Generic Language in Parent-Child Conversations

Susan A. Gelman,
Department of Psychology, University of Michigan

Peggy J. Goetz,
Communication Arts and Sciences, Calvin College

Barbara W. Sarnecka, and
Department of Cognitive Sciences, University of California – Irvine

Jonathan Flukes
Department of Psychology, University of Michigan

Abstract

Generic knowledge concerns *kinds* of things (e.g., birds fly; a chair is for sitting; gold is a metal). Past research demonstrated that children spontaneously develop generic knowledge by preschool age. The present study examines when and how children learn to use the multiple devices provided by their language to express generic knowledge. We hypothesize that children assume, in the absence of specifying information or context, that nouns refer to generic kinds, as a default. Thus, we predict that (a) Children should talk about kinds from an early age. (b) Children should learn generic forms with only minimal parental scaffolding. (c) Children should recognize a variety of different linguistic forms as generic. Results from longitudinal samples of adult-child conversations support all three hypotheses. We also report individual differences in the use of generics, suggesting that children differ in their tendency to form the abstract generalizations so expressed.

Generic concepts—concepts of general *kinds* of things (e.g., dogs in general)—are central to human reasoning. The capacity to think about kinds as distinct from individuals underlies our abilities to make novel inferences (Prasada, 2000), explain regularities (Prasada & Dillingham, 2006), and reason about individuation and numerical identity (Carey & Xu, 1999; Macnamara, 1986; Needham & Baillargeon, 2000). Generic concepts are expressed with generic noun phrases. In English, generic noun phrases (hereafter referred to as "generics") can appear in any of a variety of forms, including bare plurals (e.g., "Knives are dangerous"), indefinite singulars (e.g., "A dog has four legs"), and mass nouns (e.g., "Gold is valuable").[1] What they all have in common is that they refer to kinds rather than individuals (Carlson & Pelletier, 1995; Lyons, 1977). For example, "Apples are juicy" refers

Address for correspondence: Dr. Susan Gelman, University of Michigan, 530 Church St., Department of Psychology, Ann Arbor, MI 48109-1043; gelman@umich.edu..
Authors' contact/affiliation information:
Dr. Susan Gelman, Department of Psychology, University of Michigan, 530 Church St., Ann Arbor, MI 48109-1043; tel.: 734.764.0268; FAX: 734.615.0573; gelman@umich.edu.
Dr. Peggy Goetz, CAS — DeVos Communication Center, Calvin College, 1810 East Beltline SE, Grand Rapids, MI 49546; tel.: 616.526.6283; FAX: 616.526.6601; pgoetz@calvin.edu.
Dr. Barbara Sarnecka, Cognitive Sciences, University of California, 3151 Social Science Plaza A, Mail Code: 5100, Irvine, CA 92697; tel.: 949.824.8495; FAX: 949.824.2307; sarnecka@uci.edu.
Jonathan Flukes, Research Evaluation and Assessment Programs, Oakland Schools, tel.: 248.209.2546; johnathan.flukes@oakland.k12.mi.us.

[1]Generics are not limited to these forms, as they can also be expressed with a definite singular noun ("The stegosaurus is extinct"), a definite article plus adjective ("The elderly need better health care"), a bare singular ("Peace on earth, good will to man"), and even a definite plural ("The Sioux Indians are a proud people.").

to the category of apples, rather than any particular apple or group of apples. Indeed, some properties are true only of a category, and not of any individual, such as "Dinosaurs are extinct" or "Rabbits are numerous."

The present study addresses two key questions: How do children learn to use the devices provided by their language to express generic knowledge? And how content-specific is generic usage? We examine these questions by analyzing natural language data in detail. Doing so allows us to examine the earliest uses of generics, and to test competing models of how generics are learned. This, in turn, allows us to understand better how young children think about kinds.

How are generics learned?

Although generics are fundamental in adult reasoning, they pose a challenging problem for learners. From a conceptual standpoint, generics are potentially difficult because they are abstract. One cannot point to a kind, one can only point to *instances* of a kind. Thus, in order to figure out what generic noun phrases refer to, the child must make an inductive inference beyond anything she can observe. Generics are thus a paradigm case of the more general observation that language acquisition requires inferential leaps (see Chomsky, 1975; Quine, 1960).

From a linguistic standpoint, generics pose a further challenge. There appears to be no single linguistic form or marker to indicate genericity, in any language (Carlson & Pelletier, 1995). Instead, each means of referring to a kind is also used (in other contexts) to refer to individuals. Consider the examples below:

"Birds lay eggs." / "Birds are flying overhead."

"A bird has hollow bones." / "A bird woke me up."

"Honey is 25% sweeter than table sugar." / "Honey fell off the spoon onto the table."

In each line, the same noun phrase (NP) is found in both sentences, but the NP is generic in the first sentence and not in the second. Thus, children cannot learn generics simply by correlating regularities in the input morphology with regularities in the world.

Given these considerations, the question of how and when generics first appear in children's speech is of great interest. We know from past work that young children have generic concepts by 3 or 4 years of age (e.g., deaf home-signers can express generic knowledge without a language model; Goldin-Meadow, Gelman, & Mylander, 2005; see "Prior research" section for more detail). However, it is unclear how children learn to use the devices provided by their language to express generic concepts, and also when this capacity first emerges.[2] When generic NPs are first used by young children, who *initiates* generic talk—children or adults? That is, do children initiate generics from their earliest uses, or are generics acquired only after an initial "apprenticeship" period, when adults initiate and scaffold children's use?

---

[2]In this context, it is important to note that count nouns always entail reference to a kind. For example, even labeling a specific individual, as in "This dog is hungry," implicitly makes reference to the kind "dogs." If children had no insight into the existence of generic categories, they would presumably treat all nouns as roughly equivalent to proper names (e.g., "dog" would only refer to the originally named instance). However, although children undoubtedly possess categories even before they learn language, and although implicit categories underlie children's use of specific reference (e.g., Waxman & Lidz, 2006), these facts do not demonstrate that children have access to a concept of dogs-in-general, to which they can refer explicitly.

The classic developmental story would suggest that generic noun phrases should emerge relatively late. On this view, children's early words and concepts are initially concrete and grounded in the "here and now" (Fenson, Dale, Reznick, et al., 1994; Inhelder & Piaget, 1964; Nelson, 1973; see Simons & Keil, 1995, for review), and only later are children able to express more abstract concepts. Recent evidence suggests that in certain tasks, children recall more detailed individuating information than adults (Fisher & Sloutsky, 2005). For example, on a memory task in which participants viewed multiple instances of various animals (e.g., cats), adults had relatively poor memory for individual instances (implying that they remembered category-level information instead), whereas children recalled individuating information much better. Sloutsky and Fisher interpret this evidence as suggesting that children are relatively poor at category-level representations, focusing instead on item-specific representations:

> "...the ability to encode the semantic level, or category information, is a product of development. ... although adults can form both category-level (or "gist") representations and item-specific representations, young children tend to form mostly item-specific representations..." (Fisher & Sloutsky, 2005, p. 595, emphases added)

Given this perspective, one should expect that young children would rarely initiate generic talk. Rather, such talk should appear primarily in response to adults who raise the topic.

In contrast, a competing position suggests that expression of generics should emerge early: On this view, kinds (not just individuals) are fundamental in children's thought (Gelman & Waxman, in press). Kinds provide the basis for children's inferences about the world and accordingly, may be reflected in children's earliest language use.

The present study also addresses the question of which theoretical model best accounts for the learning of generic noun phrases. One learning model suggests that generics are acquired by learning a limited set of forms that are associated or correlated with generic meaning in the input. Such a learning model was proposed by Smith, Jones, and Landau (1996) to account for the learning of count vs. mass nouns. In their words:

> "... children repeatedly experience specific linguistic contexts (e.g., "This is a ___" or "This is some ___") with attention to specific object properties and clusters of properties (e.g., shape or color plus texture). Thus, by this view, these linguistic contexts come to serve as cues that automatically control attention....[D]umb forces on selective attention—that is, associative connections and direct stimulus pulls— underlie the seeming smartness of children's novel word interpretations." (Smith et al., 1996, pp. 145-146)

If this model holds for the acquisition of generics, we should expect a gradual process of acquisition, by which children slowly learn to map generic meaning onto each of a variety of particular linguistic forms. For example, children might first acquire the most common form, and only later acquire less common forms.

Although the associative model may account for the acquisition of count and mass nouns (Smith, Jones, & Landau, 1996; but see Golinkoff, Hirsh-Pasek, Bloom, et al., 2000, for debate), it would seem to have difficulty accounting for the ease with which children acquire generic NPs, given that (a) generic referents are not observable, (b) the linguistic contexts associated with generic NPs vary widely, and (c) generic NPs have no morphological marker in some languages (e.g., Mandarin, Quechua).

In contrast, we propose that children learn generics by assuming that utterances *not* marked as specific are generic by default. There are innumerable ways to mark an utterance as

specific (e.g., using "this", "three", "yesterday"). All of these devices locate an utterance within an identifiable context (this place, that time). Because generics cannot be so identified, there is not a limited set of features that correspond to the set of generic utterances. If children assume that there is a distinction between generic and non-generic concepts, and if children look for markings that indicate specificity, then they could interpret all utterances that *lack* specific markers as generic. In other words, we suggest that young children assume that an utterance is generic unless that interpretation is blocked in some way (Gelman, 2003).

If the generics-as-default position is correct, we should expect three patterns in the data: (a) Children should talk about kinds early in development. Specifically, they should produce generic NPs as soon as they have mastered the relevant linguistic forms (e.g., in English, plurals, articles, and tense). (b) Parental scaffolding of generic forms should not be necessary in order for children to express generic concepts. (c) Children should quickly acquire a variety of linguistic forms to express generics.

How content-specific are generics?

Generics can be used to refer to any content, including animals (e.g., *Bats* live in caves), artifacts (*Needles* are sharp), food (*Carrots* are crunchy), inanimate natural kinds (*Gold* is shiny), social categories (*Americans* are impatient), etc. There are no formal linguistic restrictions on which content domains can receive generic expression. This is an important point, and one question we ask in the present study is whether and when children realize this linguistic fact (i.e., by using generics for a broad range of content domains).

However, a further point is that in everyday usage, parents use generics more when talking about animals (including people) than when talking about artifacts, even controlling for how often people talk about each domain (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Gelman & Tardif, 1998). Why more generics are produced for animals is not immediately clear. It is not because people are more familiar with animals than with other kinds of things, nor is it because animals are more similar to each other than other kinds of things, nor is it because animals are more thematically related to each other. Because even when all of these factors are controlled for, parents still produce more generics for animals than for artifacts (Gelman et al., 1998).

Developmental evidence can help address competing claims in the literature. On the one hand, children—like adults—may produce more generics for animal kinds. There are two different stories of why this might be the case. Atran (1998) and Pinker (1994) have proposed that humans have a distinctive, hardwired appreciation of the biological domain. If true, this would predict that generics would at first be used specifically for biological categories and gradually spread by analogy to other content domains. A related but distinct possibility is that children and adults alike have a conceptual bias to treat animal and artifact categories differently: they may more readily construe animals as kinds with an underlying essence (see Keil, 1989; Gelman, 1988; Massey & R. Gelman, 1988; see Gelman, 2003, for review). For example, for animals, transformations are judged not to influence an item's identity (e.g., a lion cannot become a tiger by wearing a tiger costume or having stripes dyed onto its fur); for artifacts, transformations can change identity (e.g., a coffee-pot can be transformed into a vase; Keil, 1989). If this is the case, then children—like adults—may construe animal kinds as having more predictive power than artifact kinds (deeper similarities, greater coherence, etc.), thereby more easily conceptualizing animal categories as abstract wholes, and hence using generics more for animals than artifacts. However, this essentialist view would suggest that children will essentialize non-biological categories as well (e.g., social kinds, such as *cowboys* and *teachers*).

A third possibility is that children may treat all concept domains as equivalent, with respect to generics. Despite evidence of animate/artifact distinctions in early childhood, children's use of generics may not link to these deeper, more essentialist concepts. For children, generics may be used simply to express regularities in the world, of which there are many for all domains. Gelman and Bloom (in press) suggest that generics may be interpreted differently over development, with young children thinking of generics as expressing regularities of *any* sort, but adults thinking of generics as expressing relatively more essential properties. An examination of the content of children's generics will help speak to these issues.

Prior research

By 2 years of age, English-speaking children hear generics produced by their parents, and they hear more generics for animals than for artifacts (Gelman et al., 1998). Similar patterns are found in the speech of Mandarin Chinese-speaking parents, despite the fact that Mandarin lacks plurality markers, determiners, or obligatory marking of aspect (Gelman & Tardif, 1998). Furthermore, preschool-aged children produce generics (Hollander, Gelman, & Star, 2002), comprehend generics as distinct from specific reference (Gelman & Raman, 2003), and are sensitive to some of the same contextual effects as adults (Gelman, Chesnick, & Waxman, 2005).

Most relevant to the current study, Goldin-Meadow, Gelman, and Mylander (2005) found that even children who receive minimal linguistic input refer to generic-like concepts. That work focused on American and Chinese children who are profoundly deaf, receive no signing input, and create their own system of communicative gestures ("home sign"). The home signs in this group include gestures that appear to refer to kinds (e.g., "squirrels eat nuts"). The fact that children produce generic-like reference in the absence of an adult language model would suggest that generics are spontaneous in children's speech. In addition, 3- and 4-year-old hearing children from two cultures (China and the U.S.) were recorded in conversation with their mothers, and were found to use generics at both ages. Finally, all four groups of children (hearing and deaf children in each of the two cultures) produced generics primarily for animates (animals and people). This last point is important, because previous studies of content differences in generic usage had focused on adults— leaving open the question of whether children, like adults, produce more generics about animals than about other kinds of things.

The present study

Goldin-Meadow et al.'s findings strongly suggest that young children possess generic knowledge. However, they leave open the question of how children learn to use the formal devices of their language to express that knowledge. How do children learn to map generic concepts onto the formal expressions in English—including bare plural NPs, indefinite singular NPs, and mass NPs? Do children map generics onto these forms one by one, suggesting that these forms are gradually acquired? Or do children acquire these forms all at once, suggesting that generic interpretations are a default, mapped onto forms that lack indications of specificity?

We examine these issues with a fine-grained investigation of the contexts of generic use, in longitudinal samples of natural adult-child speech. The extensive dataset permits us to examine children's very earliest uses of generics—children who are a full 1-2 years younger than those studied by Goldin-Meadow et al. (2005). It also permits a study of how spontaneous children's generics are, by examining the extent to which children follow their parents' lead, and by examining whether children's generics reflect more than learned facts (by looking at children's generic *questions*). Finally, the longitudinal database permits us to

see how generic use changes over development, as well as individual differences in generic use.

Methods

Participants

The data for this study were the transcripts of 8 monolingual, English-speaking children (2;0 to 3;7 at first recording, followed longitudinally through to ages 3;1 to 4;11; 6 males, 2 females) from the CHILDES database (MacWhinney & Snow, 1990). The children were: Peter (Bloom, 1970); Adam, Sarah (Brown, 1973); Abe (Kuczaj, 1976); Ross, Mark (contributed by Brian MacWhinney); Naomi (Sachs, 1983); and Nathaniel (contributed by Catherine Snow). Seven of the children were European-American and 1 was African-American; 7 of the children were from middle- to upper-class backgrounds and one was from a working-class family.

We had two criteria for inclusion of transcripts: child's age and child's MLU (mean length of utterance, in morphemes). We included only those transcripts for which children had an MLU of 2.5 or above for at least 3 transcripts in a row, so that children had adequate command of the syntactic forms necessary for generics. All transcripts from ages 2;0 to 4;11 that met the MLU constraint were included. All children were followed over at least 2 age periods (2-3, 3- 4), and half were followed over all 3 age periods. See Table 1.

Coding Procedures

Coding proceeded in 3 phases: 1) identifying generic NPs, 2) coding sequences of discourse surrounding the generics, and 3) coding the content and form of generic NPs.

Phase 1: Identifying Generics—Transcripts were searched for generic NPs from all speakers, children and adults. This procedure was reported in detail in Gelman (2003). Given the size of the database, it was neither feasible nor economical to read through the entire transcripts of each child. Therefore, using a computer algorithm, we identified all utterances with plural nouns, mass nouns, or nouns preceded by "a" or "an." Each utterance was surrounded by a window of +/- 2 lines. We amended the lists to exclude inappropriate search instances (e.g., *buttons* used as a verb instead of a plural noun).

These NPs were then coded as generic or non-generic by 2 coders, with agreement of 96%. Disagreements were resolved by discussion. Generics were defined as NPs that refer to general categories and are not tied to a particular situation or point in time. They were identified by a combination of morphological, syntactic, semantic, and pragmatic cues. For example, generics could not be examples of particular individuals or instances, and so numbers, pronouns, the word "*some*", and the word "*the*" were used as indications that an NP was not generic.[3] They also usually could not be in sentences in the past or future tense or in the progressive form (e.g., "A fish was swimming" is non-generic). Our measure of generics is thus conservative, in that we count children as producing a generic only when they can do so using the appropriate forms of English. Although children may have tried to express generic concepts at an earlier age, we cannot confidently code generics as such before children use articles appropriately, and reliably distinguish singular from plural, or progressive from non-progressive.

---

[3]We excluded NPs with "the" on the basis of prior research showing that such forms were almost never used to express generics in child-directed speech (Gelman et al., 1998).

Generics were also rechecked following the coding of discourse sequences (see below). Altogether children produced 3,593 NPs that were identified as generics, and adults produced 4,863 NPs that were identified as generics. These generic NPs were then used as the target generic nouns which were coded further.

Phase 2: Coding of Discourse Sequences—After generics were initially identified, we examined the transcripts in which the generics were used, in order to find all references to the same category (including non-generic referents) in the same discourse sequence. For example, for the generic "Can owls crawl?" the transcript was searched, by computer and by hand, for all references to "owl," regardless of form (generic or non-generic, noun or pronoun).

We defined a "sequence" as all references to a target concept (whether generic or non-generic, whether noun or pronoun), with the constraint that no more than 12 lines could occur between successive references (see Appendix A). This constraint was imposed in order to avoid including words that happen to refer to the target topic but appear in a different conversation (e.g., two different conversations about dogs at different points in the same transcript). For example, if the child produced a generic reference to dogs, all references to dogs were coded until there were 12 sequential utterances with no reference to dogs (generic or non-generic) before or after the sequence. Such utterances could include, for example, "a dog," "the dog," "chihuahua," "Fido," "it," etc. Because the conversations with Ross and Mark were only partially transcribed, generic NPs for them were included only if there were at least 12 transcribed utterances on each side of the target generic. Due to this constraint, 112 child generics and 156 adult generics were excluded from Ross and Mark's transcripts, representing 12.3% and 8.4% of child and adult generics in their transcripts, respectively.

Finally, all utterances identified in the sequences above were rechecked for genericity. For each utterance, all available aspects of context were used to determine whether an utterance was generic or not. Because of this, some forms were included that were not used in the initial search for generics in Phase 1 (e.g., generics preceded by "the," as when Adam's mother said "Why are ***the Indians*** bad?" after Adam [age 3;4] said "***Indians*** be bad"[4]; or generics expressed with "it" or "he," as when Sarah's mother said, "***it*** has two legs", referring to "an R").

We also identified the speaker (adult or child) who produced the first generic in each sequence, thus initiating the generic topic in conversation.

Phase 3: Content and Form Coding—All relevant NPs in the target sequences were coded for 3 things: content, form, and utterance type.

Content consisted of 4 categories: <u>animate</u> (referring to a person, animal, or person/animal part; e.g., "***Girls*** are bad"); <u>artifact</u> (referring to human-made objects or substances; e.g., "If you play with ***cords***, dat's very dangerous, if you play with ***cords***"); <u>food</u> (natural or prepared food or drink; e.g., "I don't like ***bread***"); and <u>other</u> (including a wide range of content domains, such as plants, non-artifact human constructs [e.g., jokes], and objects found in nature; e.g., "What do ***thorns*** do to you"). Agreement on content coding was 93%. Disagreements were resolved by discussion.

---

[4]These examples illustrate the unfortunate tendency for generics to be used to express stereotypes.

Form included four coding categories: <u>plural</u> ("little kids"; [Nathaniel]); <u>singular</u> ("a airplane" [Adam]); <u>mass</u> ("peanut butter" [Peter]); and <u>other</u> ("what kind of fish" [Abe]; "a spooky furniture", "night" [Sarah]).

Utterance type included 3 coding categories: <u>question</u> (e.g., "How come **girls** are more fun?" [Ross's father]); <u>response</u> (response to a question within the prior two lines of dialogue; e.g., in response to the question above, "Because when I see **them** I just, I just, I think, I just talk with my heart" [Ross]); and <u>neither</u> (e.g., "That's what **mothers** do" [Sarah, after her mother commanded her to shut the oven door). Agreement on utterance type was 96%. Disagreements were resolved by discussion.

Results

Altogether, the data set consisted of 3,624 sequences, each containing one or more generics (*M* = 2.33 generics per sequence). The sequences ranged in length from 1 to 158 relevant NPs. A total of 19,626 relevant NPs were in the target sequences (including both generic NPs and non-generic references to these same topics) and of these 8,456 (43%) were generic.

As mentioned earlier, two major questions guide this research: How and when do children learn to use devices of their language for referring to generic concepts? And how content-specific is their usage? To address how generic forms are learned, we examined (a) the age at which children start to produce generics in their speech, (b) the forms in which generics are produced, and (c) who initiates generic talk at different points in development, children or adults. To examine content-specificity of generics, we looked to see (d) how often children and adults produce generics for animals, artifacts, and other domains. Additionally, because of the richness of this longitudinal data set, we also were able to examine (e) whether there are stable individual differences in generic use. Prior analyses of questions (a) and (d), for the children's portion of the data only, appeared in Gelman (2003). Those data are reproduced here to enable comparison with the adults.

Emergence of Generics

We begin with the question: When do children first produce generics? All 8 children produced generics, including all 6 children with transcripts available in the youngest age period (2 years old). Table 2 shows some of the earliest examples of generics from each of these children; Figure 1 shows the rates of production as a function of speaker and age. We conducted a series of planned comparisons (paired t-tests) within each speaker tier (children, adults) and within each age group (2, 3, and 4 years) separately. The dependent measure was the proportion of total utterances with a generic. ("Total utterances" included all utterances produced by that speaker in the transcripts, regardless of whether they contained generics or were in the targeted sequences.) Children's production of generics increased significantly between 2 and 3 years of age, $t(5) = 3.16$, $p < .05$, and between 2 and 4 years of age, $t(3) = 4.70$, $p < .02$, but not between 3 and 4 years of age, $p > .20$. Adults produced more generics when speaking to 3-year-olds than to 2-year-olds, $t(5) = 4.39$, $p < .01$. When comparing children and adults within each age period, we see significantly more generics produced by adults than by 2-year-olds, $t(5) = 2.82$, $p < .05$, and a non-significant tendency in the same direction when the children are 3 years old, $t(7) = 2.29$, $p = .056$. However, by 4 years of age, children produce as many generics as adults, $t(5) = .10$, $p > .90$. In summary, production of generic NPs increases markedly between 2 and 4 years of age. Nonetheless, despite the developmental increase, the primary finding in these analyses is that children produce generics as soon as they have mastered the needed syntax.

Forms of Generic Expression

We also examined the forms of generic NPs produced by children at each age. The primary question was whether children would start out using only a subset of the forms that adults use to express generic concepts (suggesting a gradual mapping of generic meaning onto different NP forms), or whether they would, from the start, use the full range of forms. Evidence clearly supports the latter. Every speaker (including all six 2-year-olds) produced generics in all 3 primary forms examined: plural, singular, and mass forms. Although generics were most typically expressed with plural NPs (56% of children's generics; 58% of adults' generics), a sizeable number were expressed with singular NPs (20% of children's generics; 23% of adults' generics) and mass NPs (21% of children's generics; 18% of adults' generics). These rates were largely constant across the 3 ages.

A secondary question is whether generics and non-generics would show different distributions of form. Because indefinite singular generics are more restricted than bare plural or mass generics (Cohen, 2001), we predicted that singular NPs would be relatively less common for generics. We therefore examined all of the children's searched NPs, and coded the non-generics as either singular or non-singular (including mass, plural, and other). As predicted, non-generic NPs were significantly more likely to be singular (35%) than were generic NPs (20%), t-paired(7) = 5.46, $p < .001$.

Finally, we examined whether children used all three linguistic forms to express generics in each primary content domain. If we find that all forms are used in all content domains, this would further support the notion that children readily express generics with a range of forms, and understand that generics are not tied to any particular form. Indeed, this is what we found. Of the 12 possible cells (4 domains × 3 forms) that could be used by each of 8 children, only 2 of the 96 cells were empty: one child never used bare plural NPs to express generics about food; another child never used indefinite singular NPs to express generics about food. (Given that many kinds of food are only expressible with mass nouns, e.g. rice, milk, it is perhaps not surprising to find these occasional gaps.) Except for these two exceptions, every child expressed generics in each of the 12 possible form-domain combinations (4 domains × 3 forms). We therefore conclude that children honor no constraints on the forms or domains used to express generics.

We also conducted a 3 (domain: animal, artifact, food) × 3 (form: mass noun, bare plural, indefinite singular) × 2 (speaker: adult, child) ANOVA, obtaining a main effect of form, $F(2,28) = 33.85$, $p < .001$, indicating that plural generics were more common than mass or singular generics. We also found a domain × form interaction, $F(4,56) = 145.62$, $p < .001$. Not surprisingly, mass noun generics were relatively more common for food than for animates or artifacts, ps < .001, Bonferroni's. This is to be expected, given that any reference to food is more likely to be expressed with mass nouns, whether generic or non-generic. A more interesting result is that plural generics were most common for animates, whereas singular generics were most common for artifacts, ps < .01, Bonferroni's. Given that both animate and artifact generics *can* be expressed with either plural or singular form, this distribution is not in any sense forced by the structure of the language, but rather is a choice on the part of the participants (both children and adults). This finding suggests that children may be sensitive to some of the semantic distinctions between plural and mass generics (Cohen, 2001).

We therefore conclude that children readily map generic concepts onto a range of linguistic forms from the start. Nonetheless, generics are more often expressed in non-singular (plural, mass) form than non-generics, and generic form maps non-randomly onto generic content, suggesting early sensitivity to semantic distinctions between plural and singular generic forms (a point we consider in more detail in the section titled "Content Specificity"). We

next turn to the key question of whether young children initiate conversations about generic concepts, or simply follow up on generic topics that have been initiated by adults.

Initiating Generic Talk

The analyses in this section examine who initiates conversations containing generic NPs. We approached this issue by partitioning the conversations into "sequences" (see Methods), for which the initial generic was produced either by a child (child-initiated sequences) or by an adult (adult-initiated sequences). Below are two examples of *child-initiated* generic sequences, the first when Adam was 3;6; the second when Naomi was 3;2. Generic NPs appear in italic boldface type, NPs referring to individuals in the category named by the generic appear in italics.

Example 1:

Adam: Why *Paul* keeps going over here?

Mother: *He* thinks *he*'s grown a little

Adam: Why *he* can't play with children?

Mother: H'm?

Adam: Why **babies** can't play with children?

Mother: Because **they**'re just **babies**.

Example 2:

Naomi: *I*'m not ticklish.

Mother: *You*'re not ticklish?

Naomi: No.

Mother: I didn't know *you* weren't ticklish. Well, I won't tickle *you* any more if *you*'re not ticklish.

Naomi: **Naomis** are not ticklish.

Mother: **Naomis** aren't?

Mother: Oh.

In contrast, the example below illustrates an *adult-initiated* generic sequences (child is 4;1):

Mark: Happiness! (Giggling and throwing toy)

Father: I don't care whether it means happiness or not, we don't throw **pretty ponies**. *She*'s too heavy.

Ross: I know.

Father: You understand, *she*'s not soft. *She* really shouldn't be in that soft animal crowd.

We first compared the rate of child-initiated versus adult-initiated sequences. Each sequence had exactly 1 initiation (namely, the first generic produced in that sequence). For each child,

we counted up the number of child initiations and the number of adult initiations at each age, then converted these to percentages. Because these percentages are inverse to one another (i.e., at each age, the rate of child-initiated sequences and the rate of adult-initiated sequences add up to 100%), the statistical analyses focused on one speaker only (the child). As shown in Figure 2, the rate of child initiations increases significantly between 2 and 3 years of age, $t(5) = 2.99$, $p < .05$, but not from 3 to 4 years of age. We also compared the rate of initiations at each age to 50%, as a way of determining whether children or adults were significantly more likely to initiate a generic sequence. On this analysis, 2-year-olds initiated sequences significantly below 50%, $t(5) = 3.01$, $p < .05$; 3-year-olds showed a non-significant tendency to initiate sequences below 50%, $t(7) = 2.11$, $p = .073$; but 4-year-olds did not differ from 50%, $t(5) = 0.22$, n.s. Thus, at age 2 and (to a lesser degree) age 3, adults initiate most of the sequences, but by 4 years of age, children initiate as many sequences as adults.

We next conducted an analysis of just those sequences that included a child generic, to determine how often these sequences were generated by the children themselves (see Figure 3). In contrast to the earlier analysis, there were no significant changes with age. Furthermore, for both 3- and 4-year-olds, more than 50% of the sequences containing child generics were initiated by the child rather than the adult, ps $< .001$.

In sum, the developmental pattern of generic initiation followed that of generic production in general. Even 2-year-olds initiated generics, though at a consistently lower rate than adults. By the age of 4, however, these differences no longer existed, and children were as likely to initiate generic sequences as were adults. Moreover, if we consider only sequences in which the child produced at least 1 generic, most were child-initiated.

Content Specificity

In this section we examine content differences in generic use. Recall that in the analyses of linguistic form, above, we found that all children produced generics in multiple domains (including animates, artifacts, food, and other), thereby demonstrating sensitivity to the linguistic fact that all nouns can be used in noun phrases that refer to kinds. In this section, we focus on whether, despite this general principle, children prefer to express generics about certain domains more than others, as adults have been found to do (e.g., Gelman et al., 1998; Goldin-Meadow et al., 2005).

We collapsed over age groups, in order to enable a comparison that included all 8 corpora. As shown in Figure 4, the majority of generics concerned animates (people, animals, and animal parts or products), artifacts, and food. The remaining generics were combined into a category of "other". Every child produced generics in all four content categories (animate, artifact, food, other). Thus, generic utterances were not limited to one or another content domain, for any of the 8 children that we studied.

We conducted a 2 (speaker: child, adult) × 3 (content: animals, artifacts, food) ANOVA. The dependent measure was the percentage of generic NPs of each content domain. Because the ANOVA does not permit inclusion of all categories when the percentages add up to 100%, we excluded the "other" category, as it is least frequent and of least interest from a theoretical perspective. Results indicated a main effect of content, $F(2,28) = 10.16$, $p < .001$. Pairwise comparisons indicated that more generics were produced for animates than for artifacts or food, both ps $< .02$. Artifacts and food did not differ from one another. There was also a significant main effect of speaker, $F(1,14) = 23.23$, $p < .001$, indicating that adults produced more generics falling into the "other" category than did children. There was no significant interaction between speaker and content, indicating that the tendency to produce

more generics for animates than for artifacts or food was equally evident for both kinds of speakers (children and adults).

Non-generics—Before concluding that children have an animacy bias in producing *generics*, it is important to analyze children's baseline speech.[5] Thus, all of the NPs that were selected for coding in this study (bare plurals, mass nouns, indefinite singulars, plural pronouns; see Method), were coded for content (see Figure 5), whether or not they were generic. We conducted a repeated-measures ANOVA with genericity (generic, non-generic) and content (animals, artifacts, food) as the within-subjects factors, excluding the "other" category in order to make the analysis comparable to that of generics above. The dependent measure was the percentage of either generic or non-generic NPs produced in each content domain. Results indicated a main effect of genericity, $F(1, 14) = 40.27$, $p < .001$, and a genericity $\times$ content interaction, $F(2,14) = 8.18$, $p < .005$. Pairwise post-hoc tests with the Bonferroni correction indicated that more generics than non-generics were produced for animates, $p < .02$, and that more non-generics than generics were produced for artifacts, $p < .005$. Thus, children's tendency to produce more generics about animates cannot be attributed to children talking more about animates overall.

Sequence-initiating generics—Next we examined the content of the *first* generic in each sequence (rather than all generics in each sequence). We wished to learn about the content-specificity of generics that are spontaneous and not simply following the lead of the other speaker. This approach has a secondary benefit as well. In our coding, we noticed that some of the sequences, particularly for animate topics, were extremely long (e.g., one of Mark's sequences involved over 100 generics about cavemen), despite the fact that the modal number of generics per sequence was just 1. It is therefore possible that the larger number of generics produced about animals and people overall reflected just a few very long conversations about these things. By counting first generics only, each generic *sequence* is counted just once, so that we can determine whether the content patterns reported above hold up more generally.

Results were very similar to those of the overall analyses reported earlier. We conducted a 2 (speaker: child, adult) $\times$ 3 (content: animals, artifacts, food) ANOVA. The dependent measure was the percentage of generic NPs of each content domain (counting only the first generic NP in the sequence). Results indicated a main effect of content, $F(2,28) = 12.09$, $p < .001$. Pair-wise comparisons indicated that more generics were produced for animates than for artifacts, $p = .051$, more generics were produced for animates than for food, $p < .05$, and more generics were produced for artifacts than food, $p < .05$. There was also a significant main effect of speaker, $F(1,14) = 9.00$, $p = .01$, indicating that adults produced more generics falling into the "other" category than did children.

Generic questions—We next examined the content of speakers' generic *questions*. Questions are of special interest because they express curiosity and uncertainty. If speakers display content effects in statements only (not questions), then this could mean that content differences reflect memorized facts (e.g., "birds fly", "milk comes from cows"). In contrast, content effects in questions could not readily be attributed to prior knowledge. Overall, 743 of children's generics (21%) were questions, and 1,932 of adults' generics (40%) were questions. An example of one child's generic questions can be seen in the sample below:

Adam: Does *snake* crawl?

---

[5]Coding of non-generics was conducted for the children only, given that previous research already demonstrated that domain-specificity of parental generics cannot be attributed to baseline frequency of speech in the animate domain (Gelman et al., 1998).

Mother: Yes, *snakes* crawl.

Adam: Does *lions* crawl, I mean, walk? Does *lions* walk?

Mother: I guess *they* walk, yes.

Adam: Does *lions* crawl, huh? Does *lions* crawl?

Mother: I really don't know, Adam, if you want to be so specific.

We conducted a 2 (speaker: child, adult) × 3 (content: animals, artifacts, food) ANOVA. The dependent measure was the percentage of generic questions from each content domain. Results indicated a main effect of content, $F(2,28) = 10.03$, $p < .001$. Pairwise comparisons indicated that more generic questions were produced for animates ($M = 46\%$) than for artifacts ($M = 26\%$), $p < .02$, and more generic questions were produced for animates than for food ($M = 18\%$), $p < .001$. There were no significant differences between artifacts and food. There was also a main effect of speaker, $F(1,14) = 10.42$, $p < .01$, indicating that children produced more generics falling into these three content domains (93%) than did adults (85%). There was no significant interaction between speaker and content.

Individual Differences

In this section we examine the nature and extent of individual differences in generic usage, in order to shed light on the mechanisms by which generic language is acquired. Specifically, if there are consistent differences in the rate of generic production, we can ask whether these differences seem to be motivated by differences in the input, or instead whether these differences seem more endogenous to the child (e.g., corresponding to stable differences in children's tendency to form generalizations).

Spearman's rho correlations revealed that there were strong consistencies over time for the children: *rho* = .78 from ages 2 to 3 ($N = 6$, $p < .05$, one-tailed), *rho* = .97 from 3 to 4 ($N = 6$, $p < .001$), and *rho* = .95 from ages 2 to 4 ($N = 4$, $p = .051$). In contrast, adult speakers showed less consistency over time: *rho*s = .94 from child ages 2 to 3 years ($N = 6$, p < .01), .49 from child ages 3 to 4 years ($N = 6$, n.s.), and -.80 from child ages 2 to 4 years ($N = 4$, n.s.). The same patterns hold when we consider parents only (excluding adult speakers who were not parents): *rho*s = .60 from child ages 2 to 3 years ($N = 6$, n.s.), .66 from child ages 3 to 4 years ($N = 6$, n.s.), and -.20 from child ages 2 to 4 years ($N = 4$, n.s.).

Next we looked to see whether children who produced more generics were in conversation with adults who produced more generics. On the whole, there is at best a weak relationship between children and adults, in rate of generics, with Spearman's rho correlations as follows: .34 at age 2 ($N = 6$, n.s.), .65 at age 3 ($N = 8$, $p = .078$), and .64 at age 4 ($N = 6$, n.s.). In other words, children who produce relatively more generics tend not to be in conversation with adults who produce relatively more generics. When correlating children's speech with that of *parents only* (excluding non-parent adult speakers), the patterns are roughly the same: *rho*s = .17 at age 2 ($N = 6$, n.s.), .65 at age 3 ($N = 8$, $p = .078$), and .72 at age 4 ($N = 6$, $p = .10$). Certainly the small numbers of participants limits the power of these correlations, and may account for why correlations upward of .60 and above are non-significant; nonetheless, it is clear that the consistency of children over time cannot be attributed wholly to the parents.

Given that individual differences in children's generic usage do not directly correspond to variation in the input, we next examined whether children who produce more generics show evidence of being better at forming other sorts of generalizations in their language. First, we hypothesized that children who produced more generics might be more advanced in their

syntactic skills. To test this, we correlated rate of generics with child's MLU. For children, rate of generic production correlated with MLU as follows: age 2 (*rho* = .60, n.s., *N* = 6), age 3 (*rho* = .95, *p* < .001, *N* = 8), age 4 (*rho* = .77, *p* = .07, *N* = 6), and across all ages (*rho* = .76, *p* < .05, *N* = 8). In contrast, for adults rate of generic production did not correlate significantly with child MLU (*rho*s = .14, .59, .49, and .62 at ages 2, 3, 4, and combined across all ages, all n.s. except for the combined score, where *p* = .10). Thus, children who are linguistically more sophisticated, are more likely to produce generic utterances. Although the small number of participants in the present dataset does not enable us to examine the strength of individual differences while controlling for MLU, such analyses would be useful in the future.

Next, we examined whether children who produced more generics are more likely to generalize (or overgeneralize) morphological regularities. Thus, we correlated children's generic usage with their overregularization rates, as calculated by Marcus, Pinker, Ullman et al. (1992, Table 2, p. 36). Data were available for 6 of the 8 children we studied, excluding Mark and Ross. The findings are somewhat mixed. On the one hand, the Spearman's *rho* correlations were non-significant (.50, .37, and .20 at ages 2, 3, and 4, respectively). On the other hand, of these 6 children, one child (Abe) was an outlier in both overregularizing and forming generics. Abe made many more overregularization errors than the other children, and he produced many more generics than the other children. Indeed, Pearson r correlations of these data are highly significant (.96 at age 2, *p* = .01; .92 at age 3, *p* = .01; .86 at age 4, *p* = .14). Clearly more data are needed. However, these findings raise the provocative possibility that children who are apt to generalize about kinds (as reflected in their use of generics) may also be more apt to generalize about grammatical structures (as reflected in their MLU and overregularizations).

Finally, there is a high correlation between the rate at which children produce generics and the animacy bias discussed earlier (Spearman's *rho* = .81, *N* = 8, *p* < .02), despite the fact that children do not talk significantly more about animates as their MLU increases (*rho* = .57, *N* = 8, n.s.). In other words, children who produce the *most* generics are also most likely to produce a relatively higher proportion of *animate* generics. (Note that this effect is not due to certain children producing more generics overall, since the proportion of animate generics controls for the overall rate of generic production.) In contrast, adults showed no relationship between animacy and rate of generic production (*rho* = .24, *N* = 8, n.s.).

Discussion

Generic concepts are central to human reasoning—they are the foundation for category-based induction, explanation, prediction, and deontic judgments (Prasada, 2000; Prasada & Dillingham, 2006). Likewise, learning to express generic concepts is an important part of language learning. Although past research has demonstrated that generic concepts are available to young children by 3 or 4 years of age (Goldin-Meadow et al., 2005), little is known about how children learn the devices of their language to express generic concepts. We are interested in their initial use, change over time, and relatedness to parental speech. Accordingly, the present study examines when and how children and parents produce generic NPs in natural language. We return to the questions that we began the paper with: How do children learn to use the devices provided by their language to express generic knowledge? How content-specific is their usage?

Acquisition of generics

First, this work confirms the earlier findings by Goldin-Meadow et al., that reference to generic kinds emerges early and spontaneously. Despite the potential conceptual and linguistic challenges of acquiring generics, even the youngest children in our sample (2-

year-olds) produced generic NPs, and by 4 years of age children produced them as often as adults did. Furthermore, children often initiated conversations about generic concepts: that is, they often provided the *first* generic for a given category in a particular conversation. Even at age 2, children initiated most of the generics that they produced (in most sequences that included a child-produced generic NP, children in fact provided the initial generic in the sequence). Thus, at no age did children's generics seem to be simply following the lead of an adult.

These results have two major implications. First, they further undermine long-standing claims in the literature that young children focus on concrete, individual-level representations, and have difficulty reasoning about abstract, category-level representations (e.g., Fisher & Sloutsky, 2005; Inhelder & Piaget, 1964; see Simons & Keil, 1995, for review and critique). The evidence consistently demonstrates that even 2-year-olds express thoughts about abstract entities—namely, kinds. They do so even when not prompted by a parent or other adult; indeed, most of the generics 2-year-olds produce are self-initiated (see also Goldin-Meadow et al., 2005). Furthermore, they use generics when seeking information (i.e., in their questions), not just when repeating learned facts.

Second, these results suggest that generic concepts are already present and available when children begin to learn generic language. The data from 2-year-olds are particularly revealing on this point, because children in that youngest age group had just mastered production of the relevant linguistic forms (plurality, determiners), as indicated by their MLU (Brown, 1973). Importantly, we found no lag between the time when children acquired these forms and the time when they began to produce generics, consistent with the idea that generic concepts were already present, awaiting verbal expression.

These findings provide evidence regarding which theoretical model best accounts for how children learn to express generic noun phrases. Perhaps most importantly, children used a variety of linguistic forms (bare plurals, mass nouns, and indefinite singulars) to express generics from the start, and did so within each of the major content domains (animals, artifacts, food, and other content). Children are thus strikingly productive in the variety of forms and contents with which they express generics. We therefore suggest that generics are *not* acquired by first learning a limited set of forms that are associated with generic meaning in the input. If children were using this more bottom-up associative approach, we might expect that they would at first learn the most commonly expressed generic form (i.e., bare plurals), and only later acquire the rarer forms (e.g., indefinite singulars; mass nouns). Instead, the data clearly indicate that a variety of forms of generic expression are used from the start.

In contrast, we propose that children learn generics by assuming that utterances *not* marked as specific are generic by default. We argue that it would be highly difficult for children to acquire generics by means of learning a fixed set of form-meaning correspondences, given the complexity of forms and contexts required for interpreting an utterance as generic. Our finding that children readily make use of multiple forms to express generics supports the notion that children are not painstakingly mapping generic meaning onto a small set of fixed linguistic forms. We therefore hypothesize that in learning generics (at least in English), the child's task is not to acquire a particular form, nor to map one formal set of cues onto a set of perceptual properties. Rather, the child's task is to learn how to recognize references to specific individuals, given that the default interpretation of any NP is the generic interpretation.[6]

Two other aspects of children's formal usage of generics are noteworthy. First, generics were disproportionately used in plural and mass form, relative to non-generics (i.e., plural and

mass forms represented a higher proportion of generics than of non-generics). This finding is interesting because in adult speech, indefinite singular generics are more restricted in distribution, compared to bare plural generics (Cohen, 2001). The finding that children are less likely to express generics using the indefinite singular seems to suggest an early sensitivity to this distributional constraint. Second, although all linguistic forms were used with all content domains, there was a tendency for form and content to correlate: bare plural generics were disproportionately used for animals (e.g., "Do butterflies bite?" [Abe]), and indefinite singular generics were disproportionately used for artifacts (e.g., "Like a truck" [Nathan]). This result implies that formal differences in expression may have corresponding conceptual implications (e.g., perhaps animal kinds are construed as more coherent, and thus more likely to elicit plural expression). We next consider the issue of content-specificity more directly.

Content-Specificity

Another important finding of the present study concerns how generics are distributed across different content domains. The first point is that children are sensitive to the fact that there are no linguistic constraints on which content domains allow generic reference. As noted earlier, all content domains can be used to express generics, and each of the 8 children in this study seemed to appreciate this, appropriately expressing generics in all the content categories (animate, artifact, food, other). At no point do children seem to expect that generics are limited to a single content domain.

Nonetheless, it is interesting that both children and adults produced more generics for animates than for artifacts or food. This result is consistent with prior cross-sectional studies of parental speech (Gelman et al., 1998) and child speech (Goldin-Meadow et al., 2005). As in the prior work, the domain differences did not arise from people talking more about animates overall. Nor was it the case that children were simply following the lead of adults, as the animacy bias was found even when looking only at the first generic in each sequence. This result is also consistent with the earlier study of deaf children of hearing parents (Goldin-Meadow et al., 2005), where children produced generics in the absence of any adult language input. Thus, the tendency to produce more generic statements about animates is robust and pervasive.

The obvious question, then, is why the animacy bias emerges. That an animacy bias was demonstrated even by preschoolers suggests that it does not require an extensive knowledge base. It is possible, however, that speakers have accumulated more factual knowledge about animal categories than about artifact categories. For example, perhaps adults know more properties about dogs as a category than about chairs as a category, thereby giving them more to say about that domain. However, this seems unlikely to be the whole story, given that the animacy bias shows up even in generic *questions*, which express not factual knowledge but rather uncertainty and curiosity.

Does the animacy bias suggest, then, an early-emerging tendency to treat biological categories as special (Atran, 1998; Pinker, 1994)? At first glance, the data might seem to support such an account. But *categories* of animals and people are not necessarily biological, even if the individuals that belong to them are. A sizeable subset of the generic NPs for animates in this dataset actually refer to social, rather than biological kinds (teachers, poor people, cowboys, Italy people, strangers, good little girls, bad people,

---

[6]Hollander, Gelman, and Star (2002) provide further evidence that generics may be a default for young children, in that the youngest children (3-year-olds) interpreted "all" and "some" questions as if they were generic. See also Leslie (2006) for extended discussion of the "generics-as-default" argument.

carpenters). The categories generating the most generics are therefore more accurately described as "animate" than "biological" (at least, from an adult perspective).

Our data thus provide evidence against the idea that children treat any lexicalized item as equally good candidates for a generic, as children find some items better candidates than others. The data also suggest that children treat social as well as biological categories as good candidates for generics. Therefore, we suggest that children, like adults, form generics regarding categories they view as more richly structured.

Individual Differences

Although generics are consistently produced in natural speech, previous studies have found wide variation in the frequency of generics in productive speech. For example, in one study where mothers all read the same picture book to their children, rates of generic usage ranged from 0% to over 33% of all utterances (Pappas & Gelman, 1998). Likewise, in a project examining parent-child conversations about gender, the rate of maternal generics ranged from 0% to 67% of all on-task utterances (Gelman, Taylor, & Nguyen, 2004). However, past work was limited by recording mother-child dyads in only a single, brief conversation. We cannot tell whether differences found in a 15-minute laboratory task are stable over time, or merely reflect random variation. Another unanswered question going into the present study was whether individual differences in adults correspond to individual differences in children.

The present study revealed stable individual differences in children's rate of generic production, such that the children who produce relatively more generics at age 2 also produce relatively more generics at age 3, and those who produce relatively more generics at age 3 also produce relatively more generics at age 4. It does not appear that adult input can account for these stable differences, as there were no significant correlations between parents and children at any age. Although the present data include too few participants to warrant the conclusion that parental input has no effect (as perhaps with a larger sample the parent-child correlations would have reached statistical significance), we can nonetheless reasonably infer that the individual differences in the children are not reducible to parental input.

Why, then, do children display these individual differences? As noted earlier, children's rate of generic production correlates significantly with MLU: children who are more advanced linguistically are also more likely to produce generics. But before we conclude that these individual differences reflect wholly developmental differences, we note that children's MLU is already quite high by age 4 (with individual differences ranging from 3.26 to 7.47), suggesting considerable linguistic skills by age 4, across the sample. Therefore, developmental factors alone are unlikely to account wholly for individual variation in frequency of generics.

One highly speculative possibility is that individual differences in generic usage may reflect broad differences in children's tendency toward abstraction. Some children may more readily form bold abstractions than other children. It is interesting in this regard that children who form more generics also display greater command of early grammatical categories (as seen in MLU), *and* show some tendency to overregularize (though clearly more data are needed on both these points).

Individual differences in generic usage may also have a conceptual basis. Perhaps some people tend to focus on individuals *as individuals*, whereas others focus on individuals *as representing broader kinds*. There is already evidence of stable individual differences in children's preference for thematic versus taxonomic relations. Taxonomic relatedness is

based on shared category membership (for example, horses and cows are the same type of thing) and can readily be expressed by generics. Thematic relatedness is based on interrelatedness in the world (for example, a horse and a barn are found together) and may reflect more of a focus on individuals. Interestingly, individual differences in thematic vs. taxonomic preferences at age 3 can be traced back to individual differences in behavior at 13 and 24 months of age (Dunham & Dunham, 1995). It would be interesting to know if children with a taxonomic preference are also more likely to produce generic NPs.

More speculatively, variation in generic use could also reflect individual differences in essentialism (see Haslam, Rothschild, & Ernst, 2000, for evidence of such variation in adults). An example of individual variation in children's essentialist reasoning can be seen in their beliefs regarding intelligence. Dweck (1999) finds stable individual differences in this regard, with some children consistently endorsing an "entity" theory (that intelligence is immutable; in other words, an essentialist theory) and other children consistently endorsing an "incremental" theory (that intelligence is flexible and can be improved with practice and experience; in other words, a non-essentialist theory). These differences can be seen in children as young as first grade, and have powerful implications for children's persistence in the face of failure by fifth grade (Cain & Dweck, 1995). It would be interesting to determine whether generic rates reflect individual differences in conceptual orientation, and if so, the causal basis of such a correlation.

The current study used longitudinal, densely sampled data. The limitation of this approach is that one can examine only a small number of children. In future work, it would be valuable to examine these issues in a larger, more representative group of children. In the current study, 7 of 8 children were of middle- to upper-class backgrounds, and most were children of developmental psychologists and/or psycholinguists. Differences in class and/or education are likely to influence the patterns of individual differences. A larger and more diverse sample should provide more insights into whether the consistent individual variation in generic usage seen here reflects stable individual differences, or instead reflects developmental changes that will disappear once the children reach a certain level of linguistic skill. To this end, it would also be useful to conduct an analysis controlling for MLU, using a larger sample of participants.

Are the "Generics" Truly Generic?

One important question to consider is whether we can be confident that the NPs coded as generic in these transcripts truly express generic concepts. With natural language data such as these, we must take care that the coders do not "read in" interpretations that the children did not have in mind. Although this interpretive problem cannot be wholly countered, we believe there are several reasons to have confidence in the coding presented here.

First, we were conservative in our coding: The generic utterances were initially identified exclusively as those that had the appropriate form (either bare plural, plural pronoun, mass noun, or indefinite singular), they were checked within the context of the full transcripts in which they appeared, and both coders had to agree that the utterance was generic. If the coders were in doubt about whether an utterance was generic or not, the utterance was coded as non-generic.

Second, we know from prior experimental tasks that children can *interpret* generics appropriately by the end of the third year of life (Gelman & Raman, 2003). We also know from prior quasi-experimental studies that children (and adults) produce more responses that are coded as generic when they are placed in certain contexts (e.g., reading through a picture book vs. playing with toys; Gelman & Tardif, 1998; Gelman, Chesnick, & Waxman, 2005).

Thus, the frequency of utterances coded as generic varies sensibly as a function of the item under consideration, even when controlling for content.

Third, these interpretive issues are of greatest concern for the youngest children, whose command of the morphosyntax and pragmatics is most fragile. However, we included children only if they consistently displayed MLUs of 2.5 or above (thus guaranteeing a certain level of syntactic skill). Moreover, in the analyses of individual differences, the youngest children's generic responses correlated highly with those of the older children. We therefore believe that the utterances coded as generic were intended as such by the child and adult speakers.

Summary and Conclusions

From the time that children first master the linguistic tools to form generic expressions in English, they produce such expressions in everyday speech. Furthermore, from their earliest appearance, generics are most often used to refer to kinds of animals and people. These generic expressions are not just produced in response to what adult interlocuters say; they are often spontaneous. It is often the child who produces the first generic expression on a topic. Because prior studies have shown that children of this age *comprehend* generics (Gelman & Raman, 2003), we can infer that children possess generic concepts by age 2-1/2, and probably earlier. Although the current data cannot speak to development prior to age 2, we note that generics are constructed on the basis of remarkably little evidence. Generic referents can never be displayed: One can never see or point to *dogs* as a kind, one can only point to *individual* dogs. Likewise, linguistic concomitants of generics are always ambiguous, because cues that mark genericity in English (e.g., bare plural nouns, indefinite singular nouns) are also used to express non-generic concepts. That generics are produced so early despite these inductive challenges suggests that the generic/non-generic distinction is one that even very young children are prepared to learn.

Acknowledgments

Appendix A

Sample sequence, from Abe (age 2;9). (*italics* = all references to target noun; **bold font** = generic)

Abe: It's *a elephant*.

Mother: *A elephant*?

Abe: Uhhuh.

Mother: Do you like **elephants**?

Abe: Uhhuh we seed *one* at the zoo.

Mother: We sure did we saw *him* eating, didn't we?

Abe: Uhhuh hay, *he* ate hay!

Mother: Uhhuh *elephants* like hay.

Abe: Uhhuh and peanuts we getted some peanuts for *him*.

Mother: That's right, next time we go to the Chicago Zoo, maybe we'll see *him* again. *He*'ll probably remember you, because *elephants* never forget.

References

Atran S. Folk biology and the anthropology of science: Cognitive universals and cultural particulars. Behavioral and Brain Sciences. 1998; 21:547–609. [PubMed: 10097021]

Bloom, L. Language development; form and function in emerging grammars. MIT Press; Cambridge, MA: 1970.

Brown, R. A first language: The early stages. Harvard; Cambridge, MA: 1973.

Cain KM, Dweck CS. The relation between motivational patterns and achievement cognitions through the elementary school years. Merrill-Palmer Quarterly. 1995; 41:25–52.

Carey, S.; Xu, F. Sortals and kinds: An appreciation of John Macnamara.. In: Jackendoff, R.; Bloom, P.; Wynn, K., editors. Language, logic, and concepts: Essays in memory of John Macnamara. MIT Press; Cambridge, MA: 1999. p. 311-335.

Carlson, GN.; Pelletier, FJ. The generic book. University of Chicago Press; 1995.

Chomsky, N. Reflections on language. Pantheon; New York: 1975.

Cohen A. On the generic use of indefinite singulars. Journal of Semantics. 2001; 18:183–209.

Dunham P, Dunham F. Developmental antecedents of taxonomic and thematic strategies at 3 years of age. Developmental Psychology. 1995; 31:483–493.

Dweck, CS. Self-theories: their role in motivation, personality, and development. Psychology Press; Philadelphia: 1999.

Fisher AV, Sloutsky VM. When induction meets memory: Evidence for gradual transition from similarity-based to category-based induction. Child Development. 2005; 76:583–597. [PubMed: 15892780]

Gelman SA. Children's expectations concerning natural kind categories. Human Development. 1988; 31:28–34.

Gelman, SA. The essential child: Origins of essentialism in everyday thought. Oxford University Press; New York: 2003.

Gelman SA, Bloom P. Developmental changes in the understanding of generics. Cognition. in press.

Gelman SA, Chesnick R, Waxman SR. Mother-child conversations about pictures and objects: Referring to categories and individuals. Child Development. 2005; 76:1129–1143. [PubMed: 16274430]

Gelman SA, Coley JD, Rosengren K, Hartman E, Pappas A. Beyond labeling: The role of maternal input in the acquisition of richly-structured categories. Monographs of the Society for Research in Child Development. 1998; 63(1)

Gelman SA, Raman L. Preschool children use linguistic form class and pragmatic cues to interpret generics. Child Development. 2003; 24:308–325. [PubMed: 12625452]

Gelman SA, Tardif TZ. Generic noun phrases in English and Mandarin: An examination of child-directed speech. Cognition. 1998; 66:215–248. [PubMed: 9689770]

Gelman SA, Taylor MG, Nguyen S. Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. Monographs of the Society for Research in Child Development. 2004; 69(1)

Gelman SA, Waxman SR. Looking beyond looks: Comments on Sloutsky, Koos, & Fisher. Psychological Science. in press.

Goldin-Meadow S, Gelman SA, Mylander C. Expressing generic concepts with and without a language model. Cognition. 2005; 96:109–126. [PubMed: 15925572]

Golinkoff, RM.; Hirsh-Pasek, K.; Bloom, L., et al. Becoming a word learner: A debate on lexical acquisition. Oxford University Press; New York, NY: 2000.

Haslam N, Rothschild L, Ernst D. Essentialist beliefs about social categories. British Journal of Social Psychology. 2000; 39:113–127. [PubMed: 10774531]

Hollander MA, Gelman SA, Star J. Children's interpretation of generic noun phrases. Developmental Psychology. 2002; 38:883–894. [PubMed: 12428701]

Inhelder, B.; Piaget, J. The early growth of logic in the child. Norton; New York: 1964.

Keil, F. Concepts, kinds, and cognitive development. MIT Press; Cambridge, MA: 1989.

Kuczaj, S. –Ing, -s, and –ed: A study of the acquisition of certain verb inflections. Ph.D. dissertation, University of Minnesota; 1976.

Leslie, S. Generics: Cognition and comprehension. Unpublished doctoral dissertation, Princeton University; 2006.

Lyons, J. Semantics. Vol. 1. Cambridge University Press; New York: 1977.

Macnamara, J. A border dispute. MIT Press; Cambridge, MA: 1986.

MacWhinney B, Snow C. The child language data exchange system. Journal of Child Language. 1990; 17:271–296.

Marcus GF, Pinker S, Ullman M, Hollander M, Rosen TJ, Xu F, Clahsen H. Overregularization in language acquisition. Monographs of the Society for Research in Child Development. 1992

Massey C, Gelman R. Preschoolers' ability to decide whether a photographed unfamiliar object can move itself. Developmental Psychology. 1988; 24:307–317.

Needham A, Baillargeon R. Infants' use of featural and experiential information in segregating and individuating objects: A reply to Xu, Carey and Welch (2000). Cognition. 2000; 74:255–284. [PubMed: 10640572]

Pappas A, Gelman SA. Generic noun phrases in mother-child conversations. Journal of Child Language. 1998; 25:19–33. [PubMed: 9604567]

Pinker, S. The language instinct. W. Morrow; New York: 1994.

Prasada S. Acquiring generic knowledge. Trends in Cognitive Sciences. 2000; 4:66–72. [PubMed: 10652524]

Prasada S, Dillingham EM. Principled and statistical connections in common sense conception. Cognition. 2006; 99:73–112. [PubMed: 16443448]

Quine, WVO. Word and object. MIT Press; Cambridge, MA: 1960.

Sachs, J. Talking about the there and then: The emergence of displaced reference in parent-child discourse.. In: Nelson, KE., editor. Children's language. Vol. 4. Erlbaum; Hillsdale, NJ: 1983.

Simons DJ, Keil FC. An abstract to concrete shift in the development of biological thought: The insides story. Cognition. 1995; 56:129–163. [PubMed: 7554792]

Smith LB, Jones SS, Landau B. Naming in young children: A dumb attentional mechanism? Cognition. 1996; 60:143–171. [PubMed: 8811743]

Waxman, SR.; Lidz, J. Early word learning.. In: Kuhn, D.; Siegler, R., editors. Handbook of Child Psychology. 6th Edition. Vol. 2. Wiley; Hoboken, NJ: 2006. p. 299-335.
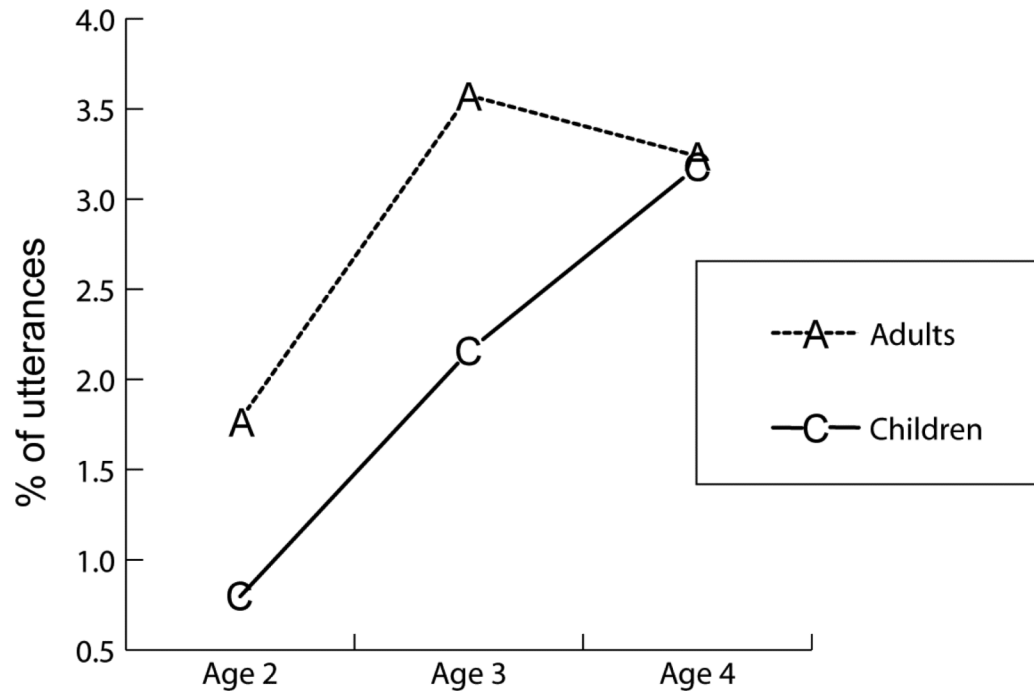
Figure 1.
Mean percentage of utterances that were coded as generic, as a function of speaker and child age.
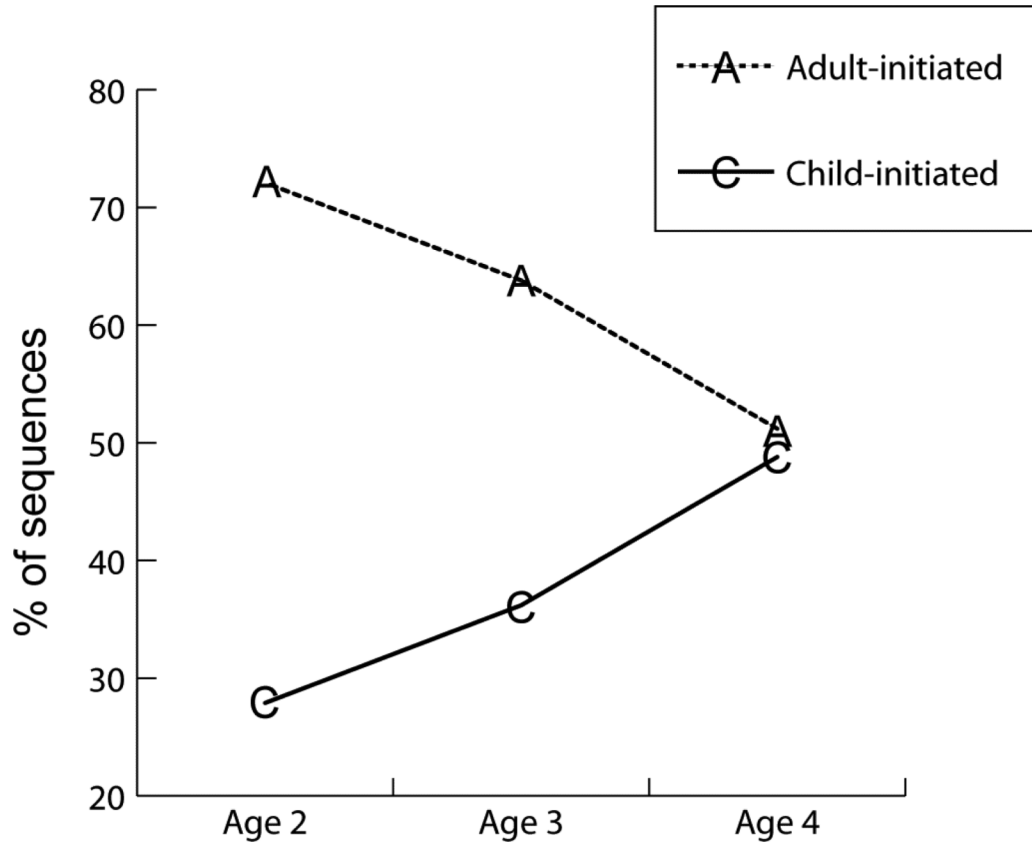
Figure 2.
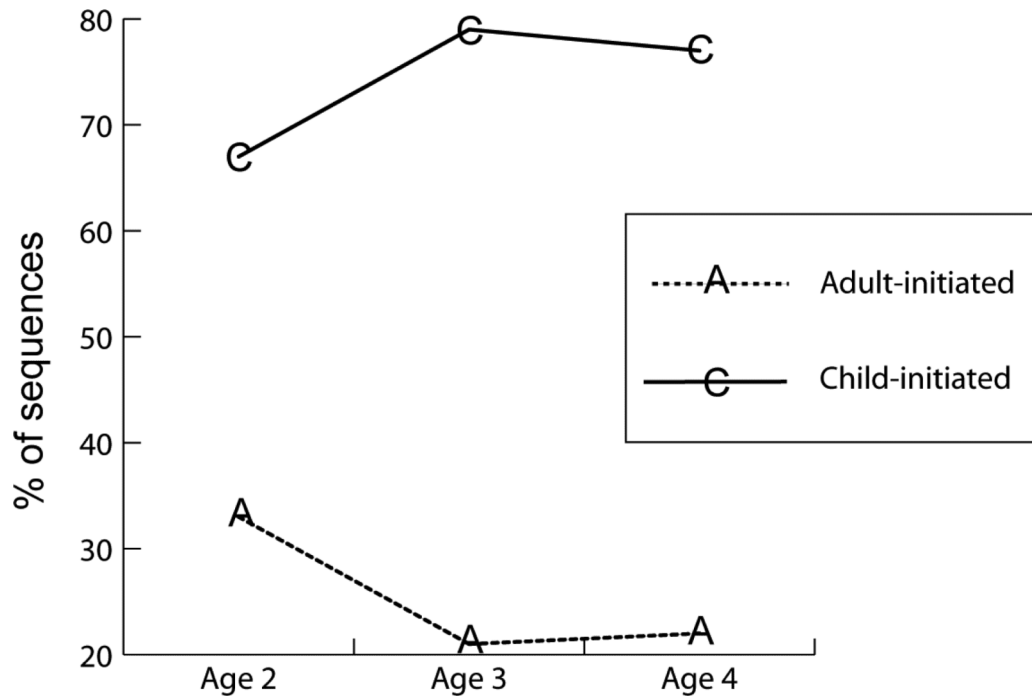Generic initiation rates as a function of speaker and child age.

Figure 3.
Initiation rates of sequences in which a child generic was produced, as a function of speaker and child age.
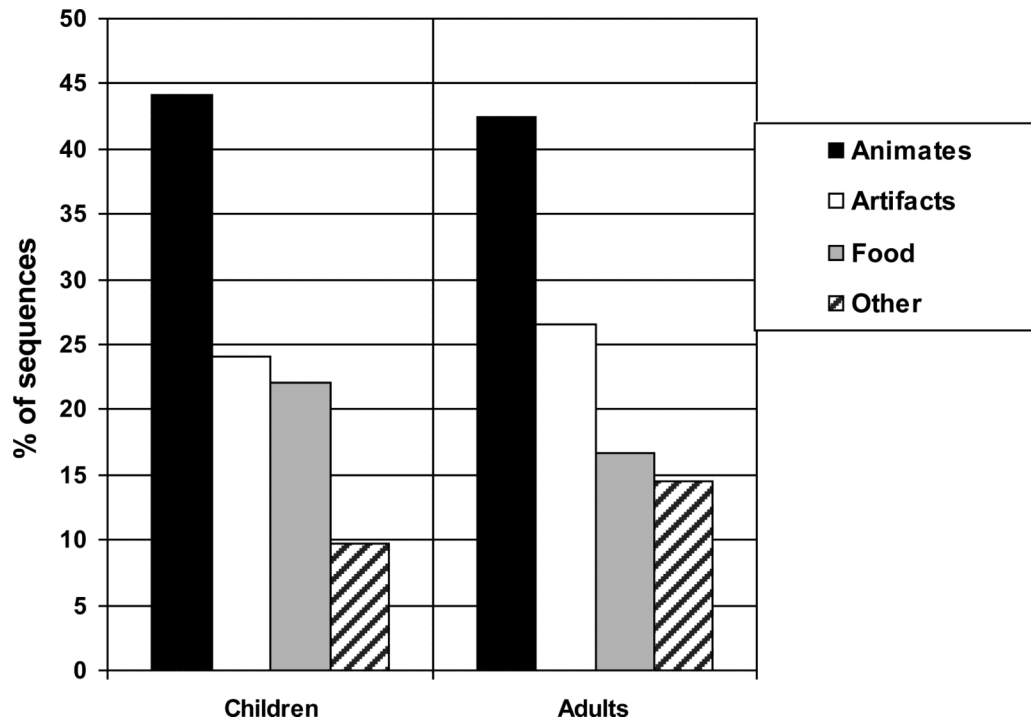
Figure 4.
Percentage of generic utterances that correspond to each of 4 domains (animate, artifact, food, other), as a function of speaker and child age.
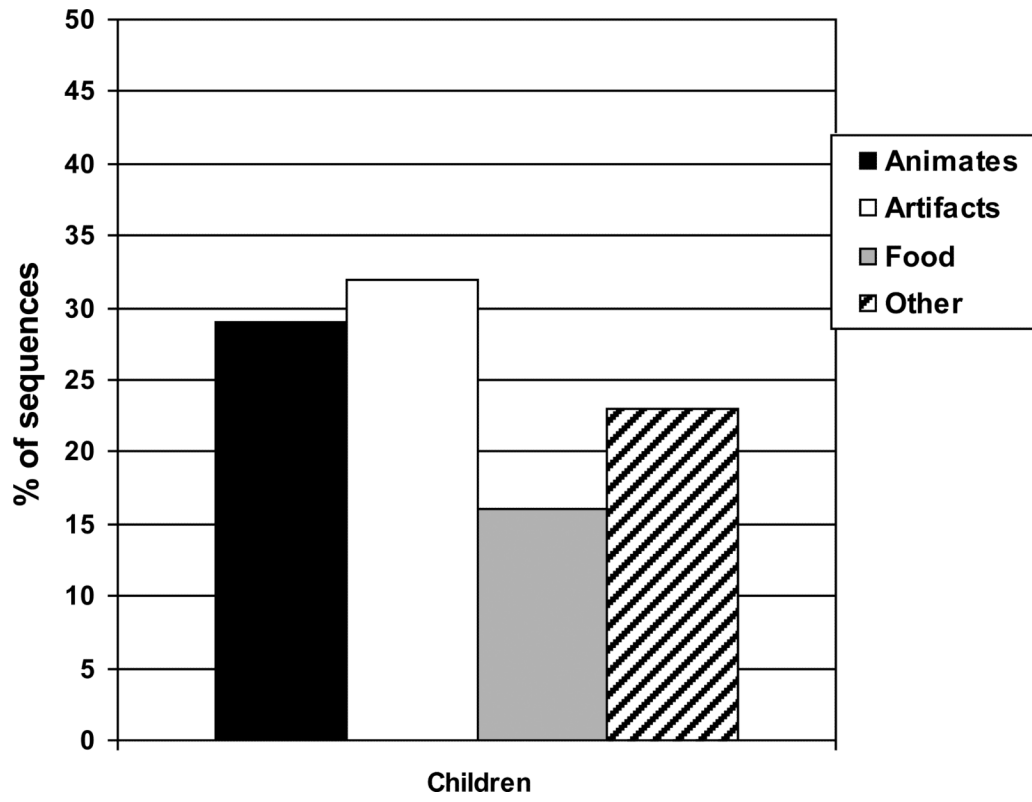
Figure 5.
Children's non-generic utterances as a function of domain.

Table 1

Summary of Data Set

Child	Researcher	Number of transcripts	Number of child utterances	Number of parent utterances	Number of other adult utterances	2 years	3 years	4 years
Abe	Kuczaj (1976)	203	22,182	21,440	319	X	X	X
Adam	R. Brown (1973)	45	46,468	20,129	5,966	X	X	X
Naomi	Sachs (1983)	42	11,408	7,215	0	X	X	X
Ross	MacWhinney (1991)	32	16,931	19,583	0	X	X	X
Nathaniel	Snow	32	13,361	19,251	1,185	X	X	
Peter	L. Bloom (1970)	14	25,366	2,658	22,109	X	X	
Mark	MacWhinney (1991)	22	12,494	22,783	0		X	X
Sarah	R. Brown (1973)	85	24,113	20,702	9,088		X	X

Table 2

Examples of generics produced by 2-year-olds

Abe: (2;5.14) "I want a *cranberry muffin*. I like **them**."

Adam: (2;8.0) "Don't be (a)fraid, (a)fraid a **horses**."

Naomi: (2;8.23) "**Glasses** go in the dishwasher."

Nathaniel: (2;5,18) "**Birds** fly away the sky?"

Peter: (2;1.0) "It goes on **lips**."

Ross: (2;6.14) "I kill **monsters**."