LANGUAGE LEARNING AND DEVELOPMENT, 2(1), 51–76

Aspectual Bootstrapping in Language Acquisition: Telicity and Transitivity

Laura Wagner
Department of Psychology
The Ohio State University

Two studies looked at whether children used transitivity as a structural cue to telicity semantics. Telicity comprehension was tested using an event-counting task (Wagner & Carey, 2003). Children watched animated movies depicting a goal (a dog pushes a ball into a can) achieved via 2 spatiotemporally distinct steps (the dog pushes the ball twice). The movies were described with either a telic or an atelic predicate, and the dependent measure was whether children counted the event's goals or steps. The predicates conveyed telic meaning with either a transitive structure (*The girl painted a flower*) or an intransitive structure (*The vase broke*), or atelic meaning with either a transitive (*The dog pushed the ball*) or an intransitive structure (*The vase moved*). Experiment 1 showed that 3- and 5-year-old children successfully understood telic and atelic descriptions, regardless of their syntactic form. Experiment 2 showed that 2-year-old children were successful overall, but showed a bias toward linking transitive structures to telic meanings. The results support the idea that telicity is among the semantic features that are linked to the transitive sentence frame and can be signaled by a structural cue.

The telic–atelic distinction is one of a handful of semantic contrasts that structure the linguistic encoding of events. Cross-linguistically, telicity has a variety of syntactic reflexes, including case marking, verbal morphology, adverb selection, auxiliary selection, argument structure variations, and word order. Despite the range of domains in which telicity information is relevant, no language has a direct explicit marking of the telic–atelic distinction; for example, no languages simply add a morpheme to a verb or a sentence as a signal that the predicate has telic (or atelic) meaning. Thus, despite the general semantic importance of telicity, learning the range of syntactic implications associated with it poses some obvious challenges.

Correspondence should be addressed to Laura Wagner, Department of Psychology, The Ohio State University, Columbus, OH 43210. E-mail: wagner.602@osu.edu

The primary aim of the studies reported here is to investigate children's understanding of telicity. Children were tested on one semantic entailment of telicity, namely, the fact that telic and atelic predicates provide different criteria for event individuation. The coding of telicity in English (discussed below) is decidedly nontransparent and relies on the semantic composition of the verb, its arguments, and even adjuncts present in the sentence. Thus, a secondary aim of this work is to investigate how children acquire the argument structure reflexes of telicity in English. In particular, these experiments look at the possibility that children begin with a simple structural heuristic (transitive sentences have telic meaning) from which they bootstrap themselves into the adult system. The results suggest that 2-year-old children do use a transitivity heuristic for telic semantics, but that older children do not—at least with the familiar verbs tested here.

THE TELIC–ATELIC DISTINCTION

Telicity is a lexical semantic property of predicates; it indicates whether or not an event has an inherent endpoint or boundary. I use the term *telic* to refer to all events that possess such an endpoint, including both accomplishments and achievements in the Dowty–Vendler system (Dowty, 1979; Vendler, 1967). I use the term *atelic* to refer to all events without such an endpoint, including activities, semelfactives, and states.

Telic and atelic predicates can be distinguished from each other by a variety of linguistic tests (cf. Dowty, 1979; Smith, 1991; Vendler, 1967, for an overview). For example, atelic predicates have only a single interpretation with the adverb *almost,* whereas telic predicates have two. Thus, (1a) means that Kelly did not even begin to eat, and (1b) may mean either that Ruben did not begin to eat the pizza, or that he began the pizza but did not finish it.

1 (a) Kelly almost ate.
  (b) Ruben almost ate a pizza.

Telic and atelic predicates combine naturally with different adverbs of time indicating the duration of an event. Atelic predicates (2a) combine naturally with *for X time,* whereas telic predicates (2b) combine naturally with *in X time.* Note that it is certainly possible to get a sensible interpretation out of the non-natural alternatives (marked with a #) but that these alternative interpretations do not signal the duration of a single instance of the event.

2 (a) Clay sang for 5 minutes/#in 5 minutes
  (b) Fantasia sang the song in 5 minutes/#for 5 minutes

The imperfective paradox highlights the different completion entailments of telic and atelic predicates. Atelic predicates do not specify an inherent boundary point, and the imperfective version of an atelic predicate (3a) entails the perfective version (3b). By contrast, telic predicates do specify an inherent endpoint, and the imperfective version of a telic predicate (3c) does not entail the perfective version (3d). The difference in entailments between the imperfective and perfective versions of a telic predicate can be seen quite clearly if the continuation in (3e) is added to each. The continuation sounds perfectly natural after the imperfective version (3c), but following the perfective version (3d), the continuation sounds like a contradiction.

3 (a) Tamyra was singing ⇒
(b) Tamyra sang
(c) LaToya was singing the song -/->
(d) LaToya sang the song
(e) … when Simon cut her off.

One further test which distinguishes telic and atelic predicates is their ability to be naturally counted (Bach, 1986). Because telic predicates specify inherent endpoints for their events, those events can be naturally counted in terms of how many endpoints were achieved (4a). Because atelic predicates lack the endpoint specification, counting them is decidedly more awkward (4b).

4 (a) Ryan fell asleep 3 times
(b) ?? Ryan slept 3 times

Bach (1986) argued that atelic predicates could not be naturally counted at all, but that does not seem to be strictly true. It is more difficult to count atelic events than telic ones because atelic descriptions do not provide criteria for counting—such as an inherent endpoint—the way that telic ones do. Nevertheless, criteria can be found outside of the description itself, in the form of classifier phrases (5a) or even from general contextual knowledge (5b). The real difference between telic and atelic predicates with respect to counting, therefore, is that telic predicates provide a specific criterion, whereas atelic predicates are flexible about where their counting criterion comes from.

5 (a) Paula and Simon had 3 bouts of fighting
(b) Ali and Foreman fought 2 times

The studies reported here exploit this particular difference between telic and atelic predicates: Children are asked to use the different individuation (i.e., counting) criteria provided by telic and atelic predicates to decide how many times an

event happened. For telic predicates, correct individuation requires children to use the criteria specified in the meaning of the predicate itself, namely, its inherent endpoint. For atelic predicates, the studies rely on the fact that we require extra predicate information to decide on a plausible individuation boundary. The stimuli used will provide a strong contextual cue for individuation, namely, spatiotemporal pauses in the event's action. Thus, although in principle, one can supply a wide variety of individuation criteria for atelic predicates (because atelic predicates provide none of their own), in practice, these studies stack the contextual deck to encourage the use of a spatiotemporal criteria for individuating the atelic predicates.

COMPOSITION OF TELICITY

Fundamentally, telicity is a semantic property, namely, the specification (or lack thereof) of an inherent boundary. Syntactically, telicity has no simple direct realization. That said, transitive structures—that is, those containing an internal direct object—appear to be a particularly good way to specify the semantic notion of boundedness. Thus, it is the presence of the direct object *a pizza* that makes (1b) telic, and its absence that keeps (1a) atelic.

Direct objects make a predicate telic only to the extent that the objects themselves provide a clear boundary for an event (Krifka, 1992; Tenny, 1994; Verkuyl, 1993). Some kinds of objects—mass nouns and bare plurals, for example—do not quantify specific amounts and therefore do not identify any specific boundary for an event (6a); by contrast, quantified count nouns do pick out specific quantities and therefore can contribute to telic meaning (6b).

6 (a) Ruben ate/ate peanuts/ate peanut-butter
  (b) Ruben ate a peanut-butter sandwich

The direct object can serve to measure out the event (Dowty, 1991; see also, Krifka, 1992). The progress of the eating event in (6b) can be measured by the incremental disappearance of the sandwich—as the object goes, so goes the event. If the object does not identify a specific boundary point, such as a specific quantity of peanuts or amount of peanut butter, then neither does the event. The event effectively inherits its boundedness property from the boundedness of the object.

The importance of direct objects, and by extension, transitive structures, for creating event boundaries has made them a centerpiece of several theoretic accounts of telicity (e.g., Tenny, 1994; van Hout, 1996; Verkuyl, 1993). However, transitivity is neither necessary nor sufficient for conveying telic semantics. For example, there are intransitive predicates that bear telic semantics as in (7):

7 (a) Paula entered
  (b) The singer's voice broke

Such predicates are not really problematic, though, as they are unaccusative and require a special thematic analysis that treats the subject of the verb as if it has the thematic role typical of an internal argument. Thematically, therefore, (and on some accounts, syntactically as well) these predicates do contain a direct object.

Other examples better demonstrate the limitations of the link between transitivity and telicity. There are transitive predicates that remain resolutely atelic despite the presence of a well-specified count-noun direct object (8a); there are structures that are consistently ambiguous between telic and atelic interpretations (8b). Moreover, telic interpretations can arise from means completely unrelated to the presence of a direct object, as with prepositional phrases (8c).

8	(a)	The dog pushed the ball	(for an hour/#in an hour)
	(b)	The tailor lengthened the pants	(for an hour/in an hour)
	(c)	The butterfly flew to a tree	(#for an hour/in an hour)

Events can be measured out by elements other than direct objects. In (8c), the path defined by the prepositional phrase measures out the event; Hay, Kennedy, and Levin (1999) have argued that ambiguous sentences such as (8b) are measured out—when they are telic—by implicit measure phrases (e.g., *The tailor lengthened the pants 3 inches*). Indeed, Hay et al. suggest that all telic interpretations may depend on the presence of degree phrases, whether they are implicitly or explicitly represented. Moreover, measuring out may not even be a necessary condition for creating telic interpretations (Jackendoff, 1996).

Nevertheless, despite these examples, transitive structures play an important, recurring role in accounts of telicity, even in accounts that do not focus on it (cf. Hay et al., 1999; Jackendoff, 1996; Pustejovsky, 1995). Regardless of whether transitive structures are the only way to convey telicity, or whether their ability to do so actually depends on deeper semantic or syntactic properties, transitivity does seem to be, at least, a paradigmatically good way to exemplify telicity.

THE ACQUISITION OF TELICITY

In children's early production of verbs, telicity appears to be an important organizing category. Although in the adult language, verbs alone are not sufficient to determine the telicity of the entire predicate, very young children often produce incomplete predicates, so assessment of knowledge of telicity has been based solely on the available information—that is, the verb. Children acquiring a variety of languages have been found to distribute their verbal morphology according to the telicity value of the verb. For example, very young children acquiring English generally restrict their past-tense marking to telic verbs (*broke, made*) and their progressive marking to atelic verbs (*riding, laughing*; Bloom, Lifter, & Hafitz, 1980; Shirai & Andersen, 1995). This phenomenon, moreover, has been documented

across a wide range of diverse languages, although obviously the particular morphology varies by language (e.g., Antinucci & Miller, 1976; Berman, 1983; Bronckart & Sinclair, 1973; Weist, Wysocka, Witkowska-Stadnik, Buczowska, & Konieczna, 1984).

There are, however, many limitations to these production studies. The phenomenon itself is not categorical, but merely a trend, and it is not completely stable cross-linguistically (cf. Stephany, 1981, showing that children acquiring Greek distribute their early morphology according to stativity). However, regardless of whether telicity is the only aspectual notion children are prepared and able to mark, it is certainly the case that it is a potent force within the language acquisition process. Its strength within the production data alone suggests that children have some kind of representation of telicity from very early on in their linguistic system.

Although there has been a fair amount of investigation into how telicity may (or may not) guide the acquisition of verbal morphology for tense and perfectivity, there has been relatively little attention paid to children's knowledge of how telic and atelic meanings are actually conveyed linguistically. One study (van Hout, 2000; see also, Schulz & Penner, 2002, for a replication with German children) looked at children's understanding of how direct objects and particles influence children's interpretations of telicity in Dutch. In this study, children were shown a story in which, for example, one mouse ate a whole wheel of cheese, but another one ate only a half a wheel of cheese. They were then asked, "did the mouse eat/eat cheese/eat his cheese/eat up his cheese."[1] The telic predicates (*eat his cheese, eat up his cheese*) specified an endpoint and were therefore true only of the mouse who had reached the endpoint; the atelic predicates (*eat, eat cheese*) did not specify any ending and were therefore true of both mice. Van Hout (2000) found that children could not reliably match the predicates to the right mice until as late as 5 years of age. This finding suggests there is quite a long gap between the age at which children notice something about the telic–atelic distinction (cf. the production studies noted above) and the age at which they know how to syntactically encode it.

Given the complexities of telicity composition discussed above, there are good reasons for believing that learning how telicity is coded should be difficult. The semantic feature of boundedness may lean heavily on transitive structures for realization, but it also requires (at least) comprehension of the mass–count distinction in nouns, how nouns are integrated with verbs to form coherent predicates, systematic alternations of thematic structure as with unaccusatives, as well as idiosyncratic lexical properties of certain verbs. We might, therefore, reasonably expect that children would find the complete adult system for marking telicity to be difficult to master and take time to accomplish.

---

[1]The Dutch forms actually used were *heeft de muis gegeten/kaas gegeten/zijn kaasje gegeten/zijn kaasje opgegeten.*

However, there are reasons to believe that van Hout's (2000) studies have underestimated children's knowledge of telicity. Some elements of the studies themselves are problematic, such as the fact that only two predicates were tested (eat and drink), and these might not be the most representative predicates;[2] and that children were asked multiple questions on the same pictures, which may have caused some confusion about the task. Moreover, Penner, Schulz, and Wymann (2003) found that children as young as 2;10 would affirm the perfective telic question "*Hat Sie'se aufgemacht?*" ('Did she open it?') only for depictions of a completed act of opening a box. They also found that language-impaired children were significantly worse than normally developing children in this task. Finally, Wagner and Carey (2003) found that children as young as 3.5 years of age demonstrated a solid knowledge of telicity, using an event-counting task. In Wagner and Carey, children were shown short animated movies in which a goal was achieved via two or more spatiotemporally distinct steps. For example, a girl might paint a flower (the goal) via two distinct brush strokes. Children heard the movie described with either a telic predicate ("The girl paints a flower") or an atelic one ("The girl paints") and were asked to count how many times it happened. Children were significantly more likely to count the number of goals achieved when given a telic predicate than when given an atelic predicate. There were several differences between the studies of van Hout (2000) and Wagner and Carey, including most notably the task used and the range of predicates used. Nevertheless, it appears that under some circumstances and with some kinds of predicates, children do understand how telicity is coded in English by 3 years old.

The studies presented here will use the event-counting task of Wagner and Carey (2003) because it appears to be more sensitive to children's knowledge of telicity. Moreover, these studies will expand both the range of predicates tested as well as the range of ages: The knowledge of 2-year-old children will be investigated. The expansion of this work to this younger age range is especially important because such children can reveal the general linguistic strategies children adopt while they are still comparatively ignorant of their target language's idiosyncrasies in marking telicity.

BOOTSTRAPPING TELICITY

Regardless of just how early children know something about the English argument structure reflexes of telicity, language-specific knowledge has to be learned, and this opens the question of how children learn it. One such learning tool is syntactic boot-

[2]Moreover, in their replication of van Hout (2000), Schulz and Penner (2002) found that German-speaking adults agreed with their children's judgments. In particular, the adults restricted the transitive sentence (*did the mouse eat his cheese*) to a telic interpretation only about half the time.

strapping. Syntactic bootstrapping refers to the general process in which structural cues (e.g., number and position of arguments) can serve as a guide for determining semantic meaning (Gleitman, 1990; Landau & Gleitman, 1985). Children can use this information to identify a general semantic classification of a verb (Is it causal? Does it involve mental states?), which thereby allows for more efficient hypothesis testing to determine the specific meaning of the verb. Although the existing evidence supports the idea that syntactic bootstrapping is part of the child's language acquisition toolkit (e.g., Fisher, 2002; Naigles, 1990; Naigles, Fowler, & Helm, 1992), no previous studies have examined whether telicity is among the semantic elements that can be usefully signaled by argument structure cues.

The studies reported here investigate children's willingness to use transitive structures as a cue to telic meaning and intransitive structures as a cue to atelic meaning. Transitivity is not an unproblematic structural cue to telicity semantics, as has been discussed above. Nevertheless, it has at least two properties that make it a viable candidate for a syntactic bootstrapping account: First, it is an easy structure to find, requiring a child only to count arguments and possibly identify one as an internal argument; second, it is importantly (if not simply) connected to telicity semantics. These properties may be enough to help children at least break into the complete coding of telicity in English argument structure. An additional reason to expect children to be sensitive to structural cues in this domain is the fact that telicity itself depends on whole structures. As noted previously, telicity is not a property of verbs themselves but depends on entire predicates for interpretation.

Both experiments use the event-counting task of Wagner and Carey (2003). This task exploits the semantic fact that telic and atelic predicates provide different criteria for individuation. As discussed above (see examples in 4 and 5), telic predicates allow events to be counted in terms of their endpoints, whereas atelic predicates require some contextual specification to determine the individual units for counting. Thus, if one wants to count (i.e., individuate) how many times Paula built a house, one does not increment the count each time she pauses to review her handiwork; instead one waits until the entire house has been built. By contrast, if one wants to count how many times Paula laughed, one could reasonably increment the count each time there was a measurable pause between bouts of laughter. In these studies, children were shown short animated movies that contained a salient goal (the right individuation unit for a telic description) achieved by two or more actions separated by a spatiotemporal pause (plausible individuation units for an atelic description). Children were given either a telic or an atelic description of the movie and asked to count what happens. The dependent measure was how many events children counted.

Because the structural cue of interest was transitivity, a full range of transitivity–telicity pairings were tested. The first pair was called *canonical* and matched a transitive structure to a telic meaning (9a) and an intransitive structure to an atelic meaning (9b).

9 (a) The girl painted a flower
(b) The girl painted

The *all-transitive* pairs kept the structural cue constant but varied the telicity value: (10a) is transitive and telic but (10b) is transitive and atelic.

10 (a) The bird popped the balloon
(b) The bird poked the balloon

The *all-intransitive* pairs were the complementary set: (11a) is intransitive and telic, whereas (11b) is intransitive and atelic. Note, that the all-intransitive pairs adopt a particularly crude structural heuristic that relies simply on number and position of arguments instead of considering deeper syntactic properties such as the unaccusative nature of (11a).

11 (a) The door closed
(b) The door slid

Finally, the mass–count pairs examined whether children were sensitive to the properties of the direct object in a transitive sentence. Although both sentences in the pair are transitive, (12a)—with a count-noun direct object—is telic, whereas (12b)—with a mass-noun direct object—is atelic.

12 (a) The girl drank a glass of juice
(b) The girl drank juice

If children are using transitivity as a structural bootstrap to help them find telicity semantics, then these four kinds of sentence pairs should be unequal in their difficulty. Sentences from the canonical pairs should be easy, as the transitivity heuristic supports the correct analysis of telicity. By contrast, children are expected to make errors with the atelic items in the all-transitive and mass–count pairs and with the telic items in the all-intransitive pairs because in these cases, the transitivity heuristic provides the incorrect analysis of telicity.

EXPERIMENT 1

Method

Participants. Two groups of children were tested; there were 16 children in each age group. Approximately half of the children tested were girls. The 3-year-old group had a mean age of 3;7 (ranging in age from 3;0 to 3;11), and the

5-year-old group had a mean age of 5;2 (ranging from 4;11 to 5;6). The data from an additional 7 participants were discarded because these children failed to cooperate on more than half the trials. Participants were children in the New York City and Cambridge, Massachusetts, area who were either brought into the lab to participate or else tested in a local day care center. According to parental report, all participants had English as their primary (and in most cases, only) language. In addition, 16 adult native speakers of English participated.

Stimuli. The event stimuli consisted of 10 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 (closing a door, drinking a glass of juice). Each instance of an event was achieved through two or three temporally distinct subactions (e.g., the boy slides the door in three stages in closing it; it takes three sips to drink the glass of juice). To prevent a simple strategy in which children simply counted the number of objects involved (e.g., the number of flowers painted or the number of empty juice glasses), the relevant objects disappeared at the end of the event instance; moreover, no trace of the event remained on the screen when children were asked their final target query. The movies lasted from ≈8 to ≈20 sec each. See Appendix A for a list of the events along with the two linguistic descriptions (telic and atelic) used for each.

Procedures. The children were told that they were playing a counting game. All children began by first counting pictures of objects.[3] Children were then told that 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 that targeted either the goal of the event (a telic predicate) or the temporally discrete process actions (an atelic predicate). 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),[4] and the participant'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

[3]The results from this object task are reported in Wagner and Carey (2003), Experiment 2.

[4]Note that the predicates were always uttered with morphology (simple present and past perfective) that would preserve the entailment of completion on the telic predicates.

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 four items all received the same type of description and the second four items received the opposite type of description. Which type of description came first was counterbalanced across participants and orders.

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 or timings of children's points were used to resolve any ambiguities in their verbal answers.

Coding. Children's counting was classified according to whether the individuals specified either a kind–goal criterion or a spatiotemporal criterion. The vast majority (over 94%)[5] of answers that children gave were clearly of one of these two types. Minor errors in one-to-one correspondence (e.g., counting four slides to the door when there were actually three) were ignored. Similarly, children with nonstandard count lists (e.g., they regularly skipped the number 3) were credited with the equivalent standard count.

Results

The mean use of the goal-based individuation criterion for the two kinds of linguistic descriptions for the different age groups is shown in Figure 1. These results show a strong replication of Wagner and Carey (2003): Overall, children (and adults) select their individuation criteria between spatiotemporal and goal based, depending on the type of linguistic description provided; however, children are much more likely than adults to count spatiotemporally determined events.

An analysis of variance (ANOVA) was performed over the percentage of times participants used a goal-based counting criterion with age group (3-year-olds, 5-year-olds, and adults) as the between-subject 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, $F(2, 45) = 11.22$, $p < .001$, and of description type, $F(1, 45) = 71.2$, $p < .001$, and also a significant interaction between the variables, $F(2, 45) = 4.87$, $p < .012$.

[5]In the vast majority of the remaining answers (about 5.8% of the total trials), children counted every spatiotemporal unit but restarted their count for each separate kind or goal. Thus for the movie in which the dog pushes the ball into a can, children might count "1, 2" for the first two pushes, and then count again "1, 2," for the second iteration of the event. Because this pattern is sensitive to both the event-goal construal as well as the spatiotemporal breaks within each of these, items receiving this pattern were given half credit in each category. In the remaining 0.2% 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.

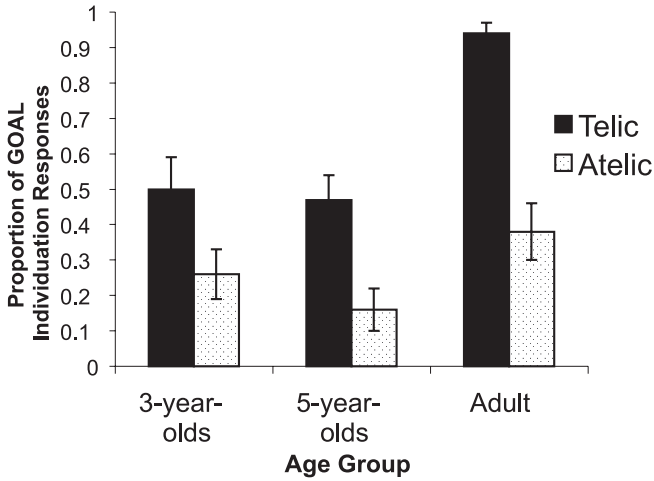

FIGURE 1 Percentages of goal responding for Experiment 1.

Post hoc analyses (Tukey's HSD, $p < .05$) on the age groups revealed that there were no significant differences between the child age groups but that children at both ages were significantly different from the adults. The overall rate of goal-based individuation for 3-year-olds was 38%, for 5-year-olds 32%, and for adults 66%. The main effect of description type reflected the fact that participants were more likely to use a goal-based criterion with telic descriptions than with atelic descriptions (mean goal response given telic descriptions was 64%, and with atelic descriptions was 26%). Moreover, planned comparisons within each age group for the effects of linguistic description show that every age group uses a significantly different counting criterion as a function of description type: 3-year-olds, $t(15) = 2.96$, $p < .01$; 5-year-olds, $t(15) = 4.27$, $p < .001$; adults, $t(15) = 7.61$, $p < .001$. 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 when using a goal-individuation criterion, whereas the children are not. As can be seen from Figure 1, the children did not differ from adults in the likelihood of counting spatiotemporally 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% ($M = 48\%$). 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 participants' 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 bi-

modal distribution among the participants 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 participants does not depend on age. To explore whether a bimodal distribution was responsible for this pattern, the responses on the telic description items only were examined, collapsed across the age groups. These responses clearly pass the Shapiro–Wilk test for being a normal (and not a bimodal) distribution ($W = .66$, $p < .001$).

A more focused analysis was also conducted on the effects of the syntactic forms of the descriptions on participants' individuations. This analysis was conducted on the combined children's data.[6] For each syntactic form pair (canonical, all-transitive, all-intransitive, and mass–count), a difference score was calculated consisting of the mean goal score given a telic description minus the mean goal score given an atelic description. These difference scores could range from positive 1 to negative 1. A score of positive 1 indicates perfect differentiation between the telic and atelic descriptions. A score of 0 indicates that participants made no distinction in their responses between the telic and atelic descriptions.[7] The difference scores for all four syntactic pairs were positive; in fact, the difference scores for all eight items were positive. The difference scores for each of the four pair types were the following: canonical, $d = .18$; all-transitive, $d = .28$; all-intransitive, $d = .50$; mass–count, $d = .16$.

An ANOVA was conducted with this difference score as the dependent variable and syntactic form as the independent variable. The results were not significant, $F(3, 124) = 2.5$, $p > .1$; moreover, a comparison of the form types with each other were also nonsignificant by Tukey's HSD. The lack of significant differences is consistent with the possibility that children are able to differentiate the telic from the atelic meaning equally well for all syntactic forms tested, at least as it pertains to individuation.[8]

---

[6]The adult data were omitted because their high rates of overall success would likely overshadow any small item effects. The children's data were combined because the previous analysis showed no differences between the groups. However, just to be sure, the analysis discussed here was also conducted on the data from each of the age groups (3-year-olds, 5-year-olds, and adults) separately. No differences between the syntactic types were found.

[7]Note that negative scores are also possible. A score of negative 1 indicates a systematic confusion in which children systematically choose the goal ending given an atelic description and the spatiotemporal ending given a telic one.

[8]One reviewer suggested that children might attempt an object-based rather than an event-based counting strategy, in which case events that involved the creation or destruction of discrete objects (individual peanuts being eaten, flowers getting painted) might be easier for children than events that involved changes of state over a single object (the door being slid; the ball being pushed). As noted in the Methods section, such a strategy was discouraged by erasing the evidence of the objects at the end of each event instance, but that does not guarantee it was not attempted. The lack of any item effects—and there are none, whether the items are grouped syntactically as in the main text, or grouped on the ad hoc basis of multiple versus single objects—strongly suggests that children who succeed on this task are not using an object-based counting strategy.

Discussion

The results overall replicated Wagner and Carey (2003): By the age of 3 years, children can reliably distinguish between telic and atelic predicates and use this information to inform their individuation of events. Moreover, as was found previously, children as old as 5 years continue to show a general spatiotemporal bias when individuating events. That is, when these children made a mistake in this task, the mistake was to adopt a spatiotemporal individuation strategy when a goal-based strategy was called for.

For all structure types, including the hypothesized problem cases (all-transitive and all-intransitive), children made numerically more goal responses for telic than for atelic sentences. It appears, therefore, that at the age of the children tested here, transitivity is doing comparatively little work as a cue to telicity. Other information (most notably, one presumes, the known meanings of the verbs and nouns in the sentences themselves) takes precedence over the structural cue. To determine whether the structural cue is losing to more potent lexical cues or whether it is never used as a cue at all, we must examine the connection between transitivity and telicity in children for whom the lexical cues are less potent.

EXPERIMENT 2

The aim of this experiment was to see if younger children—2-year-olds—who have less entrenched lexical knowledge would use the structural cue of transitivity as a cue to telicity. The task was essentially the same as it was in Experiment 1, but three modifications were undertaken to make it more age appropriate. First, in an effort to cater to the attention span of the younger children, every movie was slowed down slightly, and where necessary, movies were modified so that each one depicted only a single goal ending achieved in two spatiotemporally distinct steps.

Second, some individual items were changed. Prior to undertaking this study, it was unclear how difficult this task would be for 2-year-olds, and the item switches were introduced as a safeguard. The hardest pair type from the previous experiment (the mass–count pairs) was replaced by a pair type that older children robustly succeeded on in Wagner and Carey (2003), namely, an intransitive–prepositional phrase pair. Thus, the atelic description consisted of an intransitive sentence ("The butterfly flew"), and the telic description consisted of an intransitive sentence plus a prepositional phrase ("The butterfly flew to a tree"). In addition, two items that adult participants had found particularly awkward ("The squirrel ate/ate a peanut" and "The girl played a drum/song") were replaced with syntactically equivalent items that had worked well in Wagner and Carey (2003; "The girl built a house/worked" and "The bird popped/poked a balloon"). The remaining descrip-

tions were all identical to those used in Experiment 1. The full set of descriptions is in Appendix B.[9]

The third and final modification concerned the coding procedures. Although 2-year-olds often know their numbers, their understanding of counting appears to be quite shallow (Wynn, 1990). Children in the current study were asked to count, as in the previous experiment, but they were in addition asked to indicate when the event happened. As a response to the *when* question, pointing and verbal indications (e.g., "There!") were accepted. As will be discussed in the Methods section below, these responses were coded by two independent observers for reliability.

Methods

Participants. Sixteen 2-year-old children were tested; approximately half the participants were girls. Their mean age was 2;10 (ages ranged from 2;9 to 3;0). The data from an additional 21 participants was discarded for the following reasons: failure to answer any questions including the object warmup (2), failure to answer any questions about the event movies (7), parental interference and experimenter error (5), and failure to provide clear responses for a counterbalanced set of questions about the movies (7). Participants were children in the Cambridge, Massachusetts, area who were either brought into the lab to participate or else tested in a local day care center. According to parental report, all participants had English as their primary (and in most cases, only) language.

Stimuli. The stimuli were the same as those used in Experiment 1, with the exception of the changes already noted.

Procedures. The children were told that they were playing a counting game. All children began by first counting pictures of objects. Children were then told they would count what happened in movies. Pilot testing suggested that 2-year-olds generally did not provide numerical responses in this task, and they typically did not respond after watching the movie a single time. Nevertheless, to maintain as much parallelism as possible with the previous experiment, before be-

[9]Two of the new verbs used in this study are in fact semelfactive (*poke* and *hop*). I have identified the intransitive sentences using these verbs as atelic, as indeed they pass most of the tests for being atelic discussed in the introduction. However, as a reviewer noted, semelfactives differ from other atelic predicates in that they can be felicitously counted. The individuation of these items might, therefore, be more on a par with the individuation with the telic items. It is unclear how this fact, however, should interact with children's performance. The results in fact show that these two items are quite divergent from each other: *Hop* appeared in an intransitive sentence, and the dominant individuation strategy used for it was spatiotemporal; *poke* appeared in a transitive structure, and children frequently used a goal-oriented individuation strategy for this item (i.e., they counted when the balloon popped, not each time it was poked).

ing presented with each movie, children were asked "How many times X? Let's watch!" with the *X* term as an appropriate description of the event that targeted either the goal of the event or the temporally discrete process actions. The question was repeated again at the end of the movie (with the tense of the auxiliary changed to the past tense). Participants who failed to respond at this point were given a double set of instructions before being shown the movie again. The initial question was repeated ("How many times X?") but children were now told that they should "Show me when X." Children who failed to respond on the second screening were given extra encouragement and occasionally extra instructions telling them how to show: "You can point when X" or "You can say 'now' when X." If a child failed to provide a response after watching a movie three times, the experimenter went on to the next movie.

There were two fixed orders of presentation of the movies. The descriptions were blocked so that the first four items all received the same type of description and the second four items received the opposite type of description. Which type of description came first was counterbalanced across participants and orders.

Coding. The 2-year-old children in this task rarely provided numerical counts (only 8% of the total trials in the final analysis). It was therefore necessary to develop a more sensitive coding method. The key to the current method was intentional targeting. A point in the event was considered to be targeted if the child (a) incremented a numerical count at that point, (b) pointed with a finger to the screen at that time, or (c) verbally pointed to the screen at that time (i.e., said "there" or "painting"). This notion of targeting is essentially a temporal one: Children's responses had to be tightly coordinated with the timing of the movie. The emphasis on temporal information made this a comparatively difficult criterion for some children to meet: Children who were even mildly inattentive or who were slow to execute their pointing behavior typically provided unclear targeting. Trials with unclear targets were thrown out of the analysis, and as noted above, this led to a comparatively high rate of participant loss. The decision to adopt a temporal criterion was derived from the behavior of the older children in the previous experiment. Some of those children incremented their counts as they watched the movies, and in those cases, there was a tight temporal lock between children's counting and the moments in the movie that defined the individuation criterion they were using. Using a temporal targeting criterion in this study, therefore, is important for allowing plausible comparisons across the two studies.

More specifically, children were credited with having adopted a goal-individuation criterion if they targeted the endpoint of the event exclusively and with a spatiotemporal individuation criterion if they targeted either the first spatiotemporal break of the event or both the first spatiotemporal break and also the event's endpoint. Thus, to get credit for goal individuation, a children had to wait until the event had ended to target the movie; similarly, to get credit for

spatiotemporal individuation, the child had to target the first spatiotemporal break in the movie. Of those trials classified as showing spatiotemporal individuation, participants targeted just the first spatiotemporal break on 63%, whereas on the remaining 36% of these trials, they targeted both the first break and the ending of the event. In terms of targeting types, children provided actual numerical counts on 8% of trials, they pointed with their fingers on 43% of trials, and provided time-locked verbal points on the remaining 49% of trials. All responses were coded by two independent research assistants who agreed on 87.5% of the trials. Disagreements were resolved by the author.

Results. The performance of the 2-year-olds is first analyzed on its own, and then it is compared to the performance of the older children and adults.

Two-year-old data. The mean use of the goal-based individuation criterion for the two kinds of linguistic descriptions is shown in Figure 2. These 2-year-old children showed a strong effect of linguistic description: They were significantly more likely to adopt a goal-individuation criterion with a telic description (M = .77) than with an atelic description (M = .31), $t(15) = 8.88$, $p < .001$.

This result is quite robust: Every one of the 16 participants had a positive difference score, and every item pair had a positive difference score. To determine if some sentences types were easier than others, an ANOVA was conducted with sentence pair types as the independent measure (canonical, all-transitive, all-intransitive, intransitive–PP) and the difference in goal responding between telic and atelic versions as the dependent measure. The results showed that there was no significant effect of sentence pair type, $F(3, 60) = 1.0$, $p > .4$; the mean difference score

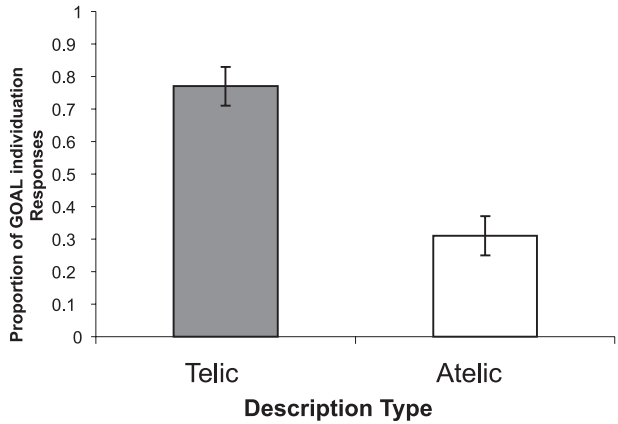

FIGURE 2 Percentages of goal responding for Experiment 2 (2-year-olds).

for each pair type was the following: canonical pairs, $d = .39$; all-transitive, $d = .26$; all-intransitive, $d = .59$; intransitive–PP, $d = .53$.

Comparison of old and young children and adults. Because there were only three types of items that participants in both experiments were tested on, the analyses conducted in this section will be restricted to all-transitive, all-intransitive, and canonical items.[10] An ANOVA was conducted on data from across the two experiments. The dependent variable was the rate of goal responding, and the independent variables were description type (telic vs. atelic) and age (2-year-old, 3-year-old, 5-year-old,[11] and adult). This analysis showed a main effect of description, $F(1, 61) = 89.0$, $p < .001$, with telic descriptions being much more likely to lead to goal responses than atelic descriptions, a main effect of age, $F(2, 61) = 12.3$, $p < .001$, but no significant interaction between the two variables. Post hoc examination of the age effect revealed that the youngest children were performing no differently overall than the adults, $F(1, 30) = 3.42$, $p > .05$, but were performing significantly differently from the older children, $F(1, 46) = 7.72$, $p < .008$.

This pattern of data was rather surprising because it appears that the 2-year-old children's performance is intermediate between that of the older children and the adults. Most notably, the 2-year-olds appear not to share the spatiotemporal bias in their responses to the same degree that the older children do. A closer examination of the data, however, revealed that the 2-year-olds' comparative lack of a spatiotemporal bias is restricted to transitive sentences.

An ANOVA was conducted on the percentage of goal-individuation responses with age as a between-subject factor, and telicity and transitivity as within-subjects factors. Note that this analysis differs slightly from the ones conducted previously. Instead of examining the difference in goal responses for each half of a sentence pair used to describe the same movie, this analysis groups sentences that share both syntactic and semantic features, averaging responses to all sentences that were telic + transitive, telic + intransitive, atelic + transitive, and atelic + intransitive. This analysis, therefore, does not highlight the differential responding between telic and atelic items (as the previous ones have done), but instead focuses on how the different age groups respond to particular form-meaning pairings. This analysis found main effects for age, $F(2, 61) = 9.98$, $p < .001$, and telicity, $F(1, 61) = 93.7$, $p < .001$. In addition, significant interactions were found between age and telicity, $F(2, 61) = 4.13$, $p < .02$, age and transitivity, $F(2, 61) = 18.3$, $p < .001$, as

[10]It might be objected that this cross-experiment analysis is unfair because two individual items used differ between the groups. However, the switched items used with the 2-year-olds had already been tested with 3- and 5-year-olds in Wagner and Carey (2003), and the performance of the older children in Experiment 1 here was virtually identical to the performance of children of the same ages in Wagner and Carey.

[11]Because Experiment 1 found no differences between the 3- and 5-year-olds, their data have been combined in these combined analyses.

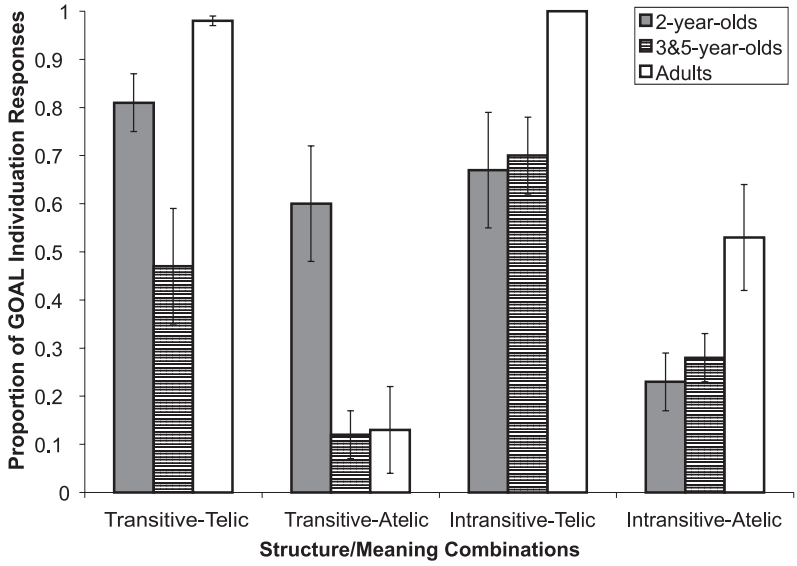

FIGURE 3 Comparison across both experiments, percentage of goal responding broken down by age and structure–meaning combination.

well as a three-way interaction between age, transitivity, and telicity, $F(2, 61) = 4.64$, $p < .013$. The graph in Figure 3 shows what is happening: With transitive sentences—regardless of telicity value—2-year-olds make more goal-individuation responses than older children; with intransitive sentences, the 2-year-olds behave just like the older children (that is, they show a spatiotemporal bias relative to adults). One-way ANOVAs on the goal-individuation responses were conducted for each grouping in Figure 3, with age as the independent variable, and all were significant: telic + transitive, $F(2, 61) = 17.9$, $p < .001$; telic + intransitive, $F(2, 61) = 3.8$, $p < .03$; atelic + intransitive, $F(2, 61) = 4.4$, $p < .02$; atelic + transitive, $F(2, 61) = 10.3$, $p < .001$. Post hoc analyses (Tukey's HSD at a $p < .05$ level) showed that for all intransitive sentences (telic and atelic), the age effect was the result of the adults being different from all the children. For the transitive and telic sentences, the 2-year-olds and adults were the same, and both groups were different from the older children; for the transitive and atelic sentences, the 2-year-olds were different from both the older children and adults.

Discussion

Overall, the 2-year-olds showed that they can use the knowledge of whether a sentence is telic or atelic to choose their criterion for the individuation of an event.

Like the older children, the 2-year-olds showed a spatiotemporal bias when dealing with intransitive sentences, erring—when they erred—on the side of oversupplying a spatiotemporal individuation. However, unlike the older children, the 2-year-olds also showed a transitivity bias: When the describing sentence was transitive, the 2-year-olds were significantly more likely than the older children to adopt a goal-based individuation strategy. This transitivity bias appears to override the spatiotemporal bias and leads to the following pattern of data: When the transitivity bias leads to the right answer (i.e., when the transitive sentence is telic—*The girl painted a flower*) the 2-year-olds do better than the older children on this task; however, when the transitivity bias leads to the wrong answer (i.e., when the transitive sentence is atelic—*The dog pushed the ball*), the 2-year-olds do markedly (and significantly) worse than the older children.

One data pattern we would expect if these children have a transitivity bias is that they should do worse at choosing their individuation criterion with the all-transitive sentence pairs; because both sentences in those pairs are transitive, we expect children to treat them all as telic. The results found here do weakly instantiate that pattern. The participants were able to distinguish between the telic and atelic transitive items (and there are no statistically significant differences among the pair types), but the weakest individuation difference score is in fact for the all-transitive pairs, as predicted.

One potential concern about the 2-year-olds' pattern of data is that it might reflect a task or age-related response bias. That is, 2-year-olds in this task might just find goal-individuation responses easier to make. Indeed, as noted by a reviewer, the 2-year-olds in the cross-age comparison show consistent goal individuation for three of four cases; only in the atelic + intransitive situation are they clearly using a spatiotemporal individuation criterion. Although a response bias account cannot be discounted entirely, there are several reasons why it is not a plausible account of the full pattern of data.

First, there is no obvious sense in which goal responses are easier for the children to make. Although a goal response always only required a single target response, spatiotemporal responses were also possible with only a single response (and in fact, two-thirds of the spatiotemporal responses did only involve a single response). Moreover, providing a goal response required children to delay their response for longer which, a priori, one would expect to be more difficult. Second, although it is true that in the cross-age comparisons, 2-year-olds provide goal-oriented responses on three of four cases, this analysis left out the noncomparable items contrasting intransitive sentences with intransitive + prepositional phrase sentences. (These items were omitted because the older children did not receive them in this study.) The 2-year-olds showed a pattern virtually identical to the older children's, including providing spatiotemporal individuation responses for the intransitive sentences. Third, although the 2-year-old's do have a higher rate of goal individuation than the older children, it is still the case that they lag far below

the ceiling rates of the adults. Most notably, in the intransitive + telic case (as well as the telic intransitive + prepositional phrase case) 2-year-olds' rate of goal responses is equivalent to the older children's and shows the equivalent spatiotemporal bias. It does not appear, therefore, that the cause of the younger children's increased goal responding with transitive items is simply the result of a general response bias toward the goal in this task.

More important, however, is the fact that children's bias appears to be syntactically conditioned: 2-year-olds fail to show the spatiotemporal bias only when the description sentences are transitive. If the increase in goal responding was a result of age or task-related demands, we would expect it to affect the younger children's performance across the board; we certainly would not expect to see it varying by the transitivity of the sentences as was found here.

GENERAL DISCUSSION

When do children know how telicity is marked in the language they are acquiring, and how do they come to know it? The experiments here focused on one important semantic difference between telic and atelic predicates, namely, that they provide different criteria for individuation. The results showed that children as young as 2;10 use the telicity of a description to choose their individuation strategy and demonstrate, therefore, that children of this age know at least some ways of marking telicity in English. Moreover, the results here suggest that children initially use transitivity as a structural cue to telicity semantics and that reliance on this cue wanes with age. In particular, the youngest children tested—but not the older groups of children—linked transitive structures to telic semantics.

The fact that the power of transitive structures to signal telicity wanes with age is to be expected—as children get older, they understand the meanings of the specific elements in the sentence better, as well as how to put those meanings together. In the terms of the syntactic bootstrapping literature (cf. Naigles, 1990; Naigles et al., 1992), the children are shifting from being "frame compliant" to being "verb compliant."

However, older children's performance is still highly subject to a spatiotemporal bias, suggesting that their specific linguistic knowledge does not wholly determine their performance. The key difference between the performance of the younger and older children is not that older children make fewer mistakes; it is that younger children's mistakes conform to a transitivity bias, whereas older children's mistakes conform to a spatiotemporal bias. Moreover, younger children are themselves not immune to the spatiotemporal bias—they show one when the sentences are intransitive. The fact that strategies come into play so strongly in this task—whether they be spatiotemporal or linguistic—suggests that counting events is a comparatively difficult thing for children to do. The origin of the spatio-

temporal strategy is most likely cognitive, and it may be linked to other, similar biases found in object individuation (see Wagner & Carey, 2003, for discussion). The transitivity bias, however, must have its source in some level of linguistic analysis. The pattern of data presented here raises two important questions: First, why do children adopt a transitivity bias, and second, why do they give it up?

With respect to the second question—why do children abandon the transitivity bias—the answer may well be that they don't really. The older children in this study are (pace the spatiotemporal errors) providing appropriate interpretations of fully grammatical English sentences containing familiar words. For a child who understands the test sentences, there is no reason to adopt an independent linguistic strategy; following known rules of semantic composition is a better idea. A more telling test would be if older children would resort to the transitivity bias in cases where the sentences contained unfamiliar linguistic elements such as novel words. In fact, the 3- and 5-year-old children in Experiment 1 were presented with a brief novel word posttest. They saw two additional movies (a man frosting a cake in two swipes and a bird popping a balloon in two pokes), one of which was described with a novel verb in a transitive structure (e.g., *The man gorps the cake*), whereas the other was described with a novel verb in an intransitive structure (e.g., *The bird geeds*). Naturally, which structure went with which movie was counterbalanced across participants. Across both the age groups, the results showed a marginal effect of transitivity, with the transitive sentence being more likely to lead to individuation on the basis of the goal, and not the intermediate substeps.[12] There are many problems in interpreting this small study—not the least of which is the fact that it followed eight trials of the same task using known words and structures designed to frequently violate the transitivity bias. Nevertheless, it raises real hope that the transitivity bias is not abandoned with age, so much, as superseded by specific linguistic knowledge of how to interpret particular verbs and predicates.

This leads next to the question of why participants at any age suspect that transitive structures are telic. Even the youngest children tested understand that transitive structures are not necessary for telic meaning. These children had no difficulty understanding that *The butterfly flew to a tree* was telic despite the lack of a direct object in the sentence. Moreover, they also seem to know that transitivity is not a sufficient cue: The youngest children do interpret transitive sentences such as *The dog pushed a ball* as atelic often enough to allow them to succeed overall on all the item pairs presented. However, if children know that transitivity is neither necessary nor sufficient to code telicity, why do they adopt it as a cue in the first place?

Two facts appear to be relevant. The first is that transitivity is in fact a reasonably good cue to telicity. As discussed in the introduction, direct objects can define event boundaries and serve to measure them out; they may not be the only way to

[12]The mean goal score for the transitive sentences was .45, and for the intransitive sentences it was .26. A two-tailed t test comparing these means was marginally significant, $t(30) = 1.99$, $p < .056$.

accomplish this, and if they lack certain properties they will not do it at all, but on the whole they are very effective at bounding events. The second factor is that the acquisition of meaning is hard (cf. Naigles, 2002), and an imperfect structural cue to meaning is better than no cue at all. Children need not be committed to a simplistic account of telicity that requires it be exclusively coded by a transitive structure; they only need to be enthusiastic about a cue that will help them make an educated guess. Transitivity is a good enough cue to telicity to accomplish that. It should be noted, moreover, that children's heuristic is not particularly comprehensive: Whereas it assumes that transitive structures signal telic meanings, it does not conversely assume that intransitive structures signal atelic meanings. Children are quite successful at assigning telic meaning to (surface) intransitive sentences.

Just how good of a cue does transitivity (or any structural cue) have to be for children to base semantic judgments on it? In the original conception of syntactic bootstrapping (Gleitman, 1990; Landau & Gleitman, 1985) it was proposed that semantic cues could be extracted from the range of frames in which a verb was used; seeing a verb in one frame alone does not typically provide enough information to make an informed judgment about the verb's semantics (see Fisher, 2002, for further argument). Because there is no necessary or sufficient link between transitivity and telicity, seeing a verb in a transitive frame is not really enough of a reason to assume telic semantics. Nevertheless, children in the studies reported here seem to have jumped readily to such a conclusion. Moreover, this is not the only study to show that children are eager to make semantic inferences from transitive structures.

Naigles and colleagues (Naigles, 1990; Naigles et al., 1992) showed that 3-year-old children use transitivity as a cue for causal action. In these studies, children were asked to act out transitive sentences containing intransitive verbs ("Noah comes the giraffe to the ark") and intransitive sentences containing transitive verbs ("Noah brings"). The 3-year-old participants enacted events that were consistent with the verb frame and not the verb itself. That is, they treated *Noah comes the giraffe* as a description of a causative event (which is actually coded by *bring* in English). Thus, despite the fact that transitive sentences need not include causal semantics (*Diana saw the mountain* is not causative), children were nevertheless willing to override the verb's semantics to preserve the link between causality and transitivity. In addition, Lidz, Gleitman, and Gleitman (2003) found, using a very similar task, that children acquiring Kannada treated transitivity as a cue to causal meaning, despite the fact that Kannada has an explicit (and wholly reliable) morphological marker of causality.

Accounts of this behavior have focused on children's apparent appreciation of the grammatical links between structure and meaning, but the story for linking transitivity to causality is about as complicated as the story linking it to telicity (see Lidz et al., 2003, for some discussion). This is not to say that the grammatical links are not true, just that they are not perfect or simple. What is remarkable is how

quickly children seem to find these complicated and imperfect links and how willing they are to exploit them in language acquisition. What's more, children are willing to link multiple semantic features onto transitivity—so far, at least causality and telicity have been documented. It seems likely, therefore, that something supports children's focus on transitivity as a cue to meaning that goes beyond mere linguistic analysis—something that makes these structure-meaning links especially noticeable and potent.

At this point, what that something might be is largely a matter for speculation. It could rest in the statistical patterns of the input. It could rest in our nonlinguistic conceptions of events. Alternatively, it might rest somewhere deep in the grammar itself. For example, functionalist accounts of language have argued that there are natural connections between certain clusters of semantic features and certain syntactic realizations of those features. More particularly, Hopper and Thompson (1980; and see also, Dowty, 1991, for a related proposal in a similar spirit) proposed a functional–semantic account of transitivity in which several semantic properties—including causality and telicity, as well as agentivity of the actor, completion of the event, and number of event participants—collectively form a functional–semantic nexus that they called *transitivity.* Language-specific realizations of Hopper and Thompson's transitivity can be complicated by a variety of factors, but the nexus itself is, for all intents and purposes, part of our Universal Grammar. Children's facility in assigning semantics, including telicity semantics, to transitive structures may, therefore, reflect their dependence on the organization of the language faculty itself.

ACKNOWLEDGMENTS

This research was funded by NIH NRSA Postdoctoral Fellowship MH12225. The manuscript preparation was funded by NIH R03 HD048533. My thanks go to Grace Lai for help in creating the movies and especially to Rachael Fuchs Melniker for her help in collecting and analyzing the data. Thank you also to Susan Carey for discussions, advice, and support.

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), 5–16.

Berman, R. (1983). Establishing schema: Children's construals of verb-tense marking. *Language Sciences, 5,* 61–78.

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

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

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

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67,* 547–619.
Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition, 82,* 259–278.
Gleitman, L. (1990). The structural sources of verb meaning. *Language Acquisition, 1,* 3–55.
Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in "degree achievements." In T. Matthews & D. Strolovitch (Eds.), *Proceedings of SALT IX* (pp. 127–144). Ithaca, NY: CLC Publications.
Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language, 56,* 251–299.
Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory, 14,* 305–354.
Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In I. Sag & A. Szabolsci (Eds.), *Lexical matters* (pp. 29–53). Stanford, CA: Stanford University Press.
Landau, B., & Gleitman, L. (1985). *Language and experience: Evidence from the blind child.* Cambridge, MA: Harvard University Press.
Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect.* New York: Mouton.
Lidz, J., Gleitman, L., & Gleitman, H. (2003). Understanding how input matters: Verb learning and the footprint of Universal Grammar. *Cognition, 87,* 151–178.
Naigles, L. R. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17,* 357–374.
Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition, 86,* 157–199.
Naigles, L. R., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb meanings. *Cognitive Development, 7,* 403–427.
Penner, Z., Schulz, P., & Wymann, K. (2003). Learning the meaning of verbs: What distinguishes language-impaired from normally developing children? *Linguistics, 41,* 289–319.
Pustejovsky, J. (1995). *The generative lexicon.* Cambridge, MA: MIT Press.
Schulz, P., & Penner, Z. (2002). How you can eat the apple and have it too: Evidence from the acquisition of telicity in German. In J. Costa & M. J. Freitas (Eds.), *Proceedings of the GALA 2001 Conference on Language Acquisition* (pp. 239–246).
Shirai, Y., & Andersen, R. (1995). The acquisition of tense-aspect morphology: A prototype account. *Language, 71,* 743–762.
Smith, C. S. (1991). *The parameter of aspect.* Dordrecht, The Netherlands: Kluwer.
Stephany, U. (1981). Verbal grammar in modern Greek early child language. In P. S. Dale & D. Ingram (Eds.), *Child language: An international perspective* (pp. 45–57). Cambridge, England: Cambridge University Press.
Tenny, C. (1994). *Aspectual roles and the syntax-semantics interface.* Dordrecht, The Netherlands: Kluwer.
van Hout, A. (1996). *Event semantics of verb frame alternations.* Unpublished doctoral dissertation, Tilburg University, Tilburg, The Netherlands.
van Hout, A. (2000). Event semantics in the lexicon-syntax interface. In C. Tenny & J. Pustejovsky (Eds.), *Events as grammatical objects* (pp. 239–281). Stanford, CA: CSLI Publications.
Vendler, Z. (1967). *Linguistics in philosophy.* Ithaca, NY: Cornell University Press.
Verkuyl, H. J. (1993). *A theory of aspectuality.* Cambridge, England: Cambridge University Press.
Wagner, L., & Carey, S. (2003). Individuation of objects and events: A developmental study. *Cognition, 90,* 163–191.
Weist, R., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E., & Konieczna, E. (1984). The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language, 11,* 347–374.
Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155–193.

APPENDIX A

TABLE A1
Complete List of Event Stimuli for Experiment 1

Linguistic Pattern	*Item*	*Telic Description*	*Atelic Description*
		"How many times … "	"How many times … "
Canonical: Transitive–intransitive	Flower	*does the girl paint a flower? (3)*	*does the girl paint? (6)*
Canonical: Transitive–intransitive	Peanut	*does the squirrel eat a peanut? (2)*	*does the squirrel eat? (4)*
Both transitive: Count–mass	Snow	*does the girl shovel the sidewalk? (1)*	*does the girl shovel snow? (4)*
Both transitive: Count–mass	Juice	*does the girl drink a glass of juice? (2)*	*does the girl drink juice? (6)*
All intransitive	Closet	*does the closet door close? (2)*	*does the closet door slide? (6)*
All intransitive	Vase	*does the vase break? (1)*	*does the vase move?(2)*
All intransitive	Drum	*does the girl play a song? (1)*	*does the girl play a drum? (3)*
All intransitive	Dog	*does the dog push a ball into a can? (2)*	*does the dog push a ball? (4)*

Note. Correct answers are in parentheses.

APPENDIX B

TABLE B1
Complete List of Event Stimuli for Experiment 2

Linguistic Pattern	*Item*	*Telic Description*	*Atelic Description*
		"How many times … "	"How many times … "
Canonical: Transitive–intransitive	Flower	*does the girl paint a flower?*	*does the girl paint?*
Canonical: Transitive–intransitive	House	*does the girl build a house?*	*does the girl work?*
All intransitive	Closet	*does the closet door close?*	*does the closet door slide?*
All intransitive	Vase	*does the vase break?*	*does the vase move?*
All transitive	balloon	*does the bird pop the balloon?*	*does the bird poke the balloon?*
All transitive	Dog	*does the dog push a ball into a can?*	*does the dog push a ball?*
Intransitive–PP	Butterfly	*does the butterfly fly to a tree?*	*does the butterfly fly?*
Intransitive–PP	Rabbit	*does the rabbit hop into a hole?*	*does rabbit hop?*

Note. PP = prepositional phrase. Correct answer is 1 for all Telic descriptions and 2 for all Atelic descriptions.