

Shifting Standards: Children's Understanding of Gradable Adjectives*

*Kristen Syrett, Evan Bradley, Christopher Kennedy, Jeffrey Lidz
Northwestern University*

Two studies demonstrate that children have knowledge of scalar distinctions between three sub-classes of gradable adjectives: relative (*big, long*), absolute with a maximal standard (*full*), and absolute with a minimal standard (*spotted*). Performance on these adjectives is compared with controls (shape, color, mood). Children appropriately shift the standard of comparison with context-dependent, relative gradable adjectives, and do not do so for the others. Reasons for non-adult-like performance with *full* are discussed. Evidence is presented that children know about the presuppositions of singular definite descriptions, suggesting that children as young as three have an accurate semantic representation of *the*.

1. Introduction

Children's acquisition of adjectives has received a great deal of attention in the field of psychology. The vast majority of this work has focused on a core set of topics: children's ability to distinguish properties from kinds of objects (cf. Gelman and Markman 1985, Waxman and Booth 2001); the influence of exemplars or reference to object category on children's ability to extend adjectives to properties (cf. Klibanoff and Waxman 2000, Mintz and Gleitman 2002); the role of comparison and contrast in adjective learning (cf. Clark 1972 and 1973, Ehri 1976, Gentner and Rattermann 1998, Ryalls 2000); and the distinctions children make between adjectives of color, dimension, and property (cf. Bartlett 1976, Nelson and Benedict 1974, Sandhofer and Smith 1999). The current investigation adds to this body of research by presenting findings regarding children's comprehension of the semantic distinctions within the class of *gradable adjectives*.

*Portions of this work were presented at the LSA Annual Meeting in Oakland, CA, 2005. This study was funded in part from an NSF Grant (BCS-0094263) to Kennedy, an NSF Grant (BCS-0418309) to Lidz, and an NIH grant (HD30410) to Sandra Waxman for the Project on Child Development at Northwestern. Special thanks go to the children and staff of Chiaravalle Montessori School and Warren W. Cherry Preschool in Evanston, IL, and to Erin Leddon and Josh Viau for help running the experiments. We are grateful to members of the Northwestern Acquisition Lab Group for helpful comments.

2. Linguistic Background

Gradable adjectives (often referred to as scalar, dimensional, size, or spatial) are adjectives such as *big* or *tall*, which describe properties of objects that hold to different degrees. For this reason, gradable adjectives can appear in comparative constructions, while non-gradable adjectives cannot. This observation is illustrated in example (1).

- (1) Neanderthal man is [more **primitive** than / *more **extinct** than]
Homo sapiens.

A standard semantic analysis of gradable adjectives is that they include as a core part of their meaning a function that takes as its input an object and returns as its output a measure of the extent to which that object possesses a property denoted by the adjective (see Klein 1991). This measurement can be formally represented as values, or *degrees*, in an ordered set, or *scale*. The values of the scale correspond to the dimension labeled by the adjective (e.g., height, weight, age, etc.) (see Kennedy 1999, Kennedy and McNally 2005).

On this view, sentences constructed out of gradable adjectives express orderings between degrees. This is most clearly illustrated by comparatives like (1), which express *explicit* orderings between degrees on a scale. Although the positive (unmarked) form of a gradable adjective lacks overt comparative morphology, it, too, describes a relation between degrees on a scale, albeit an implicit one (cf. Sapir 1944). In this case, the degree to which an object possesses some property is related to a context-dependent *standard of comparison*, which represents the cutoff point for what ‘counts as’ having the property in question in the context of utterance. The context sensitivity of the standard of comparison is illustrated by an example like (2).

- (2) The coffee in Rome is expensive.

(2) might be judged true if it is asserted as part of a conversation about the cost of living in various Italian cities (*In Rome, even the coffee is expensive!*), but false in a discussion of the cost of living in New York versus Rome (*The rents are high in Rome, but at least the coffee is not expensive!*). The price of coffee in Rome does not change in the two contexts; what differs is the standard of comparison.

However, not all gradable adjectives depend on context for the standard. While ‘relative’ gradable adjectives (e.g., *expensive, big, long, old*) are context dependent, ‘absolute’ gradable adjectives (e.g., *full, straight, spotted, bumpy, open, closed*) are not (Unger 1974, Rotstein and Winter 2004, Kennedy and McNally 2005). For example, what counts as being full does not depend on the kind of container, the kind of material used to fill the container, the location of the container, etc. Rather, any container is full only when it is maximally filled. Kennedy and McNally (2005) explain this difference in terms of the scale structures associated with the two classes of adjectives. Absolute adjectives have scales with minimal or maximal endpoints, which

provide default values for the standard of comparison. For example, if an object has any spots, it is spotted (minimal); likewise, if a container is filled to its upper boundary, we can say that it is full (maximal). In contrast, relative adjectives have open-ended scales, and so must fix their standards on the basis of contextual information.

The semantic distinctions between gradable and non-gradable adjectives on the one hand, and relative and absolute gradable adjectives on the other lead up to clear questions for language acquisition. First, can children incorporate relevant contextual information and appropriately shift the standard of comparison for relative gradable adjectives, and not do so for absolute gradable adjectives? Second, is children's knowledge about the differences in scalar structure for these adjectives reflected in their scalar judgments? Third, if children pattern in a non-adult-like way with respect to the treatment of these adjectives, what factors can we identify that account for this difference or guide them towards adult-like knowledge?

3. Previous Research

While gradable adjectives have been addressed in a number of experimental investigations over the years, there are still a number of questions remaining about children's comprehension of these adjectives. To begin, while a number of studies (see references listed above) investigated these adjectives in comparison with color or categorical terms, they did not focus on distinctions *within* this class of adjectives. Furthermore, a number of studies (cf. work by L. Smith and colleagues, and work by Ebeling and Gelman) have investigated children's ability to incorporate a contextually-defined standard. Combined, these studies provide important evidence that children can make use of a non-egocentric, comparison-class oriented standard of comparison and that their ability to appropriately do so improves with age. However, these studies did not compare gradable and non-gradable adjectives within the same task, and focused mainly on perceptual or conceptual—not linguistic—factors.

Finally, a number of tasks have used a similar methodology, in which children are presented with at least two items and asked to give or point to *the X one*, where X stands for an adjective such as *big* (Nelson and Benedict 1974, Eilers, Kimbrough Oller, and Ellington 1974, Bartlett 1976, Ebeling and Gelman 1988 and 1994). However, there is an unfortunate confound in the findings: it is not clear if children were simply treating the positive, non-comparative form as an implicit comparison, or if they were appropriately shifting the standard of comparison. That is, when asked to distinguish between two objects of unequal size, children could have understood *the big one* to mean *the bigger one* rather than as *the one that is big relative to this context*. This distinction is subtle for an adjective like *big*, but is clear when we look at absolute gradable adjectives. For adults, the definite description *the spotted one* cannot be used to distinguish between two objects with unequal (but non-zero) amounts of spots (*the more spotted one*), and *the full one* cannot be used to distinguish between two containers of unequal (but non-maximal) contents (*the fuller one*). The present work takes these differences into consideration. We show that children do not simply interpret the definite

description as an implicit comparison and do make a distinction between relative and absolute gradable adjectives with respect to the standard of comparison.

4. Experiment 1

4.1 Method

Three groups of children were tested: three-year-olds (*range* = 3;5 to 3;11, *M* = 3;8), four-year-olds (*range* = 4;1 to 4;11, *M* = 4;5), five-year-olds (*range* = 5;1 to 5;8, *M* = 5;5). There were 10 children in each age group. In addition, twenty-four adult native speakers of English participated. The number of males and females for each age group in each condition was approximately equal. The data from one child were discarded because the child responded incorrectly on more than half of the control items, and the data from one adult were discarded due to experimenter error.

Children were invited to play a game to help a puppet learn how to make requests. They were told that they would be shown two objects at a time, and that every time a pair appeared, the puppet would ask for something. Their task was to determine if they could satisfy his request.¹ This task differed, however, in a crucial way from similar tasks administered in previous experiments: the puppet's request was not always a felicitous one. This pragmatic manipulation was accomplished in the following way. Each request was worded using a definite description (e.g., *Please give me the X one*, where X is a target adjective). The singular DP is headed by a definite determiner and therefore presupposes both *existence* (there must be a red object) and *uniqueness* (there should only be one red object). The pairs were constructed so that they would either satisfy or violate one or both of these presuppositions. For some pairs, there was exactly one object able to be modified by the adjective; for others, both members of the pair fit the description; and for others, neither member of the pair fit the description. Participants were expected to either accept or reject the request appropriately. For this reason, we refer to the task as the Presupposition Assessment Task. The target adjectives were four gradable adjectives, two relative (*big*, *long*) and two absolute (minimal: *spotted*, maximal: *full*). These were compared with control (non-gradable) adjectives referring to shape (*square*, *round*), mood (*happy*, *sad*), and color (*red*, *yellow*, *blue*).

Each experimental session began with a short training session, in which the participant was presented with four pairs described by control adjectives. Two satisfied the presuppositions of the definite description, and two did not. Once the experimenters were satisfied that the participant was comfortable with the task and could accept or reject the puppet's request, they proceeded to the test session, which included 17 pairs (9 control pairs and 8 gradable pairs). There were 4 relative pairs (2 sets of cubes for *big*, 2 sets of rods for *long*) and 4 absolute pairs (2 sets of containers for *full*, 2 sets of disks

¹ Adult participants did not interact with a puppet; the experimenter made the requests.

for *spotted*). All stimuli could be easily held in one hand by the experimenter. Participants were randomly assigned to one of two orders of presentation.

In the case of relative gradable adjectives, if the participant shifts the standard of comparison, the definite description should always pick out one member of the pair (e.g., the cube that is *big* for the context), even if the items might normally be described by the negative polar adjective (e.g., two *small* cubes). For the absolute pairs, this is not the case. Given two containers where only one is full, or two disks where only one is spotted, the choice is clear. However, when given two non-full containers, or two spotted disks, even if the relevant property holds of one member of the pair to a greater degree, the request is infelicitous, because both presuppositions are not met. The participant should, then, reject the request. However, if participants treat the request as an implicit comparison across the board (e.g., interpreting *the X one* as *the Xer one*), then they should accept the request and give the puppet the fuller or more spotted member of the pair.

4.2. Results and discussion

Responses were coded relative to the expected response for each pair. It was predicted that participants would give 0, 1, or 2 objects, depending on the pair. The control pairs were constructed to elicit all three responses. For the relative pairs, it was expected that the participants would always give the member of the pair that held the property to the greater degree. For the absolute pairs, it was predicted that the participants would give the appropriate member of the pair when the request was felicitous; however, for the non-full/non-full pair we expected the participants to reject the request and give neither, and for the spotted/spotted pair, we expected the participants to reject the request and give either both or neither.

The results are presented in Table 1.

Table 1. Percentage expected responses

Age	Control Items			Test Items				(nf/nf, f/nf)
	color	shape	mood	<i>big</i>	<i>long</i>	<i>spotted</i>	<i>full</i>	
3	92	95	93	95	90	90	60	(40, 100)
4	98	100	100	95	95	85	65	(30, 100)
5	100	100	100	100	100	95	65	(30, 100)
adult	100	100	100	98	98	98	94	(88, 100)

From these results, we observe that children and adults were either at or near ceiling for the control and relative gradable pairs. They also performed well with the *spotted* pairs. It is clear, then, that both children and adults were able to accept or reject the puppet's request as appropriate. The pattern with the full pairs deviates, though. The means for the children on these pairs drop significantly. Looking at the two *full* pairs separately (as indicated in the parentheses) is illustrative. Participants unflinchingly gave the full container when there was one; it is with the non-full/non-full pair that children behaved differently from adults.

In order to find an explanation for these findings, the results from the two orders of presentation were examined. Here, an interesting pattern of interaction emerged. Every child who was shown the non-full/non-full pair early on in the sequence of object pairs gave the puppet the fuller of the two when he requested the full one. Those children who saw this pair much later in the sequence were more likely to reject the request and say that there was not a full one: 10 children rejected the request, while only 5 children accepted it and gave the fuller one, compared with 15 acceptances in the other condition. In addition, the three adults who gave the puppet the fuller container when the full one was asked for were also in the condition in which the non-full/non-full pair appeared early in the sequence. A close inspection of the orders of presentation revealed a confound in the stimulus presentation and suggested that at least one of three possible explanations could account for this pattern.

In the order in which the non-full/non-full pair appeared early on, this pair was preceded by only three pairs: two control pairs, and a relative gradable pair (*long*). It is possible that the presence of this relative pair influenced the way that children responded to the absolute gradable pair that followed soon afterwards, causing them to inappropriately use the context to define the standard of comparison.² It is also possible that children simply needed more experience with exemplars of gradable pairs in order to correctly reject the request. Finally, the children may have benefited from the presentation of a maximal standard (contrasted with lesser degree) provided by the full/non-full pair. To address the first possibility, we ran a follow-up condition, in which the order of the *long* pair and the non-full/non-full pair was simply reversed. We hypothesized that if children continued to accept the request and give the puppet the fuller member of the non-full/non-full pair in this new condition, too, then the relative gradable pair could not be held responsible for this pattern of responses.

4.2.1 Follow-up condition

17 children, representing the same three age groups were tested, along with ten adult controls. The ages of the children were as follows: six three-year-olds (*range* = 3;1 to 3;11, *M* = 3;5); six four-year-olds (*range* = 4;2 to 4;11, *M* = 4;6); and five five-year-olds (*range* = 5;2 to 5;10, *M* = 5;4). The data from four additional children were discarded because these children had difficulty with the task and/or responded incorrectly on more than half of the control items. The procedure was the same, but there was one difference in the order of presentation of the stimuli pairs—namely, the order of the *long* and non-full/non-full pairs were switched.

² This effect could be explained by a comparison process such as the structure mapping engine proposed by Gentner and Markman (1996, 1997). Under this account, features of the semantic representations for the relative and absolute adjectives would be aligned. Local matches would then be coalesced into structurally consistent mappings, and candidate inferences would be made from the base (e.g., *long*) to the target (e.g., *full*). Of course, if children use this kind of analogical learning for gradable adjectives in general, we would expect them to collapse the sub-classes of gradable adjectives into a single category and therefore not to learn the words appropriately. Thus, to the extent that this kind of structure mapping exists in this task, it must be a task effect and not indicative of the nature of adjective learning in general.

The results for this follow-up condition are presented in Table 2.

Table 2. Percentage expected responses

Age	<i>big</i>	<i>long</i>	<i>spotted</i>	<i>full</i>	(nf/nf, f/nf)
3	100	83	67 ³	50	(0, 100)
4	92	83	92	59	(17, 100)
5	100	90	90	70	(40, 100)
adult	100	100	95	85	(70, 100)

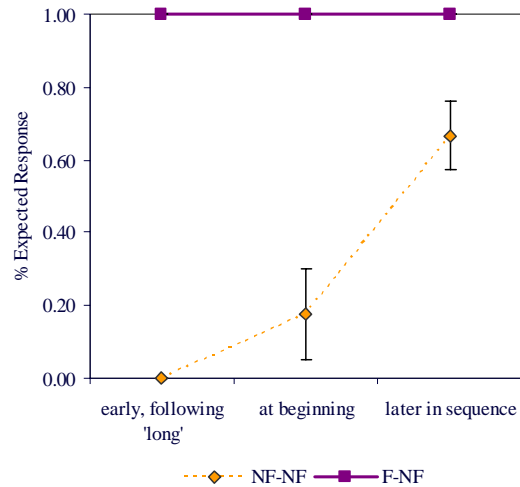
In this condition, too, children were inclined to give the puppet the fuller of the two non-full containers when he requested *the full one*. These results suggest that in the absence of a maximal standard, children (and some adults) can make room for context in their treatment of *full*. The data also present an argument against the hypothesis that it was the prior observation of a relative adjective that drove the original effect. Children's pattern of responses for the two *full* pairs for the three different ordering conditions is presented in Figure 1 on the following page. While the difference between the first two orders is not significant, the difference between either of these conditions and 'later' is (*early* (0 of 13, 0%)-*beg.*(3 of 17, 17.6%) $p=.24$, *early-later* (10 of 15, 66.7%) $p<.001$, *beg.-later* $p=.01$ by Fisher's Exact Test). Thus, seeing this pair later in the sequence (after the full/non-full pair) led to more accurate responses.

We also find evidence that those children who gave the puppet the fuller container when presented with two non-full containers have the proper semantic representation for *full*, but are influenced by the experimental situation. The first piece of evidence comes from measures of the subjects' reaction time. The experimental sessions were digitally videotaped. Where possible, videotapes were analyzed to determine reaction times.⁴ Although these measurements are still in progress, we refer to them here, because they are informative about these results. Initial measurements indicate that children were not only less accurate with this pair, but also took longer to make their decision than in the full/non-full case. The fact that reaction times increased for the non-full/non-full pair suggests that children's non-adult-like behavior with this pair is driven by an attempt to accommodate the situation to their adult-like representations. We would expect no difference in RTs between the two pairs if children had a representation for *full* that deviated from that of adults. The second piece of evidence that children have the correct lexico-

³ The reader may note that the percentage of expected responses for *spotted* is lower here than in the two original conditions. This difference may be explained by the fact that in this condition, the experimenter did not encourage the child to give the puppet both objects when both fit the description. The child was encouraged to give the puppet verbal feedback, but was not explicitly instructed to give the puppet both objects. Thus, a child might give the puppet one of the spotted disks, after noting both were spotted. Children patterned similarly with the control red-red pair.

⁴ Measurements include the difference between the onset or offset of the adjective in the puppet's request and one of three physical measurements from the child: the child's look to the targeted object (when one was given), the child's reach towards the object, and the child's touch of the object. Reaction times for the felicitous absolute, relative, and control pairs appear to be similar.

semantic representation for *full* comes from the results of a second study, presented in the next section, which investigated subject's judgments along a



scale.

Figure 1. Children's responses for two *full* pairs in each condition

5. Experiment 2

The purpose of the second study was to use children's scalar judgments as a means to assess their knowledge about the types of gradable adjectives. If children indeed make a semantic distinction between the three sub-classes of gradable adjectives, then their judgments about the extent to which the relevant properties hold of objects along a scale should reflect these distinctions.

5.1 Method

As in the first study, three groups of children were tested: three-year-olds (*range* = 3;3 to 3;11, *M* = 3;8), four-year-olds (*range* = 4;1 to 4;11, *M* = 4;5), and five-year-olds (*range* = 5;0 to 5;11, *M* = 5;5). There were 12 children in each age group. In addition, twenty-eight adult native speakers of English participated. The number of males and females for each age group in each condition was approximately equal. Eighteen children who participated in the first study also participated in the second study.⁵ For these children, the second task was administered no sooner than three weeks (and no more than four weeks) following the first study. The data from three additional children were discarded (two because these children had difficulty with the task, and one because of experimenter error).

⁵ Twelve children participated in only the first study. Seventeen children participated in only the second study.

Participants were presented with four sets of seven objects each, corresponding to the four gradable adjectives discussed above: a set of seven cubes (*big*), seven rods (*long*), seven containers ranging from full to empty (*full*), and seven disks ranging from plain to very spotted (*spotted*). For each set, the experimenter started at the positive pole and asked the participant, *Is this X?*, X being the target adjective. The predictions are clear. For the absolute maximal adjective *full*, there is only container that fits that description. For the absolute minimal adjective *spotted* all disks should be considered spotted except the plain one. Finally, for the relative sets, there should be a shift somewhere along the middle of the scale; at some point it should feel more appropriate to describe the cubes or rods as small or short instead of big or long.

5.2 Results and discussion

The scalar judgments for the four sets are presented in Figure 2 (adults) and Figure 3 (children). Participants’ judgments along the scale clearly reflect a distinction among the three sets of gradable adjectives.

Figure 2. Adults’ scalar judgments

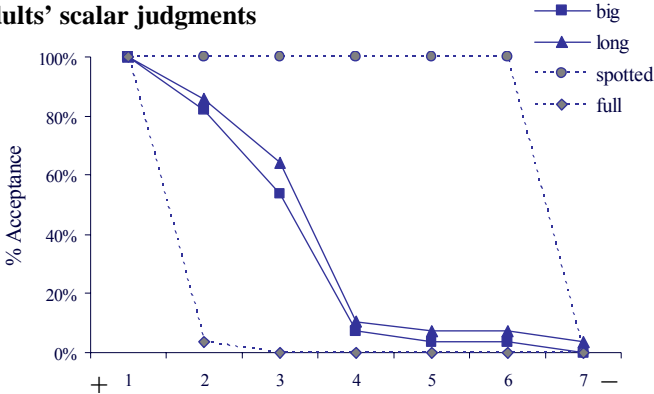
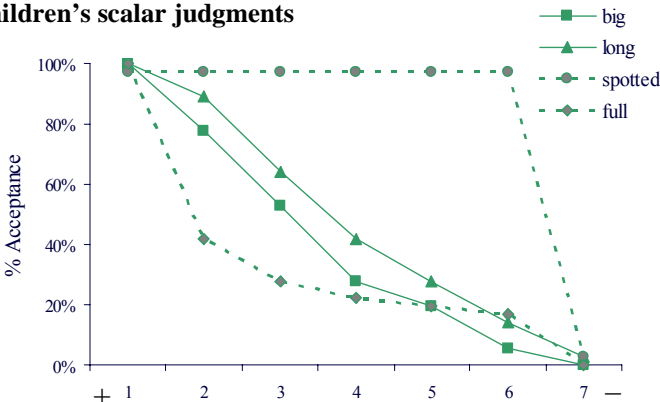


Figure 3. Children’s scalar judgments



This three-way distinction can be highlighted by focusing attention on the second through the fourth items. Whereas for both populations, the first item (at the positive pole) was judged to possess the relevant property and the last item (at the negative pole) to lack it, between the poles, the treatment of the adjectives diverges.

Adults recognized that the absolute adjectives were oriented towards the minimal or maximal endpoint, judging all but the last disk to be spotted and only the first container to be full. For the relative adjectives, there is a clear (almost categorical) shift in acceptance between the third and fourth items. Children pattern with adults with respect to *spotted*. While the shape of the curve is similar for *full*—children are much less likely to judge the containers towards the negative pole to be full, many think that *full* means something like *is filled (to a sufficient degree)*. Finally, their treatment of relative adjectives is similar but suggests a developmental difference: children's rate of acceptance for *big* and *long* also declined in the middle of the scale, but the slope between items three and four is not nearly as steep as for adults, suggesting that the distinction sharpens during the course of development.⁶

Why do children pattern differently from adults with respect to *full*? Recall that a subset of children participated in both the Presupposition Assessment Task (PAT) and the Scalar Judgments Task (SJT). This subset of 18 children included six three-year olds, seven four-year-olds, and five five-year-olds. Their results were further analyzed in order to make a cross-study comparison. 11 of these 18 children (61.1%, across age groups) incorrectly gave the puppet the fuller container in the PAT when shown the non-full/non-full pair. It could be that these children did not fully understand the word *full*. However, the results from the SJT indicate that this is not the case: 8 of these 11 children (72.7%) indicated that only the first container (the full one) was full. 2 additional children (18.2%) said that only the first two containers were full, and the last child said that the first three containers were full. Now, the members of the non-full/non-full pair were items #4 and #6 in the scale; none of the children who gave the puppet the fuller container in the PAT judged container #4 as full in the SJT. The 7 remaining children (of the 18) who correctly rejected the request when shown the non-full/non-full pair did not differ in their estimation of fullness: 4 of these 7 (57.1%) said the first container was full, 2 (28.6%) said that the first two containers were full, and the last child said that the first three containers were full. Therefore it cannot be that judgments seen in the SJT can be responsible for the performance observed in the PAT.

The SJT is also illuminating with respect to effects of presentation order. In this study, too, subjects were randomly assigned to one of two orders. In one order, the relative gradable sets were presented first (*big, long*), followed by the absolute gradable sets (*full, spotted*); in the second, the sets

⁶ A one-way repeated measures ANOVA evaluating the difference between items 2-3 and 4-5 confirms both the three-way distinction and the differences in the categorical nature of judgments between children and adults. For *spotted*, both adults and children $F=0$ and $p=1$. For *full*, adults $F=.994$ $p=.326$, but children $F=8.6644$, $p=.006$. For *big* and *long* respectively, adults $F=68.4878$ and 72.9093 , children $F=38.042$ and 41.6667 , all four $p<.0001$.

were reversed (*full, spotted*, followed by *big, long*). 11 of the 15 children (73.3%, across age groups) who judged a container other than the first one to be full saw the relative sets first. Put another way, only 4 children (26.7%) who judged a non-full container to be full were presented with the *full* set first. This fact is consistent with what we observed in the follow-up condition in the PAT, but still leaves open the possibility that the relative sets had an influence. We can therefore conclude from these data that children make a distinction between the sub-classes of gradable adjectives and that their judgments are influenced by what else they have seen in the task.

6. General Discussion

We have demonstrated that children make a distinction between sub-classes of gradable adjectives. They recognize that determining the standard of comparison for relative gradable adjectives such as *big* and *long* is context-dependent in a way that assigning the standard for absolute gradable adjectives is not. Furthermore, children make a distinction between absolute gradable adjectives with a minimal standard (e.g., *spotted*) on the one hand, and those with a maximal standard (e.g., *full*) on the other. These differences can be accounted for by positing an adult-like semantic representation for these adjectives. For *full*, we found that children are not completely adult-like, and that different factors might account for their performance. It appears that they have the correct corresponding scalar structure, but have difficulty assigning the maximal standard. Current experimentation with other maximal absolute gradable adjectives (e.g., *straight*) should prove informative.

The combined results from the two studies provide evidence confirming the existence of a relative/absolute for gradable adjectives in adult grammar. Previous psycholinguistic work (cf. Rips and Turnbull 1980) has suggested that adults make reference to a contextually-determined standard when processing sentences containing a relative adjective. The current work demonstrates that adults distinguish between two sub-classes of gradable adjectives—one that allows for a standard to be shifted relevant to the context at hand, and another that does not.

Finally, the results of the first experiment also offer enlightening evidence about children's knowledge of the presuppositions of existence and uniqueness for singular definite descriptions. When the puppet asked for *the X one*, they systematically gave him one object for the relative gradable pairs, and rejected his request when shown the pairs in which neither or both objects fit the description. Furthermore, their verbal responses provide anecdotal evidence that they knew this request highlighted *one* object. Children made comments such as, *The very big one?*, *Oh, but I have TWO red ones!*, *He thinks there must be two different colors!*, and *What red one? He should say what shape!* These observations suggest the presence of an adult-like semantic representation of *the* for children as young as three years of age—which earlier production-oriented methods had difficulty uncovering, but for which a judgment task, designed specifically to assess the presuppositions of definiteness, is ideal.

References

- Bartlett, E. (1976). Sizing things up: the acquisition of the meaning of dimensional adjectives. *Journal of Child Language* 3(2): 205–219.
- Clark, E. (1972). On the child's acquisition of antonyms in two semantic fields. *Journal of Verbal Learning and Verbal Behavior* 11: 750–758.
- Clark, E. (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 Acquisition of Language*. New York: Academic Press, pp. 65–110.
- Ebeling, K. and S. Gelman. (1988). Coordination of size standards by young children. *Child Development* 59(4): 888–896.
- Ebeling, K.S. and S. Gelman. (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(3): 369–384.
- Eilers, R. E., Kimbrough Oller, D. and J. Ellington. (1974). The acquisition of word-meaning for dimensional adjectives: the long and short of it. *Journal of Child Language* 1: 195–204.
- Gelman, S., and E. Markman. (1985). Implicit contrast in adjectives vs. nouns: implications for word-learning in preschoolers. *Journal of Child Language* 6(2): 125–143.
- Gentner, D. and A. Markman. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition* 24(2): 235–249.
- Gentner, D. and A. Markman. (1997). Structure mapping in analogy and similarity. *American Psychologist* 52(1): 45–56.
- Gentner, D., and M.J. Rattermann (1998). Deep thinking in children: The case for knowledge change in analogical development. *Behavioral and Brain Sciences* 21: 837–838.
- Kennedy, C. and L. McNally. (2005). Scale Structure, Degree Modification, and the Semantics of Gradable Predicates. *Language* 81 (2): 1–37.
- Kennedy, C. (1999). *Projecting the adjective: the syntax and semantics of gradability and comparison*. New York: Garland. (1997 Doctoral Dissertation, UC Santa Cruz)
- Klein, E. (1991). Comparatives. In A. von Stechow and D. Wunderlich (eds.) *Semantics: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter, pp. 673–691.
- Klibanoff, R.S. and S.R. Waxman. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development* 71: 649–659.
- Mintz, T. and L. Gleitman (2002). Adjectives really do modify nouns: the incremental and restricted nature of early adjective acquisition. *Cognition* 84(3): 267–293.
- Nelson, K. and H. Benedict. (1974). The comprehension of relative, absolute, and contrastive adjectives by young children. *Journal of Psycholinguistic Research* 3: 333–342.
- Rips, L and W. Turnbull. (1980). How big is big? Relative and absolute properties in memory. *Cognition* 8: 145–174.
- Rotstein, C. and Y. Winter. (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12: 259–288.

- Ryalls, B. (2000). Dimensional adjectives: factors affecting children's ability to compare objects using novel words. *Journal of Experimental Child Psychology* 76: 26–49.
- Sandhofer, C. and L.B. Smith. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology* 35 (3): 668–679.
- Sapir, E. (1944). Grading: a study in semantics. *Philosophy of Science* 11:93–116.
- Sera, M. and L.B. Smith. (1987). Categorical and relative interpretations of “big” and “little.” *Cognitive Development* 2: 89–112.
- Smith, L.B., Cooney, N. and C. McCord. (1986). What is “High”? The development of reference points for “High” and “Low.” *Child Development* 57: 583–602.
- Unger, P.K. (1974). *Ignorance*. Oxford: Clarendon Press.
- Waxman, S.R. and A.E. Booth. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology* 43: 217–242.

Department of Linguistics
Northwestern University
2016 Sheridan Road
Evanston, IL 60208
USA

k-syrett@northwestern.edu
<http://www.ling.northwestern.edu/~syrett/>