Developmental Psychology
2004, Vol. 40, No. 4, 621–632

Copyright 2004 by the American Psychological Association
0012-1649/04/$12.00 DOI: 10.1037/0012-1649.40.4.621

Reexamining the Vocabulary Spurt

Jennifer Ganger
University of Pittsburgh

Michael R. Brent
Washington University in Saint Louis

The authors asked whether there is evidence to support the existence of the vocabulary spurt, an increase in the rate of word learning that is thought to occur during the 2nd year of life. Using longitudinal data from 38 children, they modeled the rate of word learning with two functions, one with an inflection point (logistic), which would indicate a spurt, and one without an inflection point (quadratic). Comparing the fits of these two functions using likelihood ratios, they found that just 5 children had a better logistic fit, which indicated that these children underwent a spurt. The implications for theories of cognitive and language development are considered.

Typically developing children utter their first words between 8 and 14 months of age. At this time, they add words to their repertoire at a slow rate. As they get older and their vocabulary increases, their rate of learning new words also increases—it has to if they are to reach an average vocabulary level of 300 words by 24 months (Fenson, Dale, Reznick, Bates, & Thal, 1994) and 60,000 words by 18 years (Aitchinson, 1994). In addition, it is widely held that children's rate of vocabulary acquisition does not simply increase but undergoes a discrete transition at approximately 50 words. At this time, children putatively switch from an initial stage of slow vocabulary growth to a subsequent stage of faster growth. This transition has been referred to as the *vocabulary spurt*, the *vocabulary burst*, or the *naming explosion* (L. Bloom, 1973; Dromi, 1987; McCarthy, 1954; Nelson, 1973).

The spurt concept was described nicely in a recent article by Dapretto and Bjork (2000), who wrote, "Following the onset of expressive language, the rate of word acquisition is initially rather slow, with children learning only a few new words per month. Toward the end of the second year, children typically display a sudden spurt in vocabulary growth, roughly after their productive lexicons have reached 50–100 words" (Dapretto & Bjork, 2000, p. 635). This idea is adhered to explicitly or implicitly in dozens of articles and has become common knowledge in our field. For example, in her popular textbook of child development, Laura Berk wrote, "Young toddlers add to their vocabularies slowly, at a rate of 1 to 3 words per month. Over time the number of words learned accelerates . . . a spurt in vocabulary often takes place . . . [after which] many children add 10 to 20 new words a week" (Berk, 2003, p. 371).

The goal of this article is to provide a clear, testable definition of the vocabulary spurt and to look for evidence of it in the vocabulary growth patterns of children. Although it is uncontroversial that a child's rate of word learning increases during the 2nd year of life, the existence of two stages of word learning (slow prespurt learning and rapid postspurt learning) needs to be tested more carefully. Investigations of the vocabulary spurt to date have accepted almost any increase in learning rate as a spurt without evidence for an initial slow rate followed by a sustained faster rate. The hypothesis that two different stages exist must be distinguished from the very different hypothesis that children become more competent learners continuously as the result of a long series of uncoordinated improvements in their overall cognitive and linguistic abilities.

The difference between these two characterizations of vocabulary growth—a spurt versus a gradual increase—is of more than mathematical importance. The vocabulary spurt has been cited as evidence for a number of cognitive changes, some of which may no longer be supported if the spurt does not exist. We turn next to a consideration of some of these theories.

Theories of the Vocabulary Spurt

One of the most influential claims about the vocabulary spurt is that it coincides with a major cognitive change: the *naming insight*, which is the realization that words refer to things or that all things have names (Dore, 1978; Dore, Franklin, Miller, & Ramer, 1976; Gillis & De Schutter, 1986; Kamhi, 1986; Reznick & Goldfield, 1992). This insight marks a shift in the use of words from vocalizations associated with specific routines to true adult words that may be used to refer (Lock, 1978; McShane, 1980). Once children have this insight, the theory holds, they begin to acquire words at a rapid pace.

Jennifer Ganger, Department of Psychology, University of Pittsburgh; Michael R. Brent, Department of Computer Science, Washington University in Saint Louis.

Data in this study were originally collected as part of the Massachusetts Institute of Technology Twins Study and are partially reported in Jennifer Ganger's PhD dissertation. An earlier version of this work appeared in the *Proceedings of the 25th Annual Boston University Conference on Language Development*. This study was supported by National Institute of Child Health and Human Development Grant HD18381 to Steven Pinker and Grant HD08493 to Jennifer Ganger and by National Institute on Deafness and Other Communication Disorders Grant DC03082 to Michael R. Brent. We thank the parents of the children in this study for their dedicated record keeping. We also thank Sonia Chawla, Allison Baker, Emily Wallis Lilly, and Jeff Birk for their assistance in compiling and coding those records and Chaochun Wei for his help with data analysis.

Correspondence concerning this article should be addressed to Jennifer Ganger, 210 South Bouquet Street, 3405 Sennott Square, Pittsburgh, PA 15260. E-mail: jganger@pitt.edu

Another theory is that the spurt marks a change in children's object concepts. In one version of this theory (Lifter & Bloom, 1989), infants begin to use words when they enter Piaget's Sensorimotor Substage 6 and then show a vocabulary spurt as their object concepts become more detailed and differentiated, around the close of Substage 6. Similar ideas have been advanced by Corrigan (1978) and Bates, Benigni, Bretherton, Camaioni, and Volterra (1979).

Another theory about the relationship between object concepts and word learning comes from Gopnik and Meltzoff (1987). They proposed that the child's ability to sort objects into groups based on category membership (using a task known as two-category sorting) improves around the time of the vocabulary spurt (Gopnik & Meltzoff, 1987; Mervis & Bertrand, 1994; Poulin-Dubois, Graham, & Sippola, 1995). (This correlation has been disputed, though—see Gershkoff-Stowe, Thal, Smith, & Namy, 1997; Schafer & Plunkett, 1998; Woodward, Markman, & Fitzsimmons, 1994.) Gopnik and Meltzoff's (1987) explanation is that categorization reflects the child's realization that objects can be sorted in basic-level categories and that such an understanding is necessary for learning basic-level names. Once this understanding is in place, the child's acquisition of words can proceed at a more rapid pace, allowing a vocabulary spurt. Mervis and Bertrand (1994) added that the novel name–nameless category principle (N3C), a word-learning constraint specifying that novel names are extended to novel objects, also emerges at the same time as the vocabulary spurt and is a prerequisite both for it and for two-category sorting.

The vocabulary spurt has also been related to advances of a more linguistic nature. Plunkett (1993) suggested that the vocabulary spurt occurs when word segmentation has been solved, implying that the ability to pick the words out from running speech opens the floodgates to producing many new words. Walley (1993) argued that the segmental representation of words changes as the vocabulary spurt occurs, leading to the ability to represent words in more accurate detail as more and more words are learned (though see Swingley & Aslin, 2002, for conflicting ideas). Ninio (1995) provided evidence that advances in pragmatics are correlated with the vocabulary spurt, suggesting that gains in social cognition permit the acquisition of words at a higher rate. Finally, Dapretto and Bjork (2000) contended that the spurt in productive vocabulary is driven by the development of word retrieval abilities.

The Vocabulary Spurt as a Developmental Milestone

In addition to the literature dedicated to uncovering the meaning of the vocabulary spurt, there are many studies in which the phenomenon has been used simply as a milestone of linguistic and cognitive development (e.g., L. Bloom & Capatides, 1987; Choi & Gopnik, 1995; Gershkoff-Stowe & Smith, 1997) and as a marker of developmental change (e.g., Fischer, Pipp, & Bullock, 1984, cited in Lifter & Bloom, 1989). Although researchers have not reached a consensus either on the underlying cause of the vocabulary spurt or on which cognitive measures are correlated with it, there seems to be little doubt about the existence of the vocabulary spurt itself or its ultimate importance to cognitive development. Even Elman et al. (1996), who argued that the vocabulary spurt may not reflect any underlying cognitive change, assumed that there is an observable spurt in the rate of word learning (see also Bates & Carnevale, 1993; van Geert, 1991).

The existence of the vocabulary spurt seems confirmed not only by eyeballing data from studies such as Fenson et al. (1994), in which a large, cross-sectional sample of children appeared to show a spurt trend, and Dromi (1987), in which 1 child studied longitudinally also appeared to show a spurt, but also by the studies cited in the previous section, which identified a spurt in children's vocabulary development and tested whether it was correlated with some cognitive or linguistic change. However, as is shown in the next section, the methods currently used to identify the spurt cannot provide definitive evidence about whether the spurt exists.

Identifying the Spurt

To determine when a child undergoes the vocabulary spurt, one of three methods is typically used. The first is simply to estimate a child's overall vocabulary size and age. If the size is substantially greater than 50 words, or if the child's age is greater than 20 months, it is assumed that the child has undergone a spurt. The size/age test is a quick and dirty measure, used by researchers whose main focus is not the spurt itself but some other process that is thought to be related to the spurt (e.g., Schafer & Plunkett, 1998). Those who use this metric assume that there is a spurt and that most children experience it. A variant on this approach is to consider a ratio of comprehension and production vocabulary (Dapretto & Bjork, 2000).

A second method is to produce a graph of vocabulary size over time, such as that given for the child Keren by Dromi (1987), and to eyeball it for a spurt. Although eyeballing data for trends is an important step in all kinds of research, doing so may be misleading when it comes to finding a vocabulary spurt, for reasons that will be elaborated later.

Finally, the most common method for identifying when a child has a vocabulary spurt is what we shall call the *threshold approach*. In this approach, a threshold of words per unit of time must be crossed. A survey of the literature revealed the following thresholds: Mervis and Bertrand (1994, 1995) required that children learn at least 10 new words in a period of 14 days before they were considered to have entered a vocabulary spurt; Gopnik and Meltzoff (1987) required at least 10 new object words in 3 weeks; Lifter and Bloom (1989; see also L. Bloom & Capatides, 1987) required a 20-word minimum plus 12 new words in 1 month; Choi and Gopnik (1995) required 10 new object words in 3–4 weeks for a noun spurt and 10 new verbs in the same period for a verb spurt; Poulin-Dubois et al. (1995) required 15 new object words in a month; Goldfield and Reznick (1990; see also Reznick & Goldfield, 1992) required 10 new words in 2½ weeks; and Ninio (1995) identified the spurt as the 2-month period in which a child showed the biggest increase in vocabulary relative to the previous measurement. Ultimately, all children must cross all of these thresholds to reach 60,000 words by the age of 18; it is just a matter of whether or not they cross the threshold during the period of a particular study. So perhaps unsurprisingly, none of these authors had trouble finding children with a spurt, confirming intuitions that the spurt is a relatively widespread phenomenon. For example, in their influential 1990 article, Goldfield and Reznick argued that the majority of children (13 of the 18 they studied) showed a vocabulary spurt. Although Goldfield and Reznick's study has been taken as evidence that some children do not have a spurt, readers could nonetheless remain unshaken in their belief that the majority of children do.

Defining the Spurt Clearly

The key property of the vocabulary spurt, as described in the Dapretto and Bjork (2000) quote in our introduction, is that it consists of discrete developmental stages rather than continuous incremental improvement. However, none of the three methods just described (vocabulary level/age, eyeballing, or threshold) is sufficient to identify a vocabulary spurt consisting of discrete developmental stages and to distinguish it from continuous incremental improvement.

A true developmental spurt must be a transition between a slow learning stage and a faster learning stage. The learning rate may increase within each of the stages, but at the transition between the stages the learning rate must change faster than it does during either stage. Mathematically, a point where the rate of increase of a curve is greater than it is before or after is called an *inflection point*. Thus, if a child's learning rate undergoes a transition between a low stage and a high stage, there will be an inflection point in the learning rate curve at the point where the transition occurs.

Furthermore, the later, faster learning rate should be sustained for some period of time. This is true for two reasons. First, the rate must stay the same or increase if the child is to reach 300 words by 24 months (and 60,000 words by 18 years). Second, the vocabulary spurt is thought to be a reflection of an underlying cognitive change, such as the naming insight, an increase in object knowledge, or an advance in word segmentation and phonological representation. An increase in the rate of word learning that lasts only 1 week may not reflect a cognitive change—just a busy week. A cognitive change should persist over time.

Figures 1 and 2 illustrate these points and preview the strategy that will be used to identify a spurt. These graphs show data from 1 participant (Child 041B), slightly modified for expository purposes. The graphs depict the child's rate of word learning (new words per day) on the vertical axis, and cumulative vocabulary size, a variable of interest in many vocabulary studies, on the horizontal axis. Note that this form of representation (rate vs. size) departs from the usual vocabulary growth curve, which typically shows *cumulative vocabulary* versus time (see below for more on this point).

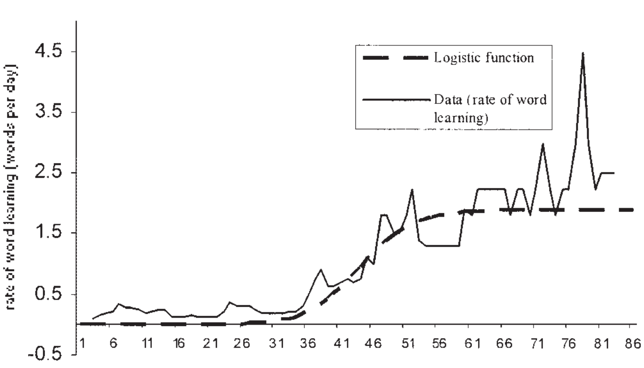

Figure 1. A spurtlike function (logistic) superimposed on slightly modified data from Child 041B.

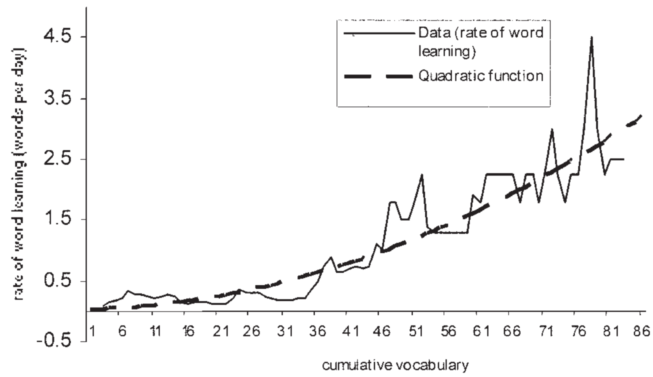

Figure 2. A nonspurtlike curve (quadratic) superimposed on the same data shown in Figure 1.

The question is, are the data in Figures 1 and 2 better described as having two distinct learning rates (i.e., a spurt) or as having a gradually increasing learning rate (i.e., no spurt)? In Figure 1, a function with two distinct levels is superimposed on the data. If this child has a vocabulary spurt, his or her data should be fit well by this function. This is what a spurt should look like. In Figure 2, on the other hand, a function with a gradual increase in rate is superimposed on the same data. The increase in word-learning rate is continual, with no point of transition. If the child has a spurt, his or her data should be fit less well by this function.

Keeping these points in mind, it should be clear now that the three methods typically used to identify a vocabulary spurt are inadequate. Ascertaining when a child's vocabulary exceeds 50 words assumes the existence of a spurt without actually checking for it. Eyeballing a curve of vocabulary size versus time is also unreliable; what looks like a takeoff point in the cumulative vocabulary curve may actually reflect a gradual increase in rate rather than a spurt, especially if one is eyeballing a cumulative vocabulary curve (total words vs. age) instead of rate of acquisition. Finally, asking whether a child has crossed a threshold rate of word learning (e.g., 10 words in 2½ weeks) is also inadequate. A child will cross this threshold whether his or her rate of word learning increases as the result of a transition between stages or as the result of a continual series of unrelated processes. Furthermore, the threshold could be crossed in a good week without heralding a permanent change.

We are not the first to point out such shortcomings with existing methods of detecting a vocabulary spurt. P. Bloom (2000) also pointed out the difference between a continuous and a discontinuous increase in word-learning rate and discussed the implications of this difference. Reviewing the available evidence, he concluded that there probably is no vocabulary spurt in the 2nd year (though he did note the possibility of a later spurt in vocabulary when children learn to read; however, because this spurt would be driven by the sudden availability of new words rather than an internal cognitive change, we do not consider it here). Regarding the original vocabulary spurt, Bloom did not have the relevant data from individual children to test his conclusion, nor did he specify precisely how to test it. To our knowledge, the present work is the first rigorous test of the vocabulary spurt hypothesis to use empirical data from individual children.

A New Method

We propose a new method for identifying the vocabulary spurt, one that tests whether the changes in a child's rate of word learning represent distinct stages with a transition in between. In the last section, we argued that the function superimposed on the data in Figure 1 represents a spurt. This plot can be generated by a logistic function, which is a function of the general form $y = a / (1 + e^{-b(x-c)})$. This function conforms to the sigmoid curve familiar from many psychological processes in which there is a transition between states. In this case, the states are two different rates of word learning. In the logistic function, there are three parameters, *a*, *b*, and *c*. The parameter *a* corresponds to the rate of learning after the transition, or the *asymptote* ($a = 1.9$ in Figure 1, for instance). The parameter *b* corresponds to the length of time over which the transition occurs, or the *slope* of the function at the transition point. It is 0.25 in Figure 1. Parameter *c* corresponds to the point at which the transition occurs, also known as the *inflection point*. The inflection point is at 45 words in Figure 1.

The logistic function is a logical choice to test the vocabulary spurt. It encompasses a family of curves that represents two steady states with a discrete transition in between. There are other functions with inflection points but none that so perfectly depict the state change that is the essence of the vocabulary spurt hypothesis. The method proposed herein is to test how well a logistic function fits each child's vocabulary growth data. Goodness of fit of the logistic function is tested by examining the amount of variance accounted for and also by comparing it to the fit of a model that does not have an inflection point, the quadratic function, shown in Figure 2.

A quadratic function has the general form $y = ax^2 + bx + c$. Here, *b* corresponds to the steady increase in word-learning rate ($b = .0017$ in Figure 2), *a* corresponds to the steady acceleration in word-learning rate ($a = .0004$ in Figure 2), and *c* is a constant defining where the function crosses the vertical axis ($c = .0239$ in Figure 2). Depending on the other parameter values, differences in *c* may affect either the time at which word learning begins or the rate at which it begins. The quadratic family of curves was chosen because it has certain properties one might expect a vocabulary curve to have—namely, it can increase or level off depending on whether it is concave up or down. However, no quadratic curve has an inflection point (i.e., a point of rapid change surrounded by points of slower change), an essential hallmark of a transition between slower and faster word learning. There are, of course, an infinite number of other functions without an inflection point, but the quadratic is most similar to the logistic in form and number of parameters, so it was a logical choice.

One might wonder why a simple linear function is not being used as a model. The answer is that a linear function is a special kind of quadratic function in which *a*, the coefficient of the quadratic term, happens to be 0. Because the family of quadratic curves contains the family of linear curves as a special case, there is always a quadratic curve that fits the data at least as well as the best linear function. To illustrate this point, imagine a set of learning-rate data that lie on a perfect straight line. The result of fitting a quadratic curve to these data would be a straight line—that is, a curve in which $a = 0$. This line would account for 100% of the variance in the data. Furthermore, the family of quadratic functions, unlike the more restricted family of linear functions, has the same number of parameters as the logistic function (three). This means that a comparison between quadratic and logistic fits is more fair than a comparison between linear and logistic fits, where the number of parameters is unequal.

One might also wonder why vocabulary rate, rather than vocabulary size, is being used as the dependent measure. Rate is used because it is a hypothesis about rate that is being examined—namely, that the learning rate undergoes a transition from relatively low to relatively high. It therefore makes sense to examine the change in rate directly. It is possible to formulate a mathematically equivalent model in terms of cumulative vocabulary—the integral of rate. This is the graph most researchers think of when imagining a spurt (see Figure 3 for an example). It is the one that should have an "elbow" if there is a vocabulary spurt. However, testing the notion of an "elbow" would require the use of terms such as "an increase in acceleration," which are more difficult to grasp intuitively. Testing the change in rate instead, which is mathematically equivalent to testing change in acceleration in the cumulative vocabulary curve, allowed us to use a simple logistic function and therefore provided the simplest and most direct test.

To summarize, both the logistic function and the quadratic function have three parameters, and both are among the most plausible representations of vocabulary rate over time, but the logistic function has an inflection point (which would indicate a transition between slow and fast word learning), whereas the quadratic function does not. Thus, comparing the fits of these two functions to the data can tell us whether a child's word-learning rate shows the signature of a vocabulary spurt.

We compared the logistic and quadratic models using R^2 and likelihood ratios, which are described in greater detail in Experiment 1. We applied this method to original empirical data from 20 children. We then applied the same method to data gleaned from 17 children in a study by Goldfield and Reznick (1990) and to the data of Keren, a child described by Dromi (1987).

The remainder of this article is divided into three experiments. In Experiment 1, we present three different ways of looking at the original data. First, the proposed spurt test was applied to 20 children as they acquired their first 90 words. Second, the words

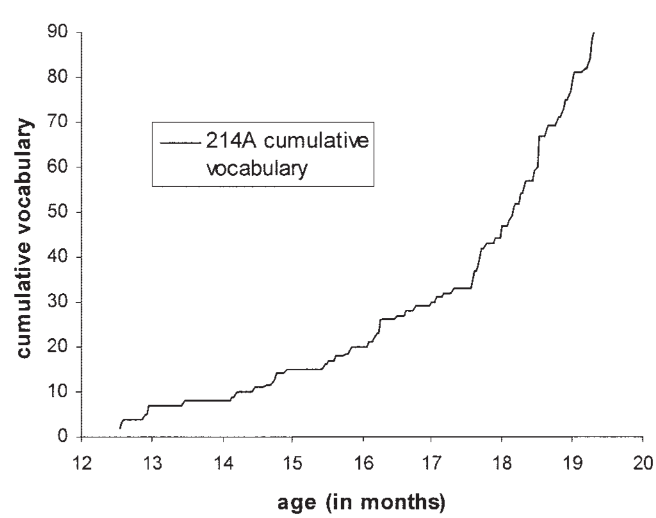

Figure 3. Cumulative vocabulary as a function of time in Child 214A.

were restricted to nouns to test the hypothesis that the vocabulary spurt is driven by noun learning. Third, another set of children from our dataset who reached 150 words during the study were tested in a search for a later spurt. It is important to note that these three analyses represent three different ways at looking at essentially the same data, not three independent tests. Experiment 2 is a test of the proposed method with Goldfield and Reznick's (1990) data. Experiment 3 is a test of the proposed method with Dromi's (1987) data.

Experiment 1

Method

Participants

Participants were 35 children selected from 98 twin pairs recruited as part of a twin study of language development (Ganger, Pinker, Chawla, & Baker, 2000). Twins were recruited from throughout the United States through Mothers of Multiples clubs, the internet, newspapers, the 1995 Twins Days Festival (in Twinsburg, Ohio), advertisements in *Twins Magazine*, and word of mouth. Twins were not accepted into the study if (a) one or both twins had a history of chronic ear infections (defined as more than three infections or the prescribed use of prophylactic antibiotics); (b) the twins were exposed to a language other than English on a regular basis; (c) there was a history of speech or language disorders in the twins' immediate family; (d) either twin's birthweight was less than 4 pounds (1.8 kg) or the twins' gestational age was less than 32 weeks; or (e) either twin had a prolonged hospital stay after birth, defined as 10 days or longer owing to factors typically associated with prematurity such as immature lung development or apnea. Because twins may be at greater risk for language delays and disorders (Hay, Prior, Collett, & Williams, 1987; Mittler, 1970; Mogford, 1993), these rather stringent criteria made it more likely that our participants experienced normal language development.

The first analysis reported below is based on 20 children (10 girls, 10 boys; 13 from identical twin pairs, 7 from same-sex fraternal twin pairs). Because of unavoidable biases in recruitment, this sample was entirely Caucasian. On a scale where 0 denotes that high school was not completed, 1 denotes graduation from high school, 2 denotes some college, 3 denotes a college degree, and 4 denotes an advanced degree, the parents in the final sample reported their educational level to be an average of 3.1. Five of the children had at least one additional sibling who was older; none had younger siblings at the onset of the study. The average age of this sample when parents began keeping journals was 15.3 months (median age = 14.5 months, $SD = 3.4$).

Children were selected for this study from the 98 pairs of twins according to the following three criteria. First, no more than 10 consecutive days could be missing from the child's vocabulary journal, because such holes might produce artifactual spurtlike jumps in vocabulary when the journal resumed. Second, no more than 20 words could be reported in the 1st week of data collection, because such a high level of starting vocabulary in the journal might indicate that the child already knew many more than 20 words and may even have already undergone a spurt. Note that excluding such children does not bias us against finding a spurt. There is no way to know whether the 20 words reported in the 1st week of data collection were learned that week or had been learned previously. Because the children were recruited into the study at a variety of ages, there is no reason to think that children who already had 20 words when their parents started keeping diaries learned them quickly and/or had undergone a spurt. Third, at least 90 words must have been learned by the child before the last 2 weeks of journal entries, because the final 2 weeks of each child's data were removed to guard against the possibility that the parents keeping the journals might have done a less careful job of it in the final weeks before they left the study.

Since co-twins are not independent sources of data, only one twin from each pair could be used. If both twins in a pair met the criteria, only the twin whose name came first alphabetically was included. All of these criteria taken together left 20 children.

Procedure for Collecting Vocabulary Data

Children's words were tracked by asking parents to keep journals of their twins' language use. Parents kept a list of all the different words used by each twin each day, even if the words had been said on previous days. They were told to note which words were immediate imitations and which were spontaneously produced (imitated words were not credited to the child) and to write down any accompanying context if it might help elucidate the word's meaning. In cases of uncertainty, parents were told to use their intuition and to be as consistent as they could from one day to the next and from one twin to the other.

The daily lists of words were kept on forms provided by the experimenters. The forms had spaces for the date, the twins' names, each twin's daily word and/or phrase list, and comments on the words and phrases. Parents mailed these journals back to the laboratory each week and were contacted by experimenters over the phone or by e-mail every 3–4 weeks to address problems and questions.

Parent report has a long tradition in studies of early language production (e.g., Bates, Bretherton, & Snyder, 1988; Fenson et al., 1994; Goldfield & Reznick, 1990; Nelson, 1973; Pine & Lieven, 1990) because it allows for larger sample sizes than does direct observation. It may also be more accurate, because parents have access to a larger range of contexts in which their children use a word or construction. One parent-report vocabulary instrument, the MacArthur Communicative Development Inventory (Fenson et al., 1994), a checklist of 680 words, has been shown to have high test–retest reliability (rs = .80 to .90), high internal consistency between different vocabulary measures within the inventory (rs = .95 to .96), and high reliability when compared with a laboratory measure of expressive vocabulary ($r = .72$; Bates et al., 1988; Fenson et al., 1994). This suggests that parents are reasonably good recorders of their children's language development.

The use of free-report journals is less common but was necessary in this study to document trends in development over time and to decrease the burden on parents of twins. The correlation between free-recall and checklist vocabulary estimates ranged from .60 to .86 in another study (Reznick & Goldfield, 1994) and was high in a test with the original twin sample of this study as well ($r = .98$; Ganger et al., 2000).

Procedure for Testing for a Vocabulary Spurt

Children's data were represented as rate of new word acquisition (specifically, new words per day) versus cumulative vocabulary. The rationale for using rate as the dependent measure, rather than vocabulary size, has been described above. The rationale for using cumulative vocabulary as an independent variable instead of time or age is that the vocabulary spurt is thought to occur around the 50-word mark in cumulative vocabulary rather than being tied to a specific age or length of time after the first word.

In order to give the reader a sense of what these data are like, Figures 3 through 6 show the relationship between different representations of the same data using plots from 1 participant, Child 214A. In Figure 3 there is a typical graph of vocabulary versus time (cumulative vocabulary vs. age in months). Figure 4 shows the same data graphed as rate of vocabulary acquisition (new words per day) versus time (age in months). Figure 5 shows new words per day versus cumulative vocabulary—the representation used in the present study. Finally, Figure 6 shows a smoothed and corrected version of these data, created in a process described next.

To smooth the data out over occasional missing days or 1-day bursts, we calculated an average rate over a span of 9 words. For each level of vocabulary, this average consisted of a window of 9 words (with the day

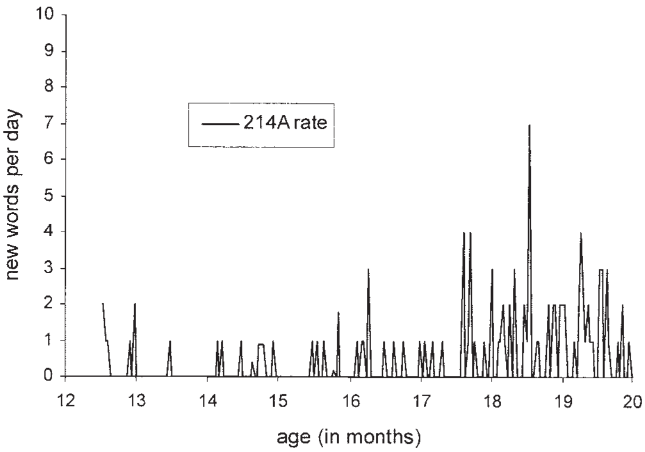

Figure 4. Rate of vocabulary learning as a function of time in the same child shown in Figure 3 (Child 214A).

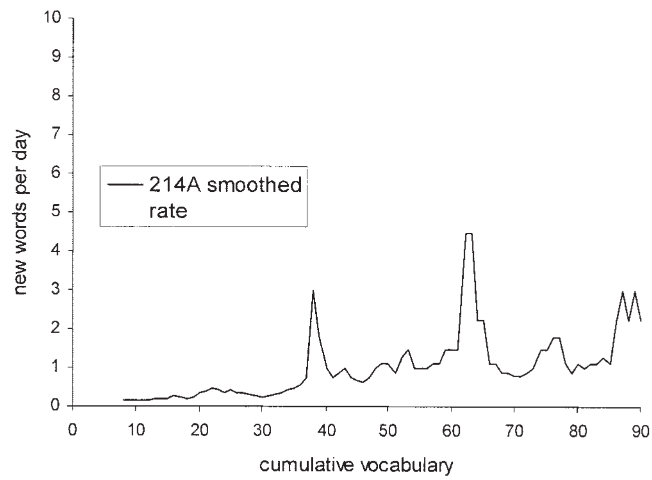

Figure 6. Rate of vocabulary learning as a function of cumulative vocabulary in the same child shown in Figures 3, 4, and 5 (Child 214A) with smoothing over a 9-word window and correction for missing days.

of interest at the center) divided by the number of days it took to learn those 9 words. This number of words (9) was chosen arbitrarily but was chosen because it looked like a large enough number to smooth out temporary jumps that were not of interest while at the same time being a small enough number that any spurting trends would still show up. True spurting trends would still show up because it generally took a few days, on average, for a child to learn 9 new words, and smoothing over a few days should not obscure a true spurt. For missing days, an average of the surrounding days was filled in. (Recall that no child in this study had more than 10 consecutive missing days.)

As described in the introduction, our method of determining whether a child had a vocabulary spurt was tied to whether his or her rate of acquiring new words could be modeled well by a logistic function. A logistic model was fit to each child's data by using the nonlinear regression function of SPSS for Windows and entering the model

$$\mathit{asymptote}/\{1 + e^{[-\mathit{slope}\,(\mathit{words} - \mathit{inflection\ point})]}\}$$

by hand. The dependent variable was new words per day. *Words*, the child's cumulative vocabulary level, was the independent variable. *Asymptote*, *inflection point*, and *slope* are parameters that are fit to the child's

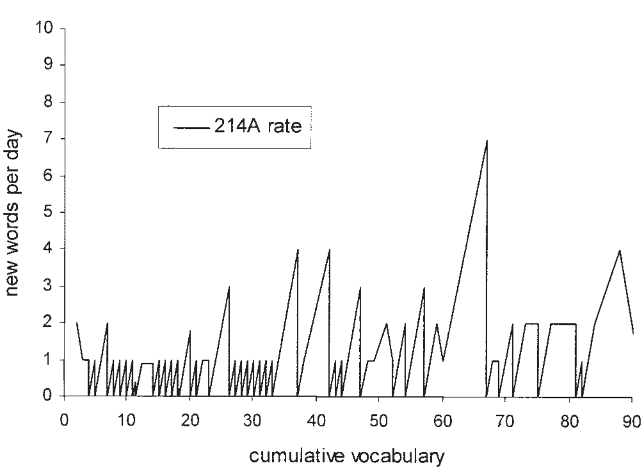

Figure 5. Rate of vocabulary learning as a function of cumulative vocabulary in the same child shown in Figures 3 and 4 (Child 214A).

data. (Their meaning is described in more detail in the introduction.) The parameters of this model were fit to each child's data independently using the parameters option of the SPSS nonlinear regression function. This means that SPSS's model-fitting algorithm tried dozens to hundreds of combinations of parameter values until it found those that best fit the data. Initial values of the parameters were set as follows: *asymptote* = 3.0, *inflection point* = −0.1, and *slope* = 0.1. These values may appear arbitrary, but they were varied during exploratory analysis, and they had little effect on the result as long as they were not extremely far from reasonable final values. Furthermore, SPSS's algorithm uses these values merely as starting points; it still tests a wide range of values and selects those that best fit the data.

The most important parameter for our purposes is the inflection point because it represents the center of the transition—if there is one—between a stage during which the rate of vocabulary growth is low and a stage during which it is high. In this analysis and in the nouns-only analysis, described below, an inflection point anywhere between 20 and 90 words was accepted, because the vocabulary spurt is thought to occur at around 50 words. A minimum of 20 words was decided upon because many investigators agree that before 20 words is too early for a true spurt (e.g., Lifter & Bloom, 1989). A maximum of 90 was chosen because data from most of our participants ended at 90 words after the last 2 weeks of data were removed. If a child did not have an inflection point between 20 and 90 words, we concluded that the child did not have a vocabulary spurt.

However, while having an inflection point at the right place is necessary for having a spurtlike curve, it is not sufficient. SPSS's logistic model-fitting algorithm will find an inflection point somewhere in the data even if the data could be better fit by a function without an inflection point. Therefore, we also tested the fit of a quadratic model (discussed in more detail in the introduction) by using the curve-fitting function of SPSS for Windows and checking the quadratic option.

The two models were compared for their goodness of fit by using R^2 and likelihood ratios. The model with the higher R^2 was considered potentially to be the better fit, but to make the comparison more rigorous, we also computed likelihood ratios for the two models. A likelihood ratio is the ratio of the probability of the observed data points under one model (the logistic) to their probability under an alternative model (the quadratic). Therefore, the larger the ratio, the better the logistic curve fits the data relative to the quadratic curve. Note that this is not the ratio of the probabilities of the models but rather the ratio of the probabilities of the data if the model is correct. Like a traditional significance value, it is a function of the sample size as well as the effect size. By convention, a ratio of 100:1 is required to be very confident of the result, a ratio of 10:1 is worth a second glance, and anything smaller is not acceptable. Likelihood ratios are typically given in (base 10) logarithms. Thus, a ratio of 100:1 is a log of 2, and 10:1 yields a log of 1. Following this convention, the log of the likelihood ratio (henceforth LLR) had to be 2 or larger to show that the logistic model fit better than the quadratic model.

When the deviation of the data from the true model is normally distributed, the LLR can be computed using the formula

$$[(\textit{Quadratic RMS Residuals})/(\textit{Logistic RMS Residuals})]^{\textit{number of observations}},$$

where *RMS Residuals* is the root mean square of the residuals, or the square root of the *MSE* of the regression. The residuals become smaller as goodness of fit increases. To test whether the logistic model fits better than the quadratic model, one puts the quadratic residuals in the numerator and the logistic residuals in the denominator. As the fit of the quadratic model improves, the residuals in the numerator will decrease, and the whole ratio will therefore get smaller. As the fit of the logistic model improves, the residuals in the denominator will decrease, and the whole ratio will instead get larger.

Results

Table 1 displays the inflection points, R^2 values, and LLRs for each of the 20 children. Nine of the 20 had an inflection point between 20 and 90 words. These 9 were therefore the only candidates for a vocabulary spurt. Six of these 9 children had a larger R^2 for the logistic model than for the quadratic model, and 4 of those had LLRs ≥ 2, indicating a better logistic fit. One additional child (037B) had an LLR that was between 1 and 2.

Table 1
Results of Experiment 1 Initial Analysis: Testing for a Spurt Between 20 and 90 Words

Subject	Inflection point[a]	Logistic R^2	Quadratic R^2	LLR[b]	Older sibling?
041B	−370	0.000	0.566		
025A	−189	0.000	0.563		
038A	−128	0.000	0.622		
239B	−127	0.000	0.512		
214A	−94	0.000	0.400		
067A	−60	0.000	0.138		
062A	12	0.145	0.148		
230B	18	0.257	0.448		yes
218A	22	0.219	0.234		
012A	25	0.328	0.337		
237A	26	0.823	0.781	≥ 2	
050A	29	0.299	0.201	≥ 2	
024A	31	0.612	0.620		
022A	33	0.783	0.755	≥ 2	
227A	44	0.545	0.508		
037B	55	0.393	0.380	≥ 1	yes
243A	56	0.888	0.831	≥ 2	yes
009A	317	0.692	0.716		yes
030A	338	0.000	0.684		yes
010A	554	0.465	0.528		

Note. LLR = log likelihood ratio.
[a] Inflection points less than zero or greater than 90 are not possible given the range of the data (0–90 words). In such cases, the program could not find a reasonable fit within the range of actual data. [b] LLRs < 1 are not given.

Discussion

Of the 20 children considered, 4 (perhaps 5, if one borderline case is included) showed evidence of a vocabulary spurt. This finding does not support the notion that the vocabulary spurt is a widespread or universal phenomenon. It is not even close to the 72% of children (13/18) that Goldfield and Reznick (1990), using a threshold approach, found to have a spurt.

There are several possible objections to this finding. One is that a number of studies have indicated that the vocabulary spurt is based predominantly on nouns (e.g., Gopnik & Meltzoff, 1987), and in fact the term *naming explosion* usually refers not to the vocabulary spurt as a whole but to a jump in noun learning alone. Similarly, Goldfield and Reznick (1990) found that spurters tended to have a preponderance of nouns in their vocabulary whereas nonspurters had a more balanced vocabulary composition.

This possibility can be addressed by comparing the proportions of nouns in the vocabularies of spurters and nonspurters. The 4 spurters had a mean proportion of .70 nouns in their vocabulary (SD = .10), whereas the 15 nonspurters had a mean proportion of .63 nouns (SD = .10). The difference was not significant, one-tailed $t(17) = 1.17$, $p = .13$, but the effect was in the right direction and, given the small number of subjects in the spurt group, should probably be considered borderline.

This finding provided the rationale for a second analysis, in which nouns alone were considered. From the sample of 20 children in the original analysis, we selected a subset with which to examine the hypothesis that the vocabulary spurt might involve only nouns. All words were first classified as nouns or non-nouns by an undergraduate research assistant and double-checked by one of the authors. Later, the process was repeated by a second research assistant (without double checking by the authors) to provide reliability statistics. All common nouns (count and mass) were included, as were proper names and brand names (e.g., Cheerios). Personal, demonstrative, and possessive pronouns (e.g., *he, this, mine,* respectively) were not included. In potentially ambiguous cases (e.g., *duck*) where the parent did not provide sufficient context to differentiate the part of speech, the word was classified following the rater's intuitions of which category the child was likely to have meant. Such decisions were based on the intuitive likelihood that a young child would say the word. In cases of uncertainty, frequency counts from Francis and Kucera (1982) were consulted, and the word was classified as the more frequent form. Both the initial classification of words into nouns and non-nouns and the resolution of difficult cases were done for a different study by raters with no knowledge of which children were classified as spurters in the present study. The percentage of agreement between the two raters was 91% (calculated by dividing the number of words both raters identified as nouns by the number of nouns identified by one rater or the other). The reliability coefficient was .82.

Children were used in the noun analysis if they had at least 90 common nouns after their last 2 weeks of data were removed and had no more than 20 common nouns reported in the 1st week of data collection. The 90-word minimum was chosen so the results would be comparable to those in the first analysis. Because excluding non-nouns reduced the total vocabulary considerably, these criteria left only 11 children with which to test the nouns-only spurt hypothesis (7 girls, 4 boys; 6 from identical twin pairs, 5 from fraternal twin pairs). The mean age at the onset of data collection for this subset was 16.0 months (median age = 15.0 months; *SD* = 4.0).

Table 2 displays the results. Five of the 11 children had noun data that were modeled by a logistic curve with an inflection point between 20 and 90 words. These 5 were therefore the only candidates for a vocabulary spurt based on nouns alone. Four of the 5 children with inflection points in the target range had a larger R^2 for the logistic model than for the quadratic model, but only 1 (Child 243A) had an LLR ≥ 2, indicating a better logistic fit. It should be noted that another child (Child 050A) had an inflection point of 19 and an LLR ≥ 2. If the inflection point search were expanded only slightly, this child would also be included. This result suggests that noun learning alone does not undergo a spurt and, likewise, that the vocabulary spurt was not being masked (in the original analysis) by non-nouns.

Why, then, have so many others concluded that noun learning is central to the vocabulary spurt? One possibility is that nouns lend themselves to being learned in bursts because groups of nouns often co-occur in the real world. A visit to the zoo, for example, may result in a dozen new animal kind names, but few or no new verbs, adjectives, or social words. If a threshold approach is used (requiring, for instance, 10 new words in 2.5 weeks), a trip to the zoo might make the child's vocabulary cross that threshold. However, just because the child can learn 10 new nouns at once does not necessarily mean he or she has undergone a significant cognitive change in his or her word-learning ability. Rather, the burst was driven completely by experiential factors (a trip to the zoo).

Although vocabulary composition does not appear to have driven the results, another possible objection to the original analysis (and to the second) might be that the cumulative vocabulary of the children has been cut off at 90 words. In fact, Mervis and Bertrand (1995), in a response to Goldfield and Reznick (1990), made exactly this argument, pointing out that Goldfield and Reznick's nonspurters only reached an average of 86 words during the study. Mervis and Bertrand presented evidence, using a threshold criterion, that some children show a spurt later in vocabulary learning, at an average of 112 words. To address this possibility, we present a final analysis that addresses the possibility that there is a spurt beyond the 90-word mark.

Table 2
Results of Experiment 1 Nouns-Only Analysis

Subject	Inflection point[a]	Logistic R^2	Quadratic R^2	LLR[b]
230A	12	0.197	0.582	
024A	16	0.106	0.108	
050A	19	0.392	0.310	≥ 2
214A	23	0.417	0.412	
037B	32	0.245	0.229	
243A	41	0.785	0.576	≥ 2
025A	54	0.860	0.855	
012A	62	0.375	0.376	
204A	321	0.379	0.384	
218A	435	0.436	0.497	
067A	1,116	0.086	0.413	

Note. LLR = log likelihood ratio.
[a] Inflection points less than zero or greater than 90 are not possible given the range of the data (0–90 words). In such cases, the program could not find a reasonable fit within the range of actual data. [b] LLRs < 1 are not given.

For this analysis, we reexamined the original set of 98 twins to find children who had at least 150 words in their lexicon before their last 2 weeks of data collection. They were included even if they had as many as 39 words in the first 2 weeks of data collection. These criteria left 18 children (10 girls, 8 boys; 9 identical twins, 9 fraternal twins). Their average starting age was 17.5 months (*SD* = 4.2). Three of these 18 children were also included in the original analysis, and another 7 are twins of children in that analysis.

Table 3 displays the results. In addition to the inflection points, R^2 values, and LLRs, words recorded in the first 2 weeks of data collection are included to give the reader an idea of whether an early spurt could have been missed. Six of the 18 children had vocabulary data that were modeled by a logistic curve with an inflection point between 20 and 150 words, but the highest inflection point was 91. All 6 had a larger R^2 for the logistic model than for the quadratic model, but only 3 of the 6 had LLRs > 2, indicating a better logistic fit. Two additional children had LLRs between 1 and 2. So, much as in our first two analyses, 3 of 18 (17%) had a spurt in this analysis of higher vocabulary levels (5 of 18, or 28%, if 2 borderline cases are included). Furthermore, because no child has a spurt after 91 words, it appears that restricting the search to the first 90 words in the original analysis did not hinder our ability to find a spurt.

A third objection to the study presented here might be that the participants are all twins. As noted in the *Participants* section, twins are more prone to language disorders and delays, partly because of prematurity and other perinatal factors. As also noted, there was a cutoff for gestational age and birthweight in this study to avoid such problems. Nonetheless, three additional problems related to the twin sample remain.

First, although some language problems in twins are due to biological factors, some are thought to be due to social factors (Mogford, 1993; Tomasello, Mannle, & Kruger, 1986). It is therefore possible that twins learn language at a slower rate than singletons, which could make them a noncomparable population and, more importantly, could bias against finding a spurt.

Second, Goldfield and Reznick (1990) reported that birth order was one major difference between their spurters and nonspurters, with spurters being more likely to be firstborn. Goldfield and Reznick argued that there is a meaningful relationship between birth order and noun learning, which is that parents with only one child have more time to play the "naming game"—the practice of incessantly labeling objects and encouraging the child to repeat and learn the labels. Poulin-Dubois et al.'s (1995) study, which reports a correlation between naming practices in parents and noun learning in children (as well as better categorization skills and earlier spurting), supports this conjecture.

As noted in the description of the participants, 5 of the 20 children in Experiment 1 had an older sibling. One of those 5 was a spurter and another was a borderline spurter (Child 037B), whereas the remaining 3 were not spurters (see Table 1), which suggests that birth order has little impact on the presence of a vocabulary spurt. However, because all of the children in this study are twins, they all have a same-age sibling. Therefore their parents presumably had less time than parents of singletons to play the naming game. Indeed, it is known that twins have less indi-

Table 3
Results of Experiment 1, Analysis 3: Testing for a Vocabulary Spurt After 100 Words

Subject	Inflection point[a]	Logistic R^2	Quadratic R^2	LLR[b]	Words reported in first 2 weeks
214A	−170	0.000	0.309		0
057A	−144	0.000	0.009		33
009B	−105	0.000	0.373		0
230A	−77	0.000	0.065		12
222B	−15	0.000	0.365		39
067A	3	0.009	0.142		0
050A	29	0.267	0.239	≥ 1	0
227B	36	0.481	0.436	≥ 2	0
239A	56	0.319	0.314		16
037A	62	0.263	0.231	≥ 1	0
025B	88	0.729	0.708	≥ 2	0
218B	91	0.426	0.372	≥ 2	15
235B	164	0.290	0.291		33
208B	574	0.449	0.438		0
001B	587	0.614	0.632		29
039A	590	0.433	0.501		39
225A	720	0.552	0.597		18
065B	1,121	0.052	0.061		30

Note. LLR = log likelihood ratio.
[a] Inflection points less than zero or greater than 90 are not possible given the range of the data (0–90 words). In such cases, the program could not find a reasonable fit within the range of actual data. [b] LLRs < 1 are not given.

vidual interaction with their mothers and fewer instances of joint attention (Tomasello, Mannle, & Kruger, 1986).

Third, the fact that the participants are twins complicates data recording. It is possible that parents of twins cannot accurately separate the expressive vocabulary of the two children when keeping written journals. Indeed, Reznick, Corley, and Robinson (1997) showed that parent ratings of expressive language in their 14-month-old twins resulted in much higher estimates of co-twin similarity than did objective observer ratings. Although the use of a daily journal in the present study (as opposed to a one-time retrospective checklist, as in Reznick et al.) should attenuate such an effect, it cannot be ruled out. If co-twins undergo a spurt at different times, or if one has a spurt and the other does not, then confusing the language of the two twins could obscure an actual spurt in one or both children. Thus the fact that the participants are twins could be responsible for the smaller number of spurters in this study relative to the number in Goldfield and Reznick (1990).

All three of these objections to using twins could be answered by applying the same technique to nontwins. This was the goal in Experiments 2 and 3. In Experiment 2, data from Goldfield and Reznick (1990) were used to test directly whether the fact that they found more spurters was due to fundamental differences in twins versus singletons or if it was, as we argue, due to our more refined decision procedure for identifying a spurt. In Experiment 3, the technique was applied to Dromi's (1987) data.

Experiment 2: Goldfield and Reznick (1990) Data

Method

Because we propose using Goldfield and Reznick's (1990) data as a test of whether our results have been biased by the use of twins, it is worth considering in some detail the characteristics of their participants and the procedures they used for collecting and compressing data.

Participants

The children are described in detail in Goldfield and Reznick (1990). The relevant details are that they were recruited through birth records and came from Caucasian, middle-class, English-speaking families. The sample consisted of 9 girls (5 firstborn, 4 later-born) and 9 boys (3 firstborn, 6 later-born). This sample is therefore quite similar to the twin sample described in Experiment 1.

Procedure for Collecting Data

The method of data collection in Goldfield and Reznick's (1990) study was also remarkably similar to that in Experiment 1. Children's word use was also assessed by mothers' diaries, and contact to discuss the records was also made with the mothers every few weeks (2.5 weeks in Goldfield & Reznick, 3–4 weeks in the present study). As described in our introduction, Goldfield and Reznick required that 10 new words be learned by a child in a 2.5-week interval in order for a spurt to be attributed to that child. This difference necessitated minor changes to our analytic procedure, which are described in the next section.

One other procedural difference was that in compressing their data, Goldfield and Reznick (1990) attributed a word to a child only if that word had been reported by the mother at least two times. In Experiment 1, words were attributed after just one parent report. It is possible (though far from certain) that this difference could make the children in Goldfield and Reznick's study less likely to show a spurt, because each child's vocabulary would undoubtedly be increased by the inclusion of words he or she said only once. If this is true, then Goldfield and Reznick's data would not be a fair test of our new technique. However, there is no reason to think, a priori, that reporting more words across the board would make a child more likely to show a spurt—this would only be true if the extra words all occurred at the same time. If they are spread across the reporting period, or if they increase steadily during the reporting period (as one might expect, because a child might use each word less often as his or her total vocabulary increased), these one-time words would not increase the likelihood of showing a spurt.

Procedure for Reanalyzing Goldfield and Reznick's (1990) Data

For each of the children reported by Goldfield and Reznick (1990), one of the authors and one undergraduate research assistant independently estimated the data from the graphs provided by Goldfield and Reznick by enlarging the graphs, affixing them to grid paper, and taping them up to a window for light. Using this method, data could be estimated for 17 of the 18 children. Estimates made by the author and those made by the assistant were highly correlated for each subject. The average Pearson *r* was .998, and the effective or composite (Spearman-Brown) reliability was 1.00 for all subjects except one, who had a composite reliability of .985.

Because the number of words learned in 2½ weeks is much higher than that learned in a day, all of the raw numbers were considerably higher for these children, and some adjustments to our analytic procedure had to be made to compensate. Specifically, all the rates were divided by 10 to bring them down to the same order of magnitude. The data from Subject 8 had to be divided by 100 instead of 10 for the program to converge. Because the numbers were still larger than those in our Experiment 1, the starting values of the parameters given to SPSS also had to be adjusted. We used the following values: *inflection point* = 5, *asymptote* = 1, and *slope* = −6. All resulting parameter estimates then had to be adjusted accordingly (i.e., multiplied by 10 for all subjects except Subject 8, whose final parameter values were multiplied by 100). These changes had no effect on the shape of the curve or the estimated location of the inflection point; they just put the data onto the same scale as the data in Experiment 1. Aside from this minor change, the same procedure outlined in Experiment 1 was followed exactly.

Results

Table 4 displays, for each of the 17 children from Goldfield and Reznick's (1990) study, inflection points, R^2 values, and LLRs, as well as Goldfield and Reznick's classification (spurt or no spurt). The subject numbering follows that given by Goldfield and Reznick.

Thirteen of the 17 children had data that were modeled by a logistic curve with an inflection point between 20 and 90 words. These 13 were therefore candidates for a vocabulary spurt. Four of these 13 had a larger R^2 for the logistic model than for the quadratic model, but only 1 of them (Subject 2) had an LLR > 2, indicating a better logistic fit. One other child (Subject 5) had an LLR between 1 and 2.

Discussion

Even including the borderline case, there are only 2 spurters (12%) among Goldfield and Reznick's (1990) 17 subjects. Both were considered spurters by Goldfield and Reznick, but so were 11 others. We therefore conclude that it is a difference in procedure, not in sample characteristics (twin vs. singleton, birth order, etc.) or data collection methods, that led to our finding a relatively small number of children with a vocabulary spurt in Experiment 1.

Table 4
Results of Experiment 2: Goldfield and Reznick's (1990) Data

Subject	Inflection point[a]	Logistic R^2	Quadratic R^2	LLR[b]	Goldfield & Reznick spurt?
6	−100	0.000	0.616		yes
1	−6	0.083	0.086	(≥ 2)	
22	19	0.865	0.714		yes
7	25	0.591	0.475		
5	26	0.995	0.985	≥ 1	yes
12	33	0.521	0.568		yes
2	44	0.999	0.980	≥ 2	yes
10	45	0.462	0.453		
13	45	0.375	0.562		
3	50	0.994	0.998		yes
15	50	0.954	0.950		yes
17	50	0.988	0.975		yes
4	53	0.993	0.999		yes
11	60	0.988	0.995		yes
9	61	0.947	0.936		yes
24	72	0.783	0.789		yes
8	170	0.628	1.000	—	yes

Note. Dash indicates that the log likelihood ratio (LLR) could not be computed because there were too few observations.
[a] Inflection points less than zero or greater than 90 are not possible given the range of the data (0–90 words). In such cases, the program could not find a reasonable fit within the range of actual data. [b] LLRs < 1 are not given.

Experiment 3: Dromi's (1987) Data

Method

Participants

In a final test of the method presented here, we used Dromi's 1987 data from her daughter Keren. The details of data collection as well as Keren's vital statistics are given in Dromi (1987).

Procedure

The data were taken from tables given in Dromi (1987), and the statistical procedure followed was the same as that outlined in Experiment 1.

Results

Logistic model fitting identified an inflection point for Keren at 95 words. The logistic $R^2 = .732$, the quadratic $R^2 = .724$, and the LLR $= .095$.

Discussion

Although Keren's data were fit well by the logistic function, they were fit just as well by a function without an inflection point, the quadratic function. Given the similarity of the two fits, we cannot conclude that Keren had a spurt.

General Discussion

We have argued that there has been a fundamental flaw with measures of the vocabulary spurt used in the past: Even the most sophisticated of them, the threshold approach, required only that a child's rate of word learning cross an arbitrary threshold, not that it exhibit an inflection point. Instead of using a threshold, we examined vocabulary rate data directly for an inflection point. The presence of an inflection point in the learning-rate curve is an inevitable consequence of a transition from a stage characterized by a low rate to one characterized by a higher rate. Although such methods had been suggested previously, this report represents the first attempt, to our knowledge, to apply them to longitudinal data from individual children. Our hope is that others will adopt this method and apply it to more data so that the field can ultimately provide an accurate estimate of how many children undergo a vocabulary spurt.

In the present report, the result was that only a minority of children had a reliable inflection point in their vocabulary rate functions. First, we considered data from a new sample of 20 children, using rate of word learning as the dependent measure and cumulative vocabulary size as the independent variable. From that sample, there were only 4 or perhaps 5 children with good enough logistic fits to indicate a possible spurt. We then reanalyzed the data in order to test whether children learning nouns alone showed a spurt, and we found that 1 or perhaps 2 out of 11 children showed a spurt under those conditions. We then expanded the inflection point search to 150 words and found that 3 or perhaps 5 of 18 children showed a spurt, none later than 91 words. We then reanalyzed data from Goldfield and Reznick (1990) and found that 1 or perhaps 2 of their 17 children showed a spurt. Finally, Dromi's 1987 data from her daughter Keren were subjected to our method and revealed no evidence for a spurt. If we take the results from the first analysis of Experiment 1 as representative of our own data and combine them with our reanalysis of Goldfield and Reznick's participants in Experiment 2 and of Keren in Experiment 3 to yield 38 children, we find that 5 (perhaps 7) of the 38 children, or 13% (perhaps 18%), had a vocabulary spurt. Readers must decide for themselves whether they wish to follow our strictest or loosest criteria or some compromise between them in order to settle on a point estimate.

Although these findings indicate that a fair number of children have a vocabulary spurt, the findings do not support the notion that it is a universal rite of passage in language development. This outcome means that we must rethink whether the vocabulary spurt reflects an important change in cognitive development.

Theories that posit a fundamental change in the way children use words appear no longer to have support. For instance, the insight that words refer to things or that all things have names (e.g., Dore, 1978; Reznick & Goldfield, 1992) was proposed specifically to account for the vocabulary spurt. If most children do not have a spurt, there is little remaining support for the naming-insight class of hypotheses.

Claims that children's object concepts are fundamentally restructured would also lose support if most children do not undergo a vocabulary spurt. Although Piaget's theory is not directly affected, the notion that advances in object knowledge during Sensorimotor Substage 6 spurred a spurt (Lifter & Bloom, 1989) is no longer necessary. Likewise, Gopnik and Meltzoff's (1987) emphasis on sorting objects as a sign of basic-level knowledge that enables fast word learning now seems misplaced; in most children, there is no evidence for an onset of faster word learning that needs to be explained

Likewise for many other claims described in the introduction. Discontinuous changes in word representation (Walley, 1993), word segmentation (Plunkett, 1993), pragmatics (Ninio, 1995), and word retrieval (Dapretto & Bjork, 2000) are no longer necessary as driving forces behind the vocabulary spurt.

On the other hand, it is possible that a continuous change in word learning causes a discontinuity in other domains. For instance, word representation changes (Walley, 1993) and naming errors (Gershkoff-Stowe & Smith, 1997) have both been reported at around 50 words. Although both are supposed to be related to the vocabulary spurt, it is possible that these changes actually occur because of some reorganization of the mental lexicon that takes place when it exceeds a certain size in order to accommodate the phonological and semantic properties of more words and their relationships. For instance, a child with a lexicon containing only one word starting with /m/, *mommy*, can be rather lazy about how to represent and produce its phonological form (e.g., "muh"). Once more words that begin with /m/ are added (e.g., *milk* and *more*), phonological representations must become more complex in order to discriminate the words. The same could be true of semantic representations, causing naming errors as the child attempts to converge on more exact meanings. All of these could be true without the presence of a spurt; a reorganization of the mental lexicon may not increase the rate of word learning—it may even be necessary just to maintain the rate.

But if there is no vocabulary spurt, we still have the task of explaining why the rate of vocabulary learning increases. One explanation is that instead of a single dramatic cognitive change, there are many small, unsynchronized changes in both higher cognitive and lower level processing abilities. These would lead to what appear to be steady increases in word-learning ability.

P. Bloom, in his 2000 book *How Children Learn the Meanings of Words*, wrote:

> It would not be difficult to see if it [a spurt] did occur. One could graph the child's vocabulary growth and look for a dramatic (or at least statistically significant) change in the slope of the line denoting rate of growth. (Note that the graph has to be of an individual child, not aggregate data, and has to be of rate, not size.) This is a simple criterion, but, as far as I know, no evidence is available that any child has ever met it. (p. 43)

The evidence is now in. We have given a detailed, quantitative method for detecting a vocabulary spurt in individual children and applied it for the first time. A minority of children—about 1 in 5—do have a spurt, but most do not.

References

Aitchinson, J. (1994). *Words in the mind: An introduction to the mental lexicon* (2nd ed.). Oxford, England: Blackwell.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Communication and cognition in infancy.* New York: Academic Press.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar.* Cambridge, England: Cambridge University Press.

Bates, E., & Carnevale, G. (1993). New directions in research on language development. *Developmental Review, 13,* 436–470.

Berk, L. (2003). *Child development* (6th ed.). Boston, MA: Allyn & Bacon.

Bloom, L. (1973). *One word at a time: The use of single-word utterances before syntax.* The Hague: Mouton.

Bloom, L., & Capatides, J. B. (1987). Expression of affect and the emergence of language. *Child Development, 58,* 1513–1522.

Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22,* 497–529.

Corrigan, R. (1978). Language development as related to Stage 6 object permanence development. *Journal of Child Language, 5,* 173–189.

Dapretto, M., & Bjork, E. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development, 71,* 635–648.

Dore, J. (1978). Conditions for the acquisition of speech acts. In I. Markova (Ed.), *The social context of language* (pp. 87–111). Chichester, England: Wiley.

Dore, J., Franklin, M. B., Miller, R. T., & Ramer, A. L. H. (1976). Transitional phenomena in early language acquisition. *Journal of Child Language, 3,* 13–28.

Dromi, E. (1987). *Early lexical development.* Cambridge, England: Cambridge University Press.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness.* Cambridge, MA: MIT Press.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., & Thal, D. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(Serial no. 242).

Fischer, K. W., Pipp, S. L., & Bullock, D. (1984). Detecting discontinuities in development: Method and measurement. In R. Emde & R. Harmon (Eds.), *Continuities and discontinuities in development* (pp. 95–121). New York: Plenum Press.

Francis, N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Ganger, J., Pinker, S., Chawla, S., & Baker, A. (2000, July). *The heritability of early language milestones of vocabulary and grammar: A twin study.* Paper presented at the Conference of the International Society of Infant Studies, Brighton, England.

Gershkoff-Stowe, L., & Smith, L. (1997). A curvilinear trend in naming errors as a function of early vocabulary growth. *Cognitive Psychology, 34,* 37–71.

Gershkoff-Stowe, L., Thal, D., Smith, L., & Namy, L. (1997). Categorization and its developmental relation to early language. *Child Development, 68,* 843–859.

Gillis, S., & De Schutter, G. (1986). Transitional phenomena revisited: Insights into the nominal insight. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech: Proceedings of an international symposium held at the Wenner-Gren Center, Stockholm, September 19–22, 1984.* New York: Stockton Press.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the spurt. *Journal of Child Language, 17,* 171–183.

Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relations to other cognitive and linguistic developments. *Child Development, 58,* 1523–1531.

Hay, D. A., Prior, M., Collett, S., & Williams, M. (1987). Speech and language development in preschool twins. *Acta Geneticae Medicae et Gemellologiae, 26,* 213–223.

Kamhi, A. G. (1986). The elusive first word: The importance of the naming insight for the development of referential speech. *Journal of Child Language, 13,* 155–161.

Lifter, K., & Bloom, L. (1989). Object knowledge and the emergence of language. *Infant Behavior & Development, 12,* 395–423.

Lock, A. (1978). The emergence of language. In A. Lock (Ed.), *Action, gesture, and symbol: The emergence of language* (pp. 3–18). London: Academic Press.

McCarthy, D. (1954). Language development in children. In L. Carmichael (Ed.), *Manual of child development* (pp. 492–630). New York: Wiley.

McShane, J. (1980). *Learning to talk.* Cambridge, England: Cambridge University Press.

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name-nameless category (N3C) principle. *Child Development, 65,* 1646–1662.

Mervis, C. B., & Bertrand, J. (1995). Early lexical acquisition and the vocabulary spurt: A response to Goldfield and Reznick. *Journal of Child Language, 22,* 461–468.

Mittler, P. (1970). Biological and social aspects of language in twins. *Developmental Medicine and Child Neurology, 12,* 741–757.

Mogford, K. (1993). Language development in twins. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances* (pp. 80–95). Hove, UK: Erlbaum.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38* (Serial No. 149).

Ninio, A. (1995). Expression of communicative intents in the single-word period and the vocabulary spurt. In K. E. Nelson & Z. Reger (Eds.), *Children's language* (Vol. 8, pp. 103–124). Hillsdale, NJ: Erlbaum.

Pine, J. M., & Lieven, E. V. M. (1990). Referential style at thirteen months: Why age-defined cross-sectional measures are inappropriate for the study of strategy differences in early language development. *Journal of Child Language, 17,* 625–631.

Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language, 20,* 43–60.

Poulin-Dubois, D., Graham, S., & Sippola, L. (1995). Early lexical development: The contribution of parental labeling and infants' categorization abilities. *Journal of Child Language, 22,* 325–343.

Reznick, J. S., Corley, R., & Robinson, J. (1997). A longitudinal twin study of intelligence in the second year. *Monographs of the Society for Research in Child Development, 62*(1, Serial No. 249).

Reznick, J. S., & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology, 28,* 406–413.

Reznick, J. S., & Goldfield, B. A. (1994). Diary vs. representative checklist assessment of productive vocabulary. *Journal of Child Language, 21,* 465–472.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69,* 309–320.

Swingley, D., & Aslin, R. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science, 13,* 480–484.

Tomasello, M., Mannle, S., & Kruger, A. (1986). Linguistic environment of 1- to 2-year-old twins. *Developmental Psychology, 22,* 169–176.

van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review, 98,* 3–53.

Walley, A. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13,* 286–350.

Woodward, A., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology, 30,* 553–566.

Received July 16, 2003
Revision received February 10, 2004
Accepted February 26, 2004 ■