Beyond Fast Mapping

Susan Carey
Harvard University

Abstract

Since the seminal 1957 studies of word learning by Roger Brown, most experimental studies of lexical acquisition have concerned fast mapping: the process through which a new lexical entry is established, and through which representations of the linguistic context of a newly heard word interact with representations of its nonlinguistic context to fix an initial partial meaning. Here I focus on the subsequent extended process through which the adult meaning is approximated. Two factors lead to an extended learning process; the size of the hypothesis space and the need, sometimes, for the creation of new semantic primitives. Sometimes lexical learning requires conceptual change. I sketch a learning mechanism through which this can be achieved. A case study of learning the meanings of verbal numerals illustrates the argument.

In this paper I provide a highly selective short history of the field of lexical development, focusing on the nature of, and significance of, fast mapping. Never did anybody believe that children typically create full lexical representations upon just one or even a few exposures to a new word. A full appreciation of the nature of the extended mapping process that follows fast mapping leads to a reevaluation of common assumptions about the learning process — the most important being that learning consists solely of hypothesis testing over a set of primitives that remain constant throughout development. Often this assumption is true, but sometimes lexical learning requires the creation of new semantic primitives, primitives not expressible in terms of those articulating the hypotheses about a word's meaning at the outset of the learning episode. I illustrate the argument with a case study of lexical learning that requires the creation of new primitives and sketch a learning mechanism that is up to the job. I end with some speculations about how common lexical learning requiring conceptual change is, and comment on how this proposal is consistent with the strongest versions of Whorfian or Quinian linguistic determinism.

A Short History

As a young psychologist interested in lexical representations, I had the good fortune to study with both Roger Brown and George Miller. Miller invited me to ponder the significance of young children's prodigious feats of word learning as they master over 6000 words by age 6. The general problem of induction requires that the child's hypotheses about word meanings be highly constrained, and the sheer efficiency of word learning early in development merely adds urgency to the problem of specifying the constraints and characterizing their provenance. From Brown I learned the role of syntactic context in constraining the meanings of newly heard words.

Unless the hypothesis space is so constrained that children can guess the meaning of any newly heard word upon first hearing it in a naturalistic context, logic requires that young children will be in the midst of working out the meanings of hundreds or thousands of words at a time. For this to be possible, there must be powerful processes that establish and maintain lexical entries of newly heard words, locating their meanings in some relevant part of semantic space, while the nuanced meaning gets worked out. This analysis was confirmed by a serendipitous observation. In one of my first stdies on lexical representations, I asked 3-

year-old children to make it so there is "tiv" green water in a container (Carey, 1978a). The apparatus made it symmetrically easy to add or subtract water, and children had been shown how it worked (by opening alternative valves). This was not a word learning experiment; it was merely a control condition needed to interpret a previous finding that children sometimes treat "less" as if it meant more. In my experiment, the children added water, just as they would if asked to make it so there is less water in the container, but surely not because they had antecedently analyzed "tiv" to mean the same as "more." Some 6 weeks later the same children were tested in replication of an experiment of Eve Clark's (1972) that showed that children never misanalyze the polarity of dimensional adjectives. The task was an opposites elicitation task: I say "up," you say "down;" I say "little," you say "big." At the end of this task I said "tiv." Most children looked at me blankly, but about a quarter said "more" or "less." From a one shot experience with the word "tiv," heard 6 weeks earlier in a context with no ostensive definition, and no use of the words "more" or "less," these children had formed a representation of this novel word and located it in an appropriate portion of semantic space. This finding was the first experimental demonstration of fast mapping, and with Elsa Bartlett, I carried out two further demonstrations of this phenomenon (Carey, 1978b; Carey and Bartlett, 1978).

Bartlett's and my fast mapping studies concerned learning a single new color word, "chromium," used to refer to olive green. In a first pilot study, the sample consisted of fourteen children in an experimental preschool run by George Miller at Rockefeller University. The word was introduced by the child's teacher in a context in which she asked the child for one of two objects that were identical except for color: "Can you get me the chromium tray, not the red one, the chromium one." All children brought the correct tray. A week later the child's comprehension of "chromium" was assessed in a totally different context (experimenter instead of teacher, different room), and 5 weeks later production was assessed, followed by a 2nd introduction by the teacher and then a few weeks later a second assessment (comprehension and production.) In the second more formal study of 20 children, the venue and introducing event were the same, but we had baseline measures of children's color word production and sorting abilities. After the baseline measures, the introduction of the word "chromium" followed. Then about a week later there was an assessment battery consisting of production and sorting again, a comprehension measure, and a hyponym assessment (is "red" a color?, is "happy" a color?, is "tiv" a color?, is "chromium" a color?, etc.) Then 10 weeks later there was another introduction event, followed about a week later by a second round of the same assessment battery.

The results from the two studies were consistent and clear. First, there was fast mapping. Over half of the children showed evidence of having added "chromium" to their lexicon, remembering it some 6 to 10 weeks after just a few introducing events, and demonstrating their knowledge in a very different context from that in which they had learned the word. All of the children whose lexicon had been affected had learned that "chromium" was a color word, most probably aided by the contrasting phrase "not the red tray, the chromium one" and by the fact that the two possible referents differed only in color.

Whereas the phenomenon of fast mapping is what is most remembered from these early papers, my own interests included equally the acquisition process after fast mapping, what we called in those papers "extended mapping." Over the two studies, only 2 of 34 children had fully mapped "chromium" to the color olive green by the end of testing, differentiating that color from the superordinate category (green for some children, the achromatics for others) into which they had placed olive green at baseline. For the remaining children, three patterns of intermediate mappings were observed. Most rarely, a few children took "chromium" to be a synonym of "green," offering focal green as well as olive green objects as "chromium." More took "chromium" to be another word for the achromatics, along with

"brown" and "grey." That is, they assimilated it to the same portion of the color lexicon as the known words they used to label olive green in the first place. Others failed to remember the word per se, but they showed that they had begun to reorganize the space of colors for the purpose of labeling. They said they didn't know or didn't remember what color word applied to olive green (in contrast to baseline measures where they labeled olive green, "green, brown" or "grey"). In spite of fast mapping, the extended mapping process is indeed protracted.

Brown's (1957) seminal study was finally being picked up by the field. Inspired by Brown's study, Katz, Baker, & Macnamara (1974) confirmed that fast mapping often relies on the child's representation of syntactic context to constrain the relevant part of semantic space into which the meaning of a newly learned word fits. Brown had introduced children to a novel word "seb." All children heard the word in the same non-linguistic context (an actor used a novel tool to perform a novel action on a novel non-solid substance.) Children took "seb" to refer to different aspects of that scene depending upon the word's syntactic context. If children heard "this is a seb," they took "seb" to mean the tool; if "this is some seb," they took "seb" to mean the stuff, and if "this is sebbing," the took "seb" to mean the action. Similarly, Katz et al. showed that a novel word heard in a syntactic context that specifies a count noun ("a dax/another dax") interpreted it to refer to any of several members of a kind whereas if used as a proper noun ("Dax"), they took it to refer to an unique individual.

These experiments vividly exemplified another logically necessary truth about fast mapping — it must be constrained by the child's representation of the situation in which the new word is used. In the above experiments, children's representations of the non-linguistic context must be articulated in terms of concepts such as *individual*, *kind*, *substance*, and *action*. Katz et al. also demonstrated that the child's success in appropriately constraining the hypothesis space on the basis of syntactic context depends upon how likely the child thinks the relevant semantic category applies to the entities in the non-linguistic context. That is, they observed that if a proper noun, "Dax," is applied to a box with ribbons, young children do *not* assume "Dax" refers to that unique individual, whereas if it is applied to a doll, they do. For young children, animate beings and their surrogates can be labeled with proper nouns; inanimate objects typically are not.

These experiments were demonstration studies, laying out very generally what the process of lexical acquisition must entail. The next 40 years of research, still ongoing, put meat on the bones. There have been countless studies in which children are taught new words with one or two exposures, and tested immediately thereafter for their interpretation of it. All such studies concern fast mapping, exploring what in the linguistic context constrains the interpretation of the novel word and how this interacts with the child's representation of the non-linguistic context to establish that initial partial interpretation. Taking just a few salient examples: Markman (1989) characterized the taxonomic, whole object, and mutual exclusivity constraints, as they inform fast mapping of nouns, and Gleitman's account of syntactic bootstrapping fleshed out many details of how verb learning proceeds (e.g., Gleitman et al., 2004). Waxman's recent studies suggest that by 12- to 14-months of age infants have mapped the conceptual distinction between kind and property to the syntactic distinction between noun and adjective (e.g., Waxman & Booth, 2003), and Xu has shown that by 9-months of age, infants take novel words to refer to kinds rather than individuals (Dewar & Xu, 2007).

Very little research has followed up on the stability over time of the structure that is the output of fast mapping, testing children weeks after a single introducing event. Markson and Bloom (1997) extended to noun representations the finding that a newly formed lexical entry, along with a representation that at least partially specifies its referent, persists for 6

weeks, and showed this to be equally true for 3-year-olds and adults. These researchers also showed that fast mapping is not limited to establishing lexical entries stable enough to support subsequent word learning —in some circumstances 3-year-olds and adults fast map novel facts about entities as well.

In sum, mappings between syntactic categories and semantic ones are in place early in development, constraining word learning. The studies described above leave open how these mappings are constructed. Some may be part of the innate language acquisition device that separates humans from other animals. Others may be abstracted from what has been learned about language through domain general learning processes. These questions are certainly not settled, and the subject of active research programs (see Smith, this volume, for example). Many other questions about fast mapping are also still open. In the 70's, the role of theory of mind in lexical learning was not even on the horizon, and the fundamental theoretical question of how, if at all, to distinguish between referential and associative mappings not yet engaged (e.g., Baldwin, 1993; Bloom, 2000; Tomasello & Akhtar, 1995).

My focus here is not on fast mapping, but on the other part of the process — the extended mapping that takes place over months or years. To motivate why this is important, I ask you to go back to the 70's again, and remember the state of the art in research on lexical development at the time. Lexical representations were taken to be formulated in terms of semantic features over which necessary and sufficient conditions for category membership, or meaning postulates that provide merely necessary conditions, or prototype structures that determine category membership probabilistically, can be expressed. On this cluster of views, the process of extended mapping consists of hypothesis testing over the space of semantic features. Hypothesis testing over structured linguistic representations, as instantiated in syntactic bootstrapping theories (e.g., Gleitman et al., 2004), or the accounts of verb learning championed by such early pioneers as Bowerman (1974) and Pinker (1984), also presupposed that the primitives needed to express the meaning of a target word were antecedently available. These views allowed, even demanded, a protracted learning process, because the hypothesis space is vast and any given piece of evidence is compatible with many different hypotheses. This view of lexical development is intuitively appealing. How else is meaning acquisition to proceed other than by hypothesis testing over some feature set — a set of developmental primitives that constitute an innate hypothesis space in terms of which word meanings can be formulated? I accept such a view of learning; what I deny is that there are no mechanisms that create representational resources qualitatively different from the representations that articulate the hypothesis space available at the outset of the learning episode.

This intuitive picture is presupposed in much current work on lexical development as well (see Smith, this volume, for example, for modern computational work on the learning process, given some input space). In the 70's, the hope, and the assumption, was that the lexical primitives that emerged in analyses of the structure of the adult lexicon would also be the developmental primitives over which hypothesis testing proceeded in development. For example, Miller and Johnson-Laird's (1976) monumental tome *Language and Perception* provided an analysis of the lexical structure of many domains of words in terms of primitives that in some cases also articulate perception. Miller opened a lexical acquisition lab in the early 70's to study the process of lexical learning formulated over the space of lexical primitives uncovered in his and Johnson-Laird's analysis of structure of the lexicon.

Indeed, one of the learning processes currently offered to address lexical learning: syntactic bootstrapping, presupposes that lexical acquisition proceeds by hypothesis testing over already existing syntactic and semantic features. I fully accept the existence and significance of syntactic bootstrapping, and its complement process, semantic bootstrapping, that plays a

role in working out the syntactic role of newly learned words. For semantic bootstrapping to occur, the child must already represent the syntactic category in question, and have some antecedent mapping to some semantic feature, so that the child can use the fact that a word in question applies to an entity with the relevant semantic feature to place that word in its correct syntactic category. For syntactic bootstrapping to occur, the child again must already represent the syntactic category in question, and have some antecedent mapping between that category and semantic features to use evidence of syntactic role to constrain the meaning of a newly heard word. These processes are not designed to create new semantic features.

Miller closed the lab just a few years after he opened it, and wrote of the spectacular failure of the assumptions that had underpinned the research effort in a charming book, *Spontaneous Apprentices* (1997). Simply put: the assumption that the definitional primitives that structure the adult lexicon and the developmental primitives over which hypothesis testing is carried out early in development are one and the same set of features is deeply mistaken. A moment's reflection shows this to be so. For example, the definition of *gold* within modern chemistry might be *element with atomic number 79*. Clearly, the primitives *element* and *atomic number* are not innate semantic features. Or take the features that determine the prototype structure of bird concepts (flies, lays eggs, has wings, nests in trees, has a beak, sings, and so on...). Subjects provide distinctive values for such features when asked to list the features of birds, and overlap in terms of these same features predicts prototypicality within the category *bird*. That is, this feature space definitely underlies adult prototypicality structure. Yet these features are not innate primitives--many are no less abstract nor less theory-laden than the concept *bird* itself).

More generally, studies on lexical development have failed to take seriously the issue of what the developmental primitives are. At least since the time of the British empiricists, many have assumed that all concept learning begins with a primitive sensory or perceptual vocabulary, but that project is doomed by the simple fact that it is impossible to express the meanings of most lexical items (e.g., "cause," "good," "seven," "gold," "dog...") in terms of perceptual features. Any detailed accounts of the process of lexical acquisition, including satisfying computational ones, require correctly characterizing the actual developmental primitives in terms of which lexical hypotheses may be formulated. As many have argued, (see Carey, 2009, for a review) developmental primitives include abstract concepts such as *object*, *agent*, *goal*, and *cause*; our conceptual repertoire need not be built from sensori-motor primitives alone. However, demonstrating this is a far cry from showing that all of the primitives that underlie the meanings of words are part of our innate representational repertoire.

What I suggest here is that extended mapping sometimes involves not only hypothesis testing over a vast space of semantic and syntactic primitives, but also constructing new features that articulate the lexicon. In what follows, I illustrate this point in a case study of the acquisition of the meaning of words that express positive integers (e.g., "two," "seven," "forty") and sketch a learning mechanism that can create new semantic primitives. Such a learning mechanism must be more powerful than, and different from, both semantic and syntactic bootstrapping processes.

Acquiring words for integers: a case study

Around age 2, English-learning children first acquire words that express integers in the adult lexicon (e.g., "one," "seven"). Young toddlers acquire a stably ordered count list, and when asked "how many" in a set, they point to each object in a set once as they recite the list. Nonetheless, they have verbal numeral words in their lexicon for almost a year and a half

before they figure out how the count list represents number — that the last number reached in a count represents the cardinal value of the set, and that for any numeral in the list, if that numeral represents *n,* then the next numeral represents *n + 1* (Fuson, 1988; LeCorre, Brannon, Van de Walle & Carey, 2006; Wynn, 1990, 1992a). Indeed, initially, they have assigned no numerical meaning to any of the numerals in their lexicon. They cannot even hand a person "one penny" (from several) if asked, or indicate which of two sets (a singleton and a set of several) has "one."

Very early in the process of learning the meanings of verbal numerals, young 2-year-olds children work out what "one" means, succeeding at both of the above tasks. But they do not yet assign a cardinal meaning to any other numeral in their lexicon. The striking phenomenon (and you can definitely try this at home if you have a handy two-year-old) is that if you ask for "two pennies" or "four pennies," they merely grab a handful, with no relation to the number requested (they do not give more for when asked for four than for two; Wynn, 1990, 1992). Notice that this shows that they know that the other words in the count sequence contrast with "one." They always grab a random number of objects greater than one when asked to hand over "two, three, four ," or "seven" (or any other numeral in their lexicon), and they also successfully point to a card with three fish when asked for "three" when that card is contrasted with a card with one, even though their choices are random when probed for "three" when a card with three is contrasted with a card with two. Such children are called "one"-knowers, for they know the meaning only of the verbal numeral "one." They have no idea which particular cardinality any of their other number words refers to. Importantly for us here, children remain "one"-knowers for 6 to 8 months before they figure out what "two" means. This is a paradigm example of extended mapping — "two" is in the child's lexicon as a word that contrasts in meaning with "one" for several months before a full interpretation is made.

Upon becoming "two"-knowers, children can construct sets of two when asked for "two," and tell which of two cards has "two fish" when given a contrast between a set of 2 fish and any other set. But they still respond randomly on "three." They are "two"-knowers for several months, and then become "three"-knowers. LeCorre and Carey (2007) showed that all children who have worked out how counting represents number have previously assigned "four" a numerical meaning. Thus, children laboriously work out numerical meanings for the numerals "one" through "four" over a 1 and 1/2 year period before they figure out how to use counting to implement the successor function. Call the children who have not yet figured out the cardinality principle (the last numeral reached in a count represents the cardinal value of the set) "subset"-knowers, for they have assigned cardinal meanings to only a subset of the numerals in the count list.

In sum, learning verbal numerals, like any word learning, exemplifies both fast and extended mapping. Very young children easily add number words to their lexicons, and initially assign them partial meanings. But they produce "four" and "seven" for over a year before they work out the adult meanings of these words.

Fast mapping —the initial meanings

Many sources of evidence suggest that just as with other words, syntactic context is a source of constraint that allows children to place verbal numerals in the right region of conceptual space. Verbal numerals are quantifiers, and Bloom and Wynn (1997) showed that from the earliest stages of numeral learning, children use them as in the same syntactic contexts as they use other quantifiers. Details of the learning process sketched above show that young children's hypotheses about number word meanings are guided by the fact that they are quantifiers. They know that "one" picks out a single individual item, that that other numerals contrast with one, designating sets with more than one. Clark and Nikatina (2009) showed

that "two" is often initially analyzed as a plural marker —e.g., children call sets of 2 balls, 8 balls, or 6 balls "two ball" or "two balls;" see also Le Corre and Carey (2007). Furthermore, Sarnecka et al. (2007) showed that children learning Japanese, a classifier language that lacks a singular/plural distinction, are delayed, relative to children learning English and Russian, both count/mass languages, at becoming "one"-knowers. One interpretation of this fact is that classifier languages, relative to count-mass languages, obscure the quantifier status of verbal numerals.

Extended mapping

Suppose we accept that the syntactic role of verbal numerals as quantifiers helps children locate the relevant part of conceptual space that underlies the meanings of these words. This leaves open two very different ways we can think of the extended mapping process. First, even having located the relevant part of conceptual space, a very large number of hypotheses as to what "two" might mean may be formulated in terms of the conceptual primitives in that space, and sorting through these takes time (6 to 8 months, on average). Alternatively, the extended mapping process reflects the construction of concepts that articulate the meanings of verbal numerals that were not available at the outset of the learning process. On this second construal, this case of extended mapping involves a representational discontinuity. Language acquisition, in this case, may require creating a representational system qualitatively different from, and perhaps more powerful than, the initial repetoire over which hypotheses are formulated..

In the adult system of representation, verbal numerals represent positive integers. They are summary symbols that represent cardinal values of sets, and the system of verbal numerals satisfies Peano's axioms. The central axiom is that each integer has a unique successor, for each integer n, the next integer is $n + 1$. Implicitly or explicitly, the adult meanings of verbal numerals must be formulated over concepts of *exactly 1*, *plus, exact numerosity, set, successor.* Within this set of primitives, "four" can be defined in many different ways. Here are two: 1) four is the cardinal value of any set that can be put into 1-1 correspondence with {x, y, z, w}, where "x," "y," "z," and "w" refer to numerically distinct individuals and { } denotes a set. 2: four is the successor of three. These two ways of defining "four" are formally equivalent, in the sense that each definition determines exactly the same concept, namely, *four*. For "four" to have the same meaning as it does in the adult representational system, its meaning must be equivalent to both of these. Of course, it is an open question just what set of primitives underlie the child's first successful representation of *four*. Whatever it is, it must provide the expressive power of the primitives listed above.

Deciding whether mapping "four" to the adult meaning requires the construction of new conceptual primitives requires correctly specifying the representations that articulate the child's hypotheses about possible meanings of number words. There are two distinct conceptual systems with numerical content that are available to prelinguistic infants. At issue, then, is whether either or both of these can support the meanings of verbal numerals.

Analog magnitude representations of number

Human adults, human infants, and non-human animals deploy a system of analog magnitude representations of number. Number is represented by a physical magnitude that is roughly proportional to the number of individuals in the set being enumerated. Figure 1 depicts an external analog magnitude representational system in which length represent number. A psychophysical signature of analog magnitude representations is that discriminability of any two magnitudes is a function of their ratio. That is, discriminability is in accordance with Weber's law.

Dehaene (1997) and Gallistel (1990) review the evidence for the long evolutionary history of analog magnitude number representations. Animals as disparate as pigeons, rats, and non-human primates represent number, and number discriminability satisfies Weber's law. In the past five years four different laboratories have provided unequivocal evidence that preverbal infants form analog magnitude representations of number as well (Brannon, 2002; Brannon, Abbot & Lutz, 2004; Lipon & Spelke, 2003, 2004; McCrink & Wynn, 2004a; Wood & Spelke, 2005a; Xu & Spelke, 2000; Xu, Spelke & Goddard, 2005). The first paper in this flurry of studies is by Fei Xu and Elizabeth Spelke, who solved the problem of how to control for other possible bases of judgment (cumulative surface area, element size, density) in a large number habituation paradigm. Xu and Spelke habituated 6-month-old infants to displays containing 8 dots or to displays containing 16 dots (Figure 2). Possible confounds between number and other variables were controlled either by equating the two series of stimuli on those variables, or by making the test displays equidistant from the habituation displays on them. Habituated to 8-dot displays, 7-month-old infants recovered interest when shown the novel 16-dot displays, while generalizing habituation to the novel 8-dot displays. Those habituated to 16-dot displays showed the reverse pattern. Subsequent studies duplicated this design (and the positive result) with 16-dot vs. 32 dot comparisons and with 4 dot vs. 8 dot comparisons.

That analog magnitude representations support these discriminations is shown by the fact that success is a function of the ratio of the set sizes. In all of the above studies, in which infants succeeded with a 2:1 ratio, they failed in comparisons that involved a 3:2 ratio (i.e., they failed to discriminate 8-dot from 12-dot arrays, 16-dot from 24-dot arrays, and 4-dot from 6-dot arrays. Also, these researchers have found that sensitivity improves by 9 months of age. Infants of this age succeed at 3:2 comparisons across a wide variety of absolute set sizes, but fail at 4:3 comparisons. Subsequent studies showed analog magnitude representations of number of different kinds of individuals (jumps, sounds), with the same profiles of sensitivity (Lipton & Spelke, 2004; Wood & Spelke, 2005).

In all of the above studies we can be confident it is number infants are responding to, because every other variable is equated either across the habituation stimuli or across the test stimuli. Of course, if the analog magnitude representations underlying performance in these habituation studies are truly numerical representations, number relevant computations other than establishing numerical equivalence should be defined over them, and indeed this is so. Elizabeth Brannon (2002) showed that 11-month-old infants represent numerical order using analog magnitude representations of sets. Koleen McCrink and Karen Wynn showed that 9-month-olds can manipulate sets of objects in the analog magnitude range to support addition, subtraction, and ratio computation (McCrink & Wynn, 2004a, 2004b). In sum, analog magnitude representations of number are available at least by 6-months of age. Preverbal infants represent the approximate cardinal value of sets, and compute numerical equivalence, numerical order, addition, subtraction, and ratios over these representations.

Parallel Individuation of Small Sets

Science moves rapidly, and the infant studies reviewed above came relatively late in the history of studies designed to show that infants are sensitive to number. The first studies, some 20 years earlier than Xu's and Spelke's studies on analog magnitude representations, concerned *small* sets —discriminations among sets of 1, 2 and 3 objects. These include many 2 vs. 3 habituation studies and Wynn's 1 + 1 = 2 or 1 violation of expectancy studies (Antell & Keating, 1983; Starkey & Cooper, 1980, Wynn, 1992b). Although some have suggested that analog magnitude number representations underlie success in these experiments (e.g., Dehaene, 1997), the evidence conclusively implicates a very different representational system (Feigenson & Carey, 2003; Feigenson, Carey, & Hauser, 2002;

Scholl & Leslie, 1999; Simon, 1997; Uller, Carey, Huntley-Fenner, & Klatt, 1999). In this alternative representational system, number is only implicitly encoded; there are no symbols for number at all, not even analog magnitude ones. Instead, the representations include a symbol for each individual in an attended set. Thus, a set containing one apple might be represented: "O" (an iconic object file) or "apple" (a symbol for an individual of the kind apple) and a set containing two apples might be represented "O O" or "apple apple," and so forth. These representations consist of one symbol (file) for each individual, and when the content of a symbol is a spatiotemporally determined object, it is an object file (Kahnemann, Triesman, & Gibbs, 1992). Infants also create working memory models of small sets of other types of individuals, such as sound bursts or events, and so I shall call the system of representation "parallel individuation" and the explicit symbols within it "individual files."

There are many reasons to favor individual file representations over analog magnitude representations as underlying performance in most of the infant small number studies (see Carey, 2009, for a more thorough review). First, and most important, success on many spontaneous number representation tasks involving small sets do not show the Weber-fraction signature of analog magnitude representations; rather they show the set-size signature of individual file representations. Individuals in small sets (sets of 1, 2 or 3) can be represented, and sets outside of that limit cannot, even when the sets to be contrasted have the same Weber-fraction as those small sets where the infant succeeds. at that age. That is, there are tasks on which infants succeed at discriminating 1 vs. 2 and 2 vs. 3 in the face of failure at 2 vs. 4, 3 vs. 6, and even 1 vs. 4 — failure at the higher numbers when the Weber fraction is the same or even more favorable than that within the range of small numbers at which success has been obtained. This is the set size signature of individual file representations.

Here I describe just one paradigm that elicits the set size signature of parallel individuation (see Carey, 2009, for others). An infant watches as each of two opaque containers, previously shown to be empty, is baited with a different number of graham crackers. For example, the experimenter might put two graham crackers in one container and three in the other. After placement, the parent allows the infant to crawl toward the containers. The dependent measure is which container the baby chooses. Ten- to 12-month-olds infants succeed at 1 vs. 2, 2 vs. 3; and 1 vs. 3, and fail at 3 vs. 4, 2 vs. 4, and even 1 vs. 4 (Feigenson & Carey, 2005; Feigenson et al., 2002). One:four is a more favorable ratio than 2:3, but infants fail at 1 vs. 4 comparisons and succeed at 2 vs. 3. Note also that 5 crackers are involved in each choice, so the total length of time of placements is equated over these two comparisons. This is a striking result. Infants could succeed at 1 vs. 4 comparisons on many different bases: putting 4 crackers into a bucket takes much longer, draws more attention to that bucket, and so on, yet infants are at chance.

Although infants could solve this problem in many different ways, apparently they are attending to each cracker, creating a model of what's in the container that contains one object-file for each cracker. As soon as one of the sets exceeds the limits on parallel individuation (apparently three at this age; see also the manual choice paradigm in Feigenson and Carey, 2003, 2005), performance falls apart. This finding provides very strong evidence that parallel individuation underlies success on this task. Although some describe what I am calling the system of parallel individuation as a "small number system," that is a misleading name. The purpose of parallel individuation is to create working memory models of small sets of individuals, in order to represent spatial, causal, and intentional relations among them. Unlike analog magnitude number representations, the parallel individuation system is not a dedicated number representation system. Far from it. The symbols in the parallel individuation system explicitly represent individuals. Figure 3 depicts several different individual file representations of two boxes. In none of these

alternative models is there a symbol that that has the content "two;" rather the symbols in working memory represent the boxes. The whole model {box box} represents two boxes, of course, but only implicitly. Furthermore, the quantitative calculations over parallel individuation models in working memory often privilege continuous variables (such as total event energy, total contour length, total surface area) over numerical equivalence (Clearfied & Mix, 1999Clearfied & Mix, 2001; Feigenson, Carey & Spelke, 2002).

If parallel individuation models do not include symbols for number, why am I discussing these models in the present context? The answer is that they are shot through with numerical content, even though that numerical content is merely implicit in the computations that pick out and index small sets to represent, that govern the opening of new individual files, that update working memory models of sets as individuals are added or subtracted, and that compare sets on numerical criteria. The creation of a new individual file requires principles of individuation and numerical identity; models must keep track of whether this object or jump, seen now, is the same *one* as that object seen before, or this sound just heard, is the same *one* as that just heard previously. The decision the system makes dictates whether an additional individual file is established, and this guarantees that a model of a set of three boxes will contain three box symbols. Computations of numerical identity are (as their name says) numerical computations. Also, the opening of a new individual file in the presence of other active files provides an implicit representation of the process of adding one to an array of individuals. Finally, working memory models of two sets of individuals can be simultaneously maintained, and when individual-file models are compared on the basis of 1-1 correspondence, the computations over these symbols establish numerical equivalence and numerical order (see Feigenson, 2005; Feigenson and Carey, 2003, for evidence of such computations).

Conceptual Discontinuity

In cases of conceptual discontinuity, no antecedent system of representation can express the concepts of a later developing one. In the adult lexicon, "seven" expresses an integer. That seven is one more than six, which is one more than five, which in turn is one more than four, and so on, is guaranteed by the counting principles mastered by age 3 or so. Once children have mastered the verbal numeral list, and use it according to the counting principles described by Gelman and Gallistel, 1972, they have mastered representations of a finite subset of the integers. As we saw above, the process of mastering the meaning of "two," let alone "seven," is a paradigm case of extended mapping. At issue for us here is whether the definitional primitives needed to express these meanings, primitives such as *one, successor, cardinal value of set,* are available at the outset of the extended mapping process.

Analog magnitude representations of number lack the expressive power of any system of representation of the natural numbers, including numeral list representations, in three crucial respects. First, they have an upper limit; they represent the cardinal values of sets that can actually be apprehended, and so the lack the capacity to represent discrete infinity. Second, because analog magnitude representations are inexact and discriminable only to a given Weber ratio, they fail to capture small numerical differences between large sets of objects. The distinction between 18 and 19, 11 and 12, 124 and 127, and so on, cannot be captured reliably by the analog magnitude representations of human adults, because as a matter of contingent fact, the human Weber ratio for sensitivity to numerical differences is 8:9 (at best). Relatedly, they are not built around, nor can they express, the successor function, the operation of adding one to a given integer in order to generate the next integer. Rather, they positively obscure the successor function. They contain no representation for exactly one. And since numerical values are compared by computing a ratio, the difference between one and two is experienced as different from that between two and three, which is again

experienced as different from that between three and four. And, of course, the difference between nine and ten, since nine and ten, like any higher successive numerical values, cannot be discriminated.

Thus, analog magnitude representations are not powerful enough to represent the natural numbers and their key property of discrete infinity. They do not provide exact representations of numbers, not even *one*, and they obscure the successor function, which is constitutive of natural number. That is, they do not contain the primitives that underlie the meanings of verbal numeral that express integers, like "seven" or "four."

The parallel individuation system does not remotely have the capacity to represent natural number, nor does it have symbols that are the primitives from which integers are defined. It is not dedicated to number representations. Number is only implicitly represented, in that computations of 1–1 correspondence are made over symbols for individuals represented in parallel in models of arrays of objects and events. The system of parallel individuation, like that of analog magnitude number representations, contains machinery for indexing and tracking sets of individuals, but it contains no symbols for cardinal values of sets, no symbols for *one* or any other cardinality. The only symbols in such models represent the individuals themselves. Also, the system of parallel individuation has an upper bound at very low set sizes indeed — three for infants. With this system of representation, infants cannot even represent 4 (even implicitly), let alone 7 or 32 or 1,345,698.

In sum, neither antecedent system with numerical content has the capacity to represent natural number, whereas the numeral list representation constrained by the counting principles does. Neither antecedent system contains the primitives from with the meaning underlying "seven" can be constructed.

Empirical Evidence for Conceptual Discontinuity

The ways in which integer representations transcend core cognition makes sense of why it takes children a year and a half or two years to figure out how counting represents number. That learning the meaning of "seven" requires the creation of new semantic primitives predicts an extremely extended mapping process. If this is correct, behavioral measures should reflect a qualitative change in representational resources during this process. We should see within-child consistency on a whole variety of tasks that reflect integer meanings. And indeed we do. There is evidence from a wide variety of measures for a qualitative shift between subset-knowers, on the one hand, and cardinal principle knowers, on the other. There is also evidence for consistency within knower-level. A "one"-knower reveals knowledge only of the numeral "one" on every task that probes for such knowledge. Ditto for "two"-, "three"- and "four"-knowers.

LeCorre et al. (2006) provide a thorough documentation of these claims; here I give just a few examples of the striking within-child consistency on several measures that suggests a qualitative shift in understanding how counting represents number upon becoming a cardinal principle knower. In Karen Wynn's (1990, 1992a) original studies, subset-knowers almost never counted to produce sets, either within their knower-level range or outside it (asked to give five apples, they merely grabbed a handful), whereas cardinal principle knowers almost always counted out large sets. Also, when simply asked to count a set of objects, children in both groups could do so with few errors, but then after counting, if asked "How many was that?" the cardinal principle knowers almost always merely repeated the last word of their previous count, whereas subset-knowers rarely did so. Rather, subset-knowers recount, or provide a numeral that does not match the last word of their count. This suggests that subset-knowers do not realize that the last word reached in a count represents the cardinal value of the set. Later studies confirmed these findings and extended them. Even cardinal principle

knowers sometimes make mistakes when creating sets of a requested number. Children are asked to count and check their answers. When the count reveals an incorrect set-size, cardinal principle knowers virtually always correct appropriately. Subset-knowers, in contrast, leave the set unchanged or correct in the wrong direction (e.g., add more objects when the count revealed that there were already too many) on more than 70% of the trials (Wynn, 1990, 1992a; LeCorre, et al., 2006).

Within knower-level, being a "one"-knower according to Wynn's "give-a-number" task predicts which pairs of sets children will succeed on when asked "which is n" (i.e., any contrast between a set of one and a set with any other number; no other contrast). Ditto for "two"-, "three-" and "four-" knowers. Similarly, knower level on Wynn's task predicts which set sizes children can successfully estimate (without counting) when simply shown a set of entities and asked how many it contains (Le Corre et al, 2006; LeCorre and Carey, 2007). A "two"-knower says "one" for sets of 1 and "two" for sets of two and uses higher numerals randomly for sets from 3 to 10.

In sum, two kinds of analyses support the claim that the representational system that is the verbal count list is discontinuous with antecedent representations. Most importantly, I have offered empirically supported characterizations of two such antecedent systems of representations with numerical content, and shown how they lack the expressive power of the count list. Secondly, I have offered evidence that the count list representation of number is, as predicted, very difficult to learn, and is systematically misinterpreted during the learning process.

The Explanatory Challenge: Quinian Bootstrapping

I turn now to the explanatory challenge: what learning processes can create representational resources with more expressive power than, or qualitatively different from, than their antecedents? What learning processes can create new representational primitives?

Many historians and philosophers of science appeal to bootstrapping processes in their accounts of specific conceptual innovations (see Carey, 2009 for discussion). Semantic and syntactic bootstrapping are not the type of bootstrapping processes described in this literature, for as detailed above, these processes account for how the child solves a difficult mapping problem, not how the child creates new representational resources. The very word "bootstrapping" is a metaphor, meant to capture the deep difficulty of the problem. After all, it is impossible to pull oneself up by one's bootstraps. Neurath's metaphor of building a boat while already in the middle of the ocean also captures the difficulty of the problem —that while not grounded one must build a structure that will float and support you. Not grounded in this case means that the planks one is building the boat with are not interpreted concepts one already represents. In other metaphors, the learner's concepts are partially grounded, as in Quine's ladder metaphor. Here, one builds a ladder grounded in one conceptual system until one has a platform that is self-sustaining, and then one kicks the ladder out from under. And in a final Quinian metaphor, one is scrambling up a chimney supporting oneself by pressing against the sides one is building as one goes along. Quine again captures that the new conceptual system that supports you is being built as you go along. This metaphor stresses, as does Neurath's boat, that the structure one builds consists of relations among the concepts one will eventually attain — it is that structure of interrelations among the to-be-attained concepts (the sides of the evolving chimney, the boat itself, the platform from which the ladder can be kicked away) that serves the crucial bootstrapping role. See Quine (1960, 1969, 1977) for his own elaboration of Quinian bootstrapping.

Although such metaphors are evocative — of both the problem to be solved and the solution —Quine never describes in detail how this learning process operates. These metaphors are

hardly satisfying to a cognitive scientist trying to understand bootstrapping mechanisms. As I illustrate here, I believe it is possible to flesh out the metaphors with appeals to processes that are well understood at a computational level.

Quinian bootstrapping processes require explicit symbols, such as those in written and spoken language or mathematical notational systems. The aspect of the bootstrapping metaphor that consists of building a structure while not grounded is instantiated as the learner initially learns the relations of a system of symbols to one another, directly, rather than by mapping each symbol onto preexisting concepts (Block, 1986). The symbols so represented thus serve as placeholders, at most only partially interpreted with respect to antecedent concepts. This is one essential component of Quinian bootstrapping. It also provides a key to the creation of new primitives: they are simply coined, not defined in terms of existing primitives, their meaning exhausted by their conceptual role in the placeholder structure. The second essential component is the process through which the placeholders become interpreted. As historian and philosopher of science Nancy Nersessian (1992) argues, these are modeling processes. Often, but not always, processes of analogical mapping are involved. Other modeling processes, such as abduction, thought experimentation, limiting case analyses, and induction, all have roles in Quinian bootstrapping. Two properties of these modeling processes are important. First, they are not deductive. There are no guarantees in bootstrapping. The structures that are tentatively posited either work, in the sense of continuing to capture the observed data that constrain them, or they do not. Second, they are all problem-solving mechanisms that play a role in thought more generally. They are bootstrapping mechanisms only when harnessed in the service of a process that happens to create a representational resource qualitatively different from its antecedents.

Bootstrapping the Numeral List Representation of Natural Number

The output of the learning process we seek to understand is the numeral list representation of natural number —an ordered list of numerals such that the first one on the list represents 1 and for any word on the list that represents the cardinal value n, the next word on the list represents n + 1. The successor function is the heart of numeral list representations of integers. That the numeral list representation of number implements the successor function is guaranteed Gelman and Gallistel's counting principles (so long as the list is stably ordered, individuals in a given count are put in 1-1 correspondence with number words, and the cardinal value of the set is determined by the ordinal position of the word in the count list, then for any numeral in the list with a given cardinal value, then the next numeral has that cardinal value plus 1).

The problem of how the child builds an numeral list representation decomposes into the related subproblems of learning the ordered list itself ("one, two, three, four, five, six . . ."), learning the meaning of each symbol on the list (e.g., "three" means three and "seven" means seven), and learning how the list itself represents number, such that the child can infer the meaning of a newly mastered numeral symbol (e.g., "eleven") from its position in the numeral list.

As mentioned above, the child first learns "one, two, three, four, five…" as a list of meaningless lexical items. There is no doubt that children have the capacity to learn meaningless ordered lists of words — they learn sequences such as "eeny, meeny, miny, mo," the alphabet, the days of the week, and so on. Indeed, nonhuman primates have this capacity, and so it is likely part of innate computational machinery (e.g., Terrace, Son, & Brannon, 2003). This step in the learning process — learning an arbitrary ordered list ("one, two, three, four, five, six…") is a paradigmatic example of one aspect of Quinian

bootstrapping: the meanings of the counting words are exhausted, initially, by their interrelations, their relative order in the list and their place in the numerically meaningless count routine. At this point in the process, the verbal numerals are placeholders with respect to the numerical meaning they will come to have.

Children do not learn the numerical meaning of "one" in the context of counting. "One" is much more frequent in speech to the child's input as a quantifier than embedded in the count routine (Sarnecka, et al., 2007). "Can you give me one?" "Would you like that one?" "I'd like one cupcake." As argued above, fast mapping of numerals is supported by their syntactic role as quantifiers (Bloom and Wynn, 1997). The syntactic role of numerals as quantifiers constrains their meaning, thus helping them learn the meaning of "one." Three, sources of evidence support Bloom and Wynn's conjecture. First of all, the partial meanings children in the subset-knower stage assign to numerals they have not yet assigned a cardinal meaning to implicate hypothesis testing over the space of quantifier meanings. For example, for "one"-knowers, "two, three, four" and so on each mean essentially "plural." Second, learners of Japanese and Mandarin, classifier languages with no singular-plural distinction, learn to count as early as do learners of English, and have equivalent number word input, but do not become "one"-knowers until several months later than do English-speakers (Sarnecka et al., 2007; LeCorre, Li & Jia, 2003). Relatively sparse number marking in syntax may slow down assigning partial meanings to verbal numerals. Finally, Palestinian Arabic has a dual marker system and also distinguishes plural morphology and collective morphology. In a study of children's learning number marking in Palestinian Arabic, Ravid and Hayek (2003) found that 3-year-olds often used the numeral translated "two" instead of the dual marker when referring to sets of two objects, whereas older children were unlikely to do this. This finding is consistent with the suggestion that the meaning of "two" is being formulated over the semantic space of quantifiers.

Before we can understand how children work out the meaning of the count list, we must establish what "one," "two", "three" and "four" mean for a "four"-knower, for it is only after being "four"-knowers that children figure out the cardinal principle. What is the format of the mental representations that underlie the numerical meanings subset-knowers have created for numerals? If "two" is a dual marker for "two"-knowers, what representations give numerical meaning to dual markers? What is the process through which a given set is assigned one numeral rather than another?

LeCorre and I proposed a system of representations that could underlie the meanings of numerals for subset-knowers that draw on the resources both of natural language quantification and parallel individuation, and thus we dubbed it "enriched parallel individuation" (LeCorre & Carey, 2007; see also Mix, Levine & Huttenlocher, 2002).

The parallel individuation system that is part of core cognition creates working-memory models of sets. The symbols in these models represent particular individuals —two boxes might be represented {*this box*, *that box*}. Even when drawing on parallel individuation alone, infants have the capacity to represent two models and compare them on the basis of 1-1 correspondence (Feigenson, 2005; Feigenson & Carey, 2003). For representations of this format to underlie the meanings of the singular determiner or the numeral "one" for subset-knowers, the child may create a long-term memory model of a set of one individual and map it to the linguistic expression "one." Similarly, a long-term memory model of a set of two individuals could be created and mapped to the linguistic expression for a dual marker or "two," and so on for "three" and "four." These models could contain abstract symbols for individuals ({*i*}, {*j k*}, {*m n o*}, {*w x y z*}) or they could simply be long-term memory models of particular sets of individuals ({Mommy}, {Daddy Johnnie}. . .). What makes these models represent "one" "two" and so forth is their computational role. They are

deployed in assigning numerals to sets as follows: The child makes a working-memory model of a particular set he or she wants to quantify {e.g., cookie cookie}. He she then searches the models in long-term memory to find that which can be put in 1–1 correspondence with this working-memory model, retrieving the quantifier that has been mapped to that model. When the child has built this structure, she has created a summary symbol "one," a new semantic primitive that can underlie the meanings of symbols for exact cardinal values of sets.

All of the computational resources required for enriched parallel individuation are known to be available to prelinguistic infants (see Carey, 2009, for full evidence for this claim). Prelinguistic infants create working-memory models of at least two separate sets and compare these on the basis of 1-1 correspondence. They also treat sets as objects, quantifying over them as required by natural language quantifiers. Still, it is important to stress that the long-term memory models that support the meanings of singular, dual, and triple markers, as well as the child's first numerals, are not themselves part of core cognition. These must be created in the course of language learning, and for English-learning children this process unfolds for a period of over a year. This why LeCorre and I designate the hypothesized system of representation "enriched parallel individuation."

The two important planks of the bootstrapping process are constructed in parallel, largely independent of each other. The child learns the explicit numeral list and the count routine as a numerically meaningless game. The child also creates numerical meanings for some numerals —in this case "one" through "four." These meanings are supplied by enriched parallel individuation. So far no bootstrapping has occurred.

The stage is set for the completion of the bootstrapping processes. Children use the placeholder structure to model sets of individuals in the world. They must note the identity of the words "one, two," "three," and "four" which now have numerical meaning, and the first words in the otherwise meaningless counting list. Also, in the course of counting, children discover that when an attended set would be quantified with the dual marker "two," the count goes "one, two," and when an attended set would be quantified with the tripal marker "three," the count goes "one, two, three." The child is thus in the position to notice that for these words at least, the last word reached in a count refers to the cardinal value of the whole set, where that exact cardinal value is represented by enriched parallel individuation.

At this point, the stage is set for the crucial induction. The child must notice an analogy between next in the numeral list and next in the series of mental models ({i}, {$j\ k$}, {$m\ n\ o$}, {$w\ x\ y\ z$}) related by adding an individual. Remember, core cognition supports the comparison of two sets simultaneously held in memory on the basis of 1-1 correspondence, so the child has the capacity to represent this latter basis of ordering sets. This analogy licenses the crucial induction: if "x" is followed by "y" in the counting sequence, adding a individual to a set with cardinal value x results in a set with cardinal value y. This generalization does not yet embody the arithmetic successor function, but one additional step is all that is needed. Since the child has already mapped single individuals onto "one," adding a new individual is equivalent to adding one.

This proposal makes sense of the actual partial meanings children assign to number words as they try to fill in the placeholders. The semantics of quantifiers explain these facts. It makes sense of the fact that subset-knowers acquire the cardinal meanings of "one" "two" "three" and "four," and no other numeral, on the assumption that by age 3 the limits on parallel individuation have reached the adult level of 4 (see Oakes, Ross-Sheey, & Luck, 2006, for evidence that the adult level is evident even in infancy under some circumstances). Only sets

of these sizes are representable by models of the sets of individuals held in parallel in working memory, thus to be matched via 1-1 correspondence to long-term memory models of sets of one, two, three, and four individuals.

We sought an answer to several questions, including how children assign numerical meanings to verbal numerals and how they learn how the list itself represents number. The bootstrapping proposal provides answers to both. The meanings of "one" through "four" are acquired just as quantifiers in natural languages are — as quantifiers for single individuals, pairs, triples, and quadruples. These words, as well as higher numerals, also get initial interpretations as part of a placeholder structure, the count list itself, in which meaning is exhausted by the fact that the list is ordered. The bootstrapping process explains how children learn how the list itself represents number, which in turn explains how they assign numerical meaning to numerals like "five" and "seven." When children first become cardinal principle knowers, the meaning of "five" is exhausted by the child's mastery of counting. The counting principles ensure that the content of "five" is one more than four, and the meaning of "seven" is one more than six, which is one more than five, which is one more than four.

Conclusions

I have argued here that the numeral list representation of number is a representational resource that transcends any single representational system available to prelinguistic infants. It has more representational power — indeed it allows for representing an infinity of concepts not available to the infant (the positive integers.) Also, it is articulated in terms of primitives (*one, exact cardinal value*) not expressed in any of the systems of representation that are the available prelinguistically. When children, at around age 3 , have mastered how the count sequence represents number, they can represent any exact cardinality expressed in their count list. Before that, they have only the quantificational resources of natural languages, parallel individuation representations that implicitly represent small numbers, and analog magnitude representations that provide approximate representations of the cardinal values of sets.

Additionally, I have taken on the challenge of specifying a learning mechanism that can underlie specific developmental discontinuities — Quinian bootstrapping. Quinian bootstrapping involves, but is not exhausted by, familiar learning processes: association, the mechanisms that support language learning that does not require Quinian bootstrapping, and so on. In addition it involves noticing analogies and making inductive and abductive leaps. It is a constraint satisfaction process, the sources of constraint being provided by the placeholder structure and the systems of representation the child is attempting to model. The specific bootstrapping proposal proposed here depends on the analogy between next on the numeral list and next state after additional individual has been added to a set.

In Quinian bootstrapping, an explicit structure is learned initially without the meaning it will eventually have, and at least some relations among the explicit symbols are learned directly in terms of each other. The list of numeral words and the counting routine are learned as numerically meaningless structures. Whereas order is essential to numerical representations, ordered relations in themselves are much more general and thus not uniquely numerical. The ordering of the number words exhausts their initial representational content within the counting routine and plays a role in the mappings and inductions through which counting comes to have numerical content.

Quinian bootstrapping depends upon integrating previously distinct representations. This is where the new representational power comes from. The concepts *set, individual, singular,*

plural, dual, and *triple* are explicitly available to support the learning of quantifiers, but are only implicit or absent in parallel individuation and analog magnitude representations. In analog magnitude representations, numerical distinctions are explicitly symbolized that are unmarked in natural language quantifiers or parallel individuation (e.g., 35 vs. 40), and although analog magnitude representations may play no role in the child's learning how counting represents number, they are integrated with counting some six months later (LeCorre & Carey, 2007). And the representations that articulate the parallel individuation system contain computations that embody the successor function, whereas neither of the other systems does. The bootstrapping process (which depends on analogical mapping) creates an explicit representational system with all of these properties, a representational system that maps onto each of its sources and thus serves to integrate them.

The above analysis leaves open how much of the extended mapping process involved in learning the adult lexicon requires Quinian bootstrapping. Episodes of conceptual development that require conceptual change or increases in expressive power are not the rule, but they are common. They arise in the context of theory changes that involve incommensurability (examples from the history of science and from conceptual change in childhood are discussed in Carey, 2009) and they arise in the context of the development of mathematical cognition (Carey, 2009, takes up the case of rational number as well as the case of natural number I have sketched here). I believe they are much more common still; human beings have the capacity to extend their conceptual repertoire in thinking about the natural, social and the supernatural world. The representational resources required to think the thoughts we can think, and to express these throughts in language, are cultural constructions that often require Quinian bootstrapping. When the child encounters lexical items from such constructed domains, extending mapping always transpires over years.

Notice, at issue here is the construction of the conceptual resources to express the meanings of words: "seven" or "alive," or "matter" or "heat" or "a half" or "mass" or "God." Although acquiring information about and changing our beliefs about the entities referred to by such words unfolds over a life-time, that is not the question at hand. Rather, what I am concerned with is how people come to have concepts that express the meaning of such words; I assume a concept/conception distinction (see Carey, 2009). The story I have told here is unabashadly Whorfian. External language plays a crucial role in Quinian bootstrapping: the placeholder structures are expressed in public symbols (mathematical as well as linguistic). The reanalysis of the extended mapping process to involve, at least sometimes, the creation of new semantic primitives, and the role of language in the bootstrapping mechanisms that make this possible, implies that language plays a crucial role in the capacity to think many of the thoughts articulated in terms of the concepts expressed in the adult lexicon.

References

Antell S, Keating DP. Perception of numerical invariance in neonates. Child Development. 1983; 54:695–701. [PubMed: 6851716]

Baldwin DA. Infants' ability to consult the speaker for clues to word reference. Journal of Child Language. 1993; 20:395–418. [PubMed: 8376476]

Block, N. Advertisement for a semantics for psychology. In: French, PA., editor. Midwest Studies in Philosophy. Minneapolis: University of Minnesota; 1986. p. 615-678.

Bloom, P. How Children Learn the Meanings of Words. Cambridge, MA: MIT Press; 2000.

Bloom P, Wynn K. Linguistic cues in the acquistion of number words. Journal of Child Language. 1997; 24(3):511–533. [PubMed: 9519584]

Bowerman, M. Learning the structure of causative verbs: a study in the relationship of cognitive, semantic and syntactic development. In: Clark, E., editor. Papers and reports on child language development; Stanford University Committee on Linguistics; 1974. p. 142-178.

Brannon EM. The development of ordinal numerical knowledge in infancy. Cognition. 2002; 83:223–240. [PubMed: 11934402]

Brannon EM, Abbott S, Lutz DJ. Number bias for the discrimination of large visual sets in infancy. Cognition. 2004; 93:B59–B68. [PubMed: 15147939]

Brown R. Linguistic determinism and the part of speech. Journal of Abnormal and Social Psychology. 1957; 55:1–5.

Carey, S. Less may never mean more. In: Campbell, R.; Smith, P., editors. Recent Advances in the Psychology of Language. New York: Plenum Press; 1978a. p. 109-132.

Carey, S. The child as word learner. In: Bresnan, J.; Miller, G.; Halle, M., editors. Linguistic Theory and Psychological Reality. Cambridge, MA: MIT Press; 1978b. p. 264-293.

Carey, S. The Origin of Concepts. New York: Oxford University Press; 2009.

Carey S, Bartlett E. Acquiring a single new word. Proceedings of the Stanford Child Language Conference. 1978; 15:17–29.

Clark EV. On the child's acquisition of antonyms in two semantic fields. Journal of Verbal Learning and Verbal Behavior. 1972; 11:750–758.

Clark EV, Nikatina TN. One vs. more than one: Antecedents to plurality in early language acquisition. Linguistics. 2009; 47:000–000. (page numbers to be tracked down).

Clearfield MW, Mix KS. Number versus contour length in infants' discrimination of small visual sets. Psychological Science. 1999; 10:408–411.

Clearfield MW, Mix KS. Amount versus number: Infants' use of area and contour lenght to discriminate small sets. Journal of Cognition and Development. 2001; 2(3):243–260.

Dehaene, S. The Number Sense. Oxford: Oxford University Press; 1997.

Dewar K, Xu F. Do 9-month-old infants expect distinct words to refer to kinds? Developmental Psychology. 2007; 43:1227–1238. [PubMed: 17723047]

Feigenson L. A double-dissociation in infants' representation of object arrays. Cognition. 2005; 95:B37–B48. [PubMed: 15788156]

Feigenson L, Carey S. On the limits of infants' quantification of small object arrays. Cognition. 2005; 97(3):295–313. [PubMed: 16260263]

Feigenson L, Carey S. Tracking individuals via object files: evidence from infants' manual search. Developmental Science. 2003; 6(5):568–584.

Feigenson L, Carey S, Hauser MD. The representations underlying infants' choice of more: object files versus analog magnitudes. Psychological Science. 2002; 13:150–156. [PubMed: 11933999]

Feigenson L, Carey S, Spelke E. Infants' discrimination of number vs. continuous extent. Cognitive Psychology. 2002; 44:33–66. [PubMed: 11814309]

Fuson, KC. Children's Counting and Concepts of Number. New York: Springer-Verlag; 1988.

Gallistel, CR. The Organization of Learning. Cambridge, MA: MIT Press; 1990.

Gleitman LR, Cassidy K, Nappa R, Papafragou A, Trueswell JC. Hard Words. Language Learning and Development. 2005; 1(1):23–64.

Katz B, Baker G, Macnamara J. What's in a name? On the child's acquisition of proper and common nouns. Child Development. 1974; 45:269–273.

Kahneman D, Treisman A, Gibbs B. The reviewing of object files: Object-specific integration of information. Cogntive Psychology. 1992; 24:175–219.

Le Corre M, Brannon EM, Van de Walle GA, Carey S. Re-visiting the competence/performance debate in the acquisition of the counting principles. Cognitive Psychology. 2006; 52(2):130–169. [PubMed: 16364281]

Le Corre M, Carey S. One, two, three, four, nothing more: How numerals are mapped onto core knowledge of number in the acquisition of the counting principles. Cognition. 2007; 105:395–438. [PubMed: 17208214]

Li, P.; Le Corre, M.; Shui, R.; Jia, G.; Carey, S. Effects of plural syntax on number word learning: A Cross-linguistic study. Paper presented at the 28th Boston University Conference on Language Development; Boston. 2003.

Lipton JS, Spelke ES. Origins of number sense: Large number discrimination in human infants. Psychological Science. 2003; 15(5):396–401. [PubMed: 12930467]

Lipton JS, Spelke E. Discrimination of large and small numerosities by human infants. Infancy. 2004; 5(3):271–290.

Markman, EM. Categorization and Naming in Children: Problems of Induction. Cambridge, MA: MIT Press; 1989.

Markson L, Bloom P. Evidence against a dedicated system for word learning in children. Nature. 1997; 385:813–815. [PubMed: 9039912]

McCrink K, Wynn K. Large-number addition and subtractrion by 9-month old infants. Psychological Science. 2004a; 15(11):776–781. [PubMed: 15482450]

McCrink, K.; Wynn, K. Ratio abstraction by 6-month-old infants. Paper presented at the International Conference on Infant Studies; Chicago, USA. 2004b.

Miller, GA.; Johnson-Laird, P. Language and Perception. Cambridge, MA: Harvard University Press; 1976.

Mix KS, Huttenlocher J, Levine SC. Do preschool children recognize auditory-visual numerical correspondences? Child Development. 1996; 67(4):1592–1608. [PubMed: 8890501]

Nersessian, N. How do Scientists Think? Capturing the dynamics of conceptual change in science. In: Giere, R., editor. Cognitive Models of Science. Minneapolis: University of Minnesota Press; 1992. p. 3-44.

Pinker, S. Language Learnability and Language Development. Cambridge, MA: Harvard University Press; 1984.

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

Quine, WVO. Ontological Relativity and Other Essays. New York: Columbia University Press; 1969.

Quine, WVO. Natural kinds. In: Schwartz, SP., editor. Naming, Necessity and Natural Kinds. Ithaca, NY: Cornell University Press; 1977. p. 89-108.

Ravid D, Hayek L. Learning about different ways of expressing number in the development of Palestinian Arabic. First Language. 2003; 23(1):41–63.

Ross-Sheehy S, Oakes L, Luck SJ. The development of visual short-term memory capacity in infants. Child Development. 2003; 74(6):1807–1822. [PubMed: 14669897]

Sarnecka BW, Kamenskaya VG, Ogura T, Yamana Y, Yudovina JB. From grammatical number to exact numbers: Early meanings of "one," "two," and "three," in English, Russian, and Japanese. Cognitive Psychology. 2007; 55(2):136–168. [PubMed: 17070794]

Scholl, BJ.; Leslie, AM. Explaining the infant's object concept: Beyond the perception/cognition dichotomy. In: Lepore, E.; Pylyshyn, Z., editors. What is Cognitive Science?. Oxford: Blackwell; 1999. p. 26-73.

Simon TJ. Reconceptualizing the origins of number knowledge: A "non-numerical" account. Cognitive Development. 1997; 12:349–372.

Starkey P, Cooper R. Perception of numbers by human infants. Science. 1980; 210:1033–1035. [PubMed: 7434014]

Terrace HS, Son LK, Brannon EM. Serial expertise of rhesus macaques. Psychological Science. 2003; 14(1):66–73. [PubMed: 12564756]

Tomasello M, Akhtar N. Two-year-olds use pragmatic cues to differentiate reference to objects and actions. Cognitive Development. 1995; 10:201–224.

Uller C, Carey S, Huntley-Fenner G, Klatt L. What representations might underlie infant numerical knowledge. Cognitive Development. 1999; 14:1–36.

Waxman SR, Booth AE. Mapping words to world in infancy: Infants' expectations for count nouns and adjectives. Journal of Cognition and Development. 2003; 4(3):357–381.

Wood J, Spelke E. Infants' enumeration of actions: Numerical discrimination and its signature limits. Developmental Science. 2005a; 8(2):173–181. [PubMed: 15720375]

Wynn K. Children's understanding of counting. Cognition. 1990; 36:155–193. [PubMed: 2225756]

Wynn K. Children's acquisition of the number words and the counting system. Cognitive Psychology. 1992a; 24:220–251.

Wynn K. Addition and subtraction by human infants. Nature. 1992b; 358:749–750. [PubMed: 1508269]

Xu F, Spelke ES. Large number discrimination in 6-month old infants. Cognition. 2000; 74:B1–B11. [PubMed: 10594312]

Xu F, Spelke E, Goddard S. Number sense in human infants. Developmental Science. 2005; 8:1. [PubMed: 15647061]

Number	Analog Magnitude Representation
1	—
2	——
3	———
4	————
7	———————
8	————————
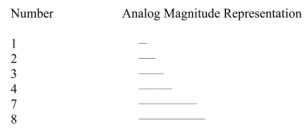
Figure 1.
External analog magnitude representation of number in which number is represented by line length.
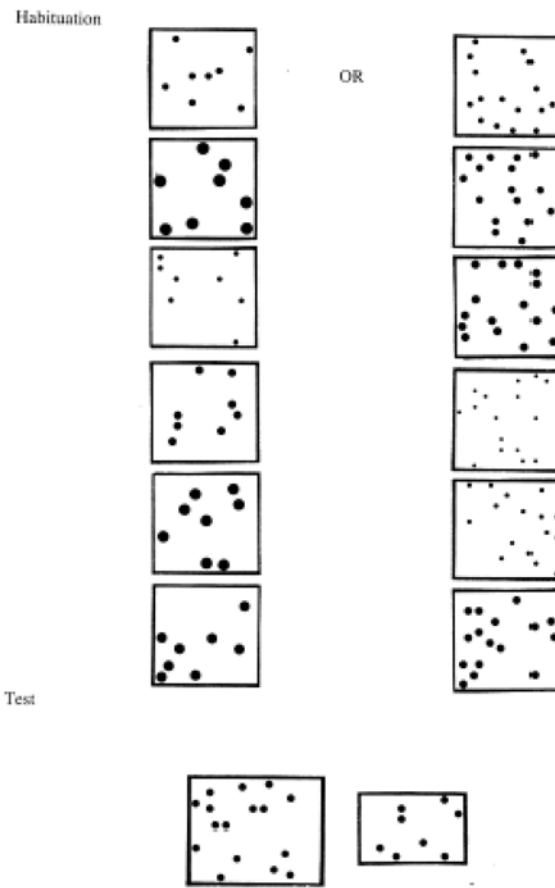
Figure 2.
Sample habituation sequences and test stimuli from Xu & Spelke (2000).
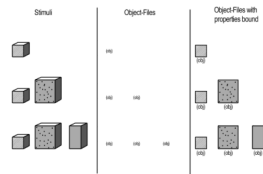
Figure 3.
Two versions of the memory structures that might subserve parallel individuation of small sets of objects. In one, each object is represented by an object-file that abstracts away from specific features (OBJ). In the other, each object is represented by an object-file on which shape, color, texture and spatial extent features have been bound.