Available online at www.sciencedirect.com
SCIENCE DIRECT®

COGNITION

Cognition 90 (2003) 163–191
www.elsevier.com/locate/COGNIT

Individuation of objects and events: a developmental study

Laura Wagner*, Susan Carey

Harvard University, Cambridge, MA, USA

Received 28 August 2002; revised 23 March 2003; accepted 30 June 2003

Abstract

This study investigates children's ability to use language to guide their choice of individuation criterion in the domains of objects and events. Previous work (Shipley, E. F., & Shepperson, B. (1990). Countable entities: developmental changes. *Cognition*, *34*, 109–136.) has shown that children have a strong bias to use a spatio-temporal individuation strategy when counting objects and that children will ignore a conflicting linguistic description in favor of this spatio-temporal bias. Experiment 1 asked children (3-, 4-, and 5-year-olds) and adults to count objects and events under different linguistic descriptions. In the object task, subjects counted pictures of familiar objects split into multiple pieces (as in Shipley, E. F., & Shepperson, B. (1990). Countable entities: developmental changes. *Cognition*, *34*, 109–136.) and described either using an appropriate kind label (e.g. "car") or the general term "thing". In the event task, subjects watched short animated movies consisting of a goal-oriented event achieved via multiple, temporally separated steps. The events were described either with an appropriate telic predicate targeting the goal (e.g. "paint a flower") or with an atelic predicate targeting the steps in the process (e.g. "paint") and the subjects' task was to count the events. Relative to adults, children preferred a spatio-temporal counting strategy in both tasks; there was no difference among the three groups of children. However, children were able to significantly change their counting strategy to follow the linguistic description in the event but not the object task. Experiment 2 extended the object task to include counting of other types of non-spatio-temporal units such as sub-parts of objects and collections. Results showed that children could use the linguistic descriptions to guide their counting strategy for these new items, though they continued to show a bias for a spatio-temporal individuation strategy with the collections. We suggest potential cognitive origins for the spatio-temporal individuation bias and how it interacts with children's developing linguistic knowledge.

Keywords: Individuation; Objects; Events; Developmental study; Aspect

* Corresponding author. William James Hall, 33 Kirkland Street, Harvard University, Cambridge, MA 02138, USA.
E-mail address: lwagner@wjh.harvard.edu (L. Wagner).

0022-2860/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0010-0277(03)00143-4

1. Introduction

The world provides us with an undifferentiated stream of experience, full of sounds and sights, surfaces and motions. Making sense of it requires (among other things) breaking that stream into individual units such as tables, dinners and events of eating. There are many bases for creating units from the stream of experience. One can divide the world into spatio-temporally defined units such as bounded coherent objects that trace spatio-temporally continuous paths or units of sound segregated from the background by the equivalent of edge discontinuities (e.g. a clap). Alternatively, one can discover units on the basis of transitional probabilities formulated over a lower level of analysis (as in Saffran, Aslin, and Newport's (1996) demonstrations that words are segmented from a spatio-temporally continuous stream on the basis of co-occurrences of syllables or phonemes). Finally, one can also impose units on the basis of conceptually mediated kind or causal analysis. Events of building a house are complete when the house is finished. An animal ceases to exist when it dies, in spite of the spatio-temporal continuity of its body.

Languages both reflect and impose criteria of individuation: successful communication requires that speakers of a given language individuate the world into the same units as do the other speakers of that language. This observation raises several related developmental questions: how do children come to individuate the undifferentiated stream of experience into the same units as do adults, when and how do they learn their language's means of doing so, and under what circumstances are children and adults guided by others' linguistic descriptions to focus on one possible unit of experience over another? This study focuses on the third of these questions as an initial wedge into the first two, building on earlier seminal studies by Shipley and Shepperson (1990) and extending them to event representations.

Objects and events are generally encoded by different sorts of linguistic units (typically nouns and predicates, respectively). Object individuation is implicated in the count/mass distinction whereas event individuation is implicated in aspectual marking on verbs. Nonetheless, there are deep parallels between the ways humans individuate objects and events.

1.1. Previous work on object individuation

By late in the first year of life, infants control various different criteria for object individuation, and some aspects of language already influence the infants' division of the stream of experience into units. A wealth of data suggest that spatio-temporal criteria are important to individuation of objects, even very early in infancy. Young infants analyze spatio-temporal continuity in establishing representations of one object or two. For example, objects seen emerging, one at a time, from opposite sides of two screens but never appearing in the middle are analyzed as two objects, not one (Spelke, Breinlinger, Macomber, & Jacobson, 1992; Xu & Carey, 1996). In addition, by 12 months of age, at least, infants are able to bring kind membership to bear on the problem of object individuation. Xu and Carey (1996) showed that 12-month-old infants shown objects of two different kinds emerging one at a time from opposite sides of a single screen established representations of two numerically distinct objects. In this case, there was no spatio-temporal evidence for distinct individuals. Twelve-month-olds were similarly successful at using kind information to guide the number of times they reached into an opaque box for objects they had seen emerge from and return to the box one at a time (Van de Walle, Carey, & Prevor, 2000). Xu (2002) has shown an effect of labeling on object individuation in 9-month-olds. In the paradigm of Xu and Carey (1996), if the objects are labeled with contrastive labels as they are taken from and replaced behind the screen ("Look a duck, look a car"), 9-month-olds establish representations of two objects, whereas if labeled with a single label ("Look a toy, look a toy"), infants fail, as they do with no label or with contrastive associated sounds.

Shipley and Shepperson (1990) used counting as a wedge into preschool children's criteria for individuation of objects, asking whether language influenced what units children counted. They presented children with displays of objects, some of which were broken. For example, one display might be three intact cars and one car broken into two distinct pieces. Even when asked to count "the cars", children as old as 5 years would often separately count each separate piece of the broken car, along with each intact car. This is a surprising result; children have known the words in Shipley and Shepperson's study (e.g. "car") for years by age 5, and the difference between cars and other things has been a cue for individuation by 12 months of age. Still, when considering a scene consisting of whole cars and split cars, children broke it into spatio-temporally determined units, not units determined by the basic level kind term "car". Further, children were not influenced by the label – the same pattern emerged whether asked to count "the cars" or "the things". Thus, spatio-temporal cues persist as a powerful basis of object individuation into childhood. Preschool children who have mastery of several bases of object individuation will default to spatio-temporal cues in at least one kind of difficult counting situation. Adults, in contrast, did not count pieces of cars as "cars", and were significantly influenced by the label.

1.2. Previous work on event individuation

There has been much less work on event individuation than on object individuation. There is some evidence, though, for a spatio-temporal specification of individual events that is roughly parallel to the spatio-temporal specification of individual objects. Wynn (1996) showed that infants habituated to two jumps of a puppet dishabituated when shown three jumps, and vice versa. She controlled for a variety of co-occurring factors such as total event duration and inter-jump intervals, strengthening the case that it was in fact the number of individual jumps that the infants were enumerating. The spatio-temporal information infants used was subtle, going beyond mere temporal pauses between motion events: infants succeeded if the puppets were still, and if the puppets continued to move in a waggling fashion between jumps. Apparently infants analyze discontinuities in kinds of motion. Later studies by Sharon and Wynn (1998, 2000) have argued for other factors such as cyclicity of motion and tangent discontinuity of motion as critical for parsing events into individual units. This information is still spatio-temporal, of course.

As in the case of objects, events are also individuated on the basis of higher level conceptual criteria, such as the intentions and goals of the actor. There is strong evidence that infants represent events in terms of goal states by the end of the first year (Gergely, Nadasdy, Csibra & Biro, 1995; Csibra & Gergely, 1996; Biro, Gergely, Koos & Csibra, 1996; Csibra, Gergely, Brockbanck, Biro, & Koos, 1998; Woodward, 1998), but it is unclear when infants use this information for individuation purposes. In one suggestive study, Baldwin, Baird, Saylor, and Clark (2001) habituated 10-month-old infants to short video clips of a woman performing a goal-directed action, such as picking a towel up off of the floor. During the test trials, pauses were inserted into the movie, either at the moment the goal was achieved or else at some point before or after the goal. Infants looked longer when the breaks did not coincide with the goal state, suggesting that the goal was a natural break point in the infants' representation of the stream of experience.

1.3. Language and Individuation

The world may be ambiguous, but linguistic descriptions commit the speaker to one among all of the available possible interpretations. This generalization applies to ambiguity regarding individuation as well as ambiguity regarding categorization. Within the object domain, the role of language in individuation is relatively straightforward. Count nouns provide criteria for individuation and numerical identity (see Macnamara, 1986). Whereas Xu (2002) found that infants inferred from contrasting labels that there were two individuals in an event, the role of count nouns in providing cues for individuation is considerably more subtle. Different kinds of count nouns provide different kinds of individuation criteria – contrast "cow" and "body" and "calf" and "pet" and "herd" and "leg". A calf is not a different individual from the cow it becomes, and a cow does not cease to exist when it ceases to be a pet, but a cow does cease to exist when it dies, in spite of the continued existence of its body. A herd is a collective individual; spatio-temporally individuated cows make up its parts. A leg is an individual part of a spatio-temporally specified cow. Experiment 2 begins to explore the flexibility with which preschool children impose different criteria of individuation upon objects, groups of objects, and parts of objects, and whether Shipley and Shepperson's finding that units preschool children count are uninfluenced by the "fork/thing" contrast extends to other linguistic contrasts ("cow/herd" and "cow/leg").

The linguistics of event individuation differs greatly, on the surface, from the linguistics of object individuation. The event equivalent of a count noun is a telic predicate (Bach, 1986; see also Smith, 1991). Telic predicates specify the end-point (or *telos*) of the event they describe and this end-point defines the criteria for individuation and numerical identity for that event. Thus, the telic predicate "build a house" specifies that the event ends with a house; an event that stopped short of a completed house would not constitute an entire instance of the event described. Telic predicates can be contrasted with atelic ones, which do not specify the result of an event and tend to describe actions or processes, such as "work" or "play". Events described with atelic predicates may be individuated into units spatio-temporally; some sort of pause between bouts of playing may serve to establish individuals. The linguistic marking of telicity depends on the interaction of many elements in a predicate, including the verb ("The vase breaks" is telic while "The vase moves" is not), the arguments of the verb ("The girl ate an ice-cream cone" is telic while "The girl ate" and "The girl ate ice-cream" are not), as well as adjuncts in the predicate, such as prepositional phrases ("The butterfly flew to a tree" is telic while "The butterfly flew" is not).

Given the greater linguistic complexity of describing event individuation, relative to object individuation, one might expect that children would be even less likely to be influenced by the linguistic description of an event in deciding how to break it into countable units than they are in the case of objects. However, there is some evidence that children know from early on that the telic/atelic distinction ought to have some sort of linguistic implications. Examinations of children's earliest productions of verb morphology (such as past tense marking and progressive "-ing" marking in English) show that children typically use these markers in a restrictive way, and that the restriction depends on the probable telicity of the verb in question. Verbs which, for adults, typically form the center of a telic predicate ("break", "make") appear with one type of morphology (in English, the past tense) while verbs which typically yield atelic predicates ("play", "ride") appear with the other (in English, the progressive) (Antinucci & Miller, 1976; Bloom, Lifter, & Hafitz, 1980; Bronckart & Sinclair, 1973; Shirai & Andersen, 1995). However, the morphology that children restrict in this way is not in fact dependent on the telicity of the predicate in the adult language. Thus, in English, we use past tense and progressive markers freely with both telic and atelic predicates regularly (i.e. "broke", "breaking", "played" and "playing" are all perfectly grammatical).

Investigations into children's knowledge of how telicity is actually marked in natural language have been very limited. There has been one set of studies which looked at telicity knowledge in children acquiring English and Dutch (van Hout, in press). van Hout and colleagues presented children with completed and incomplete versions of the same scenario (e.g. one elephant who drank all the water in her bucket and one who left the bucket half full) and then asked children whether telic or atelic descriptions could be used (e.g. "Did this elephant drink?", "Did this elephant drink her bucket of water?"). Children as old as 5 years did not reliably apply the telic description only to the completed version of the event.

These results suggest that it may take children years between the time that they understand something about telicity and when they understand the details of how their native language marks it. Some caution is required in interpreting these results, however. van Hout used a very small set of predicates (all involved eating and drinking), and children were asked within an item to assess both scenarios with respect to a particular sentence. Children's difficulties may be restricted to just these types of verbs, or arise from some general confusion about what to do when asked the same questions repeatedly within a trial. It remains to be seen if children might succeed with telicity marking given a simpler task and different sorts of events. Moreover, it also remains to be seen when children understand the implications of the telic/atelic distinction for event individuation.

Wynn (1990) showed that 2- and 3-year-old children can count entities that are not objects (claps, beeps and jumps). These units were spatio-temporally defined. It is not known whether young preschoolers can also count more conceptually determined event units, specified by goals, for example. Nor is it known whether they can use the language with which an event is described as a basis for one construal over another. If the results with events parallel those with objects, then preschoolers should be able to count spatio-temporally defined units only, no matter how the events are described. In Experiment 1 we replicate Shipley and Shepperson's object counting studies and extend the methodology to begin to explore preschool children's event individuation and their mastery of linguistic devices that determine how the stream of experience is to be broken into event units.

2. Experiment 1: individuation of objects and events

In this study, participants are asked to count objects and events. One counts individuals, be they objects or events, and so the process of counting is diagnostic of what the child considers to be an individual. Note that counting is a valid measure even for children who have not fully mastered the number system; all that is required is that children assign number labels in one-to-one correspondence with the individuals being enumerated. One-to-one correspondence guides counting almost from its earliest inception in 2-year-olds. It certainly is in place by age 3, the age of the youngest children tested in Experiment 1 (Gelman & Gallistel, 1978).

Both the object and event conditions of Experiment 1 present participants with ambiguous displays, such as two forks each broken into two pieces or two events in each of which a rabbit hops two times and then hops into a hole. For each participant, the linguistic description of the events is manipulated. For half of the object stimuli, the label specifies individuals according to kind of object ("fork" in this example). For the other half, the label specifies spatio-temporal individuals ("thing", following Shipley & Shepperson, 1990). For half of the event stimuli, the language specifies a goal-defined individual ("hop into the hole"), and for the other half a spatio-temporal individual ("move"). All materials are designed to allow both sorts of individuation and to give different counts according to the different criteria (spatio-temporal vs. kind/goal). In the above examples, there are two forks but four things and the rabbit hops into the hole twice but moves six times.

The object portion of the task is essentially a replication of one of the conditions in Shipley and Shepperson (1990) – the homogenous, familiar, aligned condition. Instead of counting real objects, children are asked to count objects depicted on a computer screen. As we do not expect the change in presentation format to influence the results, we predict a straightforward replication of Shipley and Shepperson's results. Children should prefer to count the spatio-temporally defined individuals irrespective of the label used instructing them what to count. This preference should decline with age and the sensitivity to the linguistic label should increase with age.

The event condition closely parallels the object condition. Indeed, the object trials were restricted to cases where the broken objects were aligned because the temporal parts of events must be adjacent to each other, as events unfold in time. Four types of goal-defined events were used: events of creation where the goal is the created object (paint a flower, build a house), events of destruction where the goal is the elimination of an object (eat an ice-cream cone, break a vase, pop a balloon), events of directed motion where the goal is a destination (butterfly flies to a tree, rabbit hops into a hole, boy swims across a river), and events where an object changes from one state into another state which is the goal of the event (blow up a balloon, close a closet door).

Four types of syntactic devices were used to specify the goal (i.e. four kinds of telic predicates were used): transitive sentences where the event's goal is specified by the direct object of the sentence (paint a flower, pop a balloon, blow up a balloon), prepositional phrase sentences where the event's goal is specified by the object of a preposition (hop into a hole, fly to a tree, swim across the river), passive sentences where the event's goal is the subject of a passive sentence (ice-cream cone was eaten, the house was built) and unaccusative intransitive sentences in which the event's goal is specified by the subject of the sentence (the vase broke, the closet door closed). The latter two syntactic types (passives and unaccusatives) are more complex syntactically and, at least for the passives, are generally acquired later than the former two types (transitive and prepositional phrases) (Brown, 1973). They have the potential virtue of specifying the event's goal in the subject position. The syntactic realizations of the atelic predicates (the atelic descriptions) were all intransitive sentences with the exception of one transitive item ("The bird poked the balloon"). Every test sentence (telic and atelic) passed standard linguistic tests for their telicity value (cf. Dowty, 1979; Smith, 1991). For example, every telic sentence can be used with the adverbial phrase "in a minute" to indicate it took a minute to accomplish the goal; similarly, every atelic sentence can be used with the phrase "for a minute" to indicate the duration of the event.

2.1. Methods

2.1.1. Subjects

Three groups of children were tested; there were 16 children in each age group. The 3-year-old group had a mean age of 3;6 (ranging in age from 3;0 to 3;10); the 4-year-old group had a mean age of 4;4 (ranging in age from 4;0 to 4;10); the 5-year-old group had a mean age of 5;4 (ranging in age from 4;11 to 5;11). The data from an additional 12 subjects were discarded because these children failed to cooperate on more than half the trials. Subjects were children in the New York City area who were either brought into the lab to participate or else tested in a local daycare center. According to parental report, all subjects had English as their primary (and in most cases, only) language. In addition, 16 adult native speakers of English participated.

2.1.2. Stimuli

2.1.2.1. Objects. The object stimuli were created and displayed using a Microsoft Powerpoint program. Pictures of 12 familiar objects were selected from clip-art images and then the images were split into two or three approximately equal pieces. Each test item contained one, two, or three copies of an image. For eight of the items, all of the copies were split (and all were split into the same number of pieces). The remaining four items contained one split item and one or two whole items. See Appendix A for an example stimulus and a list of all 12 stimuli.

2.1.2.2. Events. The event stimuli consisted of ten short animated movies and were created with the Macromedia Director animation program and displayed using Quicktime on a laptop computer. Each movie began with a display of theatrical curtains accompanied by a short musical phrase and ended with another display of curtains and music, and the words "The End". The curtain scenes framed the test movies and defined the domain of counting for each movie (children who did not do so spontaneously were encouraged to keep counting until the movie was over). Each movie contained one or two instances of an event with a definite goal (building a house, eating an ice-cream cone). Each instance of an event was achieved through two or three temporally distinct sub-actions (e.g. the girl rests between building each half of the house, it takes three bites to eat the ice-cream cone). The movies lasted from ~8 to ~20 seconds each. See Appendix A for a list of the ten events along with the two linguistic descriptions (telic and atelic) used.

2.1.3. Procedures

The children were told that they were playing a counting game. All children began with the object task and then proceeded to the event task. For the object task, they were shown the Power-point slides on the computer one at a time and asked, "How many X are here?" The X term was either an appropriate kind label description of the object or else the word "thing". The items were presented in a fixed order. The descriptions were blocked so that the first six items all received the same type of description and the last six items received the opposite type of description. Which type of description came first was counter-balanced across subjects. When the type of description switched, subjects were told, "Now listen carefully, I'm going to change how I'm talking about these pictures". All subjects heard this warning, but it was included primarily to prevent children who heard the same label ("thing") for all of the early trials from ignoring the linguistic description through the remaining trials.

Following the object task, children were told that the game was going to change and that now they would count what happened in movies. Before being presented with each movie, they were asked, "How many times X? Let's watch!" The X term was an appropriate description of the event which targeted either the goal of the event or the temporally discrete process actions. The question was always presented before the movie was started, and it was uttered at least twice before the opening curtains ended. The question was repeated again at the end of the movie (with the tense of the auxiliary changed to the past tense) and the subject's answer was recorded. Children who were inattentive during the movie or who refused to respond at this point were shown the movie a second time and queried in the same manner as before. There were two fixed orders of presentation of the movies. The descriptions were blocked so that the first five items all received the same type of description and the last five items received the opposite type of description. Which type of description came first was counter-balanced across subjects and orders. Since every description of the events was different from every other one, children were not signaled about the change to the different description type.

Children who were reluctant to count or who had unorthodox counting systems were encouraged to also point to the computer screen during their counting. The locations and/or timings of children's points were used to resolve any ambiguities in their verbal answers.

2.1.4. Coding

Children's counting was classified according to whether the individuals specified either a kind/goal criterion or a spatio-temporal criterion. The vast majority (over 92%)[1] of answers that children gave were clearly of one of these two types. Minor errors in one-to-one correspondence (e.g. counting 5 or 7 marker pieces when there were actually 6) were ignored. Similarly, children with non-standard count-lists (e.g. they regularly skipped the number *3*) were credited with the equivalent standard count.

2.2. Results

2.2.1. Analysis of object individuation

There were two major findings in the object individuation case. First, children overwhelmingly adopted a spatio-temporal criterion for their individuation while adults used a kind-based criterion. Second, although linguistic description exerted a reliable influence on responses, children's spatio-temporal strategy and adults' kind-based strategy largely persisted across linguistic description types. The mean use of the kind-based individuation criterion for the two kinds of linguistic descriptions for the different age groups is shown in Fig. 1.

An ANOVA was performed over the percentage of times subjects used a kind-based counting criterion with age group (3-, 4-, and 5-year-olds and adults) as the between subjects factor and linguistic description (spatio-temporal "thing" and specific kind label) as the within subjects factor.[2] This analysis confirmed that there was a main effect of age group ($F(3, 60) = 113.5$, $P < 0.0001$) and of description type ($F(1, 60) = 8.37$, $P < 0.005$) but no interaction between the variables.

Post-hoc analyses on the age groups (Tukey's HSD) revealed that there were no significant differences among the three child age groups but that children at all ages were significantly different from the adults. The overall mean rate of kind-based counting was 19% for 3-year-olds, 11% for 4-year-olds, 35% for 5-year-olds, and 90% for adults. The source of the difference between the children and adults is the rate across linguistic descriptions at which the groups count based on spatio-temporal criteria: children rely primarily on the spatio-temporal cues regardless of linguistic description while adults seem to individuate based on kind, also largely regardless of the linguistic description.

The main effect for description type was caused by a very small general trend for kind descriptions to lead to more kind-based counting (the overall mean rate of kind-based

[1] The remaining answers fell into three categories: combination counting, sub-part counting, and completely uncodable. In combination counting (about 5% of total trials), children counted every spatio-temporal unit but re-started their count for each separate kind or goal. Thus, on the markers picture, children might count "1, 2, 3" pointing to the first marker and then count again "1, 2, 3" pointing to the second marker. Because this pattern is sensitive to both the object kind construal/the event goal construal as well as the spatio-temporal breaks within each of these, items receiving this pattern were given half-credit in each category. In the sub-part counting (about 1% of total trials), children counted not the spatio-temporally discrete objects or events but distinctive sub-parts, such as the spokes of the bicycle wheel, or every motion of the girl's arm when painting. As this technique seemed to depend on a kind of spatio-temporal parse (it certainly did not depend on kind label or goal information), items counted in this way were credited as using a spatio-temporal criterion. In the remaining 2% of trials, children's behavior was either totally non-numeric (e.g. "a lot") or else otherwise unclear. These trials were omitted from all analyses.

[2] In this and all other analyses, an initial analysis was conducted checking for effects and interactions with order. In no cases were any effects found and so this factor has been omitted from the analyses reported here.

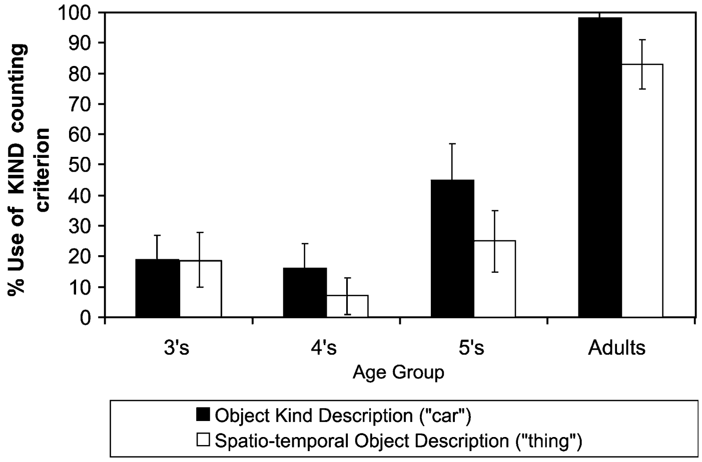


Fig. 1. Experiment 1, object trials. Mean use of the kind-based individuation criterion.

counting given specific kind labels was 45% and given spatio-temporal "thing" was 33%) and inspection of the means shows that numerically, all groups show a trend to individuate based on kind more often given kind descriptions as opposed to when given the general label "thing". Planned comparisons within each age group for the effect of description type, however, show that none of the groups counted reliably differently depending on linguistic description, although the adults and 5-year-old groups did so marginally: adults' $t(15) = 2.01$, $P < 0.06$; 5-year-olds' $t(15) = 1.87$, $P < 0.081$. The numerical trend (and the marginal effects) appears to be driven by a small number of individuals in each age group. The number of subjects (out of 16) who used a kind-based counting criterion more often (whether significantly or not) given a kind label than a spatio-temporal label was the following: 3-year-olds, $N = 3$; 4-year-olds, $N = 4$; 5-year-olds, $N = 6$; adults, $N = 5$. The dominant pattern of counting for subjects in every age group was simply to use the same criterion for individuation for every trial irrespective of linguistic description. All that changed with age was which criterion was used: children opted for the spatio-temporal analysis while adults used the kind-based analysis.

2.2.2. Analysis of event individuation

The mean use of the goal-based individuation criterion for the two kinds of linguistic descriptions for the different age groups is shown in Fig. 2. As predicted, children are more likely than adults to count spatio-temporally determined events. Adults always count goal-defined events when given a telic description of what to count, whereas children did so only about 60% of the time. However, in contrast to the object condition, the event results show that children (and adults) definitely select their individuation criteria between spatio-temporal and goal-based depending on the type of linguistic description provided. Thus, although the children show a general preference for using a spatio-temporal individuation strategy, they do overcome that preference to a significant extent when the linguistic description focuses on the goal of the event.

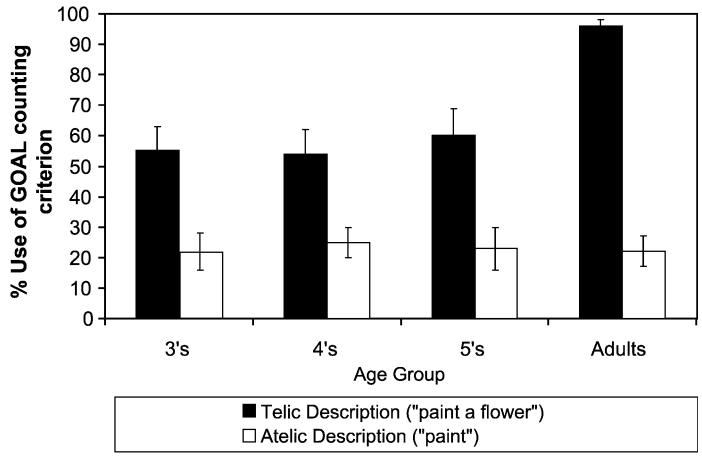


Fig. 2. Experiment 1, event trials. Mean use of the goal-based individuation criterion.

An ANOVA was performed over the percentage of times subjects used a goal-based counting criterion with age group (3-, 4-, and 5-year-olds and adults) as the between subjects factor and linguistic description (atelic and telic descriptions) as the within subjects factor. This analysis confirmed that there was a main effect of age group ($F(3, 60) = 2.77$, $P < 0.049$) and of description type ($F(1, 60) = 159.7$, $P < 0.001$) and also a significant interaction between the variables ($F(3, 60) = 9.34$, $P < 0.0001$).

Post-hoc analyses on the age groups revealed that there were no significant differences among the three child age groups but that children at all ages were significantly different from the adults. The overall rate of goal-based individuation was 39% for 3-year-olds, 40% for 4-year-olds, 41% for 5-year-olds, and 59% for adults. The main effect of description type reflected the fact that subjects were more likely to use a goal-based criterion with telic descriptions than with atelic descriptions (the mean goal responses given telic descriptions was 66%, and with atelic descriptions 23%). Moreover, planned comparisons within each age group for the effect of description show that every age group used a significantly different counting criterion as a function of description type: 3-year-olds' $t(15) = 5.0$, $P < 0.001$; 4-year-olds' $t(15) = 4.0$, $P < 0.001$; 5-year-olds' $t(15) = 4.5$, $P < 0.001$; adult's $t(15) = 15.03$, $P < 0.001$. The source of the interaction effect (which is also the source of the age effect) is that given a telic description, adults are at ceiling at using a goal individuation criterion while the children are not. As can be seen from Fig. 2, the children did not differ from adults in the likelihood to count spatio-temporally determined individuals when given atelic descriptions.

One potential point of concern in the data is the fact that children's overall rate of using the goal-based individuation criterion with telic predicates hovers near 50% ($M = 57\%$). Clearly this pattern does not reflect any sort of chance-like behavior. Recall that the classification of a response as being goal-based depends not on a forced choice between two options, but on subjects' specific numerical responses to each item. Moreover, the fact that there is a main effect of description type speaks to the fact that the experimental manipulations are determining the pattern of behavior. Nevertheless, it is possible that this pattern could arise from a bimodal distribution among the subjects such that roughly half of them succeed at the task and half do not. The fact that there is no change with age argues that such a division among the subjects does not depend on age. To explore whether a bimodal distribution was responsible for this pattern, we collapsed across all the children's age groups and examined their responses on the telic description items only. These responses clearly pass the Shapiro–Wilk test for being a normal (and not a bimodal) distribution ($W = 0.9$, $P < 0.001$).

An item analysis provides some insight into the children's performance. Table 1 presents the proportion of counts that picked out individuals specified by goals for each item, broken down by children vs. adults and by telic vs. atelic predicates. There is a natural partition among the items for the children, with five (of ten) eliciting more goal-based individuation and the remaining five eliciting more spatio-temporal individuation. Collapsing across both description types, the five goal-biased items (paint, vase, rabbit, house, pop a balloon) had an overall rate of goal-based responses of 63%. In contrast, the five spatio-temporal-biased items (closet, ice-cream cone, butterfly, blow balloon, swim) had an overall rate of goal-based responses of only 28%. Thus, some items were more naturally seen in terms of spatio-temporally specified individuals and some were more naturally seen in terms of goal specified individuals, no matter what the linguistic description. These differences were only mildly related to the general category of the event (both the creation events – painting a flower and building a house – were in the goal-biased group but there were destruction and directed motion events in both the goal- and spatio-temporal-biased groups), and bore no relationship to the syntactic form of the event (the four syntactic types were distributed equally in the two bias groups). It appears, therefore, that there are idiosyncratic properties of individual items which govern whether they are more naturally construed by children as goal-based or spatio-temporally-based individual events.

Table 1
Item effects for events of Experiment 1

	Children		Adults	
	Atelic descriptions	Telic descriptions	Atelic descriptions	Telic descriptions
Swim	13	28	50	100
Blow balloon	13	35	13	88
Ice-cream cone	25	39	50	88
Butterfly	4	52	0	88
Closet	18	52	13	100
Pop balloon	25	64	0	100
House	25	71	25	100
Vase	45	74	0	100
Rabbit	24	77	0	100
Paint flower	46	79	38	100

Percentage of goal-based individuation for each item given the atelic and telic descriptions. The first five items are spatio-temporal biased while the last five items are goal biased.

Note, however, that the item effects are independent of the question of whether the linguistic description influenced individuation. For each item, the percentage of goal-based individuation given an atelic description was subtracted from the percentage of goal-based individuation given a telic description. A score of 100 is a perfect success on this task: it indicates that 100% of the responses to the telic description were goal-based while 0% of the responses to atelic description were goal-based. A score of 0 indicates that the telic and atelic descriptions elicited equal amounts of goal-based responses.[3] The mean differentiation score for adults across items was 78 (ranging from 38 to 100) and the mean differentiation score for children was 33 (ranging from 14 to 53). Of course, these differences reflect the interaction reported above; children's differentiation based on linguistic description is much weaker than that of adults.

Using the differentiation scores to look for possible item effects, we asked whether there was a bimodal distribution in children's success with different items. We demonstrated previously that children's overall 57% goal-based response rate given goal cues was not due to a bimodal distribution among the *children* (with some being passers and others failers). However, it is still possible that there is a bimodal distribution among the *items*, with children being able to reliably differentiate the two individuation possibilities for some items but not others. This is not the case. The mean differentiation score for the goal-biased group of items was 40 and the mean differentiation score for the spatio-temporal-biased group of items was 27. These means are not statistically different from each other using a Tukey–Kramer (HSD) comparison. Thus, regardless of what children's initial biases are with respect to an event, linguistic cues describing the opposite analysis can pull children off their preference to roughly equal degrees.

Finally, an analysis of the syntactic forms of the descriptions used showed that this was not a determining factor in children's responses. The mean differentiation scores (goal-based counting given a telic description minus goal-based counting given an atelic description) for each of the syntactic forms of the telic description were the following: transitive description $M = 32$, prepositional phrase description $M = 39$, passive description $M = 31$, unaccusative intransitive $M = 30$. None of these means were significantly different from each other using Tukey's HSD. An effect of syntax, however, is most likely to be found with the youngest children. A comparison of the differentiation scores for just the 3-year-olds also revealed no differences among the syntactic forms.

2.3. Discussion

As predicted, children preferred to use a spatio-temporal strategy for object individuation, irrespective of the linguistic description used, and there was a slight (non-significant) increase in the reliance on the linguistic description with age. That is, children were more inclined to count both the pieces of the fork as well as the whole forks, regardless of whether they were asked to count "forks" or "things". These results replicate

[3] Note that negative scores are also possible. A score of -100 suggests systematic confusion, with 0% of the responses to goal cues being goal-based and 100% of the responses to the spatio-temporal predicates being goal-based.

those of Shipley and Shepperson with one exception: the adults here failed to reliably use the linguistic cues to determine individuation, opting instead to individuate based on kinds regardless of the cue used. That is, adults tend to count both pieces of the fork together as a single unit regardless of whether they were asked to count "forks" or "things".[4]

There were two important results in the event domain. First, unlike in the object domain, the individuals children chose to count were influenced by the linguistic description of the events. Children distinguished telic and atelic predicates, and, somewhat surprisingly, 3-year-olds did so as reliably as 5-year-olds. Moreover, this success did not depend on the type of event (creation vs. change of state, etc.) nor on the syntactic form of the predicate used. Children succeeded equally in all cases. Second, as in the object domain, spatio-temporal criteria of individuation were more salient for children than for adults. Both children and adults had access to both types of individuation criteria (spatio-temporal and goal-based), and when given an atelic description, both relied on spatio-temporal criteria. But when given a telic description, children were markedly more likely than adults to continue to rely on spatio-temporal criteria of individuation. In sum, children were able to select between counting a closet door's closing and units of a closet door's sliding, but they sometimes got distracted by the individual slides, whereas adults did not.

These results extend one of the findings of Shipley and Shepperson from the domain of objects to the domain of events – the greater salience of spatio-temporal criteria of individuation to children. However, unlike in the object domain, even the youngest children tested were shown to be sensitive to the linguistic description of the events in their construal of the individuals they should count.

There is, however, one troubling factor in this story: adults were just as insensitive to the linguistic descriptions in the object case as the children were. Although adults and children opted for different default individuation strategies in the object task (children used a spatio-temporal strategy while adults used a kind-based one), neither group reliably switched their strategy based on the linguistic descriptions. It is possible, therefore, that there is something particular about the items used in the object case which make the linguistic descriptions irrelevant. If so, this task would be revealing about the differences in default individuation strategies for children and adults, but it would not indicate that children make better use of linguistic cues in the verb/event case than in the noun/object case. We take up the question of whether children and adults are sensitive to linguistic descriptions under any circumstances in object individuation in the next experiment.

3. Experiment 2: more object individuation

It is odd that the children's construal of individuals in the event domain was influenced by linguistic descriptions while their construal of individuals in the object

[4] A reviewer suggested that these object items may not look split into pieces so much as occluded. This problem arises because of our pictorial form of display (in contrast to Shipley and Shepperson who used actual objects). An occlusion analysis would certainly explain the adult pattern but it makes the children's data all the more startling. We certainly agree that our 2-D images may contribute to what makes the split objects weird, but given our replication of Shipley and Shepperson's findings with children, it doesn't appear to be the whole story.

domain was not. Count nouns, after all, serve the linguistic function of providing criteria of individuation (Macnamara, 1986; Xu & Carey, 1996). The event and object conditions of Experiment 1 differed in two ways that may account for the failure of the linguistic description to have much effect on how children (or adults) construed the individuals they were meant to count. In the event condition, both descriptions were informative – both "build a house" and "work", and "hop into a hole" and "hop" are informative predicates. "Thing", however, is a particularly empty noun – it does not mean "spatio-temporally defined object", as can be seen from the adults' responses. Rather, it just means "individual", and thus leaves the listener to construe the array in terms of the most salient individual. Also, each part of each event was a perfectly well formed example of a commonly named action (swimming, swimming across a lake, eating, eating an ice-cream cone). On the other hand, a broken fork is not exactly a fork, but it isn't clear what else one might call it.

Shipley and Shepperson's suggestion that young children count spatio-temporally determined objects, no matter what they are asked to count, did not rest *only* on the split object stimuli and the linguistic contrast between "thing" and basic level kind terms. They also presented children with arrays consisting of two or three sets of different kinds of objects (e.g. three dogs and three cats). Preschool children were asked either "how many animals" or "how many kinds of animals" there were. They overwhelmingly counted the individual animals in both cases, whereas adults of course categorically distinguished their counts under the two descriptions. However, even these data may underestimate young children's flexibility in individuating objects given different linguistic descriptions. Counting "kinds" requires considerable explicit command of the logic of super-ordinate classification (see Macnamara, 1986; Markman, 1989; Piaget, 1954).

Experiment 2 was designed to test further the hypothesis that preschool children count spatio-temporally defined individual objects, irrespective of the linguistic description under which they are asked to count. Two new kinds of object trials were used (in addition to the split objects used previously): collection/object and part/object. Collective nouns, such as "forest", "army" and "family", pick out groups of whole objects as individuals. The individuals picked out by the collective nouns in Experiment 2 have familiar, readily named objects as their parts. Previous work on children's comprehension of collective nouns suggests that children's knowledge of these nouns develops over the course of a few years. Children prefer a collective interpretation (as opposed to a super-ordinate one) for a label applied to a group of objects (Markman, 1989), and the meanings of common collective nouns such as the ones used here appear to be understood only marginally at best for children aged 3 and even 4 (Huntley-Fenner, 1995). By the age of 5, however, children can use syntactic cues for collective nouns (e.g. use of singular articles and lack of plural marking) to learn novel collective terms (Bloom & Kelemen, 1995). Moreover, Sophian and Kailihiwa (1998) showed that with adequate training, 5-year-olds (but not 4-year-olds) could select their counting strategy for one collection term, "family", and correctly count both the number of "families" and the number of "rabbits in the rabbit family". Accordingly, Experiment 2 tested 3- and 5-year-olds to assess whether they would count the people when asked to count both "the people" and "the families".

The second new type of stimuli were objects with commonly named parts – e.g. a bike and its wheels, a butterfly and its wings. None of the sub-parts were spatially separated from the whole (the wheels were on the bike, the wings still attached to the butterfly) and all of the sub-parts were integral components of the whole object. If children really count only whole objects no matter what the linguistic description, they should count the bicycles whether asked to count the bicycles or the wheels, and the butterflies whether asked to count the butterflies or the wings. Body parts are among the earliest words children learn (Dale & Fenson, 1996). Previous research on children's word learning has shown that parts are salient and nameable by at least age 2. Markman (1989) showed that children this young readily learn names for novel parts, as long as they already know the name of the whole object. More directly, Giralt and Bloom (2000) showed that 3-year-old children could successfully count parts (and not wholes) at least when they knew the name of the part in question. We therefore strongly expect that even 3-year-old children should succeed at using the linguistic descriptions to select their counting criterion for these trials.

3.1. Methods

3.1.1. Subjects

Two groups of children were tested; there were 16 children in each age group. The 3-year-old group had a mean age of 3;7 (ranging in age from 3;0 to 3;11); the 5-year-old group had a mean age of 5;2 (ranging in age from 4;11 to 5;6). The data from an additional eight subjects were removed because these children failed to cooperate on more than half the trials. Subjects were children in the New York City and Cambridge, MA areas who were either brought into a lab to participate or else tested at a local daycare center. According to parental report, all subjects had English as their primary (and in most cases, only) language. In addition, 16 adult native speakers of English participated.

3.1.2. Stimuli

The object stimuli were created from clip-art images and displayed using a Microsoft Power-point program. Three categories of items were created, with four examples in each category: split objects, collections/objects, and parts/objects. The split object items were taken from Experiment 1; they were modified for the sake of generality to show a different number of complete objects of a given basic level kind. As before, each basic level object was split into two or three spatially separate units. The collection/object items showed one, two or three collections (forests, piles) each consisting of from three to six basic level objects (trees, books). The part/object items consisted of two or three objects from familiar basic level kinds (butterflies, bicycles), each of which also had easily nameable sub-parts (wings, wheels). Subjects counted a total of 12 items. Appendix B shows the items from Experiment 2, along with the two different linguistic descriptions used to instruct what to count in each case.

3.1.3. Procedures and coding

The procedures used were identical to those used in the object condition of Experiment 1. Subjects were asked to count "How many X are here?" For the split items, the noun either labeled the basic level object kind (e.g. "car, pencil") or, following Shipley and Shepperson, was the general term "thing", meant to label spatio-temporally separate objects. For the collection/object items, the noun either labeled the collection (e.g. "forest", "pile") or the separate bounded objects, labeled at the basic object level, that were members of the collections ("tree", "book"). For the part/object items, the nouns either labeled separate bounded objects, labeled at the basic object level ("bike", "butterfly"), or parts of the objects ("wheel", "wing"). For the collection items, the members of the collection were considered the spatio-temporally specified individuals, because these individuals (a grape, a person, a book, a tree) each constitute a bounded object on its own. However, it is worth noting that in the actual depiction of these items, the collections themselves were the unified bits of clip-art used. For example, each tree in the collection forest was not itself spatially discrete on the display screen – there was some overlap among the trees to help create a viable looking forest. The three types of object items were presented interspersed among each other in a pseudo-random order.

In 94% of the cases, children either counted the individuals as specified by the basic level object terms, or they counted the spatio-temporally defined "things", the parts, or the collections.[5] The coding scheme for the split objects stimuli was the same as in Experiment 1. In spatio-temporal counts, children counted each piece as "one"; in kind counts they combined broken pieces into individuals for the purpose of counting. For the collection/object stimuli, in spatio-temporally specified object counts, children counted each individual object; in collection counts they counted the whole group as "one". For the part/object stimuli, in spatio-temporally specified object counts, they counted each basic level object and in part counts they counted each part.

3.2. Results

The data for the split object, collection/object, and part/object stimuli have been graphed separately in Fig. 3a–c (respectively). Overall, the results with the split objects replicate the results with these objects from Experiment 1. Children prefer to individuate with a spatio-temporal criterion while adults prefer to use a kind-based criterion, largely irrespective of the linguistic description. By contrast, with the part/whole and collection items children (and adults) were able to switch their individuation criterion based on the linguistic description. The responses to the collection items were quite parallel to the event results from Experiment 1 in that children clearly showed a bias for an individuation criterion that picked out spatio-temporally defined objects, but not to the extent that this criterion over-rode the influence of the linguistic descriptions. With the part/whole cases, the 5-year-olds relied exclusively on the linguistic descriptions and were indistinguishable from adults, whereas 3-year-olds did occasionally count the spatially specified whole objects when asked to count the parts.

An ANOVA was performed over the percentage of times subjects used a kind, collection, or part individuation criterion (as opposed to a spatially individuated object criterion), with age group (3- and 5-year-olds and adults) as the between subjects factor,

[5] Virtually all of the remaining 6% of responses were combination counts in which the child counted every object in the collection or every spatio-temporal portion of the split object but restarted their count back at "1" for each collection or kind object. These items were again given half-credit for each counting strategy. Fewer than 1% of all the trials were unclear and uncodable.

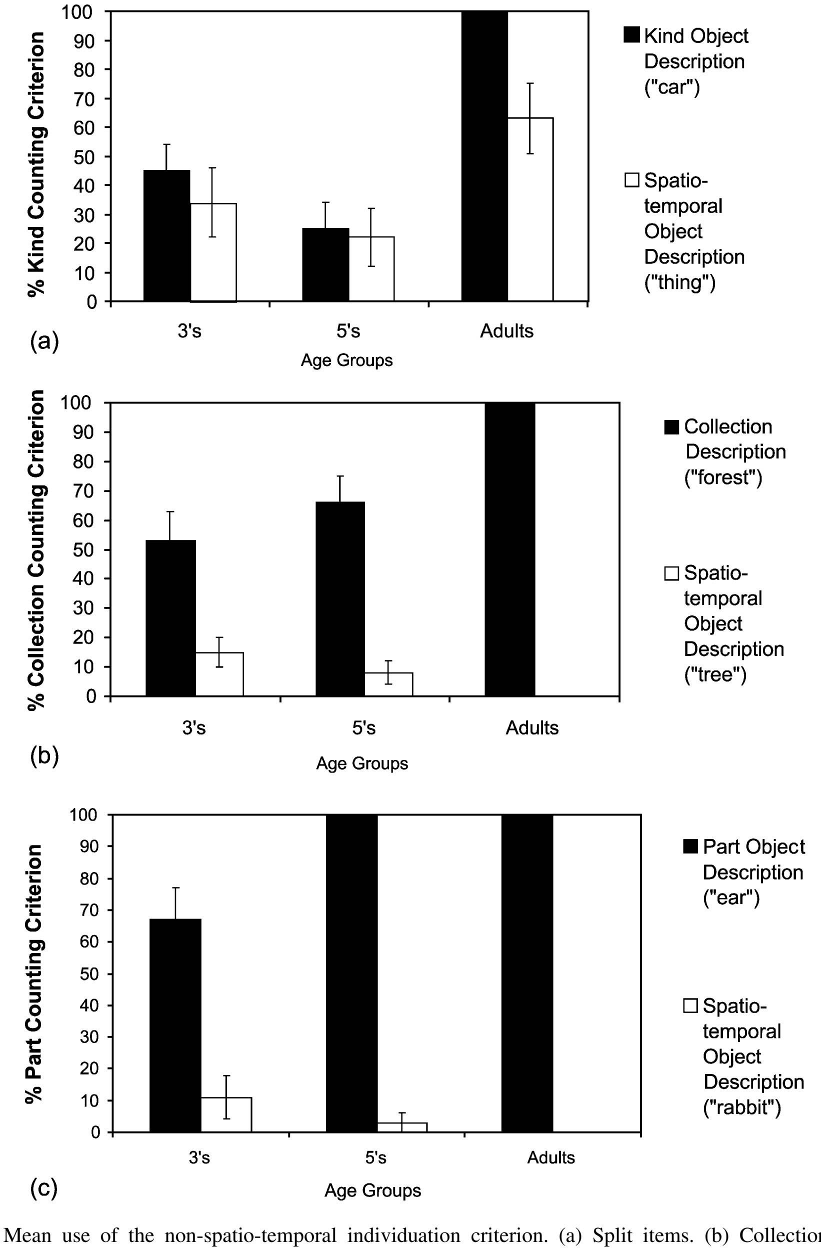


Fig. 3. Mean use of the non-spatio-temporal individuation criterion. (a) Split items. (b) Collection items. (c) Part/whole items.

and linguistic description type (spatial individual and non-spatial individual) and stimulus type (split object, collection/object, part/object) as the within subjects factors. This analysis confirmed that there was a main effect of age group ($F(2,45) = 15.3$, $P < 0.0001$); post-hoc analysis (Tukey's HSD) revealed that both child groups were different from adults, but neither was different from each other. The overall mean rate of kind/collection/part counting was 45% for 3-year-olds, 37% for 5-year-olds, and 60% for adults. There was also a significant main effect of linguistic description ($F(1,45) = 22.29$, $P < 0.0001$) showing that subjects were more likely to use a kind/collection/part individuation criterion when the stimuli were described by kind terms, collection terms, or part terms. The overall mean rate of kind/collection/part counting given a spatio-temporal description was 20% whereas given a non-spatio-temporal description it was 75%.

In addition, there was an interaction between linguistic description and age ($F(2,45) = 15.53$, $P < 0.0001$). There was no main effect of object type, but there were significant interactions between object type and age group ($F(4,90) = 7.86$, $P < 0.0001$) and between object type and linguistic description ($F(2,90) = 48.12$, $P < 0.001$). These two-way interactions are best seen in the light of the three-way interaction among all three variables ($F(4,90) = 2.96$, $P < 0.024$). Basically, the pattern of results was different for each of the three stimulus types.

Given the three-way interaction, we analyzed each stimulus condition separately. The data from the split object stimuli almost perfectly replicated what was found in Experiment 1: children prefer to use a spatio-temporal criterion irrespective of the linguistic description while adults prefer to use the kind-based criterion overall. The only difference is that in Experiment 2, the adults did significantly select their counting criterion based on linguistic description ($t(15) = 3.2$, $P < 0.006$), replicating Shipley and Shepperson (1990). As in Experiment 1, the ANOVA on the percentage of times subjects used a kind-based count criterion showed a main effect of description (Kind description vs. Spatio-temporal "Thing"): $F(1,45) = 7.52$, $P < 0.009$; and age (3- and 5-year-olds and adults): $F(2,45) = 5.68$, $P < 0.001$; but no interaction between these two. Post-hoc analysis showed that the age effect resulted from both child groups being different from the adults, but not from each other.

The pattern of results in the collection/object and the part/object conditions differs markedly from that found in the split object condition. The pattern with the collection/object items is precisely the same pattern found for events in Experiment 1: the children succeeded at using the linguistic descriptions to determine what they counted but they periodically counted objects in lieu of collections, while adults were at ceiling across the board. An ANOVA on the percentage of times participants counted collections confirmed that there was a massive effect of linguistic description ($F(1,45) = 219.8$, $P < 0.0001$), a main effect for age ($F(2,45) = 3.45$, $P < 0.04$) and an interaction between the two ($F(2,45) = 17.07$, $P < 0.001$). Post-hoc tests also confirmed that the age effect was due to both child groups being different from the adults, but not from each other. Also similar to Experiment 1, the overall success rate with the collections was around 60%. To determine if this rate was the result of a bimodal distribution, with some children being passers and others failers, we examined children's responses on the collection items only. These responses pass the Shapiro–Wilk test for being a normal (and not a bimodal) distribution ($W = 0.81$, $P < 0.001$).

The pattern with the part/object condition is very similar to that of the collection/object condition, except that in this case, only the 3-year-olds make errors with the non-spatio-temporal object items (the parts); both the 5-year-old children and the adults are at ceiling in this condition. An ANOVA on the percentage of times subjects counted parts showed another massive effect for linguistic description ($F(1, 45) = 312.4$, $P < 0.0001$), a main effect of age ($F(2, 45) = 3.71$, $P < 0.32$) and an interaction between the two ($F(2, 45) = 8.71$, $P < 0.001$). Post-hoc comparisons (Tukey's HSD) on the age groups showed that 5-year-olds were no different from adults; 3-year-olds were different from both the adults and the 5-year-olds.

3.3. Discussion

Contrary to Shipley and Shepperson's data, and contrary to the results of Experiment 1, linguistic descriptions can indeed influence preschool children's object individuation. It is not the case that bounded, spatio-temporally defined whole objects are the only individuals preschool children will count. When asked to count the wings, young children rarely count the butterflies, and when asked to count the families, young children do not always count the people. Apparently the failures in Experiment 1 and in Shipley and Shepperson's studies were due to the unclear ontological status of a broken fork, the difficulties in construing "kinds" as individuals, and the ambiguity of the empty noun "thing".

As in Experiment 1, subjects in the split object condition of Experiment 2 again largely ignored the linguistic descriptions in favor of their preferred spatio-temporal individuation strategy. However, with the collection/object items and the part/object items, both children and adults selected their individuation criteria as a function of linguistic description. In the object/part condition, there was almost no interference from the spatio-temporally defined whole objects: only the 3-year-olds ever counted whole objects when asked to count parts. This result replicates previous findings in this area (Giralt & Bloom, 2000). However, in the collection/object condition, as in the events of Experiment 1, there was robust evidence that spatio-temporally defined individuals are particularly salient to preschool children. When asked to count the collections (families, piles), the children counted the basic level objects (people, books) fully 40% of the time. This result replicates and extends the general findings of Sophian and Kailihiwa (1998).[6]

4. General discussion

One counts individuals. Thus, what children and adults can count provides evidence concerning what individuals they represent. The present experiments found that children can use both spatio-temporal and causal/kind criteria for individuation, both in the domain

[6] Our results differ from Sophian and Kailihiwa (1998) in some details, most notably in the fact that they found a bimodal distribution among the children while we did not. The cause of this difference is unclear. It is probably *not* the result of the additional collection terms we used – in the present experiment, performance with the term "family" was no different from performance with the other three collection terms used ("forest", "pile", and "bunch"). It could be the result of Sophian and Kailihiwa's training procedure, which may have divided the children according to meta-linguistic skills.

of events and objects. Children as young as 3 years of age can count spatially separated objects, whether members of kinds or not, and they can also count collections and parts of objects. They can count spatio-temporally bounded events (episodes separated by pauses), and they can count events individuated by achieved goals. Furthermore, linguistic descriptions guide the construal of both events and objects in terms of countable individuals in children as young as 3. However, 3- to 5-year-old children differ from adults in their likelihood of adopting spatio-temporal criteria of individuation in counting contexts, even when these are at odds with an explicit linguistic description of what they are asked to count. Preschool children (ages 3–5) are markedly more likely than adults to count bites when asked to count events of eating an ice-cream cone, to count each piece of a split car when asked to count cars, to count trees when asked to count forests, and (3-year-olds only) to count bunnies when asked to count ears. Finally, both children and adults were much less sensitive to the linguistic description in the split object conditions of Experiments 1 and 2 than in any of the other conditions (events, collection/object, part/objects).

We discuss the two central findings from this study in turn: (1) the salience, in the preschool years, of spatio-temporal criteria for individuation, both in the case of objects and events; and (2) the sensitivity to linguistic descriptions by even the youngest children, in both the object and event domains.

First, consider the finding that, relative to adults, children preferred to count spatio-temporally defined individuals. Perhaps this finding reveals more about how children learn to count than it does about criteria for individuation. During the counting routine, young children often touch each individual as they recite the count list, and they do so even before they have worked out how the number words represent number (Wynn, 1990). Most usually, the individuals to which this routine applies are bounded, spatio-temporally defined, whole objects. Perhaps this overlearned routine interferes with counting other entities. While we do not doubt that this could be a factor in these results, it cannot be the whole story, for the simple reason that within the object domain the likelihood to count spatio-temporally defined units varied from stimulus condition to stimulus condition – most for the split objects, intermediate for the collection/objects, and least for the part/objects. Also, the count routine is not practiced on events, yet the results for the event stimuli were identical to those of the collection/object stimuli. We conclude that the results reflect the relative salience of different criteria of individuation.

It appears that the spatio-temporal bias is quite robust, as it is revealed in other stimulus/linguistic contexts than those originally studied by Shipley and Shepperson. It may derive from the fact that spatio-temporal criteria of individuation are privileged in a mid-level object-tracking system available to both infants and human adults. This system has been studied under the description "object-based attention" (Cavanagh, Labianca, & Thornton, 2001; Kahneman, Treisman, & Gibbs, 1992; Scholl, 2001; Scholl & Pylyshyn, 1999). It is called "mid-level" precisely because information about object properties and object kinds is not used for object individuation – rather objects are individuated and tracked on the basis of spatio-temporal information. Moreover, there is a growing body of evidence supporting the idea that infants make use of this object-tracking system in most individuation tasks (Carey & Xu, 2001; Leslie, Xu, & Tremoulet, 1998; Xu & Carey, 1996). The point is argued in more detail in Carey and

Xu (2001) but the intuition is that infant individuation studies are really just very simple object-tracking studies. The evidence that the adult and infant studies tap into the same system comes from several common properties of processing, including the fact that both manifest a set-size limitation of about three on the number of objects that can be tracked in parallel, and both manifest insensitivity to property or kind information under conditions of difficulty.

The combined adult and infant data suggest, moreover, that this system has several properties associated with modularity (cf. Fodor, 1983): it operates automatically, provides relatively shallow representations (containing possibly only spatio-temporal information under some circumstances), and is largely encapsulated from other kinds of knowledge about objects. The spatio-temporal bias found here could be the natural result of these modular properties. Because the process is automatic, the spatio-temporal individuation criterion will always be an available analysis of the stimuli presented in the object conditions of Experiments 1 and 2. Because the process is informationally encapsulated, considering non-spatio-temporal information (such as the kind identity of an object) will always require a separate step of processing. The spatio-temporal bias, therefore, may simply reflect the fact that spatio-temporal individuation is easier to do.

Note that our claim here is not that spatio-temporal individuation is a step on the way to other, more complex, kinds of individuation. As discussed in Section 1, kinds cannot be spatio-temporally defined (e.g. a single spatio-temporal unit may also be a passenger several times over; a dead cow continues to be a spatio-temporal body, but is no longer a proper cow), nor can collections (e.g. a family is still a family no matter how spatio-temporally disparate its members). Kind- and collection-based individuation of objects requires a qualitatively different analysis of the situation.[7] It is an empirical question whether the output of the spatio-temporal analysis acts, heuristically, as an input to the kind/collection analysis, or whether the two sorts of analyses operate in parallel over objects completely separately. What is clear is that different types of linguistic descriptions help us choose among multiple, independently available analyses. Regardless, our position makes several empirical predictions. For example, we predict that subjects will be able to identify spatio-temporal units more quickly than kind or collection units. We also predict that children's memory for spatio-temporal units will be good, even when they are asked to identify other types of units; in addition, we predict that the reverse will not be the case – memory for kind or collection units should suffer when subjects are asked to focus on spatio-temporal units.

The finding that children show a comparable spatio-temporal bias in event individuation suggests that there may be a parallel preattentive mechanism that automatically individuates events into spatio-temporal units and that is uninfluenced by higher cognitive analyses based on causal interpretations of the events. Wynn's infant results (Sharon & Wynn, 1998, 2000; Wynn, 1996) support the existence of a mechanism

[7] We take it to be an open question whether sub-parts of objects might be spatio-temporally constitutive of their host objects. It seems that at least a reasonable argument can be made for a hierarchical analysis of parts and wholes, while it cannot be made in general for collections, and can never be made for kinds.

for analyzing the actions and motions around us that outputs a crude set of event units defined largely in spatio-temporal terms. Our knowledge of the causal features of events (e.g. the goal of the event, the intentions of the actors, etc.) would require a separate step of processing to get linked into this spatio-temporal parse.[8] Thus, we offer a parallel account for the preschool children's greater reliance on spatio-temporal information in the event counting task to our mid-level object-tracking account of the object data. Because the system provides the spatio-temporal individuation of events automatically, this analysis will always be available to our subjects; because the system is informationally encapsulated, the goal-oriented analysis will require distinct processing which will not always be adequately undertaken by the subjects. At this point, the existence of such a system is largely speculative, but based on the parallelism with the object case, we can make several predictions about the properties such a system should have. As noted above, the system should operate automatically and be informationally encapsulated. Moreover, the spatio-temporal event system should be cognitively primary in the sense that it is available early in development (earlier even than the system that allows us to analyze events in terms of goals and intentions) and is part of our event processing throughout our lives. In addition, the specific predictions noted above in the object case can also be made for events: subjects should be faster to identify spatio-temporally defined events compared to goal-based events; subjects should retain better memory for the spatio-temporal analysis even when asked to focus on goals but not the reverse. Clearly this is an area that requires a great deal of further research.

This line of reasoning can explain the children's spatio-temporal bias, but something more needs to be said to explain the adult reversal of this bias in the case of the split objects. Recall that in the split object case adults prefer to ignore the spatio-temporal criterion in favor of a kind criterion, combining pieces of forks together into a single unit to count them under any linguistic description. Most likely, the adult pattern of results reflects two factors. First, the processing demands may well be different for adults than for children. We have in general an extraordinary amount of experience at identifying object kinds by the time we are adults. To the extent that additional processing is needed to go beyond the spatio-temporal analysis, it has been so well practiced that the cost to performing it is trivial. In a task such as this one, where adults counted low numbers of familiar objects with no time pressure, any processing costs associated with a kind-based analysis would be unnoticeable – though perhaps if we increased the task demands by requiring adults to verbally shadow during the task we might see the effects of the increased processing. However, the very fact that we do have such extensive experience with kinds may mean that kind processing itself may be automatized in much the way that the process of reading can become automatized for proficient readers. If adults have developed automatic input analyzers for object kind identification, then we might even predict that adults would default to a kind interpretation. Indeed, adults are more willing to count a transformed example of a kind (e.g. a crushed or cut-up cup) as an instance of that

[8] And, just as in the object case, the goal or result analysis of an event is not composed out of the spatio-temporal analysis. Although in the real world, results are typically achieved via some spatio-temporally defined process, they need not be: one can die, for example, without the event having any spatio-temporal sub-parts at all. Indeed, the linguistic category of achievements (cf. Dowty, 1979; Vendler, 1967) is defined by this property.

kind than are 4-year-olds (Gutheil, Valderama, & Freedman, 1999). The second factor is the oddity of the split object stimuli. As can be seen in Appendix A, the split objects are very obviously halves of specific wholes; indeed, the wholes are still aligned and separated only by a small gap. Real cars would never be split like that. We assume that adult expertise with these object kinds led them to ignore the small gap internal to each object.

We turn now to the second main finding from these studies: the fact that even the youngest children are sensitive to linguistic descriptions in both the object and event domains. We tested whether children could use different linguistic descriptions to select among competing individuation criteria. Success with this task presupposes two prior pieces of knowledge: first, that children are able to use different individuation criteria to parse experience into objects and events, and second, that they know how language signals which criterion is relevant.

In the object case, linguistic descriptions led children to modify their counting criteria for the collection and part/whole objects but not for the split objects. We have discussed above why we believe that the split objects are not a fair test of children's individuation or linguistic abilities. We take the success with the collections and part/whole objects as indicating that by the age of 3 children can create individuals in the object domain both that coincide with independently moveable 3-D objects as well as that do not. Also, by the age of 3, they know how their language refers to the different kinds of individuals, and they are guided by linguistic descriptions in parsing objects in terms of individuals. One additional point that requires discussion is the fact that children do better earlier with the part/whole items than with the collection items (see also Giralt and Bloom (2000) for early success with part/whole counting and Sophian and Kailihiwa (1998) for comparably late success with one collection term, "family"). Recall that despite their overall success, both 3- and 5-year-olds showed a strong bias to count spatio-temporal units as opposed to collections (e.g. around 40% of the time children counted people when asked for families, and books when asked for piles) but only the 3-year-olds ever counted spatio-temporal units when asked for sub-parts (e.g. counted bunnies when asked for ears). The difference between these items has to do with the different conceptual status of the non-spatio-temporal units (i.e. the sub-parts and the collections). In the part/whole condition, the sub-parts are very nearly spatio-temporal units themselves. Although the bunny's ears are in fact attached firmly to the bunny, they are spatio-temporally unified and it is easy to see the point where they could be detached from the bunny to create a separably moveable object. These properties are common to nameable sub-parts of objects in general and are clearly the case for all of the objects used here. By contrast, identifying a collection as an individual requires an understanding of why the elements should be specially grouped as a unit. The reasons a group of individual objects are grouped into a collection can be quite complex and depend on a wealth of world knowledge (consider the collection terms "family", "flock", "army" and "marching band").[9] Even children as old as 5 years probably have an incomplete understanding of why we consider some collections to be units and the fragility of their conceptual knowledge may lead them to resort more frequently to their default spatio-temporal individuation strategy.

[9] Moreover, as noted by a reviewer, our static pictures provide especially poor cues for discerning these complex relationships.

In the case of events, children also succeeded at selecting their individuation criteria as a function of the linguistic descriptions used. Their success here again indicates that these children can individuate events both in terms of spatio-temporal or goal-based criteria. More surprisingly, these results argue that by the age of 3, children have figured out how their language marks telic predicates. They have mastered the telic/atelic distinction as it is marked in English. As discussed above, this marking makes use of varied syntactic devices and requires the integration of information from all parts of a predicate. Moreover, children's success did not depend on any particular sentence type used; children succeeded equally well no matter how the test question marked telicity. Previous research (van Hout, in press) had found children's knowledge of telicity marking in English and Dutch to be quite marginal, even for children as old as 5 years, but these studies had restricted themselves to predicates involving food ("eat" and "drink"). Interestingly, one of the weakest items (that is, the difference score was small, though in the right direction) from the current study also involved food ("eat an ice-cream cone"). We might speculate that the source of the children's problem is in their understanding of terms referring to food items. Food is one of the few domains in which reference freely switches between mass and count terms. As the telicity of the sentence depends on whether the sentence object is a mass or count noun ("eat ice-cream" is atelic, "eat an ice-cream cone" is telic), any confusion about the mass/count status of the food noun would lead to confusion about the telicity of the sentence as a whole. Regardless, 3- to 5-year-old children's current success with so many different semantic and syntactic types strongly suggests that they have general knowledge about how telicity is marked in English.

Of course, it is also the case here that children were far from ceiling in counting the goal units given telic predicates. Despite an overall success, children are still counting the spatio-temporal breaks significantly more often than the adults are with the telic predicates. We analyzed the parallel finding with the collections as reflecting children's conceptual difficulties in understanding why a collection is a viable unit. For the event case, a conceptual story does not appear to be in order. Recall that in the event case, some events were biased towards a goal interpretation in general while others were biased towards the spatio-temporal interpretation. The fact that for some items the goal is therefore more salient to the children than the spatio-temporal breaks suggests that there is nothing hard per se about thinking about events in terms of their goals. Instead, we suspect that the original intuition put forward in Section 1 – that the linguistic marking of telicity is complicated – is the culprit. Children's overall success speaks of their general knowledge of how telicity is marked, but the process of actually integrating the relevant information from all the parts of the sentence may still be a hard one for them. As the going gets tough, children resort to their default individuation strategy which is the same for events as it was for objects, a spatio-temporal one.

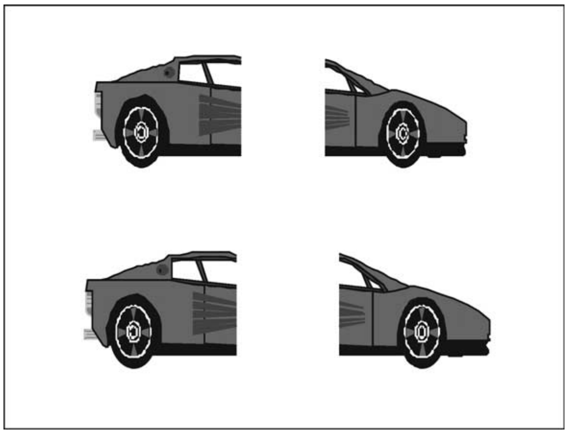
Acknowledgements

This research was funded by NIH grant HD38338 to Susan Carey and NRSA Post-doctoral Fellowship MH12225 to Laura Wagner. We thank Ximena Acevedo and

Rebecca Chase for assistance in collecting data, and Grace Lai for her artistic aid with the movies.

Appendix A. Stimuli for Experiment 1

Example object stimulus: two cars/four things



Complete list of object stimuli used in Experiment 1 (correct answer under each description given in parentheses)

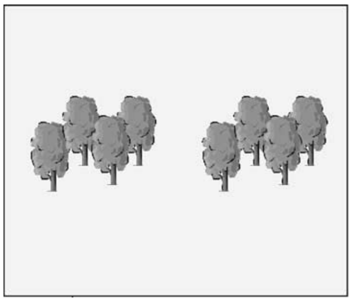
Item	Type	Object kind description	Spatio-temporal object description
Airplane	Split	*Airplane (1)*	*Thing (2)*
Apple	Split	*Apple (1)*	*Thing (2)*
Bicycle	Split	*Bike (1)*	*Thing (2)*
Car	Split	*Car (2)*	*Thing (4)*
Cupcake	Split	*Cup cake (2)*	*Thing (4)*
Marker	Split	*Marker (2)*	*Thing (6)*
Pencil	Split	*Pencil (1)*	*Thing (3)*
Watermelon slice	Split	*Watermelon (2)*	*Thing (6)*
Donut	Whole and split	*Donut (1 or 2)*	*Thing (3)*
Fork	Whole and split	*Fork (1 or 2)*	*Thing (4)*
Light bulb	Whole and split	*Light bulb (2 or 3)*	*Thing (5)*
Tree	Whole and split	*Tree (1 or 2)*	*Thing (3)*

Complete list of event stimuli for Experiment 1 (correct answer under each description given in parentheses)

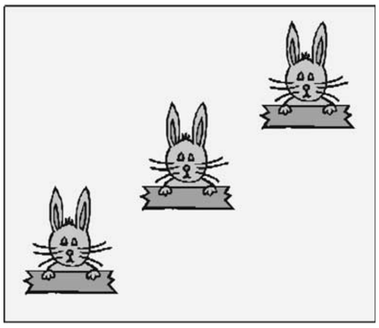
Item	Telic description	Atelic description
	"How many times…"	"How many times…"
Blow balloon	*does the boy blow up the balloon? (1)*	*does the boy blow? (2)*
Butterfly	*does the butterfly fly to the tree? (1)*	*does the butterfly fly? (3)*
Closet	*does the closet door close? (1)*	*does the closet door slide? (3)*
House	*was the house built? (1)*	*does the girl work? (2)*
Ice-cream cone	*was the ice-cream cone eaten? (2)*	*does the girl eat? (6)*
Paint	*does the girl paint a flower? (2)*	*does the girl paint? (4)*
Pop balloon	*does the bird pop the balloon? (1)*	*does the bird poke the balloon? (2)*
Rabbit	*does the rabbit jump into the hole? (2)*	*does the rabbit move? (6)*
Swim	*does the boy swim across the river? (2)*	*does the boy swim? (4)*
Vase	*does the vase break? (2)*	*does the vase move? (4)*

Appendix B. Stimuli for Experiment 2

Example collection stimulus: two forests/eight trees



Example part-whole stimulus: three rabbits/six ears



Complete list of object stimuli for Experiment 2 (correct answer under each description given in parentheses)

Item	Type	Collection, kind object, and part object descriptions	Spatio-temporal object descriptions
Book pile	Collection	*Pile (2)*	*Book (6)*
Family	Collection	*Family (3)*	*People (8)*
Forest	Collection	*Forest (2)*	*Tree (8)*
Grape bunch	Collection	*Bunch (1)*	*Grape (6)*
Airplane	Split	*Airplane (2)*	*Thing (4)*
Apple	Split	*Apple (2)*	*Thing (4)*
Car	Split	*Car (1)*	*Thing (2)*
Pencil	Split	*Pencil (2)*	*Thing (6)*
Bike	Part/whole	*Wheel (6)*	*Bike (3)*
Butterfly	Part/whole	*Wing (4)*	*Butterfly (2)*
Chair	Part/whole	*Leg (3)*	*Chair (1)*
Rabbit	Part/whole	*Ear (6)*	*Bunny-rabbit (3)*

References

Antinucci, F., & Miller, R. (1976). How children talk about what happened. *Journal of Child Language*, *3*, 167–189.

Bach, E. (1986). The algebra of events. *Linguistics and Philosophy*, *9*, 5–16.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*(3), 708–717.

Biro, S., Gergely, G., Koos, O., & Csibra, G (1996). *Inferring hypothetical states of affairs on the basis of the rationality assumption in infancy*. Poster presented at the 10th Biennial International Conference on Infant Studies, Providence, RI.

Bloom, L., Lifter, K., & Hafitz, J. (1980). Semantics of verbs and the development of verb inflection in child language. *Language*, *56*(2), 386–412.

Bloom, P., & Kelemen, D. (1995). Syntactic cues in the acquisition of collective nouns. *Cognition*, *56*, 1–30.

Bronckart, J. P., & Sinclair, H. (1973). Time, tense and aspect. *Cognition*, *2*(1), 107–130.

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

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: beyond object files and object tracking. *Cognition*, *80*, 179–213.

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: sprites. *Cognition*, *80*(1–2), 47–60.

Csibra, G., & Gergely, G (1996). *Can one-year-olds infer the goal of an action without seeing it achieved?* Poster presented at the 10th Biennial International Conference on Infant Studies, Providence, RI.

Csibra, G., Gergely, G., Brockbanck, M., Biro, S., & Koos, O (1998). *Twelve-month-olds can infer a goal for an incomplete action*. Poster presented at the 11th Biennial International Conference on Infant Studies, Atlanta, GA.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers*, *28*, 125–127.

Dowty, D. (1979). *Word meaning and Montague grammar*. Dordrecht: Kluwer.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Giralt, N., & Bloom, P. (2000). How special are objects? Children's reasoning about objects, parts, and holes. *Psychological Science*, *11*(6), 497–501.

Gutheil, G., Valderama, N., & Freedman, R (1999). *When is a door not a door? Shape, function, intent and history as determinants of artifact kind membership*. Poster presented at SRCD, Albuquerque, NM.

Huntley-Fenner, G. N. (1995). The effect of the whole object bias on preschoolers' understanding of collective nouns. In E. V. Clark (Ed.), (pp. 145–155). *Proceedings of the Twenty-Seventh Child Language Research Forum*, Stanford, CA: CSLI.

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: object specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Leslie, A. M., Xu, F., Tremoulet, P., & Scholl, B. (1998). Indexing and the object concept: developing "what" and "where" systems. *Trends in Cognitive Sciences*, *2*(1), 10–18.

Macnamara, J. (1986). *A border dispute: the place of logic in psychology*. Cambridge, MA: MIT Press.

Markman, E. (1989). *Categorization and naming in children*. Cambridge, MA: MIT Press.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Saffran, J., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, *80*, 1–46.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.

Sharon, T., & Wynn, K. (1998). Individuation of actions from continuous motion. *Psychological Science*, *9*(5), 357–362.

Sharon, T., & Wynn, K (2000). *Parsing motion for meaning: the role of tangent discontinuities*. Poster presented at the International Conference on Infant Studies, Brighton, UK.

Shipley, E. F., & Shepperson, B. (1990). Countable entities: developmental changes. *Cognition*, *34*, 109–136.

Shirai, Y., & Andersen, R. (1995). The acquisition of tense-aspect morphology: a prototype account. *Language*, *71*(4), 743–762.

Smith, C. (1991). *The parameter of aspect*. Dordrecht: Kluwer.

Sophian, C., & Kailihiwa, C. (1998). Units of counting: developmental changes. *Cognitive Development*, *13*, 561–585.

Spelke, E., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Cognition*, *99*(4), 605–632.

Van de Walle, G. A., Carey, S., & Prevor, M. (2000). Bases for object individuation in infancy: evidence from manual search. *Journal of Cognition & Development*, *1*(13), 249–280.

van Hout, A (in press). Acquiring telicity cross-linguistically: on the acquisition of telicity entailments associated with transitivity. In M. Bowerman, & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: implications for learnability*. Hillsdale, NJ: Lawrence Erlbaum.

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.

Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193.

Wynn, K. (1996). Infant's individuation and enumeration of actions. *Psychological Science*, *7*(3), 164–169.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, *85*, 223–250.

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. *Cognitive Psychology*, *30*, 111–153.