations along the lines of (12a) and (12b), respectively.

12.   a. $[[\text{spotted}_{\text{pos}}]]^{\text{c}} = \lambda x.x$ has some spots

   b. $[[\text{full}_{\text{pos}}]]^{\text{c}} = \lambda x.x$ has maximal fullness

Unlike the kinds of denotations we posited for relative GAs in the positive form, these denotations are not context dependent: *spotted* just denotes the property of having some spots, and *full* denotes the property of being maximally full. If this is right, then the difference between the contexts described above that result in infelicity with these adjectives, and the corresponding contexts involving relative GAs, is that the former do not license the possibility of accommodation: if *full* means 'maximally full', independent of context, there is no way to 'recalibrate' the meaning of the adjective (by manipulating a standard of comparison) to make the adjective true of an object that is not maximally full, but merely fuller than another.

(cf. Kennedy & McNally, 2005; Kennedy, to appear; Rotstein & Winter, 2004; Syrett, in preparation). In general, a GA is relative if the scalar concept it encodes is open (has no natural endpoints), and absolute if it is closed (has either a minimum value or a maximum value or both).

The initial reaction to these comments is typically to reject them. For example, there is a strong intuition that *full* merely requires that an object be 'close to full', and that in different contexts different approximations may do. For example, while it is clear that a jar that is only half full cannot be truthfully described as *a full jar*, and maybe not one that is 3/4 full either, it is not so clear what to say about one about one that is 7/8 full, or 15/16 full, or any noticeably-not-but-almost-there-amount full.[5] While it is clear that at some point we would typically be willing to start calling the jar *full* (and would therefore be willing to refer to it as *the full jar*), and that this point might be different in different contexts (based on our goals, the types of jars, the contents of the jars, etc.), what remains unclear is whether it follows from this that *full* really is context dependent in the same way as *tall* and *long*. Or does it indicate a different kind of context/meaning interaction, one that regulates *imprecision* rather than *vagueness* – tolerance of the false-but-informative application of a description to an object in contexts in which such applications are 'close enough to true'?

There are different answers to this question in the literature, with no clear consensus. However, the experiments we report below provide important new data bearing on this debate, both because they clearly support the core semantic distinction between relative and absolute GAs – the former have context sensitive denotations in the positive form and the latter do not – and because they further indicate that imprecise uses of absolute GAs are processed very differently from uses of relative GAs in which a standard of comparison is shifted to

---

[5] We intend the fractions here to be based on whatever is the conventional maximum or 'full-line' for the container in question. This may be the rim, or it may be a particular mark on the container, or it may be something less concrete. The subsequent comments apply regardless of which option we choose.

accommodate the presuppositions of a definite NP, suggesting that the two uses do not in fact involve the same kind of context/meaning interaction.

1.4   Overview of the Paper

This paper is organized as follows. First, we examine children's understanding of gradable adjectives in definite descriptions to probe their level of knowledge of the first two types of context/meaning interaction. We begin by providing some background on children's knowledge of gradable adjectives and the presuppositions of definite NPs, and then move on to describe our experiments on children's understanding of differentiation uses of definite descriptions with gradable adjective modifiers. We will see that children as young age 3 have adult-like competence in the interpretation of both relative and absolute GAs, assigning the former context-dependent and the latter fixed interpretations, and that they are sensitive to – and can correctly accommodate – the presuppositions of definite NPs.

We then focus on important differences between children and adults involving maximum standard absolute GAs. Given a request with a definite NP of the form *the A one*, where A is a maximum standard GA, and a context in which neither of the two objects exhibits the relevant property to a maximum degree, adults uniformly rejected the request, illustrating their implicit acknowledgement of the failure of the existence presupposition. Children, however, systematically accepted such uses, subject to other discourse conditions. Crucially, though, they also took significantly longer to accept such requests than those with relative GAs or true predications of maximum or minimum standard absolute GAs. We argue that this finding supports a distinction between true semantic context dependence of the sort involved in the interpretation of the positive form of relative GAs on the one hand, and judgments about contextually acceptable degrees of imprecision on the other. The former is a matter of the

conventional meaning of relative GAs, which happens automatically because of the kind of meaning that a relative GA has; the latter involves pragmatic reasoning about how much deviation from the conventional meaning of a term (a maximum standard absolute GA in this case) should be tolerated in contexts in which it does not hold.

Our study is important in two respects. First, it provides new insights on when children have achieved adult-like levels of competence in two fundamental aspects of the context/meaning relation (semantic context dependence and presupposition accommodation) – by three years of age. Second, it provides new data from child language that bears on more general questions about the nature of contextually-influenced interpretive variability in both children and adults. These data support a distinction between variability that is fundamentally *semantic* in nature (context dependence) and variability that is fundamentally *pragmatic* (imprecision).

2. Adjectives and Definiteness in Child Language

2.1 Children's Knowledge of Adjectives and Contextual Variability

Research on the acquisition of adjectives has largely focused on three questions. First, when do children recognize that adjectives license different inferences than nouns? The evidence points to two conclusions: (a) that noun learning precedes adjective learning, and (b) that well before reaching preschool, children recognize that adjectives refer to object properties and not object kinds (or said another way, that adjectives and nouns denote different kinds of properties) (cf. Booth & Waxman, 2003; Gentner, 1982; Macnamara, 1972; Waxman & Booth, 2001; Waxman & Markow, 1995, 1998). Second, to what extent does knowledge of the category to which a novel adjective is applied impact children's ability to acquire that word? Here, the evidence suggests that familiarity with the object kind makes it easier for children to acquire an

adjective referring to a property of that kind (cf. Hall, Waxman, & Hurwitz, 1993; Klibanoff & Waxman, 1998, 2000; Mintz, 2005; Mintz & Gleitman, 2002; Taylor & Gelman, 1988; Waxman & Klibanoff, 2000). Third, to what extent do modificational (prenominal) uses of adjectives license different inferences than predicative uses? In this domain, evidence suggests that adjectives in the prenominal position may have a privileged role in cueing contrast among object properties (Diesendruck, Hall, & Graham, 2006; Nelson, 1976; Prasada, 1992; Prasada & Cummins, 2001; but see Nadig, Sedivy, Joshi, & Bortfeld, 2003). Such results lay an important foundation to our questions, namely, when do children become aware of the different subcategories of GAs (cf. Graham, Welder, & McCrimmon, 2003, Graham, Cameron, & Welder, 2005) and how do these subcategories differ with respect to their context-dependent features?

With respect to context dependence, preschoolers have been shown to use a variety of information sources from the context to evaluate the use of an adjective. For example, they can use the orientation of an object and its intended reference to determine the relevant dimension for evaluating its size (cf. Coley & Gelman, 1989). Similarly, they allow the interpretation of an adjective to change across kinds (Sharpe, Fonte, & Christe, 1998), and appreciate that information provided by the noun may be relevant to the interpretation of a novel adjective (e.g., Mintz, 2005; Mintz & Gleitman, 2002).

Restricting our attention to relative GAs, two lines of research have demonstrated that preschoolers are able to shift the standard of comparison for relative GAs such as *big*, *tall*, *high*, and *low*. A series of papers by Gelman and Ebeling have evaluated how children's ability to judge the size of an object and their facility to move between standards is affected by the nature of the standard of comparison (Ebeling & Gelman, 1988, 1994; Gelman & Ebeling, 1989). Relatedly, Smith and her colleagues have probed the factors affecting children's willingness to

relabel an object once the standard has shifted (Sera & Smith, 1987) and their ability to make use

of a given range of values for assigning the standard of comparison (Smith, Cooney, & McCord,

1986). Together, these studies demonstrate that preschoolers take a range of contextual

information into account when assigning the standard of comparison for these adjectives.

However, since these studies did not include non-relative GAs as controls (see section 1.3

above), it has remained an open question whether children would consistently do this for all

adjectives (gradable or not), or would appropriately restrict this behavior to this class of GAs.

The study that has come closest to introducing control adjectives is Nelson and Benedict

(1974). Nelson and Benedict asked children aged three to six for their judgments about a series

of pictures of objects and found that children exhibited a shorter response latency to the

comparative form of adjectives such as *tall* (e.g., *taller*) than they did for adjectives such as

happy or *leafy*. The authors interpreted these results as demonstrating that the former adjective

carry an implicit comparison (cf. Katz, 1967), whereas the others do not. That is, it takes no more

time to process the comparative form of *tall* than the positive form, but it takes more time to

process *happier* than *happy*. However, because the authors imposed distinctions among the

different sets of adjectives that do not directly correspond to semantic distinctions found in

natural language, it is difficult to draw conclusions from these results. The current study starts

with the distinctions between relative and absolute GAs discussed above and offers behavioral

evidence from both children and adults as empirical support for these distinctions.

2.2   Children's Production and Comprehension of Definite Noun Phrases

Although the presuppositions of definiteness are clear for adults, the picture for children

is murkier. Children's overreliance on the definite determiner in earlier production studies (cf.

Karmiloff-Smith, 1979; Maratsos, 1976) has been interpreted as showing that children do not

recognize the presuppositions of definiteness (Wexler, 2003), while and their performance in recent eyetracking experiments (cf. Trueswell, Sekerina, Hill, & Logrip, 1999) has suggested that they have difficulty parsing definite NPs. However, while both of these findings suggest that definiteness poses a problem for children, they actually leave unresolved the question of whether children are aware of the presuppositions that a singular definite NP carries.

There is by now a well-known asymmetry in comprehension and production, so showing that children that children do not produce the definite determiner when it is felicitous to do so is more revealing about how aspects of the context at hand motivate the utterances they produce, and not how their evaluation of utterances intended to capture aspects of the context reflects their representations. Likewise, the results from eye-tracking experiments reveal that children have difficulty using information from the context to restrict the reference of the noun (cf. Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Meroni, 2006), not that they do not understand what presuppositions are involved. By using definite descriptions to draw out the distinctions between relative and absolute GAs, we demonstrate not only that children have different representations for these two kinds of GAs (one being context-dependent and the other not), but that they see the presuppositions of the singular definite NP as highlighting these differences in the manner outlined in sections 1.2 and 1.3.

3.Pre-Experiment Scalar Judgment Task

3.1   Introduction

The goal of the Scalar Judgment Task (SJT) was twofold. First, we wanted to elicit scalar judgments from participants in the hopes that these judgments would reflect a division among the three subclasses of GAs. Second, we sought to assess children's judgments of *big* and *long* on stimuli that would be used in Experiments 1 and 2.

3.2   Method

3.2.1   Participants

36 children representing three age groups participated in this task: 12 three-year-olds (6 boys 6 girls, range: 3;3 to 3;11, M: 3;8); 12 four-year-olds (5 boys 7 girls, range: 4;1 to 4;11, M: 4;5); and 12 five-year-olds (5 boys 7 girls, range: 5;0 to 5;11, M: 5;5). In addition, 28 adult native speakers of English (Northwestern undergraduates fulfilling an experimental requirement for a Linguistics course) served as controls.

3.2.2   Materials

The materials consisted of four sets of seven items each, as outlined in Table 1.

Table 1: Stimuli for Scalar Judgment Task

adjective	stimuli
relative	
big	7 wooden blocks painted blue, decreasing incrementally in size from $2^{3}''$ to $5/8^{3}''$
long	7 wooden rods painted green, .5″ in both width and height, decreasing incrementally in length from 8″ to 2″
absolute	
spotted	7 wooden disks painted red, all 3.75″ in diameter and .5″ thick, ranging from being covered with spots to having no spots
full	7 clear plastic containers with white lids, all 2″ in height and 1.5″ in diameter, and ranging from being full (of lentils) to empty

3.2.3   Procedure

Participants were presented with four sets of seven objects one after the other. For each item in each set, the experimenter asked the participant, *Is this A?*, where A was a target

adjective corresponding to the property exemplified by the objects in the set (*big*, *long*, *full*, *spotted*). The experimenter always started at the positive, rather than the negative, pole (i.e., at the big end, not the small end).[6] There were two conditions, based on the order of presentation of the sets. Half of the participants were randomly assigned to the 'Relative-Absolute' condition and saw the sets in the following order: *big*, *long*, *full*, *spotted* (i.e., relative GA, relative GA, absolute GA, absolute GA). The other half of the participants were randomly assigned to the 'Absolute-Relative' condition and saw the sets in the following order: *full*, *spotted*, *big*, *long*. The task took approximately 10 minutes.

<div align="center">3.3   Results</div>

The judgments from adults and children for the four scales are captured in Figure 1 and Figure 2, respectively. The seven items are indicated on the x-axis (greatest degree = 1, least degree = 7), and the percentage of acceptance is on the y-axis.

---

[6] Throughout the experiments reported in this paper, we make use of the increasing/polar-positive member of an antonym pair (e.g., *big*, not *small*, *full*, not *empty*), since our focus was on context sensitivity and not on the asymmetry between poles. Certainly a sizable literature has been devoted to investigating this phenomenon in child language (cf. Barner & Snedeker, 2007; Bartlett, 1976; Brewer & Stone, 1975; Carey, 1978; Clark, 1972; Clark, 1973; Eilers, Kimbrough Oller, & Ellington, 1974; Keil & Carroll, 1980; Klatzky, Clark, & Macken, 1973; Marschark, 1977; Townsend, 1976). In addition, we do not address the apparent decrement in children's understanding of the term *big* or the weight of semantic or dimensional features in evaluating the size of an object (cf. Bausano & Jeffrey, 1975; Clark, 1973; Coley & Gelman, 1989; Gathercole, 1982; Harris & Folch, 1985; Harris, Morris, & Meerum Terwogt, 1986; Lumsden & Poteat, 1968; Maratsos, 1973, 1974; Ravn & Gelman, 1984; Sena & Smith, 1990).

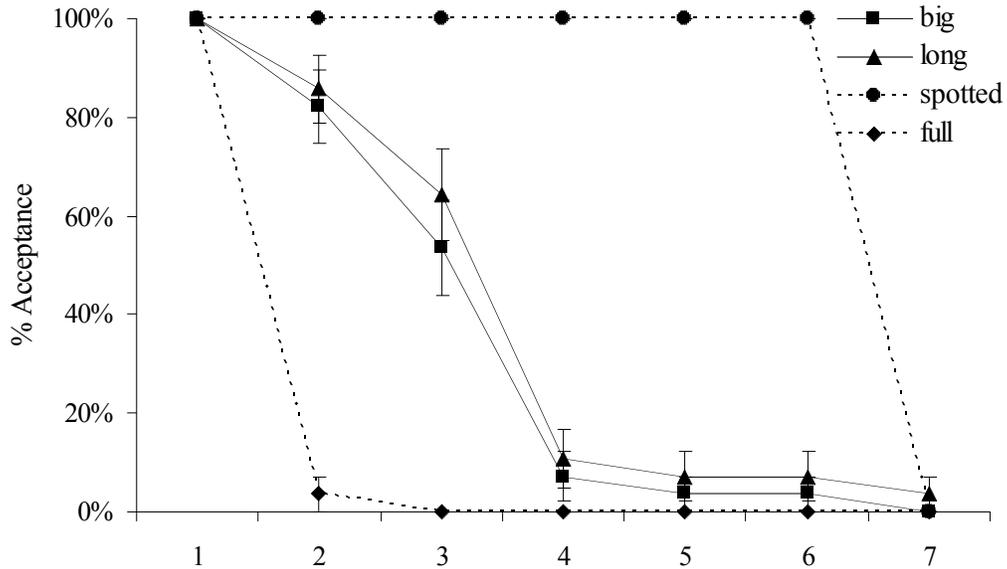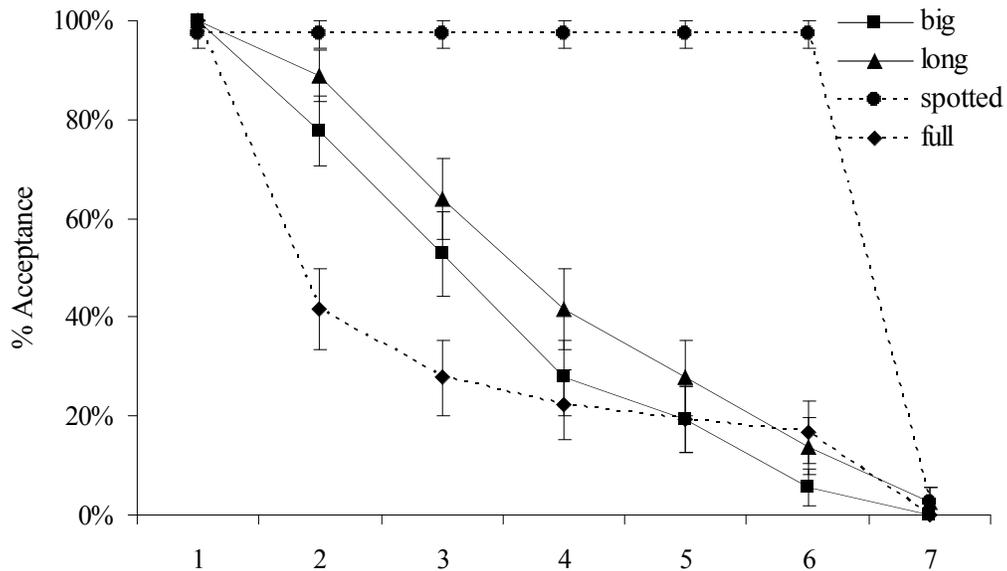Figure 1: Adults' scalar judgments

Figure 2: Children's scalar judgments

For both age groups, there was a decrease in the acceptance for both relative GAs around

the midpoint of the series. Judgments for *spotted* were clear-cut for both children and adults: both

children and adults accepted as *spotted* any disk with any number of spots on it. Judgments for

the maximum standard absolute GA *full* were noticeably different from the minimum standard

absolute GA *spotted*: adults only accepted item #1 as *full*; children's judgments, however, were

less sharp. In order to assess whether children were adult-like in their judgments of *full*, we conducted a comparison of *full* to the relative GAs *big* and *long*.

If children were adult-like in their judgments, they should have demonstrated a tendency to allow the second and third items to be *big* or *long* but not allow them to be *full*. We therefore targeted the second and third item in the series for statistical analysis. Even though the children were not at floor with *full*, as adults were, for both the second and third items, the difference between the *big* and *full* judgments and between the *long* and *full* judgments is significant in one-tailed t tests (item #2, *big/full*: t(35) = -3.654, p = 0.0004; *long/full*: t(35) = -5.596, p = 0.000001; item #3, *big/full*: t(35) = -2.485, p = 0.009; *long/full*: t(35) = -4.448, p = 0.00004.

There was also an effect of condition ('Relative-Absolute' v. 'Absolute-Relative'). Eleven of the fifteen children (73.3%) who judged a container other than the first one to be full were in the 'Relative-Absolute' condition, and therefore saw the *big* and *long* sets before the *full* and *spotted* sets. These children were scattered across age groups (three 3-year-olds, four 4-year-olds, and four 5-year-olds). Put another way, eleven of the eighteen children (or 61.1%) in the 'Relative-Absolute' condition judged the second container to be full, while only four of the eighteen children (or 22.2%) in the 'Absolute-Relative' condition did. Six children across age groups seemed to interpret *full* minimally (e.g., 'filled to some degree') and consequently judged the first six items in the set to be full. Excluding these children from the condition analysis, we observe the same asymmetry: seven children in the 'Relative-Absolute' condition, and only two in the 'Absolute-Relative' condition judged the second item to be full.

3.4   Discussion

There are three main results of the Scalar Judgment Task. First, we have evidence that both children and adults make a distinction between relative and absolute GAs, and within the

latter class, between minimum and maximum standard absolute GAs. Relative GAs such as *big* and *long* have standards that evoke a partitioning of a finite set of objects ranked along an appropriate scalar continuum somewhere around the midpoint of the continuum, while absolute GAs use standards that partition the sets at the ends of the continuum. The fact that the standard for absolute GAs such as *spotted* is minimal while the standard for absolute GAs such as *full* is maximal is reflected in judgments that separate the lowest-ranked item from the rest of the series for *spotted* and the highest-ranked item from the rest of the series for *full*.

Second, judgments of relative size along a scale appear to become more categorical with development. This trend is evident in the comparison of the percentages of the third and fourth items of the relative GA sets for the two age groups. The judgments are therefore similar in form to those reported by Smith, Cooney, and McCord (1986), who found that adults had broader categories with steeper slopes than children did for objects described as *big* or *long*. Finally, most children, excluding a small subset who interpreted *full* as 'filled to some degree', share with adults the meaning of *full* as 'maximally filled'. This finding was supported by a comparison of judgments for the *full* set with the two relative GA sets. However, this interpretation of *full* can be influenced by prior context, as seen in the ordering effect between conditions.

4. Experiment 1

4.1   Experiment 1a

4.1.1   Introduction

The goal of Experiment 1 was to use definite descriptions of the sort discussed in sections 1.2 and 1.3 to probe children's sensitivity to the kinds of context/meaning interactions discussed in the introduction to this paper. Specifically, we were interested in determining whether children would correctly shift the standard of comparison for relative GAs to accommodate the existence

and uniqueness presuppositions of the definite description, and would avoid doing so for absolute GAs.

4.1.2   Method

4.1.2.1 Participants

30 children representing three age groups participated in this task: 10 three-year-olds (5 boys 5 girls, range: 3;5 to 3;11, M: 3;8); 10 four-year-olds (4 boys 6 girls, range: 4;1 to 4;11, M: 4;5); and 10 five-year-olds (3 boys 7 girls, range: 5;1 to 5;8, M: 5;5). In addition, 24 adult native speakers of English (Northwestern undergraduates fulfilling an experimental requirement for a Linguistics course) served as controls.

4.1.2.2 Materials

The materials for Experiment 1a and 1b were the same. They consisted of a series of pairs of objects, each sharing a salient dimension (e.g., color, shape, length, etc.). The experiment was divided into a training session and test session. Each session included target stimuli, corresponding to target relative and absolute GAs, and control stimuli. The complete set of materials is outlined in the following three tables. Table 2 presents the four pairs used in the training session. Table 3 presents the target stimuli used in the test session, while Table 4 presents the control stimuli used in the test session. The column on the far right in each table corresponds to an experimental feature discussed in the Procedure section.

Table 2: Training stimuli for Experiment 1

adjective	stimuli	pragmatic status of request
happy	pictures of a happy face and an angry face	felicitous
round	pictures of a green triangle and a blue square	infelicitous
red	pictures of a red circle and a red square	infelicitous
blue	pictures of a yellow bird and a blue bird	felicitous

Table 3: Target stimuli for Experiment 1

adjective	stimuli	pragmatic status of
relative		
big	two blocks, one bigger than the other (blocks #1 and 3 from SJT)	felicitous*
big	two blocks, one bigger than the other, both smaller than the first pair (blocks #5 and 7 from SJT)	felicitous*
long	two rods, one longer than the other (rods #1 and 3 from SJT)	felicitous*
long	two rods, one longer than the other, both shorter than the first pair (rods #5 and 7 from SJT)	felicitous*
absolute		
spotted	two disks, one with a few spots, one without any (disks #5 and 7 from SJT)	felicitous
spotted	two disks, one more spotted than the other (disks #1 and 4 from SJT)	infelicitous
full	full container and container filled more than halfway (containers #1 and 3 from SJT)	felicitous
full	two containers filled somewhat, one more than the other, but neither full (containers #4 and 6 from SJT)	infelicitous

Table 4: Control stimuli for Experiment 1

adjective	stimuli	pragmatic status of request
color		
yellow	pictures of a yellow bird and a black bird	infelicitous
green	purple yo-yo and yellow yo-yo	infelicitous
red	pictures of a red square and a red circle	infelicitous
red	red poker chip and a white poker chip	felicitous
shape		
round	pictures of a red triangle and a red square	infelicitous
square	pictures of a blue square and a yellow circle	felicitous
mood		
happy	pictures of a sad face and an angry face	infelicitous
happy	pictures of an angry face and a sad face	infelicitous
sad	pictures of a happy face and a sad face	felicitous

4.1.2.3 Procedure

Participants were invited to play a game. Children were introduced to a puppet (played by a second experimenter) and were told that that the purpose of the game was to help the puppet learn how to ask for things. They were then told that they would be shown two objects at a time, and that every time they saw two objects, the puppet would ask for something. Their job was to determine if they could give the puppet what he asked for based on his request, and if they couldn't, to tell him why not. Even the youngest participants followed these directions easily. Adult participants interacted with one adult experimenter instead of the puppet.

There was an important twist in this task that distinguished it from previous forced-choice studies in which a child was asked to act on a request with an adjective such as *big* (e.g., Bartlett, 1976; Ebeling & Gelman, 1988; Eilers, Kimbrough Oller, & Ellington, 1974; Gelman & Ebeling, 1989; Gelman & Markman, 1985; Harris, Morris, & Terwogt 1986; Ravn & Gelman, 1984; Sena & Smith, 1990). In this task, the request was not always felicitous. This pragmatic manipulation was accomplished in the following way. Each request included a singular definite noun phrase of the form *the A one*, where *A* was an adjective (e.g., *Please give me the red one*). As noted in section 1.2, this type of description presupposes both existence (e.g., there must *be* a red one) and uniqueness (e.g., there must be *only one* red one). In determining whether or not they could give the puppet what he asked for, children were in essence, assessing the context with respect to the presuppositions of the definite description, and in some cases accommodating them. For this reason, we refer to this task as the *Presupposition Assessment Task (PAT)*.

Pairs of objects were designed so that they either satisfied or violated one or both of the presuppositions of the definite description. (See the far right column of the three previous tables.) For some pairs, the request (e.g., *Give me the red one*) was felicitous, because exactly one object fit the description (e.g., there was a red object and a white object). For other pairs, the request was infelicitous, either because both members of the pair fit the description (e.g., there were two red objects) or because neither member of the pair fit the description (e.g., there was a yellow object and a blue object).

For our crucial pairs – those involving relative GAs and marked as 'felicitous*' in Table 3 – whether the request was felicitous or not depended on how the adjectival modifier in the description was interpreted. If it was interpreted in the same way as in the Scalar Judgment Task (e.g., if the judgments about the size or length of items #1, 3, 5, and 7 remained constant across

tasks), then either an existence or uniqueness violation would be incurred, depending on whether

the two objects were from the lower part of the continuum or the upper part, respectively. If,

however, a new standard of comparison (distinct from the one used in the Scalar Judgment Task)

was posited in order to ensure that the adjective would be true of just one object (the bigger or

longer one; see the discussion of this point in section 1.2), the request would be felicitous. If

participants accepted requests involving relative GAs, then, this would constitute evidence both

that they know that such GAs are context sensitive in the way outlined in section 1.2, and (given

the responses in the Scalar Judgment Task) that they are sensitive to, and willing to

accommodate, the presuppositions of a definite NP.

Pairs involving absolute GAs provided crucial controls and substantiation of this

conclusion. Since absolute GAs in the positive form have fixed (maximum or minimum)

standards of comparison and are not context dependent, they should not allow the same

flexibility of use as relative GAs. For example, if *spotted* simply means 'has some number of

spots', participants should reject requests for *the spotted one* when confronted with two spotted

objects, even if one is clearly more spotted than the other, because this would involve a violation

of the uniqueness requirement of the definite description (since both objects are spotted).

Likewise, if *full* means 'is maximally full', participants should reject requests for *the full one*

when confronted with two partially filled containers, because the request violates the existence

presupposition of the definite description (since there is no full container). If on the other hand

spotted and *full* were context dependent in the same way as *big* and *long*, participants ought to be

able to shift their meanings in a way that accommodates the presuppositions of the definite

description in these examples, as with the relative GAs. Rejection of presupposition-violating

requests involving absolute GA pairs, coupled with acceptance of requests involving relative GA

pairs, therefore constitutes evidence for the relative/absolute distinction as laid out here, and shows that participants are, in fact, modulating the interpretation of relative GAs in accord with the presuppositions of the definite NP.

At the same time, this pattern of results would important evidence that participants are not treating the adjective in the definite description as semantically equivalent to the comparative form (i.e., treating *Please give me the A one* as *Please give me the more A one*). If participants were reinterpreting the request in this way, they would always accept it, regardless of adjective type, since the comparative form of any GA can be used to uniquely pick out that member of a pair that has the greater degree of the relevant property: *the more spotted one* can be felicitously used to pick out the more spotted of two disks, even though *the spotted one* cannot. Although it is unlikely that adults would reinterpret *the A one* in this way, it is a possibility that must be seriously considered for children, as their interpretation and use of comparative morphology at this young age is not fully adult-like (cf. Donaldson & Wales, 1970; Ehri, 1976; Finch-Williams, 1981; Gathercole, 1979; Gitterman & Johnston, 1983; Graziano-King, 1999; Graziano-King & Cairns, 2005; Layton & Stick, 1978; Moore, 1999). If participants reject presupposition-violating requests involving absolute GAs, however, we can be confident that they are not reanalyzing the adjectives in the test items as comparatives.

To ensure that participants understood the rules of the game and to help them feel comfortable rejecting the request, we had them participate in a brief training session before the test session began. This training session was composed of two felicitous and two infelicitous requests. (See Table 2.) Once it was evident that participants felt comfortable with the task, we proceeded with the test session. If children were still hesitant to correct the puppet after the four training items, we introduced a fifth impromptu pair accompanied by an infelicitous request.

Items in the test session included target items whose salient property corresponded to one of the target GAs, and control items. The presentation of the pair members was counterbalanced so that the object fitting the description appeared on different sides of the pairs throughout the test session. The order of the pairs was also pseudorandomized with respect to three factors: the felicity of the request, the nature of the presupposition violation, and the adjective. Participants were randomly assigned to one of two orders, a point that becomes important in the interpretation of the results. Specific details about the request, stimuli, and anticipated response for the target GA items are outlined in

Table 5.

Table 5: GA, adjective status, pragmatic status of request, stimuli, and anticipated response for Experiment 1

GA in request	GA status	pragmatic status of request	stimuli	anticipated response
big	relative	felicitous*	two blocks, one bigger than the other	give the object that exceeds the cutoff for the context (e.g., the bigger one)
long	relative	felicitous*	two rods, one longer than the other	give the object that exceeds the cutoff for the context (e.g., the longer one)

full	absolute, maximal	felicitous	two containers, one full, the other filled to some degree	give the object that manifests the property (e.g., the full one)
full	absolute, maximal	infelicitous	two containers, neither full, but one fuller than the other	give neither object
spotted	absolute, minimal	felicitous	two disks, only one spotted	give the object that manifests the property (e.g., the spotted one)
spotted	absolute, minimal	infelicitous	two disks, both spotted, but one more than the other	give neither object

4.1.3   Results

For each pair, we anticipated one of three possible responses to the accompanying request: the participant could accept the request by giving one object, reject the request and give neither object, or reject the request and give both objects. We therefore chose to capture the results in terms of percentage of anticipated responses. This method is in contrast to capturing the results in terms of the percentage of responses in which the participant gave only one object (since such a percentage would not indicate which object was given) or the percentage of responses in which the participant only gave the object with the greater degree of the property (since in this case and the former, this would not indicate what happened when the participant did not give one object, but gave both or neither). The results are presented in Table 6.

Table 6: Percentage of anticipated responses in Experiment 1a

	control items			test items				
age	color	shape	mood	*big*	*long*	*spotted*	*full*	(nf/nf, f/nf)
3 yrs	93	95	93	95	90	90	70	(40, 100)
4 yrs	98	100	100	95	95	85	65	(30, 100)
5 yrs	100	100	100	100	100	95	65	(30, 100)
adults	100	100	100	98	98	98	92	(88, 96)

These results show that both adults and children accepted and rejected the requests accordingly: both age groups were at ceiling with the control items and responded as anticipated with the test items. Recall that the pairs for the target stimuli also appeared as part of the series in the SJT. Children consistently judged items #1 and #3 in the relative sets (the bigger/longer pair in this task) as *big* or lo*n*g, and they rejected this label for items #5 and #7 in the same series (the smaller/shorter pair in this task). In spite of these judgments, participants in the PAT did distinguish between the members of these pairs, treating *big/long* as true #1 and false of #3 when confronted with this pair, and true of #5 and false of #7 when confronted with this pair. We take this as an indication that children were willing and able to shift the standard of comparison for the relative GAs. That they did not do so for the other items indicates that context-dependence is part of the lexical entry for these adjectives only and not the others.

The one instance in which children's responses did not align with those of the adults was with the items corresponding to the maximum standard absolute GA *full*. Recall that there was a pair in which one of the containers was full and the other filled only to some degree. (We will call this pair the 'full/non-full' pair, abbreviated in the table above as 'f/nf'). For another pair,

there were two containers, neither of which was full, but one was fuller than the other. (We will call this pair the 'non-full/non-full' pair, abbreviated in the table above as 'nf/nf'). Children's anticipated responses to the 'non-full/non-full' pair are significantly lower than to the other target pairs (two-tailed t-tests t(29), 'non-full/non-full' v. *big* #1/3: t = -7.62, p < 0.00001; *big* #5/7: t = -5.29, p = 0.00001; *long* #1/3: t = -5.46, p < 0.00001; *long* #5/7: t = -7.62, p < 0.00001; *spotted* #1/4: t = -4.07, p = 0.0003; *spotted* #5/7: t = -7.62, p < 0.00001; 'full/non-full': t = -7.62, p < 0.00001). When presented with the 'full/non-full' pair, children consistently gave the puppet the full container; however, when presented with the 'non-full/non-full' pair, many children did not reject the request. Instead, they gave the puppet the fuller of the two containers.

At first glance, this type of response could be taken as indication that children are, in fact, interpreting the adjective in *the A one* as a comparative, which would call into question our conclusions from the data involving relative GAs. However, children's responses to the spotted objects argue against this conclusion. Given two spotted objects with unequal numbers of spots (e.g., items #1 and 4), children consistently rejected the puppet's request for the spotted one, commenting that both objects were spotted, and did not give the puppet the more spotted member of the pair. The sharp difference between *full* and *spotted* suggests that whatever is going on here has to do with the way children understand *full*, rather than a general strategy for reanalyzing unmarked adjectives as comparatives in contexts that would otherwise result in presupposition failure.

One way in which children's responses to the 'non-full/non-full' pair are curious is that the Scalar Judgment Task discussed earlier clearly showed that children share with adults a maximum standard absolute interpretation of the adjective *full*. In particular, 18 children who participated in the Presupposition Assessment Task also participated in the SJT within an

interval of approximately three weeks. Judgments of the *full* SJT set from these children are

presented in Table 7.

Table 7: Judgments of *full* from children participating in the SJT and the PAT

		age		
response in SJT	children giving each response	3 yrs	4 yrs	5 yrs
only #1 is full	12 (66.7%)	5	4	3
#1 and 2 are full	4 (22.2%)		2	2
#1, 2 and 3 are full	2 (11.1%)	1	1	

Recall that the 'non-full/non-full' pair was composed of items #4 and 6 from the set of seven

items. *None* of these eighteen children judged item #4 to be full in the SJT; however, in the PAT,

eleven of these eighteen children (61.1%) gave this item to the puppet to satisfy his request for

the full one. (See Table 8.)

Table 8: Children in the SJT and the PAT who gave the puppet item #4

		age		
response in SJT	children giving #4 as *the full one*	3 yrs	4 yrs	5 yrs
only #1 is full	8 (72.7%)	3	3	2
#1 and 2 are full	2 (18.2%)		1	1
#1, 2 and 3 are full	1 (9.1%)	1		

Given this pattern, we decided to probe children's responses to the 'non-full/non-full'

pair further, by looking to see whether the ordering condition to which participants were

assigned had an effect. Indeed, upon further examination, we found that eight of the 11 children

(72.7%) who gave the puppet the fuller container while having judged it to be *not full* in the SJT

were in the condition in which this 'non-full/non-full' pair appeared early in the sequence. Moreover, every single one of the 15 children from Experiment 1a who were in this ordering condition gave the puppet the fuller of the two 'non-full' containers in response to his request for the full one. Only five of the 15 children in the condition in which this pair appeared later in the sequence (after the 'full/non-full' pair) responded to the puppet's request in this way (a difference between 100% and 33.3% between the two conditions).

This pattern leaves open two possibilities. The first is that children benefit from early exposure to an object that exemplifies the maximal degree of a property named by a maximum standard absolute GA in a task like this (e.g., seeing a full container). The second is that early exposure to relative adjectives in the test sequence creates a kind of priming effect, causing children to incorrectly treat *full* as relative rather than as maximum standard absolute. Whenever the 'non-full/non-full' pair appeared before the 'full/non-full' pair, it was also immediately preceded by a *long* pair; it is possible that in shifting the standard of comparison with the *long* pair to accommodate the presuppositions of the definite, children were then influenced to respond in a similar way to the 'non-full/non-full' pair that followed. This possibility is explored in Experiment 1b.

4.1.4   Discussion

The results of Experiment 1a demonstrate that both adults and children distinguish between three subclasses of gradable adjectives. They shift the standard of comparison for relative GAs, but do not for absolute GAs. Relative GAs such as *big* and *long* have as part of their semantic representation a contextually-determined standard of comparison, and so allow for the standard to shift from context to context. Accordingly, when presented with requests for *the big/long one*, participants willingly gave the bigger or longer member of the pair. In contrast,

when presented with two spotted objects, participants did not give the more spotted object; this is to be expected, since *spotted* is a minimum standard absolute GA with a meaning along the lines of 'have spots'. Adult participants similarly rejected requests for *the full one* when given two 'non-full' containers, a response that is anticipated, since *full* is a maximum standard absolute GA. However, some children gave the puppet the fuller container in this situation. This pattern contrasted with our findings from the SJT, in which we found that children did not judge the same container to be full. One possible explanation for this pattern of results is that a relative pair appearing earlier in the sequence affected children's responses to the 'non-full/non-full' pair. This possibility is explored in Experiment 1b.

4.2   Experiment 1b

4.2.1   Introduction

The goal of Experiment 1b was to determine the source of the ordering effect observed in Experiment 1a. Specifically, we wanted to know if the prior presentation of a relative GA pair in the sequence of items influenced children to erroneously treat *full* as context-dependent, causing them to give the puppet the fuller of the two 'non-full' containers in response to his request for *the full one*.

4.2.2   Method

4.2.2.1 Participants

17 children representing three age groups also participated in this task: 6 three-year-olds (3 boys 3 girls, range: 3;1 to 3;11, M: 3;5); 6 four-year-olds (2 boys 4 girls, range: 4;2 to 4;11, M: 4;6); and 5 five-year-olds (1 boy 4 girls, range: 5;2 to 5;10, M: 5;4). In addition, 10 adult native speakers of English (Northwestern undergraduates fulfilling an experimental requirement for a Linguistics course) served as controls.

4.2.2.2 Materials

The same objects from Experiment 1a were used. The only difference was in the sequence of items. In Experiment 1a, the 'non-full/non-full' pair was almost immediately preceded by a *long* pair, with only one control pair intervening. To evaluate the influence of the relative GA pair, we made a minor change in the order of presentation, simply switching the order of the *long* and 'non-full/non-full' pair so that the latter was no longer preceded by the former.

4.2.2.3 Procedure

The procedure was the same as in Experiment 1a.

4.2.3   Results

The results for Experiment 1b are presented in Table 9. The pattern of responses is similar to the one observed in Experiment 1a[7]: children and adults responded as anticipated to the training, control, and test items, with the exception of the 'non-full/non-full' pair.

Table 9: Percentage of anticipated responses in Experiment 1b

age	control items			test items				
---	color	shape	mood	*big*	*long*	*spotted*	*full*	(nf/nf, f/nf)
3 yrs	83	100	83	100	83	67	50	(0, 100)
4 yrs	96	100	100	92	83	92	58	(17, 100)
5 yrs	95	100	100	100	90	90	70	(40, 100)
adults	100	100	100	100	100	95	85	(70, 100)

Recall that the purpose of conducting Experiment 1b was to identify the source of children's non-adult-like responses to the 'non-full/non-full' pair. A direct comparison between participants' responses to this pair in both conditions of Experiment 1 is informative. As we

---

[7] The lower rate of anticipated responses to the *spotted* stimuli by the three-year-olds may be explained by the fact that in Experiment 1b, the experimenter encouraged to give the puppet verbal feedback, but, unlike in Experiment 1a, did not encourage the child to give the puppet both objects when both fit the description.

observe in Table 10, both children and adults were at their best when this pair appeared later in the sequence.[8]

Table 10: Percentage of anticipated responses to 'non-full/non-full' pair

age	pair early, after relative pair (Experiment 1a)	pair early, before relative pair (Experiment 1b)	pair later, after f/nf pair (Experiments 1a, 1b)
children	0	18	67
adults	75	70	100

It appears that the prior presentation of a maximum standard (i.e., seeing the 'full/non-full' pair earlier in the sequence) allowed participants to be more successful in their judgments of the 'non-full/non-full' pair. Children were not misled by the prior presentation of a relative GA pair to treat the fuller container as *the full one*. These conclusions are supported by the fact that while the difference between either of the two 'early' orders and the 'later' order is significant, the difference between the two 'early' orders is not (Fisher's Exact Tests, two-tailed probabilities: 'early 1a' v. 'later': $p < 0.001$; 'early 1b' v. 'later': $p = 0.01$; 'early 1a' v. 'early 1b': $p = 0.24$). However, knowing that children benefited from exposure to the maximum standard does not fully explain the results from Experiment 1. Even when given this boost, children's responses to the 'non-full/non-full' pair still deviated from their adult-like responses to the other pairs. In order to explore this difference further we therefore turned to an additional analysis, in which we evaluated children's reaction times (RTs) during the experimental session.

---

[8] Adults who gave the fuller of the two 'non-full' containers noted at the end of the experimental session without any prompting that they realized this mistake later in the experiment and wished to make clear to the experimenter that they knew what *full* means.

The purpose of analyzing the RTs for children's experimental sessions was to determine if children who gave the puppet the fuller of the two 'non-full' containers in response to his request for *the full one* took longer to do so for this pair than with other pairs. We reasoned that children might have evaluated the puppet's request given the experimental stimuli at hand and subsequently decided to treat it as being an instance of IMPRECISION: informative use of a false description to identify the object that comes closest – and close enough – to satisfying it. That is, although children might not have actually judged the fuller of the two 'non-full' containers to be *full* (maintaining an absolute maximum semantics for the adjective along the lines of 'be maximally full'), they still might have decided after some deliberation that they should tolerate some 'loose talk' on the part of the puppet and hand over the container that came closest to satisfying the description that he used in his request. Crucially, this kind of reasoning is not necessary in the case of relative GAs (such as *big* and *long*), since the presuppositions of the definite description can be accommodated, and the request therefore accepted, based entirely on the (context sensitive) semantics of the adjective. This kind of reasoning is not applicable in the case of minimum standard GAs (e.g., *spotted*), which apply to any object that has the relevant property to some degree, and so are inherently imprecise. It follows that if this reasoning is brought into play just for the purpose of deciding how to respond to a pair of objects (neither of which satisfies the semantic requirements of a maximum standard absolute GA), we should see an effect on RTs just in such cases.

Experimental sessions with child participants were videotaped using a Sony Digital8 Handycam. Videotapes were imported from the camera into Apple Inc's iMovie program as .mov files. Videos were then coded offline by research assistants in our laboratory on a Macintosh computer using SuperCoder software (Hollich, 2003). For each item in which the

child accepted the puppet's request (i.e., gave the puppet one of the two objects), the research assistants coded three measurements: the child's look to the object, the child's reach toward the object, and the child's touch of the object. These RT measurements were then coded frame by frame, where one frame is equal to 1/30 of a second.

It is necessary to provide some additional details about the coding session. First, we excluded from analysis any items in which the children's eye movements could not be coded (e.g., if the eyes were occluded), the child was already looking at or touching the stimuli before the request was uttered, or any other experimental artifact prevented the coders from obtaining measurements (i.e., there was a distraction in the background). For this reason, the total number of children whose RTs were analyzed varies from analysis to analysis. This number is always provided in a footnote. Second, rather than coding the initial look to the object, since the child could have decided to inspect the second object before deciding to give the puppet the first object, we coded the look that immediately preceded the reach to the object. A reach was a movement that ultimately resulted in touching an object. We chose to target these measurements instead of, e.g., proportion of looking time, since we wished to measure latency of response across key items. Finally, at least two coders were assigned to each experimental session, with one coder arbitrarily chosen as the default. In case there was a disagreement of more than 5 frames for any of the three measurements, a third coder was brought in as a tiebreaker for that item. There was generally very high agreement among coders.

We then targeted two sets of RT measurements for analysis. In the first, we asked if children took longer to respond to the puppet's request when it accompanied the 'non-full/non-

full' pair than when it accompanied the 'full/non-full' pair. These results are presented in Figure 3.[9]

Figure 3: Reaction times for two *full* pairs in Experiment 1a
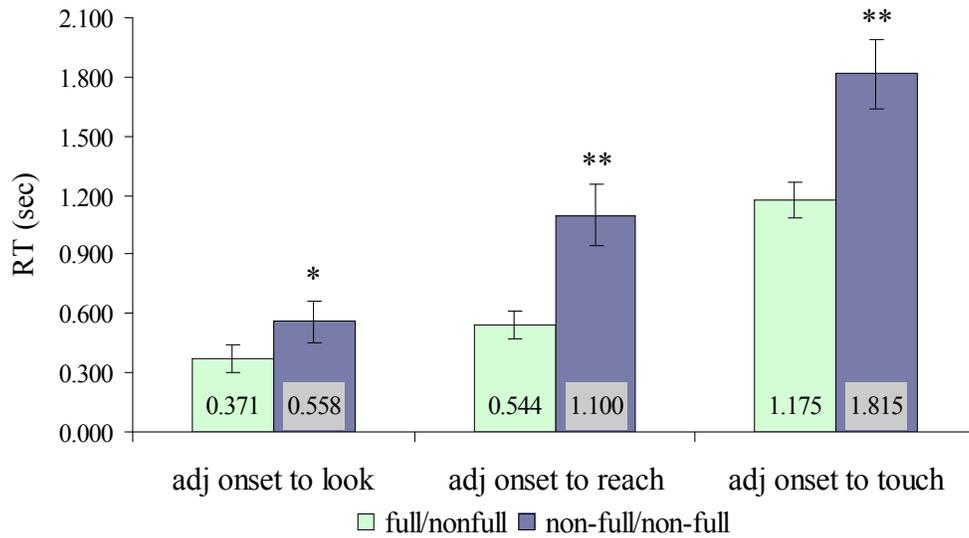
Indeed, differences between the look, reach, and touch are significant for these two pairs (one-tailed t tests: look $t(15) = 1.71$, $p = 0.05$; reach $t(15) = 3.03$, $p = 0.004$; touch $t(15) = 3.47$, $p < 0.002$).

In the second RT measurement, we compared the RTs for the 'non-full/non-full' pair to those for other key pairs in order to determine if children generally took extra time to shift the standard of comparison, or if the extra time observed with the 'non-full/non-full' pair was unique to that pair. We targeted the difference between the look and the touch of the object for four key

---

[9] We analyzed the RTs for sixteen children who gave the puppet the fuller container for each pair.

pairs, the two *big* pairs (two big blocks and two small blocks) and the two *full* pairs. These

results are presented in Figure 4.[10]

Figure 4: Reaction times for four key pairs in Experiment 1a
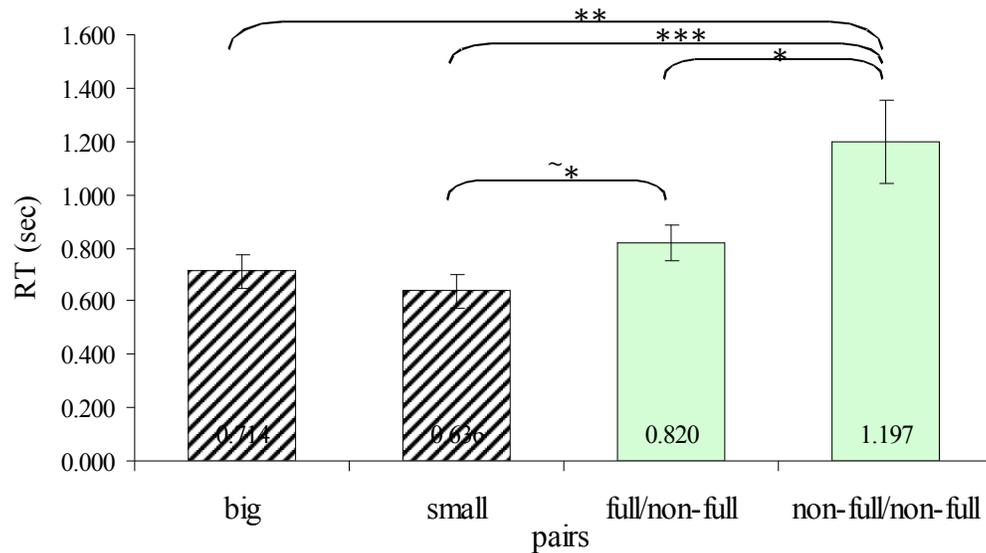
The RTs for the 'non-full/non-full' pair clearly stand out from the three others. Indeed,

while the RT for this pair is significantly longer than every other pair (two-tailed t-tests: 'non-

full/non-full' v. 'full/non-full': $t(48) = 2.42$, $p = 0.02$; 'non-full/non-full' v. 'big': $t(45) = 3.07$, p

$= 0.004$; 'non-full/non-full' v. 'small': $t(51) = 3.79$, $p = 0.0004$), the other pairs do not differ

significantly from each other (two-tailed t-tests: 'full/non-full' v. *big*: $t(53) = 1.11$, $p = 0.27$;

'big' v. 'small': $t(56) = 0.87$, $p = 0.39$; 'full/non-full' v. 'small': $t(59) = 1.96$, $p = 0.05$,

marginally significant). These results demonstrate that the shift in the standard of comparison for

the *big* pairs – even for the pair in which the small blocks were judged to be *not big* in the SJT –

---

[10] We analyzed all codable responses for these items across all children. The number of children

varied for each item: 26 for the big blocks, 32 for the small blocks, 29 for the 'full/non-full' pair,

and 21 for the 'non-full/non-full' pair.

was automatic.[11] This is expected: given that a contextually-determined standard is part of the

semantic representation for *big*, no additional time should be required to compute the standard of

comparison.

4.2.4   Discussion

The results of Experiment 1b elaborate upon those of Experiment 1a by demonstrating

that the source of children's non-adult-like performance with the 'non-full/non-full' pair (i.e.,

giving the puppet the fuller container in response to his request for the full one) is not driven by

the influence of a preceding relative GA pair. Children are significantly more likely to pattern

like adults with the 'non-full/non-full' pair if it occurs later in the sequence, some time after they

have seen the maximum standard exemplified, but even with this assistance, some children are

still inclined to allow the fuller container to count as the full one. The RT analysis shed light on

this pattern of responses by demonstrating that children take significantly longer to give the

puppet the fuller of the two 'non-full' containers, suggesting that additional reasoning beyond

that involved in assigning context (in)dependent meanings to GAs is, in fact, being brought to

bear on these examples. We will provide a more detailed analysis of the significance of the RT

data in the General Discussion section.

---

[11] One might have expected a semantic congruity effect with the small pair of blocks (i.e., longer

RTs resulting from the fact that the adjective in the request was *big*, but the blocks were judged

to be *not big*). The lack of such an effect here may be due to the fact that the difference in

magnitude between the two pairs was not extraordinarily large.

5.Experiment 2

5.1   Experiment 2a

5.1.1   Introduction

Given children's responses to the *full* pairs in Experiment 1, one question immediately surfaces: are these responses unique to *full*, or can they be generalized to a larger set of maximum standard absolute GAs? The goal of Experiment 2 was to address this question.

5.1.2   Method

5.1.2.1 Participants

30 children representing three age groups participated in this task: 10 three-year-olds (5 boys 5 girls, range: 3;2 to 3;11, M: 3;6); 10 four-year-olds (5 boys 5 girls, range: 4;1 to 4;10, M: 4;4); and 10 five-year-olds (4 boys 6 girls, range: 5;0 to 5;7, M: 5;3). In addition, 24 adult native speakers of English (Northwestern undergraduates fulfilling an experimental requirement for a Linguistics course) served as controls.

5.1.2.2 Materials

The materials were the same as in Experiments 1, with the exception of four target pairs. In place of the two *full* pairs, there were two pairs corresponding to another maximum standard absolute GA, *straight*. In place of the two *spotted* pairs, there were two pairs corresponding to another minimum standard absolute GA, *bumpy*. These pairs were designed similarly to those in Experiment 1, so that only one of the two pairs for each adjective would satisfy the presuppositions of the definite description. Details for these stimuli are presented in Table 11.

Table 11: Stimuli substituted for the absolute GA pairs from Experiment 1

adjective	pragmatic status of request	stimuli
straight (in place of *full*)	felicitous	2 wire rods, 9″ in length, one completely straight and the other bent into a "C"
	infelicitous	2 wire rods, 9″ in length, one curly and the other even curlier
bumpy (in place of *spotted*)	felicitous	two flat wooden boards, measuring 2″ x 5″ and painted orange, one with a few bumps and one with no bumps
	infelicitous	two flat wooden boards, measuring 2″ x 5″ and painted orange, one with many bumps and one with less bumps

5.1.2.3 Procedure

The procedure was the same as in Experiment 1.

5.1.3   Results

As in Experiment 1, children and adults patterned as anticipated for all items, with the exception of the maximum standard absolute GA stimuli. The same difference observed in

Experiment 1 is found between the two *straight* pairs (abbreviated in the table as 'b/b' for 'bent/bent' and 's/b' for 'straight/bent'). The percentages of anticipated results are presented in Table 12.

Table 12: Percentage of anticipated responses in Experiment 2a

age	control items			test items			
	color	shape	mood	*big*	*long*	*bumpy*	*straight* (b/b, s/b)
3 yrs	83	90	93	95	95	85	55 (70, 100)
4 yrs	95	100	97	100	95	90	85 (80, 100)
5 yrs	95	100	97	100	95	90	95 (80, 100)
adults	100	100	100	100	100	100	94 (100, 100)

The same ordering effect observed in Experiment 1 was also found. Of the seven children across age groups who gave the straighter rod for the 'bent/bent' pair, six (or 85.7%) of them saw the this pair early in the sequence, well before they were exposed to the 'straight/bent' pair. Now, while the percentage of anticipated responses to the 'bent/bent' pair was in the same direction as the 'non-full/non-full' pair, the percentage of anticipated responses to this pair was not nearly as low as in Experiment 1. In Experiment 1a, only 33% of the children patterned as anticipated with the 'non-full/non-full' pair and rejected the request, and in Experiment 1b, only 19% of the children did so. By contrast, in this experiment, 77% of the children rejected the request. However, participants' responses to this pair are still significantly lower than to most (but not all) of the other test pairs (two-tailed t-tests, t(29): 'bent/bent' v. *big* #1/3: t = -2.97, p = 0.006; *big* #5/7: t = -2.26, p = 0.03; *long* #1/3: t = -2.97, p = 0.006; *bumpy* #5/7: t = -2.97, p =

0.006; 'straight/bent': t = -2.97, p = 0.006; but v. *bumpy* #1/4: t = 1.99, p < 0.06, marginally

significant; and v. *long* #5/7: t = -1.28, p = 0.211, n.s.).

5.1.4   Discussion

The results of Experiment 2a support a generalization of the results from Experiment 1b

to a larger set of absolute GAs, since child and adult responses to *straight* and *bumpy* were

similar to their responses for *full* and *spotted*, respectively. Both age groups rejected the

infelicitous requests for the items corresponding to these GAs and did not shift standards as they

did for the relative GA pairs in Experiments 1 and 2. We observed, though, that children were

more likely to reject the infelicitous request for the 'bent/bent' pair than they were for the 'non-

full/non-full' pair. However, it is possible that the higher percentage of anticipated responses for

the 'bent/bent' pair was an effect of the stimulus design. We observed *post facto* that the objects

for *straight* shown in Table 11 seem to correspond better to stimuli for a minimum standard

absolute GA such as *bent* or *curly* than for the maximum standard absolute GA *straight*: while

item #1 was completely straight, none of the other wires had any degree of straightness. This

design differed from the *full* materials, where it was the case that while only container #1 was

full, the others in the PAT manifested some degree of being filled. This observation led us to

conduct Experiment 2b using a minimally revised set of stimuli.

5.2   Experiment 2b

5.2.1   Introduction

The goal of this experiment was to introduce a minimally revised set of *straight* materials

that more closely resembled the design of the *full* materials in Experiment 1, in order to

determine whether participants' judgments about *the straight one* when shown the 'bent/bent'

pair would resemble those in Experiment 1a and 1b more than those in Experiment 2a did.

5.2.2   Method

5.2.2.1 Participants

30 children representing three age groups participated in this task: 10 three-year-olds (3 boys 7 girls, range: 3;2 to 3;10, M: 3;6); 10 four-year-olds (4 boys 6 girls, range: 4;1 to 4;9, M: 4;7); and 10 five-year-olds (5 boys 5 girls, range: 5;1 to 5;11, M: 5;6). In addition, 24 adult native speakers of English (Northwestern undergraduates fulfilling an experimental requirement for a Linguistics course) served as controls.

5.2.2.2 Materials

The materials were the same as in Experiment 2a, with the exception of the *straight* pairs. In place of the three wire rods that were bent to varying degrees, we inserted rods that resembled more accurately the containers used for *full*. For example, a completely straight rod was paired with a rod that was straight for most of its length but which had a curl at the top, analogous to a container which is filled most, but not all, of the way. The new pairs are presented in Table 13.

Table 13: Stimuli substituted for the *straight* pairs from Experiment 2a

adjective	pragmatic status of request	stimuli
straight	felicitous	a completely straight rod and a mostly-straight rod with a curly section at the top
	infelicitous	a rod with a long straight section and a curly section at the top and a totally curly rod

5.2.2.3 Procedure

The procedure was the same as in the previous experiments.

5.2.3   Results

In Experiment 2b, we observe the same trend as in each of the previous PAT experiments, with children and adults responding as anticipated for all of the training, control, and test items, with the only exception being the 'bent/bent' stimuli. These results are presented in Table 14.

Table 14: Percentage of anticipated responses in Experiment 2b

age	control items			test items				
	color	shape	mood	*big*	*long*	*bumpy*	*straight*	(b/b, s/b)
3 yrs	88	80	87	95	90	60	70	(30, 90)
4 yrs	90	90	97	95	90	75	80	(60, 90)
5 yrs	100	100	100	100	100	80	90	(70, 90)
adults	100	100	100	100	100	94	98	(88, 100)

This time, children's responses to the 'bent/bent' pair are more like those seen in Experiment 1 than those in Experiment 2a were and are significantly lower than to all of the other test pairs (two-tailed t-tests, t(29): 'bent/bent' v. *big* #1/3: t = -5.04, p < 0.0001; *big* #5/7: t = -4.47, p = 0.0001; *long* #1/3: t = -5.39, p < 0.00001; *long* #5/7: t = -3.27, p = 0.003; *bumpy* #1/4: t = -1.88, p < 0.07; *bumpy* #5/7: t = -4.10, p = 0.0003; 'straight/bent': t = -3.53, p = 0.001).

In this experiment, 53% of the children rejected the request for *the straight one* when presented with the 'bent/bent' pair (compared to the 77% with the previous stimuli in Experiment 2a and 33% for the similarly-designed 'non-full/non-full' pair with the same orders of presentation in Experiment 1a). Again, there was also an effect of order of presentation: children were more likely to reject the puppet's request if they saw the 'bent/bent' pair later in the sequence; nine of the 13 children (69%) who gave the puppet the straighter of the two bent rods saw this pair appeared early in the sequence of items. The comparison presented in Figure 5 highlights the consistency of this ordering effect across Experiments 1 and 2.

Figure 5: Rate of acceptance of the non-maximal member in Experiments 1 and 2
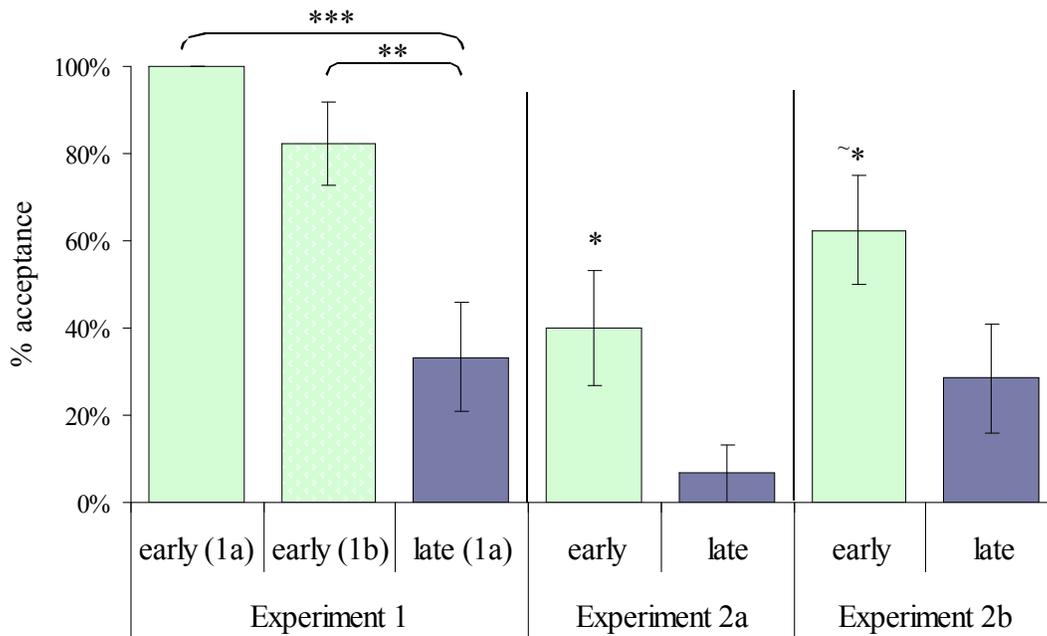
This between-experiment comparison allows us to make the following generalization: when the pair with the object exemplifying the maximum degree appears later in the sequence than the pair with both non-maximal objects, children are significantly more likely to incorrectly accept the puppet's request for the maximal object when shown the earlier pair. A series of single-factor ANOVAs for each experiment supports this conclusion (Experiment 1: $F(2, 44) = 3.21$, $p < 0.0001$; Experiment 2a: $F(1, 28) = 4.20$, $p = 0.03$; Experiment 2b: $F(1, 28) = 4.20$, $p = 0.07$, marginally significant). (See also the statistical analysis following the data presented in Table 10 for differences between the two early orders and the late order in Experiment 1.)

Recall that in Experiment 1, we conducted an analysis of children's RTs as they responded to the puppet's request. This analysis provided us with evidence that children took longer to accept the request (i.e., give the puppet the fuller container) when shown the 'non-full/non-full' pair than when shown either the 'full/non-full' pair or both *big* pairs. Although there were only five children for which the comparison between the RTs for the 'straight/bent'

and 'bent/bent' pairs could be made,[12] we observed the same trend in RTs: the differences between the adjective onset to the look, reach, and touch between the two pairs were all significant (one-tailed t tests: look $t(4) = -3.23$, $p = 0.016$; reach $t(4) = -3.15$, $p = 0.017$; touch $t(4) = -3.32$, $p = 0.015$). See Figure 6.

Figure 6: Reaction times for two *straight* pairs in Experiment 2b
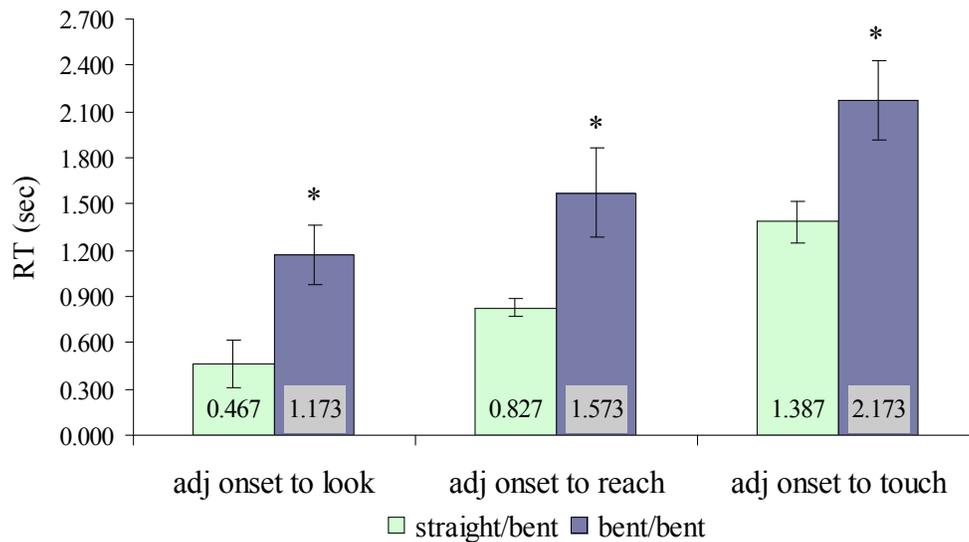
As in Experiment 1, we also compared RTs for the difference between the look and the touch for the 'bent/bent' pair to three other key pairs – the 'straight/bent' pair and the two pairs of blocks corresponding to the request for *the big one*. We again observed the same trend as in Experiment 1 (cf. Figure 7): children took significantly longer to select the straighter of the two bent pairs to satisfy the puppet's request than they did for the other three pairs (two-tailed t-tests:

---

[12] We targeted the responses from Experiment 2b only and did not combine responses across both experiments, as we did in Experiment 1, since the *straight* stimuli differed between the two experiments. Recall, too, that we were constrained with respect to which videos we could code, for reasons outlined earlier. The total number of codable sessions and items was therefore less than for Experiment 1.

'bent/bent' v. 'straight/bent: $t(23) = 2.15$, $p = 0.04$; 'bent/bent' v. 'big': $t(20) = 2.31$, $p = 0.03$; 'bent/bent' v. 'small': $t(22) = 2.75$, $p = 0.01$). And, as before, the other pairs do not differ significantly from each other (two-tailed t-tests: 'straight/bent' v. 'big': $t(31) = -0.33$, $p = 0.74$; 'big' v. 'small': $t(30) = 0.64$, $p = 0.53$; 'straight/bent' v. 'small': $t(33) = 0.86$, $p = 0.40$).[13]

Figure 7: Reaction times for four key pairs in Experiment 2b
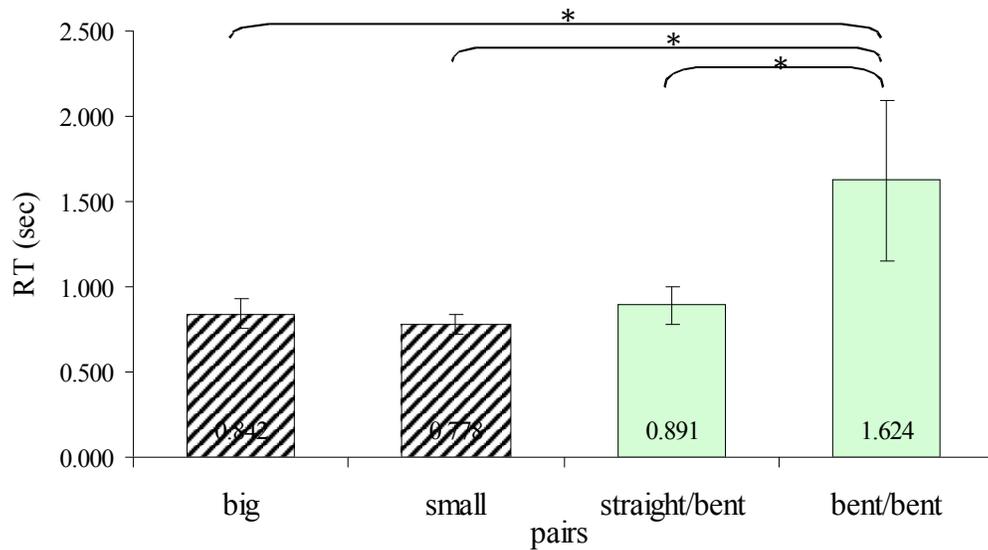
5.2.4   Discussion

The results of Experiment 2 support a generalization of the results from Experiment 1. In particular, the results show that the effect of the order of presentation and the longer RTs for the non-maximal pair are not unique to the stimuli included in Experiment 1, and therefore not unique to the lexical item *full*. Rather, the fact that we see the same pattern with the absolute maximum standard GA *straight* indicates that the interpretive processes involved in choosing an object that comes closest to satisfying a description based on a maximum standard GA are

---

[13] The data from the following number of children were analyzed for each pair: 15 for the big blocks, 17 for the small blocks, 18 for 'straight/bent', and 7 for 'bent/bent'.

different from those involved in choosing an object that satisfies a context-dependent description based on a relative GA. In short, understanding *the big/long one* as a description of the bigger/longer of two objects of unequal size is automatic, while tolerating use of *the full/straight one* as a description of the fuller/straighter of two objects when neither of them is actually full or straight requires some work.

6.General Discussion

The experiments reported in this paper provide new insights into children's competence in three different aspects of the context/meaning relation. First, our study shows that by three years of age, children have adult-like competence in assigning context-sensitive interpretations to relative GAs in the positive form. Specifically, they are able to shift the standard of comparison for such adjectives in a way that is appropriate for the context of utterance. In our case, this was determined by the pragmatic demands of the definite description in which the adjective appeared, thereby changing the extension of the predicate. Moreover, the fact that children did not assign the same kinds of context-sensitive meanings to absolute GAs in the same contexts shows that they have already made adult-like distinctions between predicates that have context-sensitive denotations and those that do not, but which are otherwise semantically quite similar. In the case at hand, both relative and absolute GAs share the fundamental feature of expressing a relation to a scalar concept – that is, both are gradable – but they differ just on the basis of the context dependence of the positive form.[14]

---

[14] According to Kennedy and McNally (2005) and Kennedy (to appear), this difference actually stems from a more basic difference between the two classes of GAs, namely the structures of the scales that represent the gradable concepts they encode: relative GAs use open scales, while absolute GAs use closed scales. If this is correct, then a more accurate way to state the

Second, our experiments show that three-, four-, and five-year-olds have adult-like competence in dealing with the existence and uniqueness presuppositions of a singular definite description. Their responses to the control items in the Presupposition Assessment Task clearly show that they are aware of these presuppositions, and that they are willing to reject as infelicitous utterances that violate them. At the same time, their responses to the test items containing relative GAs show that when the semantics of the adjective and the context of utterance allow for the possibility of presupposition accommodation – in these cases, by shifting the standard of comparison so that the adjective (and, consequently, the description) is true of just one member of a pair of objects – children, like adults, will accommodate. These two results together indicate that (at least in this domain) children are constructing the type of complex discourse models that linguistic expressions with presuppositions must be integrated into, and are willing and able to automatically update those models to allow integration when such a move is licensed.

Finally, our results provide important new data from child language that bears on fundamental questions about the nature of contextually influenced interpretive variability. The results from the Presupposition Assessment Task show that children will treat both relative GAs and maximum standard absolute GAs as 'variable': just as *the big one* was accepted as a description of the bigger of two blocks, even when it was otherwise judged to be *not big*, *the full one* was sometimes accepted as a description of the fuller of two containers, even when it was observably not full. The latter result was significantly less frequent, and was influenced by the

---

generalization is that by three years of age, children already know the mapping between scale structure and context sensitivity of the positive form. See Syrett (in preparation) for extensive discussion of this issue.

sequence of requests, but it nevertheless occurred enough to support the conclusion that some

sort of interpretive variability is at work. The question at hand is whether this variability is due to

the same factors that are involved in variable interpretations of relative GAs – semantic context

dependence (and ultimately perhaps the semantic features that give rise to vagueness) – or

whether it stems from a different source.

It is important to emphasize that this is a much more general question, and bears on both

child and adult language. While we did not see the same variability in maximum standard

absolute GAs with adults, we suspect that we would have if the materials had been slightly

different. In particular, we believe that if the fuller of the two 'non-full' containers had been

closer to full than it actually was, though still noticeably not full, adults would have behaved like

children and provided it in response to a request for *the full one*. Our belief is based on simple

observation of everyday linguistic behavior: we regularly speak imprecisely. Thus, in addition to

using absolute GAs like *full* and *straight* to describe objects that are not strictly speaking

maximally full or maximally straight, we might also describe our arrival somewhere at 3.05pm

by saying *We arrived at 3pm*, or a child's consumption of all but a few pieces of a sandwich by

saying *The child ate the sandwich*. (See Lasersohn (1999) and Sperber and Wilson (1986) for

extensive discussion of such cases.) Given that our willingness to tolerate such imprecision is

itself a matter of context (i.e., if the precise time of our arrival is important, then use of *arrive at*

3pm to describe an arrival at 3.05pm is inappropriate), it could be the case that imprecision is just

another instance of semantic context dependence. In other words, it could be the case that

interpretive variability is always fundamentally semantic, and that expressions like *full* and

straight (as well as expressions such as *arrive at 3pm*, *eat the sandwich*, etc.) have meanings

that, like *big* and *long*, require valuing some contextual parameter before they can be assigned

actual extensions. Such meanings might be slightly different for absolute maximum GAs – *full*, for example, might have a meaning along the lines of 'have a fullness that is at least as close to maximally full as required by the prevailing contextual standard of precision', but on this view, they would be of the same basic sort as the context sensitive meanings of relative GAs.

While this conclusion would be consistent with children's gross responses in the PAT (whether they gave the puppet an object to satisfy his request or not), it is not consistent with the reaction time data that we collected. Instead, the fact that children systematically took significantly longer to respond to a request for *the full/straight one* in contexts involving two objects that did not satisfy the maximum standard property (that is, which were not maximally full or straight) shows that processing of more than just the semantic content of the description was involved in their responses. To see why this is the case, consider the fact that their response times to definite descriptions based on relative GAs (e.g., *the big/long one*) were the same regardless of whether they were looking at two objects from the upper end of the continuum (both objects judged *big*/*long*, based on the Scalar Judgment Task) or two objects from the lower end (neither object judged *big*/*long* in the SJT). This has a straightforward explanation. Focusing on the case of *long* for purposes of exemplification, if children know the meaning of the relative GA, then they know that it denotes the property of having a length that exceeds the standard of comparison for length in the context. If the context contains exactly two relevant objects and is such that just one of them has this property (a constraint imposed by the presuppositions of the definite description used by the speaker), their response is straightforward: choose the longer object. It is unnecessary for them to choose some particular degree as the standard of comparison, because given two objects of unequal lengths, the transitivity of the 'exceeds'

relation ensures that if just one of the objects has a length that exceeds the standard of comparison (whatever it may be), it has to be the longer of the two.

Crucially, if the same kind of reasoning were at play in children's evaluation of requests for *the full one* in the 'nonfull/nonfull' context, we should have seen similar RTs. Specifically, if *full* meant 'fuller than the contextual standard' or even 'at least as close to maximally full as the contextual standard', then the choice between two objects of unequal fullness would again be straightforward: if context dictates that just one of the objects should satisfy the property, it will always be the fuller of the two. The fact that children took significantly longer to respond to such requests, even though their responses were the same (i.e,. they chose the object with the greater degree of the relevant property), clearly shows that the same kind of reasoning is not at work in evaluating absolute GAs in such contexts, and therefore that the kind of interpretive variability that such adjectives do occasionally display is not indicative of a context-dependent semantics.

But what kind of reasoning do the long RTs indicate? Our hypothesis is that children took longer to respond to such cases precisely because they were aware that neither of the objects in question actually satisfied the description used in the request, and that this knowledge triggered an evaluation of whether one of the objects was actually close enough to having the property in order for them to treat the speaker's description as though it were a description of that object. There are different ways to formally characterize this notion of imprecision. For concreteness, we will assume with Lasersohn (1999) that it involves (a) computation of a set of alternative denotations for an expression that are ordered relative to the actual meaning, and (b) a decision about how far down this list one can go before getting too far away from the actual denotation. In contexts in which the actual meaning of an utterance fails to apply, if there is an element among the set of its 'tolerable alternatives' that could actually be true, the utterance can be taken to

count as 'close enough to true' for the purpose of the discourse, and treated as though it had been an utterance of the tolerable alternative. Crucially, this kind of reasoning is fundamentally pragmatic in nature, as it involves judgments about communicative intent and whether a particular utterance can be used in a way that (strictly speaking) conflicts with the truth conditional requirements that it imposes by virtue of its semantics. It is this move beyond the computation of semantic content – even content that invokes contextual information – that, we claim, imposes an extra processing load and results in the longer RTs we observed.

If this is correct, then we have experimental evidence for a distinction between two types of context-dependent interpretive variability. One type – that exhibited by relative GAs in the positive form – is fundamentally *semantic* in nature, and is based on the conventional meaning of particular expressions (or combinations thereof). A second type – that exhibited by imprecise uses of maximum standard absolute GAs – is fundamentally *pragmatic*, and involves computation of a set of alternative meanings and a contextual judgment about how many of them count as 'tolerable deviations' from the actual, precise meaning of the expression. The differences between children and adults that we observed in their willingness to accept false descriptions based on maximum standard GAs can be explained by assuming that children are more willing to tolerate imprecision than adults, at least in this experimental task.[15] If our overall

---

[15] One potential objection to this interpretation of the results is that preschoolers are generally not imprecise, and can in fact be quite literal at times. (Thanks to Anastasia Giannakidou, p.c., for raising this point.) Why, then, would they tolerate imprecision in the PAT? In fact, we are not claiming that children are imprecise; indeed, the slower RTs for the 'non-full/non-full' and 'bent/bent' pairs indicate that their willingness to give the puppet the fuller or straighter member of the pair is far from automatic. Our claim is rather that their desire to maintain a high standard

interpretation of the data is correct, then we predict that smaller deviations from the absolute standard should result in adults behaving like children – accepting false descriptions based on absolute maximal standard GAs, but taking longer to do so than the time it takes to accept a description based on a relative GA.

7.Conclusion

In sum, the research described above provides evidence for four kinds of conclusions. First, by the age of three, children have fine-grained distinctions among subclasses of adjective. In particular, we have seen that children distinguish relative gradable adjectives like *big* from absolute gradable adjectives like *spotted*/*full*. Evidence for this conclusion comes from children's recognition that that standard of comparison against which a gradable adjective in the positive form is judged varies only for relative gradable adjectives such as *big*. Second, children distinguish maximal standard absolute gradable adjectives such as *full* from minimal standard absolute gradable adjectives such as *spotted*. Evidence for this conclusion comes from the Scalar Judgment task in which only objects exhibiting the maximal degree of fullness were judged as *full* and in which all objects exhibiting any degree of spottedness were judged as *spotted*. Third, we learned that children have a complete understanding of both the uniqueness and existence presuppositions associated with definite determiners in English. We would not have been able to find the context dependence of relative gradable adjectives were that not the case. It is only through these presuppositions that appropriate shifting of the standard of comparison is licensed. Fourth, we have identified novel evidence for the distinction between context dependence as a

---

of precision was, in this case, overridden by a stronger desire to actually respond to the speaker's (the puppet's) request. Evidently, the opposite is true for adults, though as we suggested above, we suspect that the results would change if the deviation from the maximum were reduced.

function of the meaning of an expression as opposed to context dependence as a function of the use of an expression in a situation. We believe that the reaction time difference between accepting the bigger of two small blocks as *big* and the fuller of two non-full containers as *full* is a direct reflection of the different computations involved in these two types of context dependence.

These results open a range of questions for future research. With respect to the subcategories of adjective, it is important to determine what kinds of evidence lead learners to classify a given adjective as gradable, relative, or absolute. To what extent do these distinctions follow directly from the mapping of an adjective to an appropriate concept as opposed to being linguistically conditioned? Even if the subclasses of gradable adjective represent limitations on the range of possible scalar meanings for adjectives, what kinds of evidence do learners use to place a novel adjective in one of these subclasses. These questions would seem to interact with the issues surrounding context dependence. If both relative and absolute gradable adjectives show some degree of context dependence, though in different ways, what information does the learner use to determine whether this context dependence is due to semantics or pragmatics. Indeed, what evidence leads the learner to recognize the distinction and to attribute some cases of context dependence to the semantics and other cases to the pragmatics?

While these results have opened as many questions as they have solved, we believe they also provide an interesting methodological lesson. In particular, while the debate concerning varieties of context dependence (in brief, one kind or two) has raged for years within linguistics and philosophy, resisting conclusive evidence, the current experiments allow us to ask about the real time consequences of the alternative approaches. Using experimental evidence from children we are able to see the alternative computations involved in semantic vs. pragmatic context-

dependence unfolding in real time. Our work demonstrates that he process of verification of the meaning of an expression can provide evidence for alternative hypotheses about the meaning itself. Moreover, such evidence can come equally well from children as from adults.

References

Abbott, B. (1999). Support for a unique theory of definite descriptions. In T. Matthews & D. Strolovitch (Eds.), *Proceedings from Semantics and Linguistic Theory IX* (pp. 1-15). Ithaca, NY: CLC Publications.

Barner, D., &. Snedeker, J. (2007). Compositionality and statistics in adjective acquisition: 4-year-olds interpret *tall* and *short* according to the typical sizes of novel noun referents. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.) *Proceedings of the 31st Annual Boston University Conference on Language Development* (pp. 81-92). Somerville, MA: Cascadilla Press.

Bartlett, E. (1976). Sizing things up: The acquisition of the meaning of dimensional adjectives. *Journal of Child Language* 3, 205-219.

Bartsch, R., & Vennemann, T. (1972). The grammar of relative adjectives and comparison. *Linguistische Berichte*, *20*, 19-32.

Bausano, M. K., & Jeffrey, W. E. (1975). Dimensional salience and judgments of bigness by three-year-old children. *Child Development*, *46*, 988-991.

Birner, B., & Ward, G. (1994). Uniqueness, familiarity, and the definite article in English. In S. Gahl, A. Dolbey, & C. Johnson (Eds.), *Proceedings of the 20th Annual Meeting of the Berkeley Linguistics Society*, *20* (pp. 93-102). Berkeley, CA: Berkely Linguistics Society.

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infants' expectations for count nouns and adjectives. *Journal of Cognition and Development*, *4*, 357-381.

Brewer, W., & Stone, J. B. (1975). Acquisition of spatial antonym pairs. *Journal of Experimental Child Psychology*, *19*, 299-307.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.

Clark, E. V. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning and Verbal Behavior, 11*, 750-758.

Clark, E. V. (1973). What's in a word? On the child's acquisition of semantics in his first language. In T. Moore (Ed.), *Cognitive development and the development of language* (pp. 65-110). New York, NY: Academic Press. Reprinted (1977) in C. Snow & B. Ferguson (Eds.), *Talking to children* (pp. 186-233). Cambridge, MA: Cambridge University Press.

Clark, H. (1973). Space, time, semantics, and the child. In T. Moore (Ed.), *Cognitive development and the acquisition of* language (pp. 27-63). New York, NY: Academic Press.

Coley, J. D., & Gelman, S. A. (1989). The effects of object orientation and object type on children's interpretation of the word "big". *Child Development, 60*, 372-380.

Cresswell, M. J. (1977). The semantics of degree. In B. Partee (Ed.), *Montague grammar* (pp. 261-292). New York: Academic Press.

Diesendruck, G., Hall, D. G., Graham, S. (2006). Children's use of syntactic and pragmatic knowledge in the interpretation of novel adjectives. *Child Development, 77*, 16-30.

Donaldson, M., & Wales, R. J. (1970). On the acquisition of some relational terms. In J. R. Hayes & R. Brown (Eds.), *Cognition and the development of language* (pp. 235-268). New York, NY: John Wiley and Sons, Inc.

Ebeling, K. S., & Gelman, S. A. (1988). Coordination of size standards by young children. *Child Development, 59*, 888-896.

Ebeling, K. S., & Gelman, S. A.. (1994). Children's use of context in interpreting "big" and "little". *Child Development*, *65*, 1178-1192.

Ehri, L. (1976). Comprehension and production of adjectives and seriation. *Journal of Child Language*, *3*, 369-384.

Eilers, R. E., Kimbrough Oller, D., & Ellington, J. (1974). The acquisition of word-meaning for dimensional adjectives: The long and short of it. *Journal of Child Language*, *1*, 195-204.

Fara, D. G. (2000). Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, *28*, 45-81. Originally published under the name "Delia Graff".

Finch-Williams, A. (1981). *Biggerest or biggester: A study in children s acquisition of linguistic features of comparative adjectives and nonlinguistic knowledge of seriation*. Doctoral dissertation, University of Kansas.

Gathercole, V. C. (1979). *Birdies like birdseed the bester than buns: A study of relational comparatives and their acquisition*. Doctoral dissertation, University of Kansas.

Gathercole, V. C. (1982). Decrements in children's responses to *big* and *tall*: A reconsideration of the potential cognitive and semantic causes. *Journal of Experimental Child Psychology*, *34*, 156-173.

Gelman, S. A., & Ebeling, K. S. (1989). Children's use of nonegocentric standards in judgments of functional size. *Child Development*, *60*, 920-932.

Gelman, S., & Markman, E. M. (1985). Implicit contrast in adjectives vs. nouns: Implications for word-learning in preschoolers. *Journal of Child Language*, *6*, 125-143.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kučzaj II (Ed.), *Language development Vol. 2: Language, thought and culture* (pp. 301-334). Hillsdale, NJ: Lawrence Erlbaum.

Gitterman, D., & Johnston, J. R. (1983).Talking about comparisons: A study of young children's comparative adjective usage. *Journal of Child Language*, *10*, 605-621.

Graham, S., Welder, A., & McCrimmon, A. (2003). Hot dogs and zavy cats: Preschoolers' and adults' expectations about familiar and novel adjectives. *Brain and Language, 84*, 16-37.

Graham, S., Cameron, C., & Welder, A. (2005). Preschoolers' extension of familiar adjectives. *Journal of Experimental Child Psychology*, *91*, 205-226.

Graziano-King, J. (1999). *Acquisition of comparative forms in English*. Doctoral dissertation, City University of New York.

Graziano-King, J., & Cairns, H. S. (2005). Acquisition of English comparative adjectives. *Journal of Child Language*, *2*, 345-373.

Hall, D. G., Waxman, S. R., & Hurwitz, W. (1993). How two- and four-year-old children interpret adjectives and count nouns. *Child Development*, *64*, 1651-1664.

Harris, P. L., & Folch, L. (1985). Decrement in the understanding of "big" among English- and Spanish-speaking children. *Journal of Child Language*, *12*, 685-690.

Harris, P. L., Morris, J. E., & Meerum Terwogt, M. (1986). The early acquisition of spatial adjectives: A cross-linguistic study. *Journal of Child Language*, *13*, 335-352.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and Philosophy*, *13*, 137-177.

Hollich, G. (2003). SuperCoder [Computer program]. Retrieved from http://hincapie.psych.purdue.edu/Splitscreen/SuperCoder.dmg.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, *29*, 597-626.

Kadmon, N. (1990). Uniqueness. *Linguistics and Philosophy*, *13*, 273-323.

Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, *57*, 129-191.

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference. Cambridge studies in linguistics 24*. New York, NY: Cambridge University Press.

Katz, J. (1967). Recent issues in semantic theory. *Foundations of Language*, *3*, 124-194.

Keil, F., & Carroll, J. (1980). The child's acquisition of "tall": Implications for an alternative view of semantic development. *Papers and Reports on Child Language Development*, *19*, 21-28.

Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland. (1997. Doctoral dissertation, University of California, Santa Cruz).

Kennedy, C. (to appear). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*.

Klatzky, R. L., Clark, E. V., & Macken, M. (1973). Asymmetries in the acquisition of polar adjectives: Linguistic or conceptual? *Journal of Experimental Child Psychology*, *16*, 32-46.

Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, *4*, 1-45.

Klein, E. (1991). Comparatives. In A. von Stechow & D. Wunderlich (Eds.), *Semantik: Ein internationales handbuch der zeitgen össischen forschung* (pp. 673-691). Berlin: Walter de Gruyter.

Klibanoff, R. S., & Waxman, S. R. (1998). Preschoolers' acquisition of novel adjectives and the role of basic-level kind. In A. Greenhill, M. Hughes, H. Littlefield & H. Walsh (Eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development* (pp. 442-453). Somerville, MA: Cascadilla Press.

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development*, *71*, 649-659.

Lasersohn, P. (1999). Pragmatics halos. *Language*, *75*, 522-551.

Layton, T. L., & Stick, S. L. (1978). Comprehension and production of comparatives and superlatives. *Journal of Child Language*, *6*, 511-527.

Lewis, D. (1968). Counterpart theory and quantified modal logic. *Journal of Philosophy*, *65*, 113-126.

Ludlow, P. (1989). Implicit comparison classes. *Linguistics and Philosophy*, *12*, 519-533.

Lumsden, E. A., & Poteat, B. W. S. (1968). The salience of the vertical dimension in the concept of "bigger" in five- and six-year-olds. *Journal of Verbal Learning and Verbal Behavior*, *7*, 404-408.

Macnamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review*, *79*, 1-13.

Maratsos, M. (1973). Decrease in the understanding of the word "big" in preschool children. *Child Development*, *44*, 747-752.

Maratsos, M. (1974). When is a high thing the big one? *Developmental Psychology*, *10*, 367-375.

Maratsos, M. (1976). *The use of definite and indefinite reference in young children: An experimental study of semantic acquisition*. New York, NY: Cambridge University Press.

Marschark, M. (1977). Lexical marking and the acquisition of relational size concepts. *Child Development*, *48*, 1049-1051.

Meroni, L. (2006). *Putting children in context: Experimental studies on children's interpretation of definite noun phrases*. Doctoral dissertation, University of Maryland, College Park.

Mintz, T. (2005). Linguistic and conceptual influences on adjective acquisition in 24- and 36-month olds. *Developmental Psychology*, *41*, 17-29.

Mintz, T., & Gleitman, L. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*, 267-293.

Moore, D. (1999). Comparatives and superlatives: Lexical before functional. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development* (pp. 474-481). Somerville, MA: Cascadilla Press.

Nadig, A., Sedivy, J., Joshi, A., & Bortfeld, H. 2003. The development of discourse constraints on the interpretation of adjectives. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (pp. 568-579). Somerville, MA: Cascadilla Press.

Neale, S. (1990). *Descriptions*. Cambridge, MA: MIT Press.

Nelson, K. (1976). Some attributes of adjectives used by young children. *Cognition*, *4*, 1-31.

Nelson, K., & Benedict, H. (1974). The comprehension of relative, absolute, and contrastive adjectives by young children. *Journal of Psycholinguistic Research*, *3*, 333-342.

Prasada, S. (1992). *Acquisition of adjective meanings: A lexical semantic approach*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Prasada, S., & Cummins, M. (2001). Structural constraints on the interpretation of novel count nouns. In A. H-J. Do, L. Dominguez, & A. Johansen (Eds.), *Proceedings of the 25th*

Annual Boston University Conference on Language Development (pp. 623-632). Somerville, MA: Cascadilla Press.

Raffman, D. (1994). Vagueness without paradox. *The Philosophical Review*, *103*, 41-74.

Raffman, D. (1996). Vagueness and context relativity. *Philosophical Studies*, *81*, 175-192.

Ravn, K. E., & Gelman, S. (1984). Rule usage in children's understanding of "big" and "little". *Child Development*, *55*, 2141-2150.

Rips, L., & W. Turnbull. (1980). How big is big? Relative and absolute properties in memory. *Cognition*, *8*, 145-174.

Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy*, *26*, 287-350.

Rotstein, C., & Winter, Y. (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics*, *12*, 259-288.

Russell, B. (1905). On denoting. *Mind*, *14*, 479-493.

Sena, R., & Smith, L. B. (1990). New evidence on the development of the word "big". *Child Development*, *61*, 1034-1052.

Sera, M., & Smith, L. B. (1987). "Big" and "little": "Nominal" and relative uses. *Cognitive Development*, *2*, 89-111.

Sharpe, D., Fonte, I.., & Christe, E. (1998). Big mice, big animals, big problems: The acquisition of adjective interpretation rules. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development* (pp. 675-683). Somerville, MA: Cascadilla Press.

Smith, L. B., Cooney, N., & McCord, C. (1986). What is "high"? The development of reference points for "high" and "low". *Child Development*, *57*, 583-602.

Soames, S. (1999). *Understanding Truth*. Oxford: Oxford University Press.

Sperber, D., & Wilson, D. (1986). Loose talk. *Proceedings of the Aristotelian Society New Series*, *86*, 153-171.

Strawson, P. F. (1950). On referring. *Mind*, *59*, 320-344.

Syrett, K., Lidz, J., & Kennedy, C. (2006a). Poster presented at the *CUNY Human Sentence Processing Conference*, March 23-25.

Syrett, K., Lidz, J., & Kennedy, C. (2006b). Paper presented at the *Main Session of the 42nd Annual Meeting of the Chicago Linguistic Society*, April 6-8.

Syrett, K., Bradley, E., Lidz, J., & Kennedy, C. (2006). Shifting standards: Children's understanding of gradable adjectives. In K. Ud Deen, J. Nomura, B. Schulz, & B. D. Schwartz (Eds.), *Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition - North* America (pp. 353-364). Honolulu, HI. Cambridge, MA: UConn Occasional Papers in Linguistics 4.

Taylor, M., & Gelman, S. (1988). Adjectives and nouns: Children's strategies for learning new words. *Child Development*, *59*, 411-419.

Townsend, D. (1976). Do children interpret "marked" comparative adjectives as their opposites? *Journal of Child Language*, *3*, 385-396.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89-134.

Unger, P. (1975). *Ignorance*. Oxford: Clarendon Press.

von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics*, *3*, 1-77.

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds'

interpretations of novel nouns and adjectives. *Cognitive Psychology*, *43*, 217-242.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel

adjectives. *Developmental Psychology*, *36*, 571-581.

Waxman, S. R., & Markow, D. (1995). Words as invitations to form categories: Evidence from

12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257-302.

Waxman, S. R., & Markow, D. (1998). Object properties and object kinds: Twenty-one-month-

old infants' extension of novel adjectives. *Child Development*, 69, 1313-1329.

Wexler, K. (2003). Determiners in development, why there is a problem: Maximal trouble.

Handout from *Semantics in acquisition* (course taught at MIT).

Wheeler, S. (1972). Attributives and their modifiers. *Noûs*, *6*, 310-334.