FROM BABBLING TO SPEECH: A RE-ASSESSMENT OF THE CONTINUITY ISSUE

Marilyn May Vihman, Marlys A. Macken, Ruth Miller
Hazel Simmons, Jim Miller

Stanford University

Controversy exists over whether there is any connection between children's babbling and the development of the adult sound system. The classic proponent of the discontinuity school is Jakobson 1941/1968, who claimed that the pairing of sound and meaning drastically alters the child's sound system. Jakobson's arguments for discontinuity are here evaluated on the basis of data on the transition from babbling to speech in a single set of children recorded weekly in two contexts: mother–child interaction and solitary play. Using the data from the mother–child context, and comparing the sound system of babbling with that of the early words in terms of the distribution of consonants, vocalization length, and phonotactic structure, we find striking parallelism between babbling and words within each child, across time and within time period. The data constitute strong evidence for continuity.*

What, if any, is the relationship between children's babbling and their early words? Are they 'two concurrent yet quite distinct forms of activity', or even 'two temporally delimited stages', as Jakobson & Waugh put it (1979:62)? Or do the early words not only follow, but evolve from, the babbled vocalizations? Must we sharply contrast 'the variety and opulence of the babbled sounds' and 'a rigorous sparseness of speech sounds' (to quote Jakobson & Waugh again)? Or is the transition a blurred one? These opposing interpretations of the transition from babbling to speech have been debated for some time. Jakobson's discontinuity hypothesis has, of course, long been known; the influence of his 1941 monograph, first published in English in 1968, can scarcely be overestimated. Textbooks and introductory courses in linguistics and psychology persistently cite Jakobson's time-honored and prestigious views. Quite possibly, a tendency to accept a sharp distinction between prelinguistic vocalization and speech per se reflects an unconscious bias toward maximum differentiation between human language, with its characteristically linguistic structure, and the non-language behavior of humans and non-humans. However, a growing literature also exists on the nature of babbling and early words which points increasingly to a rejection of the discontinuity view. Jakobson's remarks, based on the rather impressionistic diary literature of his time, were partially speculative, and conceived as just one aspect of a broader theoretical framework. The value of that framework is unquestioned; like other strong conceptual formulations, it has provided an enduring stimulus for research and

* A preliminary version of this paper was presented at the 1981 LSA Annual Meeting in New York. We would like to thank Alan Cruttenden, Charles A. Ferguson, John H. V. Gilbert, Jeri Jaeger, Lise Menn, and John Rickford for their helpful comments. We are also grateful to Lincoln Moses for suggesting the distance-function permutation test. Above all, we thank the mothers who participated in our project for their unfailing patience and enthusiasm—and, of course, the children, for the pleasure of watching their talents unfold. Our work was carried out under a grant from the National Science Foundation (BNS 7924167).

debate. We believe, however, that the data are by now sufficiently clear to allow definitive answers to the questions raised above. It will be our purpose here to re-examine the two sharply contrasting views on the relation between language and non-language by analysing a corpus of babbling and early word vocalizations collected from nine babies as part of a longitudinal study of this important transition period.

What are the specific claims of the discontinuity position? Jakobson maintained that the acquisition of the sound system proper begins when the child recognizes that sounds have distinctive linguistic value (1968:24–5): 'As soon as sound utterances "are employed for the purpose of designation" [Wundt 1912:283], the actual stage of language formation is launched.' According to Jakobson, this pairing of sound and meaning drastically alters the child's sound system. Though a short so-called silent period 'may intervene between the stage of spontaneous babbling and that of true language development', this is not crucial to the theory: 'For the most part ... one stage merges unobtrusively into the other, so that the acquisition of vocabulary and the disappearance of the pre-language inventory occur concurrently' (29). The crucial differences between 'the "permanent" speech sounds' and 'the disappearing babbling sounds' include the 'phonetic abundance of babbling' vs. the 'phonemic poverty of the first linguistic stages' (25); a lack of systematicity in the sequence of emergence or use of babbling sounds vs. the universal acquisition order posited for the development of word-based phonology; and an increasing restriction of babbling to solitary play, waking, and going to sleep, while language alone is used in social settings.

McNeill 1970 elaborates slightly on Jakobson's views, basing his arguments on Bever's 1961 re-analysis of data collected by Irwin 1947a,b. McNeill concludes (132) that 'there ... is a sharp discontinuity at both ends of the babbling period'. But around the same time, several studies based on new data begin to challenge Jakobson's hypothesis. First, Cruttenden 1970 compares babbling and speech sounds in the development of his twin daughters. With regard to babbling in the course of the transition period, he finds that 'the child's repertoire is shifted in the general direction of the language to be learnt' (114). In a more broadly based study, Oller et al. 1976 compare phonetic tendencies in the babbling of ten 6- to 13-month-old infants with reports in the literature about the phonetic tendencies of early speech. In both data sets, they find a preference for open syllables, initial stops, final fricatives etc. Similarly, Stark 1980, in her review of the phonetics of infant vocalizations in the first year of life, traces a gradual evolution toward language in the course of the period with which we are concerned. Finally, de Boysson-Bardies et al. 1981 are able to isolate, in the babbling of a French child just beginning to produce recognizable words at 18–20 months, features such as a relatively high frequency of fricatives and nasal vowels, which are specifically characteristic of adult French speech. In another study, de Boysson-Bardies et al. 1984 find that French adults are able to distinguish the babbling of non-French babies from that of French babies—again suggesting a shift in babbling toward the sound shapes of the specific language the child is hearing (but cf. Atkinson et al. 1970 for conflicting data).

However, no one so far has directly studied in detail the phonetic characteristics of both the babbling and the early words of a single set of infants—attempting, as MacNeilage suggests, 'to predict from the babbling of an individual child ... what his/her first words will be like' (1979:30). Stoel-Gammon & Cooper 1984 take a first step in that direction: They report three children's inventories and frequency of use of consonants and phonotactic shapes in the babbling of the last month preceding the onset of meaningful speech, as compared with those found in the first fifty words for each child. The three subjects prove quite similar in the basic syllabic and consonantal types used; also, the same general patterns are found in both babbling and speech (more stops than fricatives, more front consonants than back ones, more CV than CVC syllables). But in the absence of data on the babbling which occurred concurrently with the early words, they can draw no conclusions as to 'individual cases, e.g., [whether] the child who uses velars or liquids in early words produced those sounds in his or her non-meaningful vocalizations' (21).

One reason, perhaps, for the lack of a stringent full-scale test of Jakobson's position up until now is that his provocative but somewhat loose discussion leaves the way open for many possible interpretations. In particular, he fails to provide clear guidelines for distinguishing 'speech' from 'babble'. As criteria for recognizing 'embryo-words' which include 'the "permanent" speech sounds', he suggests (29) 'the persistence of the sound, the intention to express meaning by the formation in which it occurs, and the social setting of the utterance.' But frequent or repeated use of a sound is clearly not sufficient in itself for inclusion in the 'permanent speech sound' category—since, as Jakobson specifically mentions, 'the child continually repeats [certain] sounds during the babbling period' (22). Furthermore, 'exclamations and onomatopoetic formations' are said to include 'even in the language of the adult ... sounds which are inadmissible in a given language' (25–6). Yet one cannot mechanically exclude onomatopoetic forms from the 'embryo-words': as Jakobson also notes, they may be transformed into 'a simple and conventional designation', with or without adaptation of the sounds in question to the 'phonemic system of the child' (27). In another passage, Jakobson seems to suggest that one omit from the meaningful speech category those sounds 'used by the child for imitations and for chasing chickens or dogs' (30); and he further implies that occasional phonetic variants should be disregarded: 'Instead of the more usual *papa, pafa* was occasionally pronounced. If all the sound productions of the child are tossed into the same heap, it is understandable that the laws of development cannot be disclosed' (31).

As we shall see in the extended discussion below of the difficulties we encountered in our own analysis, the distinction between word and non-word can be extremely hard to draw during the transition period (cf. also Labov & Labov 1978). Thus Jakobson's 'intention to express meaning' may be clearly present in some uses of a given sound shape, but not in others. Social setting does not guarantee ease of definition either. Thus some babies are found to 'practice', or repeat over and over, newly acquired adult-based words while alone in their cribs; others continue to make heavy use of apparently non-meaningful babble in an interactional setting, even after they have used well over fifty different adult-based word types. Jakobson's remarks seem to suggest that one look only at repeated use of given sound–meaning pairings in the interactional setting, but exclude onomatopoetic forms (unless they have become the conventional designation for some referent) as well as imitation and 'rare' phonetic variants. Thus defined, the criteria for imputing word status seem so delicate as to defy objective, replicable application. In carrying out the present analysis, we have set ourselves a somewhat more manageable task, choosing to distinguish only the adult-based words from the non-adult-based vocalizations (= 'non-words' or babble). We describe below the details of our application of this distinction to the sometimes diffuse data that we encountered. We are convinced that it is in keeping with the spirit of Jakobson's position to define 'language' simply in terms of adult-based word use, even if the child's use is sometimes remote from that of the adult—and even if non-words sometimes include primitive sound–meaning pairings of the child's own creation (e.g. [ʔəh] 'I want something!', [huh] 'Look at that!') The essential issue here is the relation between babbling and speech: Do the patterns which a child uses in babble, before and during the early use of words, differ dramatically from the patterns found in those words, as Jakobson claims? Or do the two sets of patterns bear a close resemblance to one another? If the latter is true, then it can no longer be maintained that babble is unrelated to speech, though the precise nature of the relationship remains to be determined; and in this respect, at least, human language must be seen as structurally related to non-linguistic behavior.

METHODOLOGY

1. The present analysis is based on data collected from nine first-born children out of the ten that we observed over a seven-month period, beginning when each child was about nine months old. The tenth child, a girl, never spontaneously used more than four words in any given session; thus too few words were recorded to allow statistical comparison with her babbled vocalizations (the child's cumulative recorded lexicon at the close of the study included 16 different words). Audiometric screening showed that all the subjects had normal hearing.

Subjects (referred to below by pseudonyms) were located primarily through infant care classes. A primary criterion for selection was maternal enthusiasm for the project—since the mothers were to maintain a daily log, recording changes in the child's repertoire of communicative gestures and vocalizations, in addition to submitting to a long period of weekly recordings. The data were collected in two contexts, both in the child's home: a mother/child interaction setting (one-half hour weekly, both audio and video recording) and a solitary play setting (10 to 15 minutes, weekly insofar as possible, audio recording only).[1]

[1] A Uher 4000 tape recorder with a Sony Electret microphone was used for the audio recording; for video recording, we used a Sony videorecorder AV-3400 with a Sony AVC-3450 video camera, connected to a second Sony Electret microphone. The two microphones were inserted into pockets on the inside of a soft cloth vest which the children wore during the interactional session; the vest was hung near the child during solitary play.

Seven anchor points were chosen for analysis for each child, spaced at about one-month intervals. The choice of anchor-point sessions was based in part on the quantity and recording quality of data available: weeks in which we had not been able to record the child alone were avoided, if possible, as were sessions in which technical failures occurred. Beyond that, we did not use the once-monthly sessions in which the Užgiris–Hunt (1975) scales for assessment of cognitive development in infancy were administered, since we wished to analyse naturalistic unstructured data.

The first four authors prepared detailed phonetic transcriptions of the sessions, based on the IPA, supplemented by a symbology and a set of diacritic notations developed for use in transcribing child speech (cf. Bush et al. 1973) as well as by further symbols and diacritics developed in the course of this study for special use in transcribing the vocalizations of 9- to 16-month-old infants. These specially designed notational conventions were intended to moderate some of the biasing effects of an adult-language-based system of segmental transcription on the transcribing of data from children who are still developing articulatory control (cf. Oller et al. 1976, Stockman et al. 1981, Kent & Murray 1982).

All four transcribers participated in phonetic workshops and in transcription exercises, as well as in the preparation of the audio transcripts used in this study. Problem cases were discussed with a second transcriber, and were resolved by mutual agreement. To test reliability in the transcription of babbling and early words, all four authors transcribed approximately two-minute-long sections of tapes from the middle of one session for each of seven babies. The children ranged in age from 0;9.9 to 1;5.0 on these samples, and averaged 16 vocalizations each. The two authors who, between them, had transcribed 82% of the tapes used in this study then prepared a 'standard' version by retranscribing and comparing until they reached full agreement. Only the categories used in the study—place and manner in consonants, syllable shape, and vocalization length in syllables—were considered. We assigned two possible points to each stop, fricative, affricate, or nasal (to allow for agreement in place and in manner); one point to each remaining consonant; and one point to each vowel place, as a marker of the syllable count. The number of possible points for each subject's vocalizations was calculated from the standard; this number was multiplied by four and compared with the total for all transcribers (points were subtracted both for added and for omitted or substituted segments, according to the point system described above), yielding 85.8% agreement on the relevant categories (n = 126 vocalizations and 2452 possible points). Direct comparison of the original transcripts showed 82.3% agreement, using the same point system. The category of glottals, primarily in vocalization-initial and -final position, was the source of about one-third of the disagreement (33.8% for all four transcribers against the standard, 31.1% for the main transcribers); however, this category is relevant to only one of the parameters included in our quantitative and statistical analyses, viz. consonantal manner. If the differences involving glottals are disregarded, reliability reaches 90.7% for all four transcribers, 87.8% for the main transcribers.

The unit we chose to compare across babble and speech was the VOCALIZATION, based on breath group, timing, and presence or absence of a unifying intonation contour. (Reliability was 96.9% for numbers of vocalizations included in the sample transcript.) A rapid succession of word-based vocalizations united by a single intonation curve was treated as one unit—e.g. *duck-duck-duck-duck-duck*—though our awareness of the identity of the repeated word would have made division into separate units possible. Similarly, the few two-word strings that occurred for some of the children (under a single intonation curve)—e.g. *hi doggy* or *uh-oh chicken* (said as a puzzle-piece hen fell from the baby's hand)—were each treated as a single (adult-word-based) vocalization. The intent in all cases was to deal with babble and with word-based vocalizations in as comparable and consistent a manner as possible, to avoid introducing artefactual biases.

Once audio transcripts were complete, the corresponding video tapes were viewed in order to allow use of contextual cues before characterizing each vocalization as word or non-word. The chief criteria for according word status to a vocalization were consistent use in one or more contexts (including imitation) and adult-based sound shape (or range of shapes). Once an exhaustive division into words and non-words had been arrived at for each anchor-point session, counts were made of the distribution of syllable-length types, phonotactic structures, and consonantal place and manner categories.

WHAT IS A 'WORD'?

2. The critical problem was to decide, for each vocalization, whether it was or was not an adult-based word. As Oller has noted (1980:93), 'it is hard to tell, at times, whether infants are producing words or merely babbling' (cf. also Cruttenden's remark [114], 'I found it extremely difficult and perhaps rather pointless to pin down the occurrence of the first word.') Since the distinction between words and 'mere babbling' is at the heart of our analysis, we now discuss at some length the principles used in making that distinction, some of the specific problems faced in following those principles, and the reasoning applied in resolving some of the problems. The main analytic dilemmas we encountered were the following: the occurrence as babble of shapes based on the phonetic form of an adult word; the loss of adult-based meaning from what began as an adult-based sound shape; extensive homophony in the word and non-word vocalizations of a child with a highly limited repertoire of sound shapes; and the inclusion of adult-based words in strings of what was otherwise babble. By expanding now on issues associated with using 'word' as a crucial linguistic unit for this developmental period, we hope to clarify our method and also to contribute to a richer understanding of development in the transition period.

We take a word to be a conventionalized sound–meaning correspondence. We do not require that words conform to the core phonological system of the adult language (pace Jakobson), so long as they are used by the adult with a consistent conventional meaning. Thus we counted as words *uh-oh*, various forms of *yum* (such as *mmmm,* usually with a stereotyped intonation pattern), and *vroom-vroom* for the sound of cars; these were found in the vocabularies of several of our children (and their parents). Onomatopoeic renditions of animal sounds, which are great favorites with some babies but are ignored by others, vary quite a bit from one family to the next. We accepted family-specific words of this kind also, so long as they did not originate with the child. (For our treatment of words invented by the child, see the discussion of quasi- and proto-words, below.)

In the word category we included imitated as well as spontaneous vocalizations, the intent being to compare ALL adult-language-based vocalizations, or 'words', with all other vocalizations. We coded word-based vocalizations as spontaneous, repeated, imitated, or modeled; the last is a slightly delayed imitation, with possible adult (but not child) vocalizations intervening between the modeled word and the child's imitation. Repetitions of an original imitation counted as imitations, unless one or more spontaneous uses of the word preceded the adult model. Later spontaneous uses of an earlier imitated word counted as spontaneous.

Before we credited a child with a spontaneous use of a word, we required that he or she produce a phonetic form that was a recognizable attempt at the adult word, given frequent child-reduction rules (cf. Ingram 1974, 1979). In addition, the child had to use the word appropriately, with an apparently intentional meaning that was plausible in terms of the adult meaning or use of the word and commonly occurring child-semantic rules, such as over-extension of *doggy* to cats and other animals (cf. Clark 1973).

Once we had determined that a given phonetic shape represented a word used SPONTANEOUSLY during a given session, we took an important further step: We extended the classification 'word' to other apparent renditions of the same vocalization, i.e. to other highly similar phonetic shapes that could not be confidently classified as non-words on other grounds. This is our EXTENSION CONVENTION. For example, at 10 months, during Time 1 (the session that served as our first anchor point), the subject Deborah confounded both mother and observer by slipping back and forth between apparently appropriate spontaneous uses of *baby*—looking at a picture of a baby or at a mirror, reaching out to touch a baby doll—and apparently random or babbled use of the same disyllabic shape ⟨pebi⟩,[2] with no potential context. In such a case, there is no way to rule out the possibility that the child was THINKING about an appropriate context—or, for that matter, recognizing, rehearsing, or simply savoring the sound shape itself—even though there was nothing in her overt behavior to reveal it, except for the verbal act. There might have been other solutions to the problem which homophony posed for us; for present purposes, however,

[2] The notation ⟨ ⟩, taken from Scollon (1976:225) is used to indicate a 'typical' or 'standard' shape (or 'most general ... neither phonetic nor phonemic') for a given child vocalization, based on the range of recorded variants. In general, in the examples in the text, phonetic variation typical of that example type will not be represented in the phonetic shapes cited; those shapes are intended chiefly as 'labels' or names covering a range of shapes, and are used here only for clarity of reference.

the effect (if any) of our convention is certainly in the direction of making the argument for continuity HARDER to sustain rather than the reverse, since it removes just those non-word-like uses of highly word-like phonetic shapes from the category of babble.

An account of the evolution of one word, and our treatment of it, may clarify our approach. Beginning at Time 2, when he was nearly 11 months old, the subject Jonah began to use vocalizations of roughly the form ⟨dádà⟩, always with a distinctive high–low pitch sequence, in connection with the family dog, Edgar. Specifically, whenever anyone mentioned Edgar, or when hearing a dog bark outside during solitary play, or when standing at a window looking into the yard where Edgar was customarily found, Jonah would respond with a series of repetitions of his ⟨dádà⟩ vocalization, and would also show sharply focused attention to the dog's possible whereabouts. In addition to some nine such instances of evident association between Jonah's use of ⟨dádà⟩ and the name *Edgar*, as well as perhaps the word *doggy*, the same range of phonetic shapes occurred some twelve times outside these contexts—almost always with the same intonation contour, generally as an apparent expression of excitement or delight. Though Jonah had used two other words at Time 1, *Edgar* (or *doggy*) was the only word used at Time 2. Applying our extension convention, we counted all uses of ⟨dádà⟩ at Time 2 as words, regardless of context. A month later, during Time 3, the ⟨dádà⟩ vocalization scarcely occurred, though *kitty* was imitated twice as something like [títə̀], with a high–low tune. At Time 4, Jonah said [títò] as he pointed at birds passing outdoors. He shouted to Edgar, now pronouncing the name [é:gʌ̀]. And after spending several minutes trying to stand a photograph of Edgar against the window early in the session, he later returned to the photo with a series of six ⟨dádà⟩ vocalizations. Approximately the same sound pattern occurred after his mother dangled a kitty toy near him and said *kitty*, and also as she built towers with blocks while Jonah watched. Thus the general earlier meaning of excitement or delight still seemed to be present; but the word *Edgar* had branched off with a new sound shape of its own, and the connection between ⟨dádà⟩ and its presumed source word had faded. But should ⟨dádà⟩ now count as adult-word-based? We concluded that it should, even where only general excitement was being expressed, since the probable origin of the sound pattern was an adult word.

Jonah's ⟨dádà⟩ provides a good example of the problematic nature of word use during this period. For Time 4, in particular, as Jonah repeated ⟨dádà⟩ while crawling to the photograph he had propped up at the window, we have no way of deciding whether he was reacting with excitement as he remembered his intense earlier activity, naming the dog whose image is clearly recognizable in the photograph, or simply producing a babble pattern which echoed a word repeatedly used by his parents as they called to the dog from that same window. The child's meaning and the adult meaning may never have been isomorphic—though the child's meaning could plausibly be related to the adult's, given the context of use at Time 2. Certainly, the excitement the child experienced in connection with his dog seems to have been inextricably entwined with his use of the word from the beginning; and the adult meaning seems to have parted company with the word shape entirely before use of the vocalization had run its course.

Use of the extension convention meant, for most of our subjects, that a small residue of indeterminate vocalizations—which fit within the phonetic range of one of the child's spontaneous words, but were used with no recognizable context—were placed in the word category, though they may have been babbled sequences which happened to resemble one of the child's words. Most of the children produced a variety of segments and syllable types; thus their words were, in most instances, distinguishable from one another and from non-words. But one subject, Timmy, consistently produced a highly restricted number of consonant and syllable types. By Time 3, at nearly 12 months, he was using spontaneous words of the shape ⟨ba⟩ or ⟨ɸa⟩ to mean *ball, block,* and *box,* and words of the shape ⟨kaka⟩ to mean *kitty, doggy, teddy,* and *quack-quack*. In addition to 131 appropriate or plausible uses of these words, there were 12 uses of the same sound pattern with no identifiable referent—i.e., 8% of the 156 words used during that session, counting both spontaneous and imitated words. In fact, the problem was a persistent and increasing one in Timmy's data from Time 2, where 9 instances of 'unidentifiable referent' represented 23% of the total word count of 39, until Time 6, where the figure had risen to 40% (46 tokens out of 115 'word' productions). At that point, front vowels entered Timmy's system, as well as disyllabic vocalizations with contrasting consonants (C_1VC_2V). A month later, stop words became differentiated from fricative words, after which the 'unidentifiable referent' quotient dropped back to 16%. However, homophonous forms continued to figure in Timmy's lexicon at Time 7; e.g., the words *fly/flies, fire, Ruth*, and perhaps *floor* were all represented by the sound shape ⟨ʌβᵘæ⟩ (cf. Vihman 1981). Clearly Timmy's case represents an extreme in terms of the difficulty of distinguishing words from babble; by more or less arbitrarily choosing to place all such ambiguous cases in the word category (via the extension criterion), we believe we have clearly given the benefit of the doubt to the discontinuity thesis.

In one case of widespread homophony, we were able to cut down the arbitrary effect of the extension convention by separating off as a special case ('limited use') the spontaneous use of a word only within the bounds of a specific interactional episode. The problem arose at Time 5 for the subject Sean, a child who showed a strong and steadily increasing preference for monosyllabic vocalizations from Time 1 on. The progression was from 38% of all vocalizations at Time 1, both words and non-words, to 76% at Time 5. Over the same period, his use of CV and CVC shapes also increased steadily—from 12% at Time 1, when the majority of vocalizations lacked a true consonant, to 42% at Time 5. Since dentals accounted for fully 69% of all non-vocalic segments at Time 5, the shape ⟨dV⟩ was necessarily of very high occurrence. Two words emerged at this session which had not been recorded during the earlier anchor points: *down* and *thank you*. Sean's mother had reported, four weeks earlier, that Sean was using [daʊ] or [dae] to ask to be put down from his highchair, or from someone's arms or lap. During Time 5, he used [dæɪ] twice in the course of building block towers—apparently modeling this use of *down* on his mother's repeated comments, *They all fell down, It fell down,* and *You made 'em fall down*. The child's vocalizations are categorized as word uses; but the situation is not unambiguous, since Sean had earlier in the session been using ⟨dæ⟩ as a kind of completive, e.g. while placing blocks on top of one another. Because *down* was modeled, and because we have no evidence of its use otherwise except with a rather different and narrower meaning (paraphrasable as 'desire for movement downward of self'), we have not extended the label 'word' to other uses, such as to the 'completive' ⟨dæ⟩.

The *thank you* episode involved putting the blocks back into their box. Here Sean used ⟨dæ⟩ as he put the blocks in, continuing to make his action-word or completive use of this favorite sound pattern. However, after his mother had said *thanks* twice and *thank you* once, Sean began to alternate ⟨dæ⟩ with ⟨datdət⟩, sometimes with a [θ] on- or off-glide on the final [t]. Here Sean was apparently trying to accommodate his sound pattern to his mother's *thank you* pattern. We assigned this series of ⟨d⟩-vocalizations to the word category. Shortly after this episode was over, Sean walked toward the box of blocks, pointing and saying [ˀəfdatdɪt], which we counted as a repetition of *thank you*. We did not extend the word classification to earlier instances of ⟨dVt⟩, since there was no evidence that *thank you* was a word for Sean, outside the context of a single episode and the objects and activity associated with it. By withholding the extension convention just in the case of words whose contextually clear use occurred only within the bounds of a single episode where the word was modeled, we reduce the likelihood of overgenerously crediting a child with a word in situations in which a limited sound system makes chance correspondence between an adult word and babble highly likely.

We can now take up some of the further consequences of the criteria we used in distinguishing words from non-words. We chose to disregard the communicative function of the children's utterances for the purposes of this analysis. Thus what Stoel-Gammon & Cooper (250) call 'quasi-words'—'vocalizations which evidenced a consistent sound–meaning relationship but whose phonological form was not based on an adult model'—are included under the babble category. More generally, Menn had defined 'proto-words' (1976:39) as 'utterance types more word-like than expressions of affect, but different from our notion of "word" on phonological, semantic, or functional grounds.' Some of what we are classifying here as 'words', because their sound shapes derive from adult models, would certainly count as proto-words in Menn's sense (e.g. Jonah's ⟨dádà⟩). We are in fact carrying out a separate study of the proto-words in our data, starting from a different perspective. It is clear that the broad range of phenomena to be found in the vocalizations of the transition period form not one continuum but several, from least to most adult-word-like. The concept 'adult word' is itself multifaceted, of course, with a range of variability or stability along both the sound and meaning parameters, depending on lexical class or category, register, and other factors of use and context.

To make a formal comparison of the phonetic properties of words and non-words, however, it was essential to arrive at a relatively clear operational demarcation between the two—one which could be reliably applied by a team of investigators, each addressing a sample containing hundreds of vocalizations. For this reason, then, we chose to treat communicative function as irrelevant for our present purposes, once we determined whether a given vocalization did or did not have an identifiable adult model. Thus quasi-words, which form a subset of Menn's proto-words, fall into the non-word category; but vocalizations having an adult-based shape, but a use or meaning which differs from adults' (the remaining proto-words) are classed as words. The number of quasi-word types is in fact quite small in our corpus: Most of our subjects used no more than one or two—though Halliday 1975 describes his son Nigel as using over a dozen non-adult-based sound–meaning pairings, and one of Stoel-Gammon & Cooper's subjects has 13 (but nine of these were very similar in form). The number of tokens of a single invented-word type could be fairly high during a given session; but over-all, the frequency was very low. These vocalizations typically had a simple shape, such as glottal or stop consonant plus vowel.[3] In any case, given their infrequency, any classification adopted for these forms could have only a very small effect on the major point at issue here.

Several of our babies engaged in extensive communicative behavior while interacting with the mother or the observer. They made effective use of sentence-final and question intonation, of turn-taking, and of eye gaze or other body language appropriate to discourse; thus their vocalizations, while lacking any known adult source, were eloquent and quite irresistible in eliciting an adult response.[4] This type of beautifully modulated verbal behavior was counted as 'non-word', under our restrictive criteria. In contrast, adult-based words were sometimes clearly treated as babble by the babies—as when Deborah, at Time 3 during solitary play (and to some extent also during the videotaped session), produced long strings of syllables which appeared to represent three or four uninterrupted repetitions of her new word ⟨gagi⟩ or *doggy*. Because ⟨gagi⟩ seemed to be word-based, we counted these long strings as single word-based vocalizations.

One more problematic phenomenon that we encountered was the inclusion of a recognizable adult-based word in a long string of otherwise babbled syllables. In that case we included the entire string in the word category, following our customary use of the uninterrupted vocalization as a basic unit. An example, from the subject Susie, would be [adaoɹə^{m}mœmæːn], in which [mœmæːn] was one variant of her word for *mama*. Susie also embedded imi-

[3] Stoel-Gammon & Cooper note (266) that, 'Although they co-occurred, quasi-words and "real" words differed substantially in their phonetic form. The quasi-words were very simple throughout the study [i.e. for 7 months], typically consisting of CV syllables in which the consonant was [d].' The shapes of many of the other proto-words reported in the literature conform to the proposed canonical shapes CV, V, or glottal + V (cf. Leopold 1939, Vanvik 1971, Engel 1973, Halliday 1975, Dore et al. 1976, Labov & Labov 1978).

[4] Compare Menn's definition of 'jargon', based on the behavior of her subject, Jacob (37): 'Those sound patterns where the listening adult irresistibly feels that the child believes that he/she is "speaking" as other people do. Such advanced babble is not solo play, but is accompanied by the eye-contact, intonation patterns, and body posture appropriate to conversation.'

tated words in a longer string. Thus after her mother had said *A bunny, like in your book,* Susie responded with [dʌsnombʌjiˀəˀʌ], where [bʌji] appeared to derive from *bunny*. The use of words as part of a long, sentence-like string of babble was highly characteristic of this child from Time 4 on. The classification of such strings as words might increase the continuity 'look' for this baby; but one could also argue that her very creation of such hybrid strings suggests that the distinction was not a sharp one, from the baby's point of view.

QUANTITATIVE AND STATISTICAL ANALYSES

3. The phonetic structure of word and non-word utterances was analysed separately by child and by anchor point. Four phonetic parameters were examined, with three or five categories each (see Table 1). Frequency distributions were drawn separately for word and non-word utterances, for the set of categories within each parameter. The frequency values were used to examine the Jakobsonian distinction between 'variety' and 'sparseness'; changes in these values over time, and across children, were used to examine the distinction between 'lack of systematicity' and 'universal order'. No statistical analyses were conducted relevant to the third of Jakobson's criteria distinguishing babbling and speech—the gradual restriction to different contexts (but see §5, below).

PARAMETERS	CATEGORIES
Vocalization length in syllables	monosyllable
	disyllable
	polysyllable
Phonotactic structure	vowel (V)
	consonant + vowel (CV)
	(consonant +) vowel + consonant ((C)VC)
Consonant manner	glottal
	stop
	glide
	fricative
	resonant
Consonant place	labial
	dental(/alveolar/palatal)
	velar

TABLE 1.

To arrive at a statistical evaluation of our results, we sought to determine whether or not the two variables—words and non-words—were similar within a given category. If the two variables were highly dissimilar within a category, we would conclude that they were independent, and such independence would provide evidence for the discontinuity view; by contrast, within-category similarity between variables would indicate a dependent relation, and provide evidence for continuity. Figures 1–4 are graphic displays of the word and non-word trends (given as percentages) for each of the babies along the four parameters analysed, using data from the mother/child interactional setting only;[5]

[5] Numbers across the top of each figure refer to anchor-point sessions or (on Fig. 4) stages; numbers along the sides refer to percentage use of the category in question. The vertical line (Figs. 1–3) indicates the session in which four or more different word types were first used spontaneously.

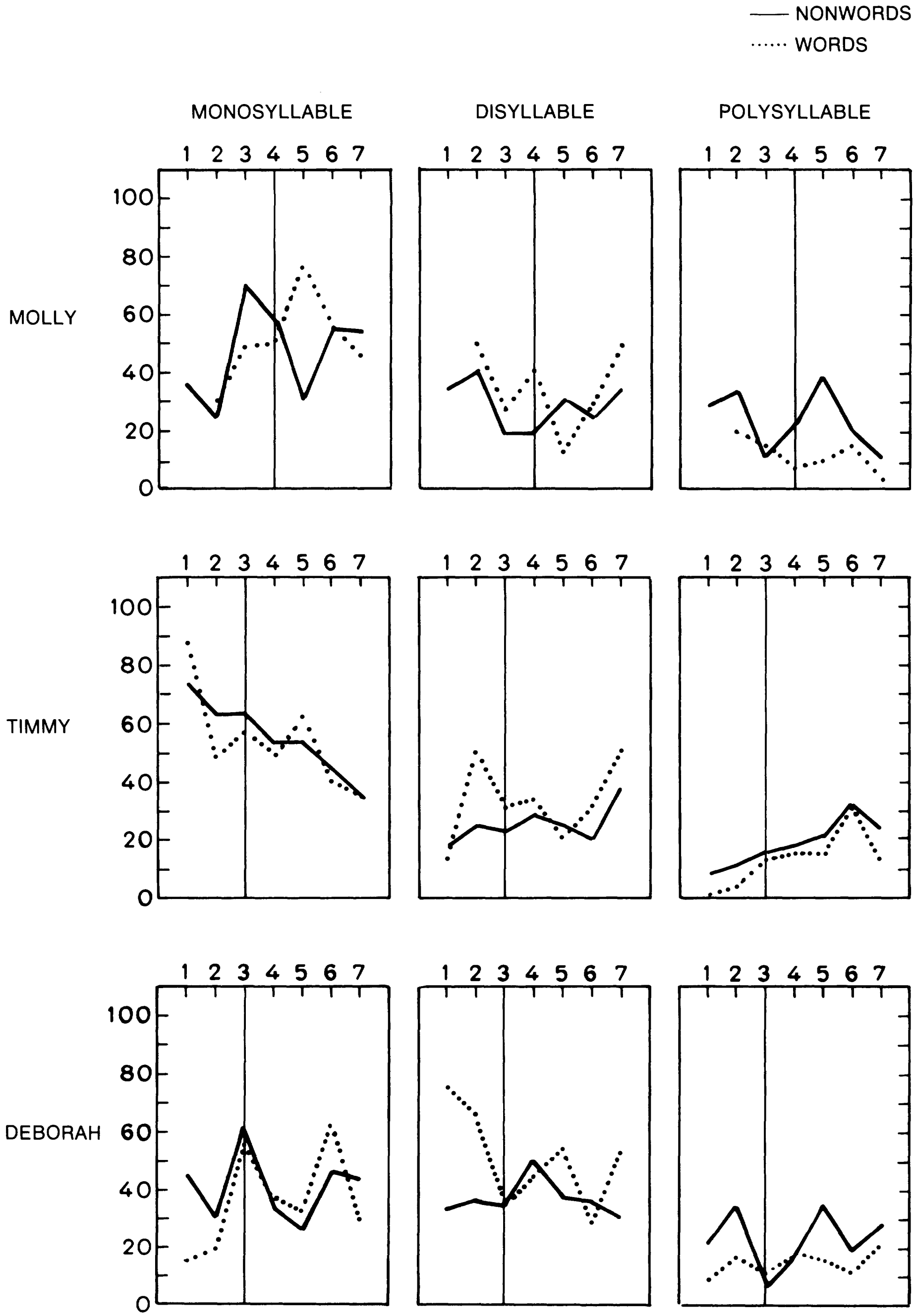

FIGURE 1. Vocalization length in syllables.

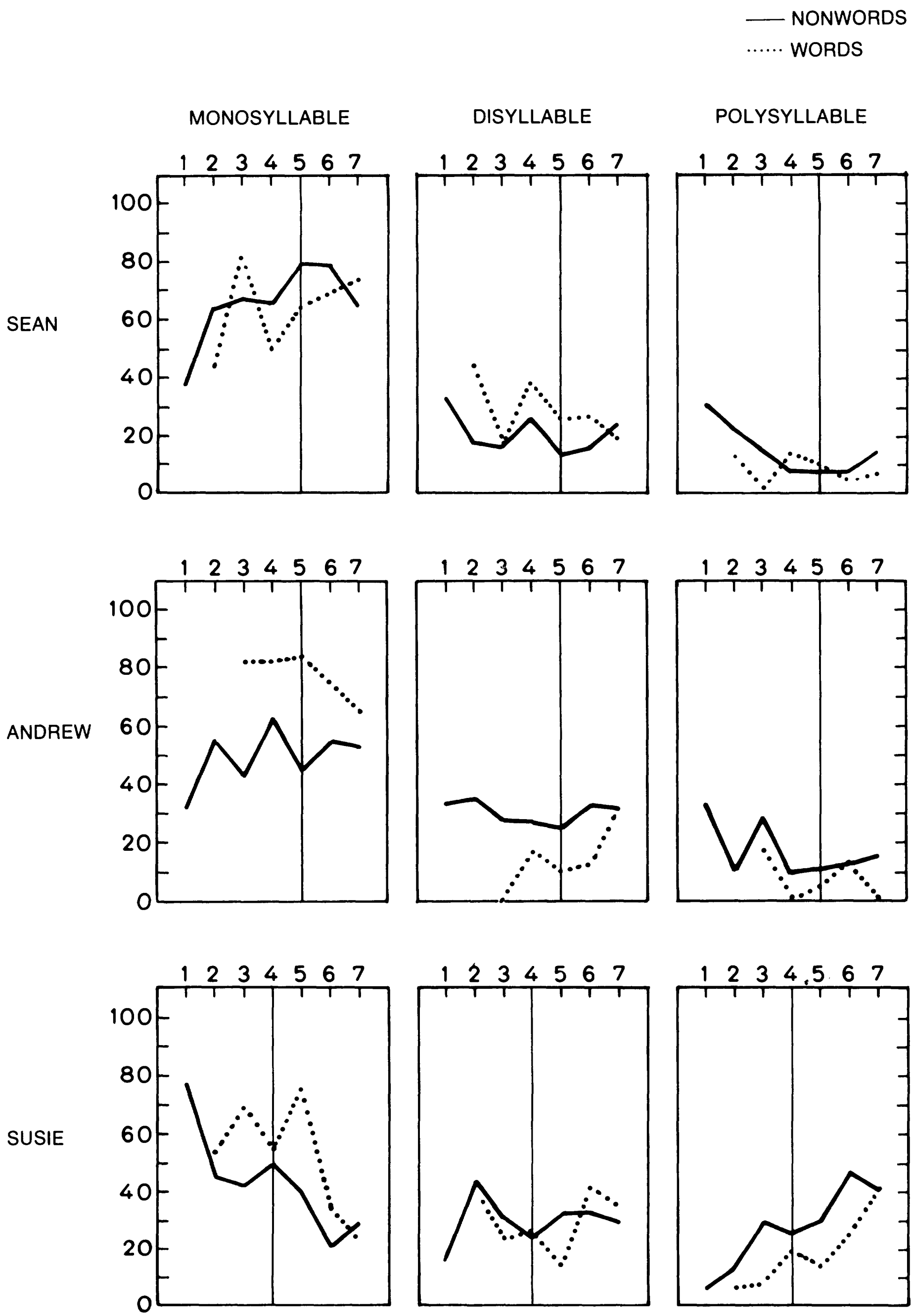

FIGURE 1. (*Continued*)

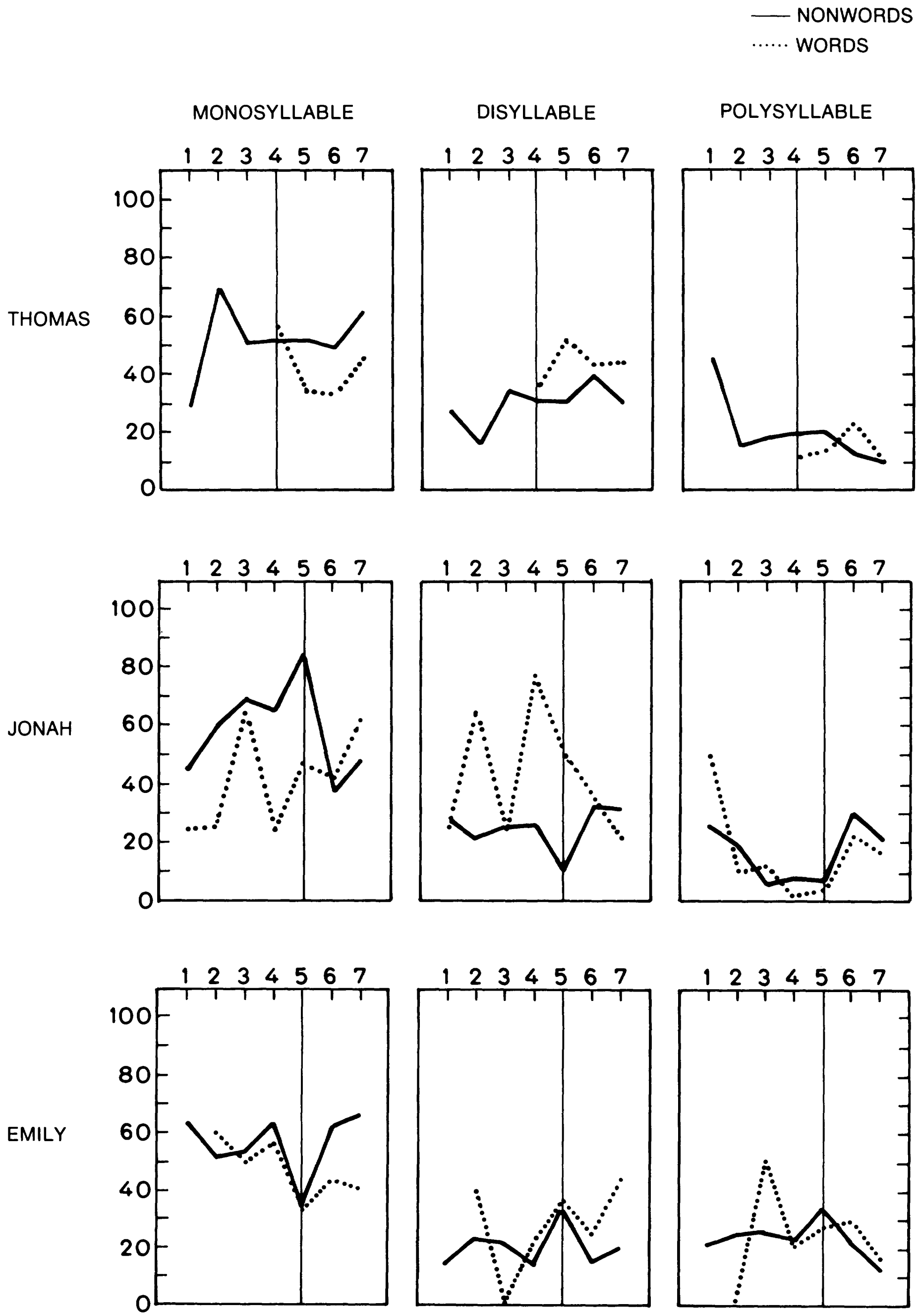

FIGURE 1. (*Continued*)

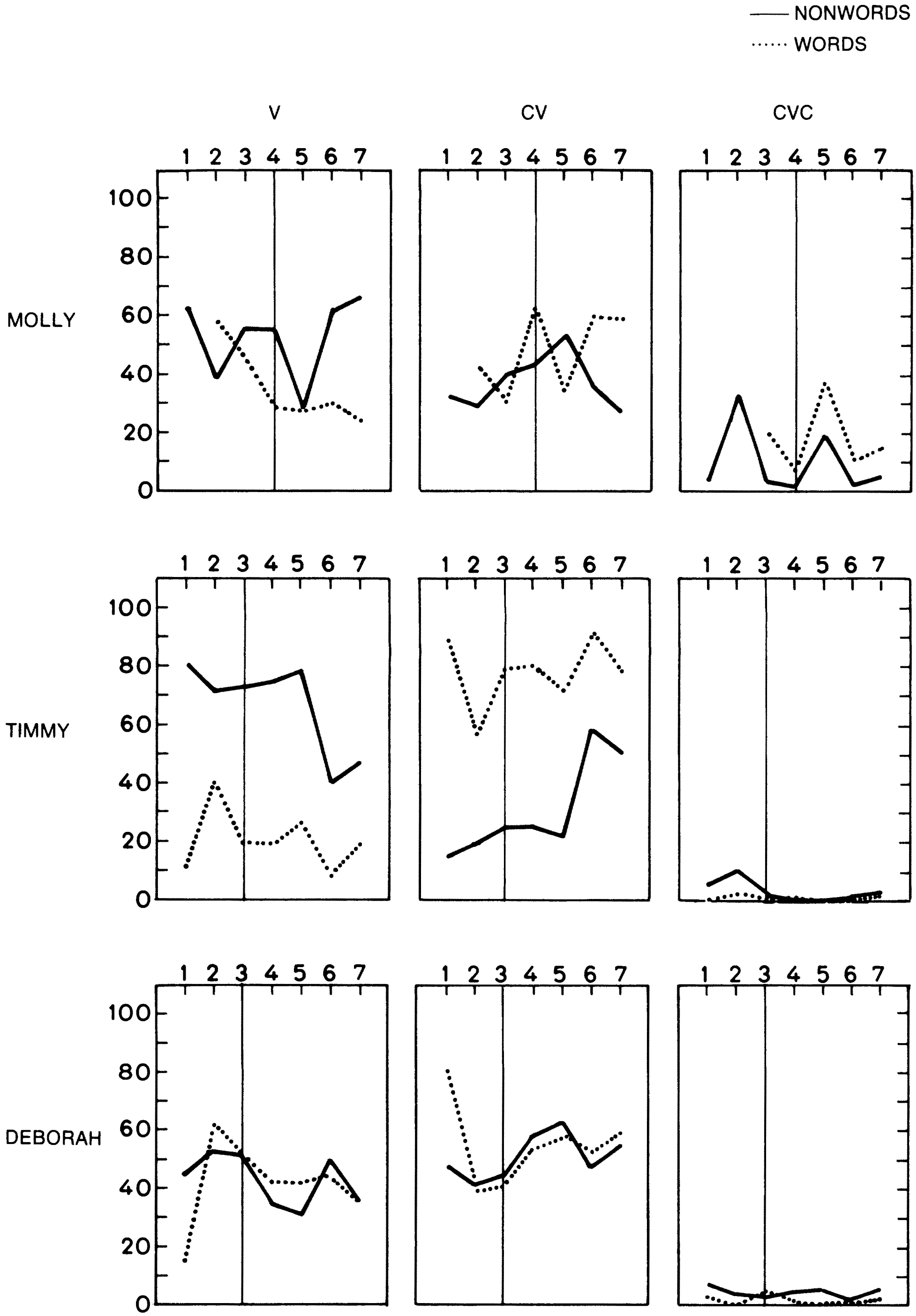

FIGURE 2. Phonotactics.

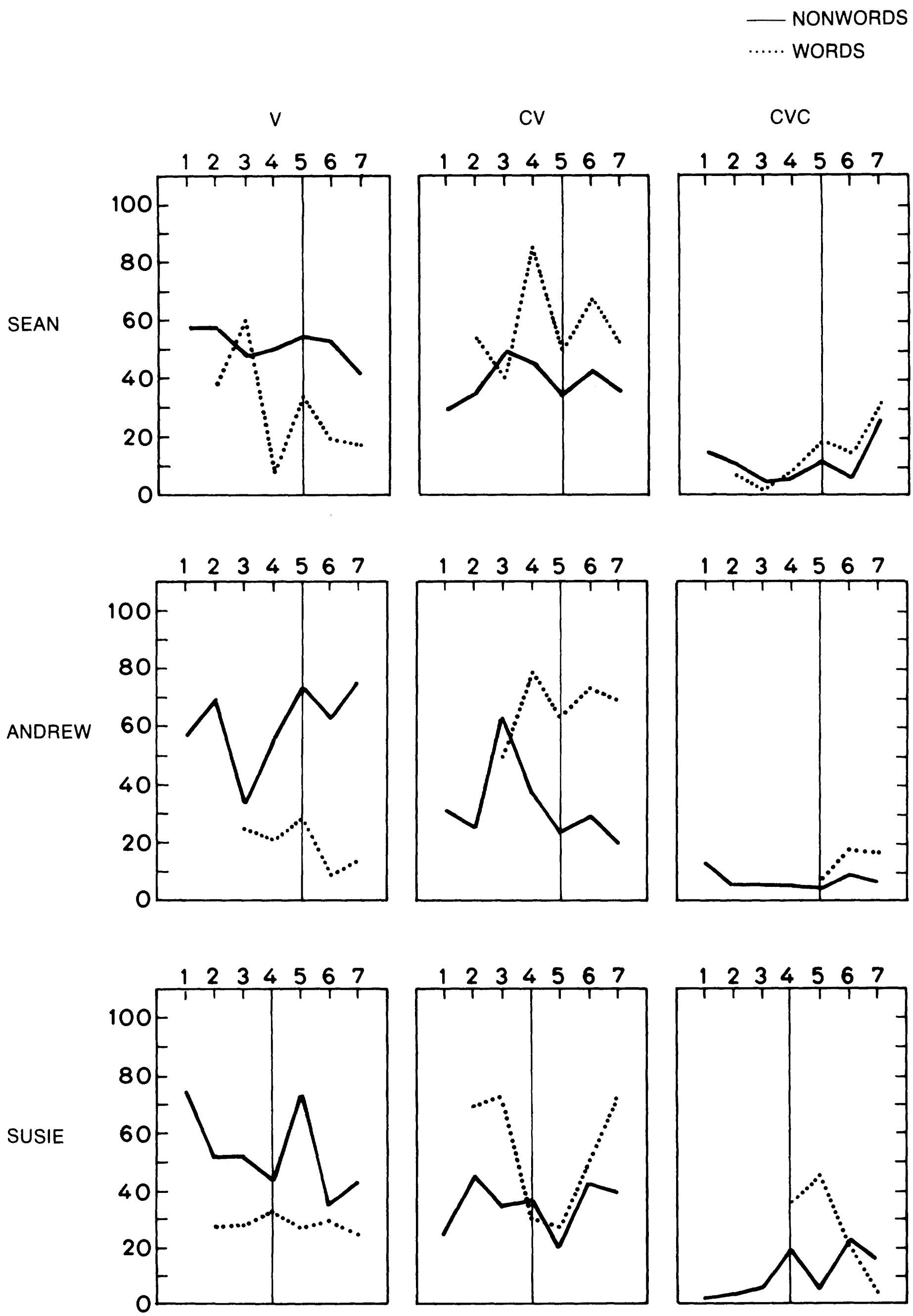

FIGURE 2. (*Continued*)

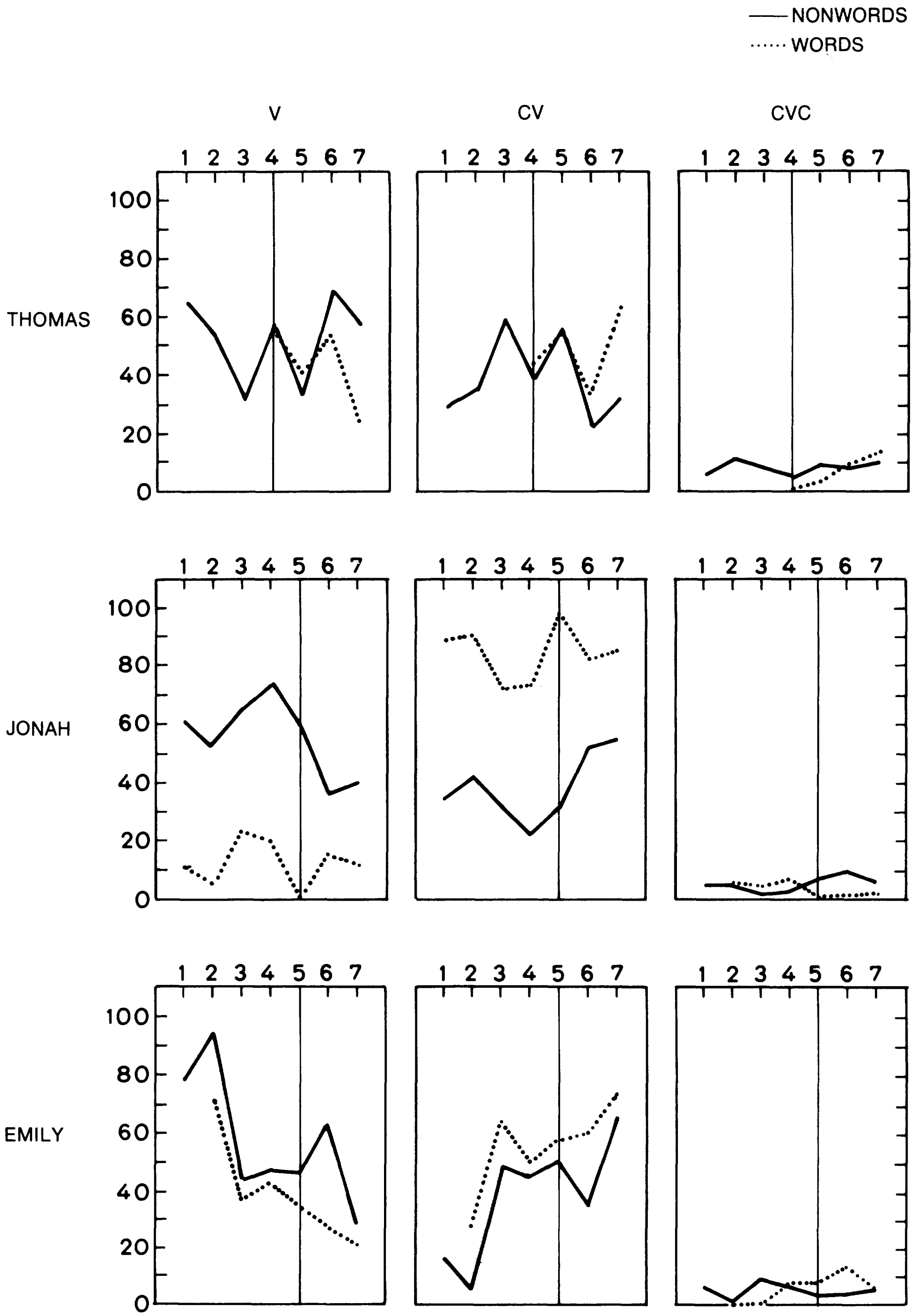

FIGURE 2. (*Continued*)

these figures represent the frequency distributions to which we will refer.[6] Data from the solitary play setting were not included in the full quantitative and statistical analyses, both because only four of the nine children had seven usable solitary play anchor-point recordings, and also because word use was more difficult to verify in the absence of video recordings.

Given numerical values for each variable and changes over time in those values, at least two ways are available to define similarity: as the closeness of fit of the data-point pairs at any one time, or as the extent to which the word and non-word curves exhibit movement together over time. We were interested in both types of similarity. It is worth noting that closeness of fit usually implies movement together (see Fig. 2, Emily, CV, for a good example). The converse is not true: Curves may move together without data-point pairs being especially close (see Fig. 2, Jonah, CV).

To measure closeness of fit, we applied a 'distance-function permutation' test. This test compares the distance between word and non-word points for each subject (intra-subject differences) with distances between word points of that subject and non-word points of other subjects (inter-subject differences). If intra-subject differences tend to be smaller, the continuity hypothesis will be supported. See Appendix A for a complete explanation of the statistical methods used in this test.

To measure movement together, we used a test based on Spearman's rank correlation coefficient. After computing the coefficient between word points and non-word points for each subject over time (and within a particular category of a parameter), we summed the coefficients for all subjects. We can test whether this sum is significantly different from zero—i.e. whether there is correlation (or movement together) within a category—by using an estimate of the sum's variance and applying a test based on the normal distribution. The continuity hypothesis will be supported if the sums of the correlation coefficients tend to be significantly greater than zero (see Appendix B).

The position of the word curve relative to the non-word curve, within a given category over time, is also basic to examining the issue in question. If babbling and speech are indistinguishable systems, then neither the word nor the non-word curve should lie consistently above the other over time for all subjects. Thus, in addition to 'closeness of fit' and 'movement together', we define similarity as the extent to which neither words nor non-words consistently dominate over time in any category.

To address the question of possible dominance over time of words or non-words within any one category, we used a test based on Wilcoxon's signed rank statistic. Essentially, the test compared the average percentages of word and non-word vocalizations within each category of a parameter, to assess whether they differed consistently. In this statistical test, unlike the others, the null hypothesis is that there are no differences between words and non-words.

[6] To avoid the most extreme case of low word use, sessions which included just one token of one word type (Sessions 1 and 2 for Thomas, and Session 1 for Molly) were omitted from the statistical analysis (see Table 2). Note that the word curve on the figures begins with the first session at which more than one word token was used by a given child.

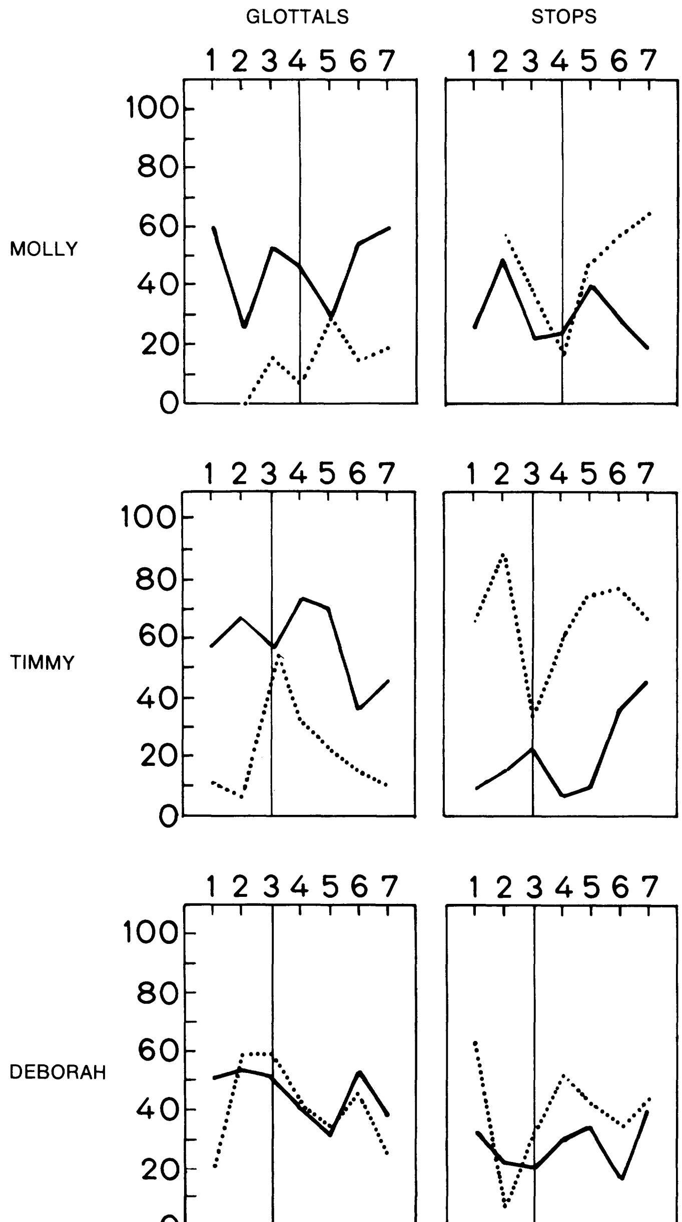

FIGURE 3. Consonant manner.

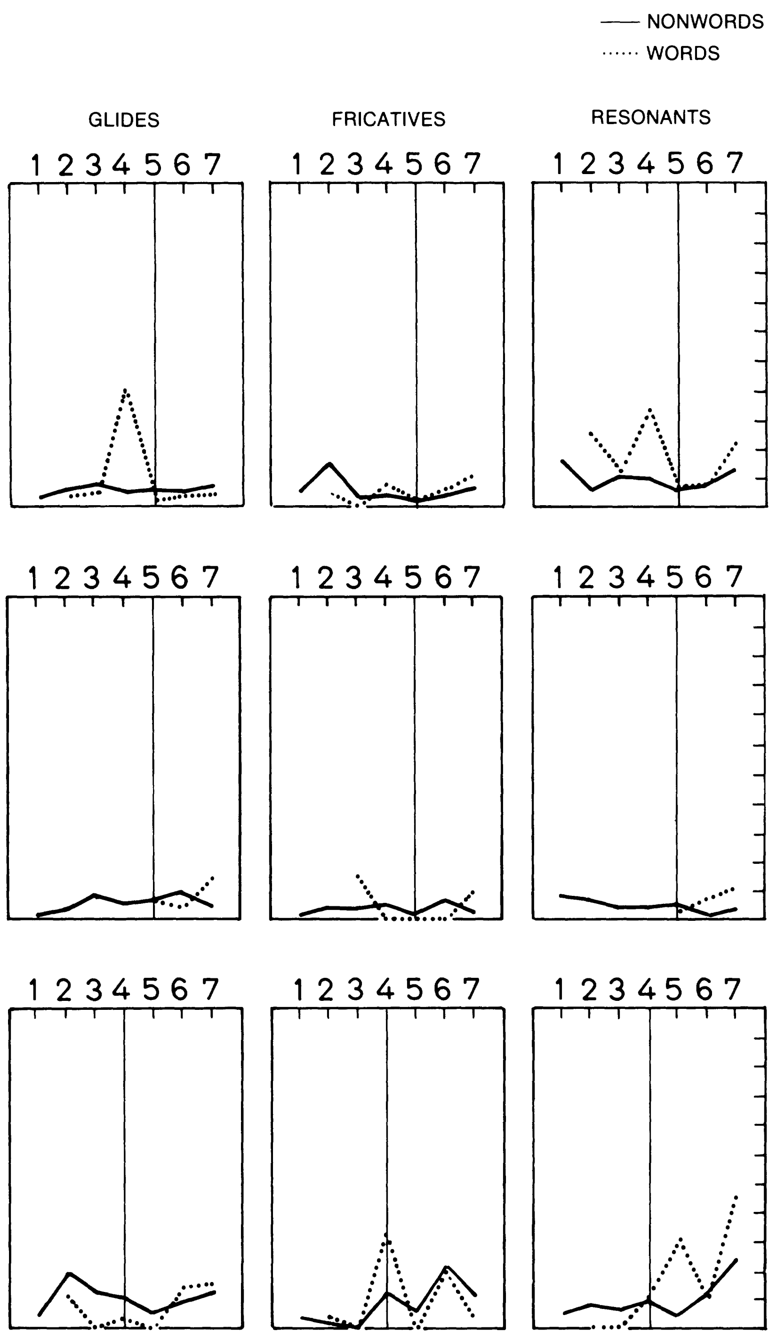

FIGURE 3. (*Continued*)

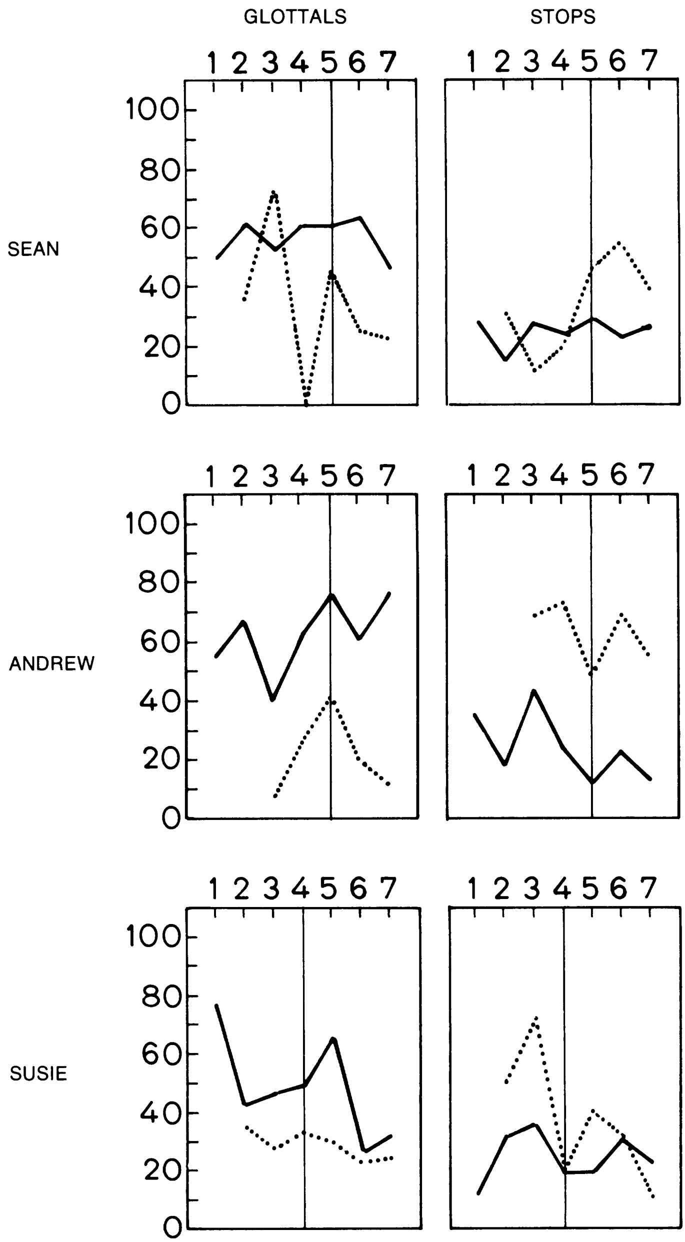

FIGURE 3. (*Continued*)

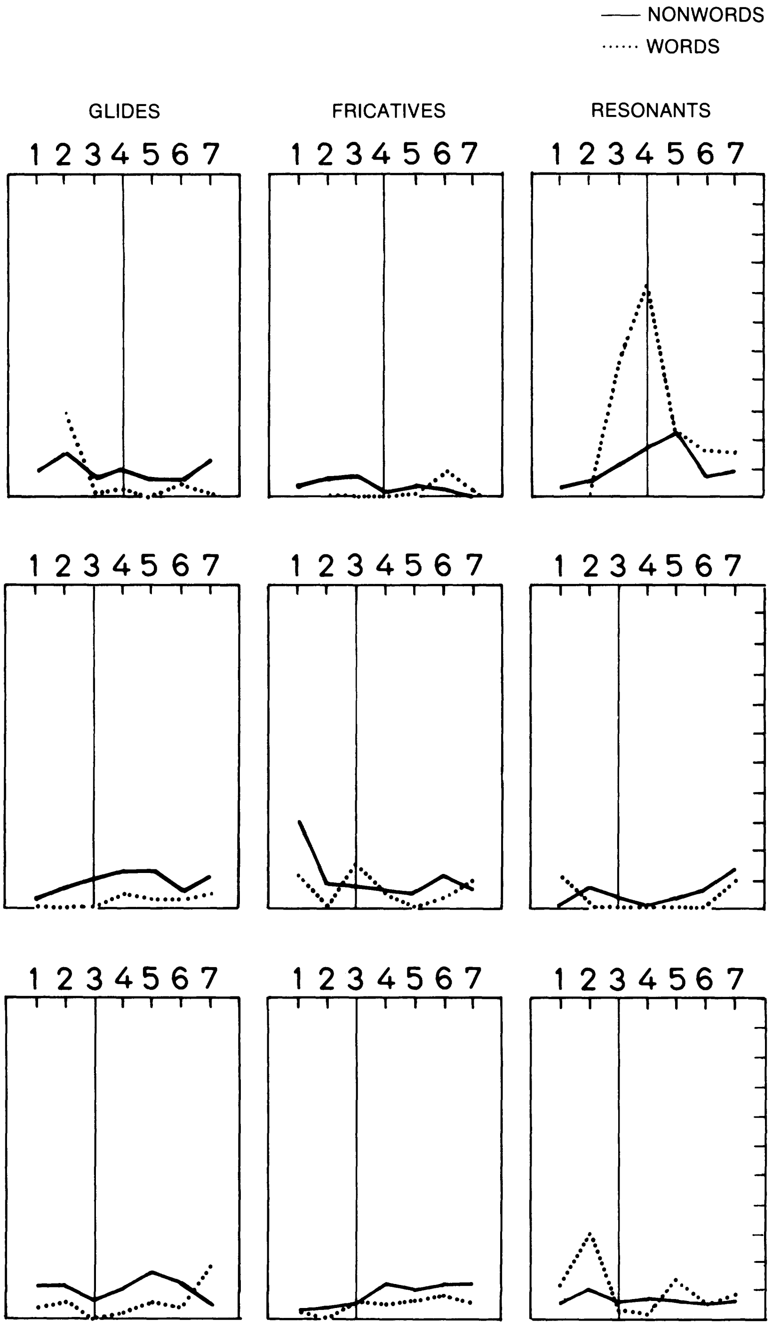

FIGURE 3. (*Continued*)

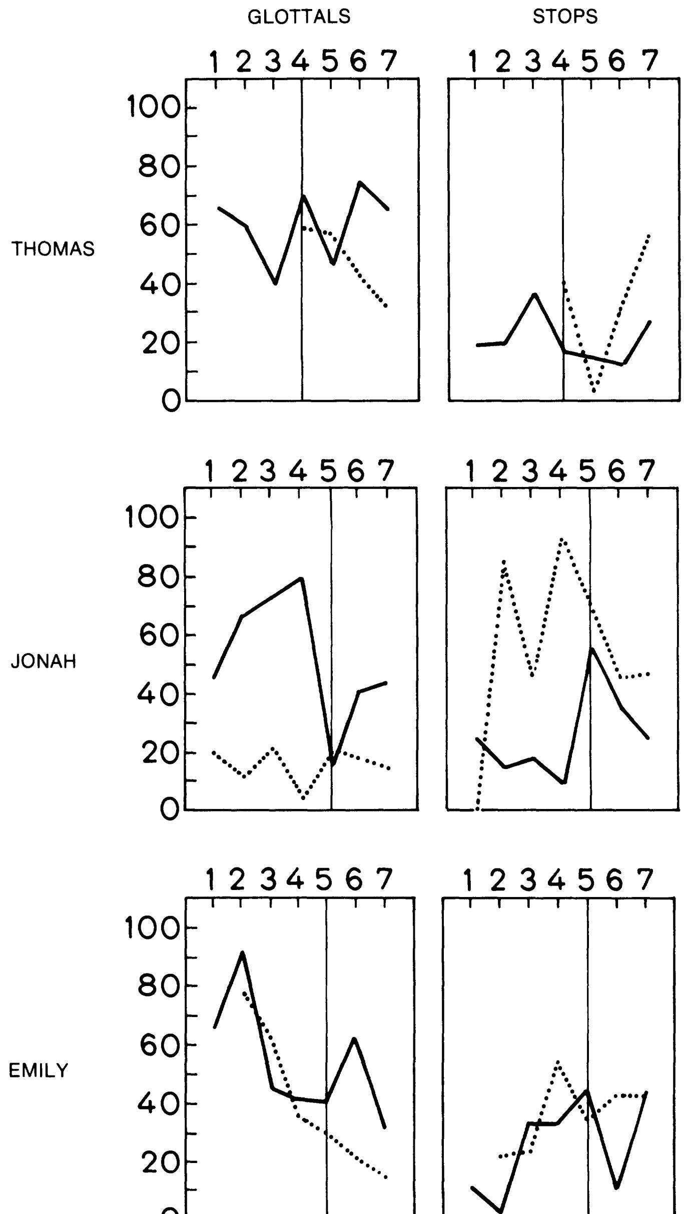

FIGURE 3. (*Continued*)

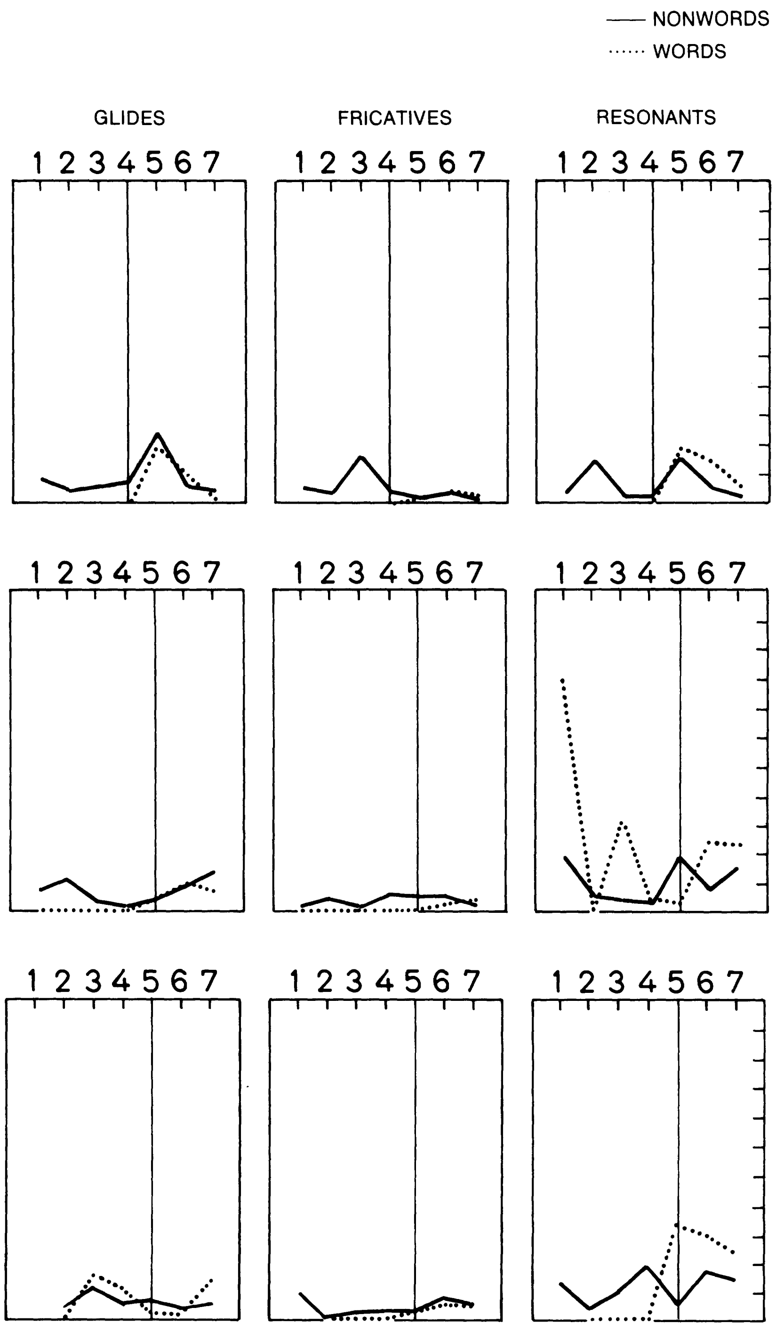

FIGURE 3. (*Continued*)

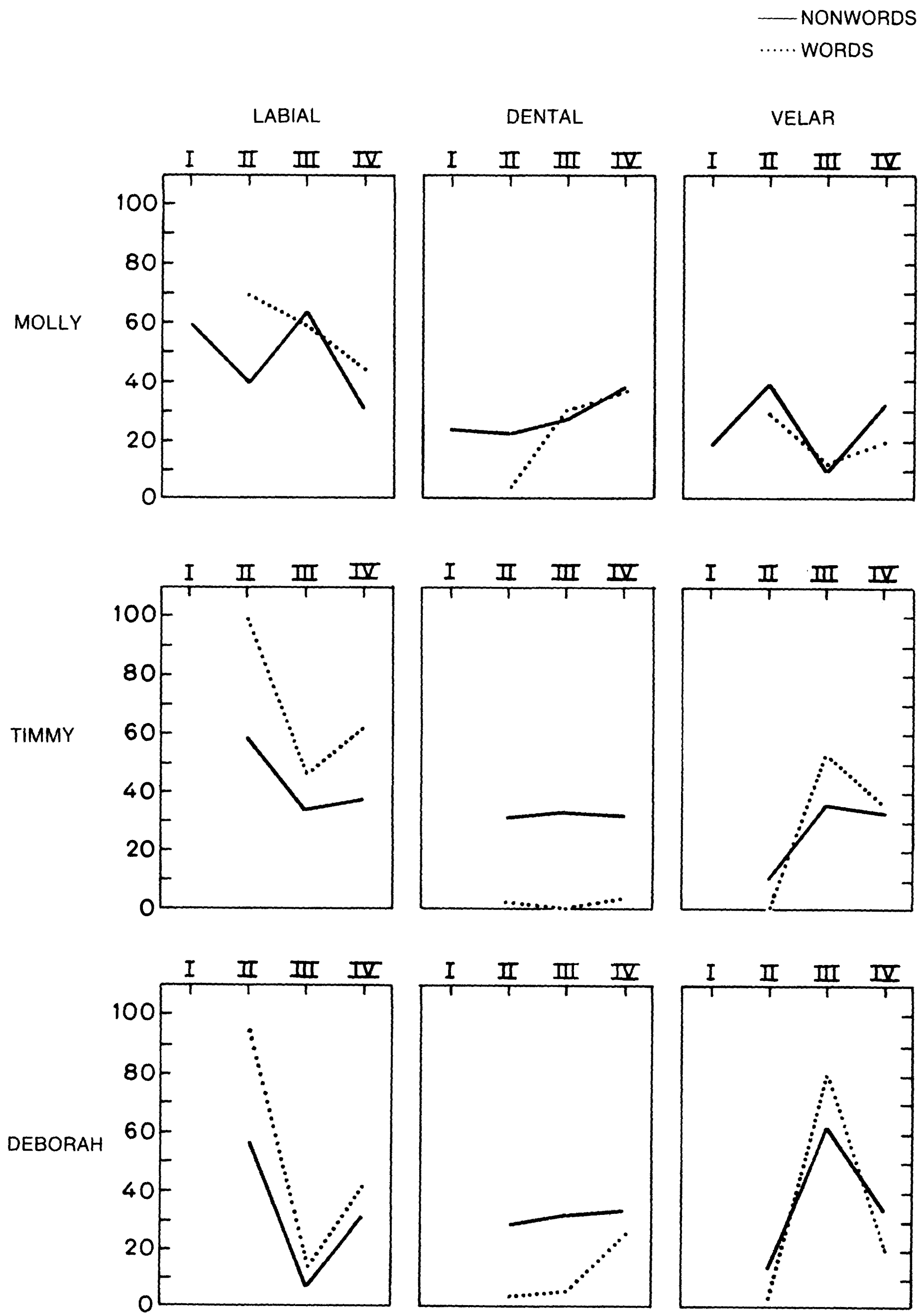

FIGURE 4. Consonant place.

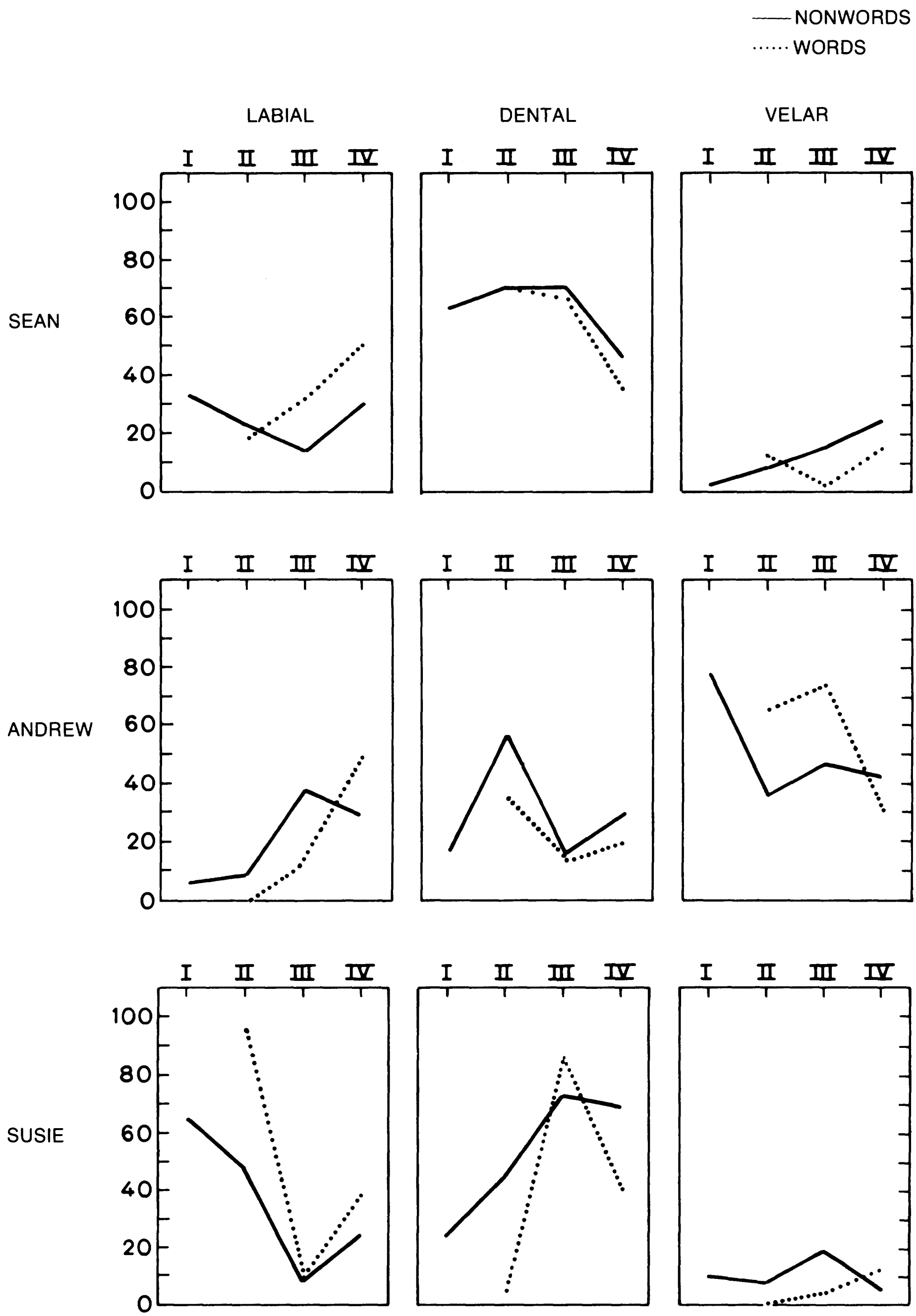

FIGURE 4. (*Continued*)

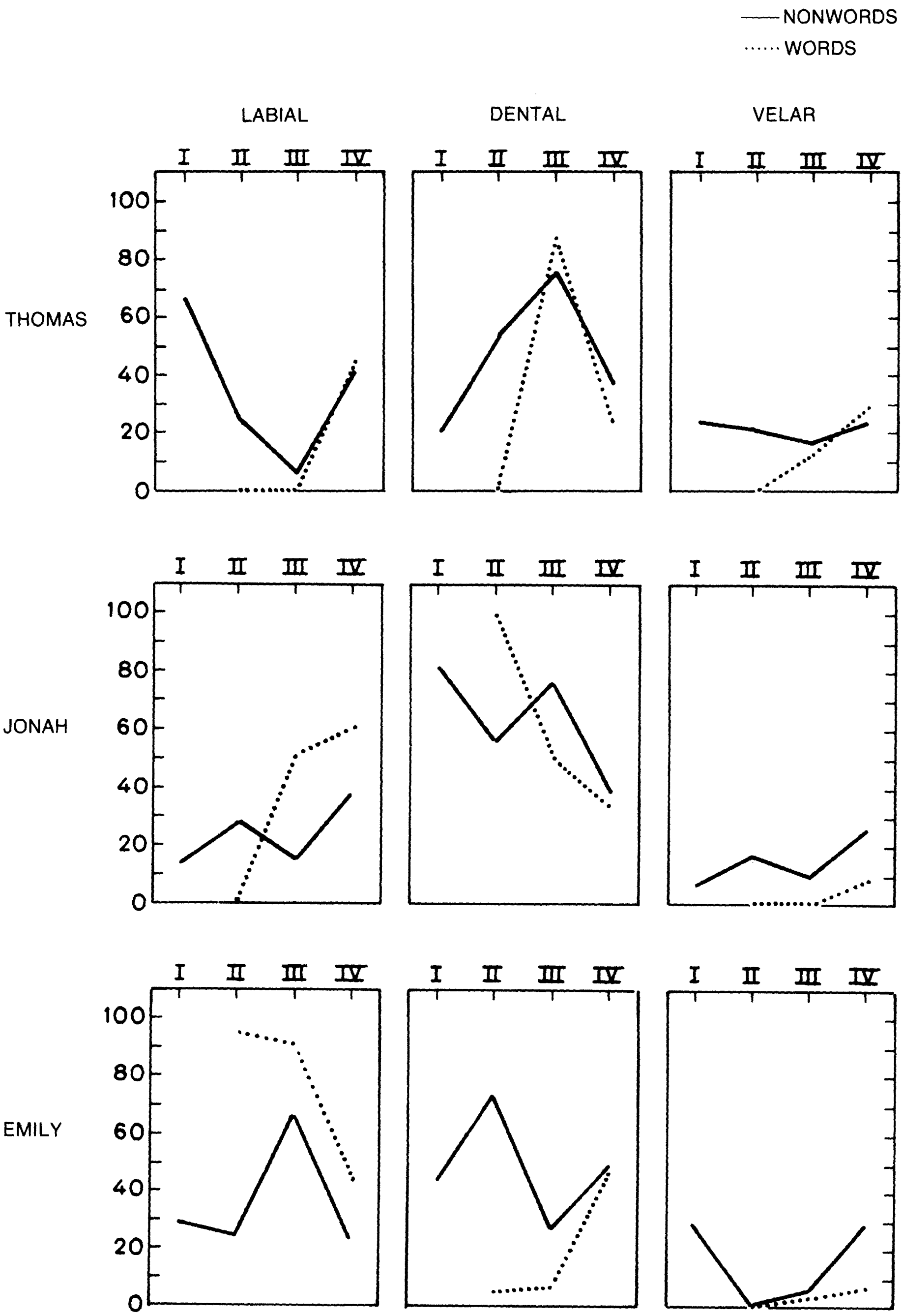

FIGURE 4. (*Continued*)

In this test, therefore, rejection of the null hypothesis is evidence for a DISCONTINUITY between trends for words and non-words (see Appendix C).

RESULTS

4.1. SUBJECT VARIABILITY. Table 2 (overleaf) is intended to give an idea of the varying progress of the nine babies toward word use over the period in question, as well as an indication of the variability in average quantity of vocalizations in each setting across subjects. Because our quantitative results are less subject to distortion (because of minor variation) for the children with higher frequency of vocalization, we have ordered the columns of Table 2—as well as the charts within each figure—from the child with the highest mean number of vocalizations in the social setting to the child with the lowest. Molly, Timmy, Deborah, and Sean are, in fact, quite close in their averages (from 265 to 241); there is then a definite drop to a lower and broader range of averages (149 to 96).[7]

It is important to point out that our word-type totals, as presented in Table 2, do not reflect our subjects' over-all level of language development at any one point. First, the totals are based on just those words which happened to occur during a single half-hour period. Though we referred to the information which the mothers reported as part of the diaries we asked them to keep, we did not include the words they noted in any of our counts, since the phonetic shapes could not be determined with any certainty. If we total, for each baby, the words we found on the anchor-point recordings plus all words reported by the mothers over the time of the study, the recorded words amount to 62% of the total, on average across the subjects; the mothers add the remaining 38% (see Table 3, p. 428). Second, we have not yet made a complete tally of the adult-based gestural symbols or gestural 'words' used by the children (e.g. a wave to mean 'bye-bye', spread hands for 'all gone'), though these are certainly closely related to language development. Third, the children often showed extensive comprehension of language well in advance of their production of adult words; we give no account of that here. Finally, to mention proto-words again, the children were often able to make clear vocal expression of their needs and interests while still remaining outside the bounds of language, as we have defined it for the purpose of this analysis.

[7] Note that only one child, Thomas, actually vocalized relatively more during solitary play than in the interactional setting (taking into account the fact that the latter was at least twice as long). Jonah and Susie vocalized almost as much alone as with mother and observer present, and Deborah and Molly about two-thirds as much; but Timmy, Andrew, Emily, and Sean produced, when alone, only half as many vocalizations (or less) per unit time.

For Emily, who vocalized least of all the children, we added several minutes of audio-only data to the half-hour videotape session, for the first four anchor points. The data include all the vocalizations that occurred on the audio tape, which ran 7 to 15 minutes longer than the video on those occasions (i.e. up to 45 minutes in all). The total vocalizations on the videotape sessions are indicated in parentheses on Table 2; the average number of vocalizations shown for Emily is based on those figures, to provide a fair point of comparison with the other subjects. From Time 5 on, the video session alone contained an adequate number of vocalizations; thus the data sets for these anchor points were not extended.

		Molly	Timmy	Deborah	Sean	Andrew	Susie	Thomas	Jonah	Emily
Time 1:	Age	0;8.17	0;9.16	0;10.2	0;9.3	0;9.18	0;9.10	0;9.26	0;9.23	0;9.9
	Spontaneous	–	1/2	3/15	–	–	–	–	1/3	–
	Imitated	1/1	2/6	2/10	–	–	–	1/1	1/1	–
	Total Vocs	153	161	204	107	103	104	58	162	33(20)
Time 2:	Age	0;9.9	0;10.6	0;10.28	0;10.0	0;11.6	0;10.23	0;10.22	0;10.27	0;10.7
	Spontaneous	–	2/24	2/26	2/17	–	1/1	1/1	1/14	2/5
	Imitated	3/10	4/15	4/17	3/15	–	8/16	–	1/9	–
	Total Vocs	279	188	160	182	63	145	160	81	130(101)
Time 3:	Age	0;10.15	0;11.6	1;0.4	0;10.29	1;0.24	0;11.27	0;11.19	0;11.25	0;11.4
	Spontaneous	3/16	6/134**	6/89**	2/8	1/6	3/6	–	1/8	2/2
	Imitated	7/27	9/22	4/10	3/9	–	4/7	–	6/9	2/2
	Total Vocs	201	267	274	243	155	75	107	154	135(83)
Time 4:	Age	0;11.9	1;0.3	1;1.3	0;11.25	1;1.29	1;1.8	1;0.25	1;0.23	1;0.1
	Spontaneous	6/52**	6/129	9/93	1/6	1/2	7/10**	9/27**	2/8	3/5
	Imitated	6/11	4/18	11/29	1/2	7/10	3/6	6/19	2/5	4/9
	Total Vocs	161	243	250	178	130	124	172	95	153(32)
Time 5:	Age	1;0.10	1;1.1	1;2.6	1;0.24	1;3.9	1;2.18	1;1.16	1;1.21	1;1.0
	Spontaneous	14/49	6/174	18/57	9/45**	5/9**	3/4	8/23	5/33**	10/56**
	Imitated	11/19	1/2	8/11	5/15	4/10	3/4	5/5	6/20	4/12
	Total Vocs	163	330	276	282	159	110	160	126	133

Time 6:	Age	1;1.8	1;1.29	1;3.24	1;2.12	1;4.22	1;3.27	1;2.21	1;2.18	1;2.9
	Spontaneous	30/208	8/101	33/98	24/77	2/3	5/10	17/21	7/19	20/58
	Imitated	18/44	10/14	17/22	20/31	4/13	14/16	3/5	10/16	7/8
	Total Vocs	424	296	276	277	233	329	209	189	126
Time 7:	Age	1;2.20	1;4.0	1;5.14	1;3.30	1;5.21	1;5.0	1;4.2	1;3.22	1;3.30
	Spontaneous	30/374	17/105	50/158	44/193	15/32	5/28	15/51	10/24	34/104
	Imitated	23/71	12/24	11/19	21/30	20/39	6/6	7/13	5/5	15/19
	Total Vocs	477	260	271	417	198	148	156	128	180
Total word types		152	88	178	135	59	62	72	58	103
Total word tokens		882	770	654	448	124	114	166	174	280
Tokens/Types*		6.5	8.75	3.7	3.3	2.1	1.8	2.3	3.0	2.7
Average Number of vocs										
mother/child		265	249	244	241	149	148	146	134	96
solitary play		87	68	84	50	46	64	110	62	29

TABLE 2. Figures for spontaneous and imitated words are given as types/tokens.

* Tokens/types calculated for spontaneous and imitated words used in mother/child interaction setting only; in most cases, some imitated words are also used spontaneously, so that the number of types over-all is smaller than the sum of the types given.

** Criterion session (session in which four or more diffferent word types were first used spontaneously).

SUBJECT	MOTHER + TRANSCRIPT		MOTHER ONLY		TRANSCRIPT ONLY		WORD TOTAL
Molly	60	43%	70	51%	8	6%	138
Sean	48	43%	42	38%	21	19%	111
Timmy	30	68%	9	20%	5	11%	44
Deborah	54	42%	44	34%	32	25%	130
Thomas	23	40%	19	33%	15	26%	57
Susie	11	44%	6	24%	8	32%	25
Jonah	14	40%	15	43%	6	27%	35
Andrew	21	26%	53	66%	6	8%	80
Emily	48	42%	45	40%	20	18%	113
Total	309		303		121		733
Mean	34	42%	34	42%	13	16%	81

TABLE 3.

Excludes imitations reported by mother only or recorded only; includes recorded imitations if mother reported spontaneous use.

Mother + transcript: Words reported by mother and also recorded during an anchor point session.

Mother only: Words reported by mother but not used during anchor point session.

Transcript only: Words recorded during at least one anchor point session but never reported by mother.

The term 'anchor-point session' on Table 3 refers only to the mother/child interactional setting. New word types could not easily be identified on the solitary play transcripts, as mentioned earlier, since we lacked both video recording and the contextual framework provided by the mother's talk. Tokens of known types could be identified, however, by using the extension convention as well as any context afforded by environmental sounds. Mean word-tokens used over all sessions by all subjects came to 7.3 per solitary play session, as compared with about 57.3 per interactional session (which was at least twice as long in each case). Mean percentage of word use over all vocalizations in solitary play was 10.3%, vs. 23.7% spontaneous word use (30.3% if imitations are included) in the interactional setting.

4.2. SIMILARITY BETWEEN WORDS AND NON-WORDS FOR EACH CHILD. The children showed differing individual sound preferences, both over time and as compared with one another; but they consistently expressed these preferences in both word and non-word categories (see Figs. 1–4). This trend is clearest for those children who vocalized most, or who used the largest numbers of word tokens. For the remaining children, the underlying phenomenon appears to be the same; but minor fluctuations in word use appear in the figures as major perturbations.

The distance-function permutation test yielded highly significant results for three of our four parameters—vocalization length in syllables, consonant manner, and consonant place—and marginally significant results for the fourth, phonotactics (see Table 4). These results show that intra-subject distances between word and non-word curves tend to be smaller than inter-subject distances; i.e., individual parallel patterns for words and non-words are stronger than the over-all similarity in patterning for the age group as a whole. The data from children with small word bases (whether resulting from the use of few tokens or the use of few types) tend to lower the apparent similarity on such

PARAMETERS	d_1		d_2	
	n	pr	n	pr
Vocalization length in syllables	18	0.009**	20	0.010**
Phonotactics	221	0.111	217	0.109
Consonant manner	2	0.001***	6	0.003**
Consonant place	0	0.000***	0	0.000***

TABLE 4.

d_1 = the average absolute difference for each word and non-word pair.
d_2 = the square root of the average squared difference for each pair.
n = the number of randomly permuted diagonals that were smaller than the original (intra-subject) diagonal (out of 2000 permutations).
pr = the proportion of randomly permuted diagonals that were smaller (i.e. $n/2000$); this figure can be interpreted as a p-value.
** $p \leq 0.01$
*** $p \leq 0.001$

a test, since use of any one form will have major effects, considerably offsetting the two curves.

For the test using Spearman's rank correlation coefficient (see Table 5), four of the categories proved significant at the .01 level, and another five at the .05 level. These results point to the same general conclusion as the distance-func-

	CONSONANT MANNER					CONSONANT PLACE		
	GLOT	STOP	GLIDE	FRIC	RES	LAB	DENT	VELAR
Molly	.406	−.029	.441	−.216	.771	.290	.143	.638
Sean	−.232	.000	.580	.106	.261	−.257	−.086	−.371
Timmy	.214	.264	.667	.436	.379	.432	.019	.173
Deborah	.664	.559	.136	.561	.218	.836	.546	.591
Thomas	.000	.800	.400	−.316	.949	1.000	.400	1.000
Susie	.600	.667	.279	.725	.406	.371	.771	−.103
Jonah	−.357	−.523	.447	.181	.182	−.143	.203	.845
Andrew	.500	.821	−.410	−.224	−.107	.667	.600	.900
Emily	.657	.343	.812	.893	−.030	.347	.088	.676
Sum	2.452*	2.902**	3.352**	2.146	3.029*	3.593**	2.684*	4.349**

	VOCAL. LENGTH			PHONOTACTICS		
	MONO	DI	POLY	V	CV	(C)VC
Molly	.200	.529	.250	−.257	−.203	.058
Sean	.203	.348	−.152	.314	.029	.812
Timmy	.764	.714	.837	.162	.373	.398
Deborah	.324	−.107	.273	.679	.573	.273
Thomas	.632	−.258	.316	.632	.400	.800
Susie	.429	.464	.771	−.403	.319	.232
Jonah	.286	−.429	.714	.357	.107	−.582
Andrew	−.308	.462	.700	−.200	−.100	.658
Emily	.029	.314	.314	.600	.714	−.088
Sum	2.559*	2.037	4.023**	1.884	2.212	2.561*

TABLE 5.

Under the hypothesis of independence of the variables of words and non-words for each of the categories above, each of the fourteen sums has mean 0 and standard deviation 1.372. The critical values for a one-tailed test of the hypothesis are then:

$$\text{*alpha} = .05 : 1.372 \cdot 1.645 = 2.258$$
$$\text{**alpha} = .01 : 1.372 \cdot 2.326 = 3.192$$

tion permutation test: consistently significant results for vocalization length, consonant place, and consonant manner, and marginally significant results for phonotactics.

Figs. 1–4 reveal certain over-all similarities among the children, such as the relatively high proportion of monosyllables and low proportion of polysyllabic strings in both word and non-word categories (Fig. 1), the still smaller proportion of (C)VC shapes (Fig. 2), and a distinct preponderance of stops over all other true consonants (Fig. 3). But the children also show clear individual patterns, with differences in the frequency of particular categories along each parameter, both in comparison to one another and over time.

For example, on Fig. 1, vocalization length, we see Sean's ever-increasing use of monosyllabic babble up to Time 5—paralleled by a less even trend, but nevertheless by consistently high frequency of monosyllabic forms for words as well. Deborah seemed to use monosyllables and disyllables about equally often; her fluctuations in use across the two categories were closely matched in babble and words from Times 2 or 3 on. For Timmy, also, a sharp parallel trend emerged in both words and non-words from Time 3 on—with monosyllables decreasing in favor of disyllables and, at Time 6, longer vocalizations.

The Fig. 1 charts for the less voluble children show comparable results: a generally parallel trend for words and non-words in every category, but with deviations from the parallel trend for one or two sessions. Thus Susie's word curve departed from the non-word curve at Time 5, when monosyllabic words leapt to 75%—but we are dealing with only eight uses of six word types as a total data base, as compared with 68 word tokens each for Molly, Deborah, and Emily at Time 5, 60 for Sean, and 176 for Timmy.

Andrew's mono- and disyllabic curves are something like what we might expect to see for all the children if words and babble were unrelated phenomena in the children's development. Only Andrew displayed such a pattern; however, we must recall that he has the smallest number of recorded word tokens of all the children in sessions 1–6 (53 vs. Susie's 80), and next-to-smallest over-all (124 vs. Susie's 114). Andrew's word curve was thus constructed out of percentages based on very few data points. In contrast, Emily, who vocalized less than Andrew on average, made far greater use of words (123 word tokens at Time 7 alone, and 157 in all the preceding sessions); her curves are in fact as strikingly congruent as those of Sean or Deborah, say, who had considerably larger data sets and similarly high numbers of word types.

Fig. 2, Phonotactics, displays curves for the phonotactics of the syllable. It was possible to divide the data exhaustively into just three syllable shapes by subsuming all consonant-final syllables under '(C)VC', and all remaining consonant-initial syllables (including those with word-initial clusters) under 'CV'. For our purposes, the glottal segments [ʔ] and [h] are treated as non-consonantal, so that 'V' includes *ʔV*, *hV*, *Vʔ*, *Vh*, *ʔVʔ*, and so on. The phonotactic data suggest a developmental trend, unlike those on vocalization length. Specifically, the mean percentage of both word and non-word vocalizations of the shape V declines over the time of the study, while the mean percentage of CV shapes tends to rise in both words and non-words (see Table 6). Since most V

		T1	T2	T3	T4	T5	T6	T7
V	NW	65	61	51	55	53	52	48
	W	12	43	37	31	29	27	22
CV	NW	29	30	43	39	39	40	42
	W	86	54	56	61	57	63	68
(C)VC	NW	6	9	5	6	12	7	9
	W	1	3	7	8	14	10	10

TABLE 6.

NW = non-words, W = words. The children who used no words at a given time have been omitted from the calculation of a mean for that time. Thus the T1 average is based on 6 subjects, T2 on 8, and T3–T7 on 9. (T1 percentages for words are based on small corpora because few subjects used words at T1, and few words were used by each of those subjects.)

utterances include one or more glottals, their early high frequency means a correspondingly high frequency of glottals. This agrees with the studies of younger infants by Irwin & Chen 1946, based on phonetic transcription (cited in Kent & Murray), in which glottals account for a steadily decreasing proportion of total transcribed consonants, from 87% at 0–2 months, and 75% at 2–4 months, to 70% at 4–8 months. However, Kent & Murray caution (362) that the large number of glottals perceived in this period may reflect the infants' instability of laryngeal control, rather than genuine consonant production.

Fig. 3 represents a summary of the consonant manner categories, collapsing fricatives and affricates in the 'fricative' category and nasals, liquids, and trills in the 'resonant' category. Stops so completely dominate the true consonant category that the stop curve virtually recapitulates the CV category of Fig. 2, while the glottals do the same for V. Thus we see a developmental trend here for glottals and stops similar to the one in Fig. 2 for V and CV, respectively (see also Table 6). The remaining categories, though uniformly at a low percentage level, do generally show a remarkably close fit between word and non-word curves: See, in particular, the curves for fricatives for Sean, Deborah, Susie, and Emily, or the curves for resonants for Timmy and Thomas.

In Fig. 4, Consonant place, we display the data grouped into developmental stages rather than by seven anchor points. This grouping allows us to make a more direct comparison across the children, and also has the effect of eliminating the distortions mentioned earlier in connection with Fig. 1—i.e. those resulting from the small data base for words, and/or from the high use of one or two word types on a particular session. The chance of distortions occurring for those reasons is increased with regard to consonantal place, since the data base is reduced by the fact that only a subset of consonants is involved. Only stops, fricatives, nasals, and affricates were included, because only within these classes does English have segments which vary across all three of the places of articulation that we charted. This presentation eliminates a large percentage of the segments used in both word and non-word vocalizations (in particular, the glottals).

We chose the spontaneous use of four or more different word types within a single video session as our criterion for assigning a word-acquisition point to each child. We chose that number because, for almost all our subjects, it marked a point after which substantial gains were made in word acquisition at each succeeding anchor point. The session at which criterion was reached is marked with a double asterisk in Table 2 and with a vertical line in Figs. 1–3. For Fig. 4, stages were defined as follows: Stage 1 means that no words at all were used during the video session. Such a stage occurred within our chosen anchor points for 6 of the 9 subjects. Stage II means that 1–3 words were spontaneously used within the session—i.e. word use short of criterion. This stage occurred for all subjects. Stage III was the session in which criterion was reached, and Stage IV combined all succeeding sessions.

The picture afforded by Fig. 4 is simplified by comparison with the other figures, but the message is essentially the same: There were differences across subjects and across time in the use of the various consonantal place categories, but a striking over-all similarity in the word and non-word trends for any given subject and category. (Note that the statistics were drawn from the raw data, not the collapsed data as presented in Fig. 4).

4.3. DIFFERENCES BETWEEN WORDS AND NON-WORDS. Table 7 gives the results of Wilcoxon's signed rank statistic, used to test whether words or non-words consistently dominated across time in any category. The results show significant associations of (a) disyllables and words, and polysyllables and non-words (within the vocalization length parameter); (b) CV and words, and V and non-

	R	R*
Vocalization length		
mono	−7	−0.39
di	39	2.28*
poly	−45	−2.64**
Phonotactics		
V	−43	−2.52*
CV	45	2.64**
(C)VC	11	0.62
Consonant manner		
glottal	−45	−2.64**
stop	45	2.64**
glide	−25	−1.45
fricative	−9	−0.50
resonant	39	2.28*
Consonant place		
labial	45	2.64**
dental	−37	−2.16*
velar	−21	−1.21

TABLE 7.

R is the sum of the signed ranks. It must lie between −45 and 45, inclusive. Extremely large or small values lead to the rejection of the null hypothesis of no difference (dependence) between word and non-word percentages within a category. Negative values of R correspond to a greater average proportion of non-words in a category than words; the opposite is true for positive values of R. R* is the standardized version of R.

Critical values for a two-sided test of the null hypothesis are:

alpha = .05, values of R greater than 1.96 .
**alpha = .01, values of R* greater than 2.576.

words (phonotactic parameter); (c) stops and resonants and words, and glottals and non-words (consonantal manner parameter); and (d) labials and words, and dentals and non-words (consonantal place parameter). The structure of our data is such that each parameter is made up primarily of tokens in only two categories—except that, in the case of vocalization length, one category has a disproportionate number of tokens, while the other two categories contain a smaller but roughly equal number. Thus any consistent relation between words and non-words in one category results in the inverse relation in the second category of similar value, thereby giving an apparent difference between words and non-words in BOTH categories. (Note that percentages sum across categories for each parameter.) We will refer to this effect as a 'cross-category interaction'.

Considering the results from all three statistical measures, we can now ask: Within a given category, how is it that the sets of values for words and non-words are very close—and change over time in highly similar ways—but yet may stand in such a relation that the values for one variable are usually greater than those for the other over time? If we were to take the latter fact as evidence of independent systems, it would be exceedingly difficult to explain the first two facts. We conclude, rather, that the closeness of fit and movement together of word and non-word curves provide evidence for dependent systems, and hence for continuity; and the consistently greater incidence of a phonetic type in, say, words as opposed to non-words can be attributed to the internal structure of the given parameter.

Let us now examine more closely the findings of the third statistical test with respect to each of our parameters. Table 7 shows a significant difference between the curves in two categories for vocalization length. As one can see in Fig. 1, the word curve for disyllables generally lies above the non-word curve ($*p = .05$); for polysyllables, the inverse relation holds ($*p = .01$). Given the cross-category interactions in our data, we hypothesize that these findings actually result from a single factor, namely the relatively greater frequency of polysyllabic non-words.

Table 7 also shows a significant difference between word and non-word curves within the categories V and CV ($*p = .05$ and .01, respectively). For V, the non-word curve tends to lie above the word curve, with the inverse relationship for the CV category. Throughout our period, once children began to use words, there was a tendency for initial syllables in words to take the shape CV. Furthermore, the differences between V and CV was greatest for words. This is, then, another cross-category interaction, with a CV/word association incidentally creating a V/non-word association as well (recall that the category (C)VC is of minor importance). We speculate that the source of the association between CV and words lies in the phonotactic structure of the language or lexicon being learned, but we would need comparable cross-linguistic data to examine this issue.

Table 7 further shows a significant difference between word and non-word curves within the consonant manner categories. The non-word curve tends to lie above the word curve for glottals ($*p = .01$; cf. the V category for phonotactics), but below the word curve for stops (*p = .01; cf. the CV category for phonotactics) and for resonants (*p = .05). Since the high-frequency glottal and stop categories mirror so closely those of V and CV, we assume that these categories are non-independent also—and, similarly, that the inverse relation between them (glottals in non-words and stops in words) results largely from a single factor, such as a tendency for the early lexicon of English-speaking children to be composed of words with stop consonants. Given the large number of categories within the manner parameter and the minor importance of the resonant category, we will not attempt to interpret the results for resonants.

Finally, a significant difference between words and non-words again appears in Table 7 for consonant place. For all nine subjects, the percentage of labials in words at Stage IV was higher than the percentage of labials in non-words; but the inverse relation was true for the dental/palatal category, with one exception (see Fig. 4, Molly).[8] When we turn to the composite lexicon of word types used by the children across all sessions, we find that words with labials made up one-third or more of each child's lexicon as recorded on our anchor points (the mean is 38%); words with dentals and palatals made up about a quarter (mean 26%). If the maternal reports are also included, labials account for 42% (on average), dentals and palatals 26%, and all other consonants 32%. The labial words were more likely to be shared by two or more babies than the dental/palatal words (48% lexical overlap, vs. 37%).[9] The labial types were also repeated more frequently by the babies (15 tokens per labial type, on average, vs. 9.6 per dental/palatal type). Word-initial *b* alone accounted for 52% of the labial word types recorded for all the subjects and for 10% of all types, including vowel-initial words. Thus, we conclude that an association of words with labials accounts for both of the significant results within this parameter.

It is not clear why labial words should have so dominated the early vocabularies of our subjects. It may be that words with labials were used more often by adults interacting with the children, but this possibility has yet to be tested. Or it may be that labials are of high frequency generally in the adult language. English word-initial /b/ and /m/ have been reported to be more frequent than /d/ and /n/, both lexically and in running conversation (cf. Locke; and cf. Macken 1980 for discussion of distributional asymmetries among stop consonants, and their significance for acquisition issues). Or labials may be preferred, once the children began to acquire adult words in larger numbers (i.e. in the period following the criterion session), because of their special

[8] We are grateful to John Locke for pointing out a general trend toward labials in words (cf. Locke 1983).

[9] The most commonly shared words in the composite lexicon were, perhaps not surprisingly, those that named the two central characters on our recordings—*baby* and *mama*, each recorded for eight babies and reported by the mother for the ninth. Nine babies also used *bird* (7 children's uses were recorded, 2 were reported only), while eight used *up* (sometimes with a labial), *daddy* and *no* ('reported only', for 2 babies), and *nose* ('reported only', for 3). Seven used *ball, block,* and *eye* ('reported only' for one baby each); six used *box, down, kitty, quack, uh-oh,* and *woof*; and five were recorded as using *moo* and *yeah*.

perceptual salience: they can be seen as well as heard, as it were. Mills' comparative data (1984) on use of initial consonants in the early phonological development of blind and sighted children supports this last interpretation.

We see, then, that certain categories show specific association with words (e.g. use of labials) or with non-words (polysyllables, glottals; see Table 7). Yet the overriding picture is one of continuity, with words and non-words exhibiting closely similar patterns, for any one subject, both at any given time and over time (see Tables 4–5). Thus, while there are some differences between words and non-words, in most cases the children can be seen to draw both word and non-word vocalizations from a single phonetic repertoire.

Discussion

5. Let us now consider again the two contrasting pictures of the transition period: that evoked by Jakobson, and that which arises from our data. First, as Table 2 clearly indicates, the gradual acquisition of words took place concurrently with the continued use of babble for all our subjects. There was a shift to a predominance of words for two of the nine children at Time 6 (Molly and Emily) and for three more at Time 7 (Sean, Timmy, and Deborah); the remaining children had 23% to 41% words at Time 7, suggesting that the shift could be expected within a month, or at most two. Jakobson did not deny the possibility of concurrent word use and babble, though he noted that a 'silent period' had been reported for some children. We did not find such a period in any of the ten children we observed.

As for setting, we noted earlier the wide variation across our subjects with respect to quantity of vocalization during solitary play. In addition, the children varied in their use of babble after they had begun to acquire words. Finally, we observed that word use, insofar as it can be reliably identified on purely audio recordings, was lower during solitary play. There is, however, no clear evidence of babble being restricted to solitary play within the period of our study—though this might be a later pattern, at least for some of the children, as babble was gradually abandoned altogether.

With regard to systematicity in the sequence of emergence of babbling sounds, Ferguson & Macken 1983 have reviewed the evidence in the literature, concluding that there is an orderly 'development of "speechiness" in babbling parallel to development in permitted phonetic complexity in speech' (143). Our data confirm this conclusion. For example, the development of phonotactics (Fig. 2), which was roughly the same for all the children, showed movement mainly from V to CV and then to (C)VC, both in babbling and in words.[10] Elbers' careful 1982 study of her Dutch son's 'repetitive babbling' (6–12 months) represents an important further confirmation of the systematicity of development in this period.

[10] For an account of systematicity in the development of vowels in the babbling period, see Vanvik 1971, Jeng 1981. Jeng in fact reports finding the SAME order of emergence of (babbled) vowels in his Mandarin-speaking subjects as Vanvik had reported for Norwegian; we have found no single, set order of emergence across subjects for consonants, but have not investigated the issue with reference to vowels.

To illustrate another point made by Ferguson & Macken (143)—'the appearance (or increased frequency) of sound types in a child's babbling which are about to appear or have just appeared in his/her speech'—consider Jonah's use of labials in babbling and in speech. At 9 months, Jonah repeatedly imitated [ba] and [ma] as [da] and [na], respectively. For the first three anchor-point sessions, Jonah used only 11% or 12% labials in babble (0–5% in words); but at Time 4, he used 56% (still 4% in words). (Between these two sessions, when Jonah was 1;0.16, his mother reported as new the use of *b* in long babbled strings.) At his criterion session a month later (Time 5), Jonah had dropped back to 17% use of labials in babble; but 51% of his word tokens now were *b*-initial (⟨ba(ba)⟩ for *bow-wow* and *rock-rock*, mainly). The same point could be illustrated by the emergence of fricatives, at a lower frequency level, first in babble and then speech, for either Deborah or Susie.

The 'variety' of babble sounds is certainly somewhat greater—as measured, say, by the number of empty cells in the full table of consonantal manner categories (stop, fricative, affricate, nasal, liquid, glide, trill, and glottal). Specifically, there are twice as many empty cells for words as for non-words. However, the number of word tokens is only one-third the number of non-word vocalizations (3717 out of 12,037). And for two of the babies—Molly and Sean, who had high numbers of word tokens—the number of filled cells is either the same (Sean) or actually greater for words than for non-words (Molly); thus it is certainly inaccurate to refer to the 'rigorous sparseness of speech sounds' (Jakobson & Waugh, 62) on the basis of these data.

Jakobson wrote of the 'wild sounds' of the babbling period, which must undergo 'a kind of deflation ... into entities of linguistic value' (25). Similarly, Oller et al. (9) confess that they 'found in the babbling data many examples (though not in high proportions) of phonetic elements which are not particularly common in meaningful child speech and for which the languages of the world have not shown preferences. Among these were syllabic nasals, bilabial trills ... and labio-lingual "consonants".' They go on to express perplexity that 'these elements might appear frequently in child babbling but not in meaningful child speech.' This paradox does not occur in our data. Though we also find 'wild sounds' of the kind Oller et al. describe, and to which Jakobson also referred, we do not find that they are restricted to non-words. On the contrary, they occur quite commonly in the onomatopoeic words, such as *yum, vroom,* or the various animal sounds. Furthermore, both Emily and Timmy contrived to incorporate an idiosyncratic trill into their renditions of animal names—*duck* and *fish*, respectively—and in Emily's case the trill was actually combined with a conventional string of segments. The notion that the 'wild sounds' occur only in babble may derive from the relative infrequency of instrumental records of the transitional period, as well as from a definition of language that excludes them a-priori.

Finally, what evidence do we have regarding the 'individual cases' about which Stoel-Gammon & Cooper wondered, with respect to the continuity from babble to the choice of sounds in early words? Let us look at the use of velars and liquids specifically, to follow up on their question. Neither of these segment types occurs commonly in early words, so it is sensible to look at their use in babble by those children who do use them in words—or at their use in words by those children who did use them in babble. From Fig. 3 we learn that, at Stage II, just before word use began in earnest, three of our nine subjects (Molly, Deborah, and Andrew) were making relatively high use of velars: Velars accounted for about one-third of the stop, fricative, affricate, and nasal consonants that occurred in their babble (the remaining children used less than one-third velars). Those same three subjects also used about one-third or more velars in their earliest recorded word productions. By comparison, five of the remaining subjects did not use velars at all in their earliest words, while the sixth had 12% velars.

In the last session of Stage II, immediately preceding the criterion session, each of the three children had a unique set of velar words, and each for the first time used at least one velar word spontaneously. Molly imitated *kitty* and made both spontaneous and imitative use of *cracker* (several tokens, some with two velar stops). Deborah used *chicken* (imitated only, two tokens, both monosyllables with initial velar stops), and made both spontaneous and imitative use of *duck, doggy,* and *Capa* (a bath toy), with several tokens each, all with initial [k] or [g]. Andrew used *chicken, cookie, clock, cow,* and *kitty* (imitated only, one token each), and *catch* (spontaneous as well as imitated, 4 tokens); all these words took the form CV, where C is a velar stop. In summary, a relatively high frequency of velars in babble did correlate with frequency in early word production.

The story with respect to liquids is considerably simpler. Of the nine subjects, only three made 10% or greater use of liquids (out of all consonant manners) in either words or non-words at any one session. For two of them, this occurred once only (Jonah, 13%—6 liquids out of 48 consonants—at Time 5, in babble; Timmy, 11%—one liquid out of 9 consonants—at Time 1, in the word *block*). For the third child, Sean, use of liquids was conspicuous in babble from the start, when he frequently used a tongue-tap series that sounded much like 'diddle-diddle ...' Sean's percentage use of liquids across time in words and babble is shown on Table 8.

TIME:	1	2	3	4	5	6	7
Non-words	15%	10%	8%	7%	4%	5%	6%
Words	–	25%	8%	27%	5%	3%	3%

TABLE 8.

We see a gradual but fairly steady decrease in use of liquids in babble over time. In words, usage is less evenly patterned; but we assume that the high-frequency use found at Times 2 and 4 must be based on the same facility or affinity for liquids that Sean exhibited somewhat earlier in his babble. Specifically, at Time 2 the liquids occurred in apparent imitative and spontaneous attempts at the word *doggy*, which took the form ⟨ʔəɹɪʰ⟩ or ⟨l̥dVlV⟩, as well as one imitation of *bottle* as [ʔɪlœʰ]. At Time 3, Sean imitated the word *ruff* (for a dog's bark), using an [r] in two tokens. At Time 4, [r] occurred in variants on (*v*)*room*, of the form ⟨(ru)bru:⟩. At Time 5, the criterion session, Sean used an [l] in his variants of *cock-a-doodle-do* and *little girl*, as well as an [ř] in *birdie*. [l] occurred in some variants of *yellow* at Time 6, and again in *cock-a-doodle-do* at Time 7 ([dəludəlu] ~ [du:dl̥du]), as well as in imitations of *blue* [blu] and *clue* [dluk].

The evidence clearly suggests that an individual child's babble repertoire will be reflected, at roughly the same frequency levels, in choice of adult words to say and in the phonetic rendition of those words. The underlying process in at least some cases may be that which is illustrated by our earlier discussion of Sean's 'limited use' of *down* and *thank you* (see §2, above). In that instance we have an excellent example of words emerging naturally from the sound system which the child has been developing in his babble, when an adult word which is perceptually salient (i.e. which occurs under sentence stress, and which relates to a cognitively salient object or event) comes close to matching the range of sound shapes the child has been producing. We take all this as concrete illustration of what MacNeilage meant when he speculated (30) that 'the child's first words can be seen as, at least partially, a matter of choosing from the babbling repertoire a set of approximations to adult word forms.'

We have not discussed the phonemic status of segments used in early words. Analysis into 'clear and stable phonological oppositions' (Jakobson, 25) is notoriously difficult for the small data sets which the early words represent (cf. Ferguson & Farwell 1975, Kiparsky & Menn 1977). If we were to attempt such an analysis, problems similar to those which confronted us in the course of the present study would immediately arise, e.g. whether to exclude imitations or onomatopoeia, and how to deal with the wide range of phonetic variants typical of the early language use of some of the children. However these problems were resolved, it would necessarily be the case that the set of segments involved in the first phonological oppositions would constitute a subset of those used in the early words—and thus of the data analysed under the rubric 'words' in this paper. For example, Deborah, the only child with as many as three spontaneous word types at Time 1, opposed *b* (*baby* [pe ~ pabi ~ peipei ~ pepɪ ~ pepewe:pipepe]) and *m* (*mama* [memi]) at this time (the third word was *uh-oh*). That *m* and *b* were also of high frequency in Deborah's babbling is suggested by the high percentage of labials used (cf. Fig. 4, II), though none of our charts directly portrays specific segments. We are confident that a reanalysis of our word-based data in phonemic terms, with a parallel re-count of segments used in co-occurrent babble for comparison, would result in no significant change in our conclusions.

Before concluding, we should briefly clarify our understanding of the meaning of our data with respect to the issues raised by behaviorists or learning theorists. We believe that no reliable inferences regarding the source of the observed similarity between word and non-word vocalizations, or babble and language, can be drawn from our analyses to date. To make such inferences, we would require additional data or analyses with a different focus. For example, to determine the validity of a concept like 'selective reinforcement' (cf. Winitz 1969), we would need to look extensively at the mothers' speech. To judge how active a role the child plays (active hypothesis tester rather than passive recipient of the caretakers' 'shaping'), we would need contrastive data of some kind—from other communities or other languages, say. We must leave such investigation for the future.

CONCLUSION

6. Let us attempt to place the evidence we have detailed here into a more coherent framework. We see the transition period as a time in which three important strands that go to make up language first come together. Those strands are pragmatics, meaning here the development of the notion of communication; phonology, or the development of the notion of language-like sound shapes and units (especially prosody, syllables, and segments); and semantics, or the notion of what it is to 'mean' (cf. Halliday 1975, who emphasizes continuity in content or 'meaning potential' across this transition period). In denying basic discontinuity between babble and the sound base of the early words, we do not intend to deny the possibility of discontinuity in other aspects of this development. An early discontinuity on the phonetic dimension, for example, is the onset of syllable production (between 6 and 9 months); most observers agree that this is a sudden development which can often be precisely dated for a given child (Oller 1980, Stark 1980, Vihman 1976). Considerably later, rule- or contrast-based phonological systematization and reorganization typically occur, though this is not likely to be a sudden development with distinguishable starting or ending points (cf. Kiparsky & Menn 1977, Vihman 1981, Macken & Ferguson 1983).

In the lexical/semantic domain, a change which occurs some time after the earliest period of word use has been seen by some observers as a sharp discontinuity. This is what Dore 1978 has called the 'nominal insight' (cf. also McShane 1980:49ff.): the 'discovery that things have names' (cf. Bates 1979:33—who does not, however, 'see the discovery of naming as a discontinuous leap'; 40). One manifestation of this is the sudden interest which some children show in establishing clearly for themselves that a given lexical type does have diverse physical referents. At least one of the children in our study, Jonah, seemed to experience a 'revelation' of the naming principle, at about 16 months. At that time, according to his mother's report, he began to point to an object, such as a ball or a butterfly, while pointing with the other hand to a picture of it in a book. (For further illustration, see Bates 1976.) Non-verbal manifestations of symbolic thought, such as 'pretend play' with stuffed animals and other objects, typically begin to occur around the same time (cf. McCune-Nicolich 1981), as does a spurt in rate of lexical acquisition.

And yet, with further research, all these developments may come to be seen as, in some sense, continuous. For example, the first two-word combinations are not currently seen as an abrupt development. Many children go through an extensive period of use of successive single words in 'vertical constructions' (Scollon 1976; cf. also Bloom 1973, Branigan 1979). They may also use associative rather than syntactic word combination principles at first, based on context and meaning without regard for language-specific word-order constraints (cf. Ramer 1976). They may incorporate unanalysed or formulaic unitphrases into their lexicon before showing evidence of free word-combination (cf. also the various 'presyntactic devices' cited in Dore et al.) The gradualness or blurriness of the transition from 'one word at a time' to word combination seems quite analogous to the continuity we find in the phonetic base for the 'first word'. The difficulty of distinguishing words from babble, the incorporation of words into a babbled string, the emptying of meaning from a word-based form (cf. also Menyuk & Menn 1979)—all these speak to the essential continuity of this transition.

Jakobson's notion of an abrupt change from babbling to speech is based on a theory of linguistic structure which requires a sharp distinction between random phonetic variation (babble, in Jakobson's terms) and tightly constrained phonemic production (words). It may also have originated in part in the diarists' tendency to disregard babble, once words begin to appear in the child's speech.[11] We do not dispute the fact of phonological structure; but we see it as emerging later for the child, after he or she has begun to make use of a sizable adult-based lexicon (cf. Ferguson & Farwell 1975, Ferguson 1978, Menyuk & Menn 1979). By contrast, we see a far earlier growth of the 'desire for communication', the development of which Jakobson appears to place at roughly the point of emergence of 'arbitrary sound distinctions aimed at meaning which require simple, clear, and stable phonological oppositions' (24–5). Certainly, for all our children, the productive pairing of sound and meaning antedated the appearance of the first adult-based words. In the time between the onset of communicative intentions (as well as the beginnings of language comprehension) and the restructuring of the early word-based phonology into the beginnings of a segmental contrast-based and rule-governed system like the adults', we see a gradual broadening of phonetic skills; the development of well-controlled sound patterns loosely based on the input language; and the use of certain patterns for the expression of rather broad, generally performative or affective meanings—as well as the acquisition of the first adult-based words. We interpret the continuity of this development as rooted in a dawning notion of what language sounds like: a somewhat different notion for each child, as expressed in the differing favorite shapes, segments, and prosodies. There were shifts and changes in these over the transition period, but they did not tend to divide words and non-words; they could and did occur within either category, and typically in both. Thus the famed division between babbling and language appears to reflect characteristics of adult perception and interpretation, rather than developmental processes manifested by the child.

Appendix A.

A test for closeness of fit of the data points for words and non-words within each of four parameters was carried out in the following way: First, for each subject on each of the four parameters, we defined p as a vector of proportions summing to 1 (or 100%), for words—e.g., p_1, p_2, p_3 = proportions of mono-, di-, and poly-syllables; and we defined p' as an analogous vector for non-words. It was then necessary to devise a distance function, defining precisely the difference between the word and non-word curves across time. Two such distance functions were implemented (omit-

[11] We thank Eve Clark for this helpful suggestion.

ting in each case those sessions in which no words were used): (1) the average absolute difference, for each word and non-word pair, for each time; and (2) the square root of the average squared difference for each pair. To assess the relative size of the distance functions so defined, a standard had to be created. In addition to the nine intra-subject comparisons, then, a distance function was also computed for each between-subject pair (e.g. non-word percentages for Subject A, and word percentages for Subject B). There are 72 of these, since each of the 9 subjects is matched with each of the other 8. The question to be tested, then, is whether the intra-subject distance functions are smaller than the inter-subject differences. In other words, we are testing whether the individual parallel patterns between word and non-word vectors (curves) are in fact stronger than the overall similarity in patterning for the age group as a whole.

For each of the four parameters—(1) vocalization length in syllables, (2) phonotactics, (3) consonant manner, and (4) consonant place—and for each of two distance functions, we have a 9 × 9 matrix (8 altogether) of the form shown in Table 9.

	W	1	2	3	4	5	6	7	8	9
NW										
1		$d(1,1)$								$d(1,9)$
2										
3										
4										
5					$d(i,j)$					
6										
7										
8										
9		$d(9,1)$								$d(9,9)$

TABLE 9. $d(i,j)$ = measured distance between non-word curve for subject i and word curve for subject j.

If the intra-subject distances are small relative to the inter-subject ones, then the sum of the entries on the diagonal of this matrix should be small, relative to the sum of entries if the diagonal were NOT composed of intra-subject entries. Thus, $d(1,1) + d(2,2) + \ldots d(9,9)$ should be small compared to $d(1,3) + d(2,7) + d(3,4) \ldots$ etc.—since the first set of individual entries should be smaller, on average, if the hypothesis of continuity is to be borne out.

For each of the 8 9×9 matrices, then, the sum of the INTRA-SUBJECT distances on the original diagonal was computed. For comparison, 2000 sums of diagonals produced from matrices with randomly permuted columns were computed (2000 being a manageable subset of the 366,000 possible permutations). Table 4 shows the results of the test. The figures given indicate the number of randomly permuted diagonal sums, out of 2000, which were smaller than the original, intra-subject sum. There were no ties. (Note that the two distance functions are strongly correlated: $r = .96$ to $.98$).

APPENDIX B.

Spearman's rank correlation coefficient is the correlation coefficient between the ranks of two variables. It is thus appropriate for data defined on at least an ordinal scale, e.g. the percentages of word and non-word utterances falling into various categories (e.g. glottal of consonant manner) over time.

Under the assumption that the two variables in question, words and non-words, are independent (our null hypothesis), the rank correlation coefficient has mean 0 and variance $1/(N-1)$.[12] If we have a number of similar correlation coefficients and wish to test them as a group for significance

[12] N is the number of observations upon which estimates of the correlation coefficient are based, i.e. the number of word/non-word point pairs for a given child. For Timmy, Deborah, and Jonah, N = 7; for Molly, Sean, Susie, and Emily, N = 6, and for Andrew, N = 5. Finally, for Thomas, N = 4, since none of his utterances were classified as words for the first three sessions.

(non-independence), we can use the fact that under the null hypothesis, the sum of the rank correlation coefficients is distributed approximately normally with mean 0 and variance equal to the sum of the variances of its component parts.

This is our strategy: Compute the rank correlation coefficient between word and non-word vectors for each subject over time and within a particular category of a measure (e.g. glottal, of consonant manner). Sum these coefficients for all subjects, and determine whether these sums are sufficiently large to warrant rejection of the null hypothesis of independence. The variance of each of the sums (shown in Table 5) is 3(116) + 4(115) + (114) + (113) = 1.883, as each correlation coefficient contributes $1/(N-1)$. The standard deviation of each of the sums is 1.883, or 1.372. Critical values for a one-tailed test can now be determined by use of a table of the normal distribution. For alpha = .05, the computations are $1.372 \cdot 1.645 = 2.258$. For alpha = .01, the computations are $1.372 \cdot 2.326 = 3.192$.

For example, the rank correlation coefficient between word and non-word vectors (curves) has been computed for each child within the category glottal of consonant manner (see upper left column of Table 5). If the word and non-word vectors were independent for all children, we would expect the sum of these correlation coefficients to be approximately normally distributed with mean 0 and standard deviation 1.372. The observed value of 2.452 is 1.79 standard deviations away from 0. Since we are interested only in the presence or absence of positive (not negative) correlation, we perform a one-tailed test of the hypothesis of independence. Thus, we would reject the null hypothesis at alpha .05 (since 2.452 is larger than the critical value of 2.258), but would not reject the null hypothesis at alpha .01 (since 2.452 is less than the critical value of 3.192).

Appendix C.

If one of two curves tends to lie above the other, then the average value of the 'higher' curve should be greater than that of the 'lower' curve. We shall address the possibility of this characteristic in our data by considering whether the average percentage of words, within a particular category of a parameter, differs significantly from that of non-words in that category. We compute weighted averages of the percentage values shown in Figs. 1–4, with weights proportional to the total number of observations (word or non-word utterances) on which a percentage is based. (We thus show the average percentages of word and non-word utterances within each category over all periods).

For each category, we will have nine pairs of averages (one pair for each subject). In order to assess whether there is a significant difference between word and non-word percentage averages for the subjects as a group, we will use Wilcoxon's signed rank statistic. Use and interpretation of this statistic are as follows.

For example, consider the word and non-word vectors (curves) within the glottal category of the consonant manner parameter. The weighted averages of the word and non-word vectors are shown in columns 1 and 2 on Table 10. Column 3 shows the differences between word and non-word averages (column 1 minus column 2). We temporarily ignore the signs of these differences and rank their absolute value from 1, for the smallest, to 9, for the largest (ties are averaged). We now affix the sign of the difference to the ranks. These signed ranks are shown in column 4.

SUBJECT	COLUMN 1	COLUMN 2	COLUMN 3	COLUMN 4
Molly	17.8	44.6	−26.8	−4
Sean	29.0	56.6	−27.6	−5
Timmy	25.9	57.0	−31.1	−6
Deborah	40.1	45.0	−4.9	−1
Thomas	46.3	61.4	−15.1	−2
Susie	27.7	43.8	−16.1	−3
Jonah	17.3	51.8	−34.5	−8
Andrew	18.0	62.1	−44.1	−9
Emily	22.9	54.7	−31.8	−7

TABLE 10.

If word and non-word averages were distributed alike (our null hypothesis), we would expect that the sum of these signed ranks should be near zero. Wilcoxon's signed rank statistic allows us to assess whether word and non-word averages seem to be distributed differently, by using the known distribution of the statistic $R^* = R \pm \frac{1}{2} / \sqrt{N(N+1)(2N+1)/6}$.[13]

In this instance, $R^* = -45 + \frac{1}{2} / \sqrt{(9 \cdot 10 \cdot 19)/6} = -2.64$. Critical values for a two-sided test of the null hypothesis are 1.96 for alpha = .05 and 2.576 for alpha = .01. Thus we would reject the null hypothesis of no difference between word and non-word curves of this category. Instead, we conclude that non-words are associated with the category glottal to a greater extent than are words.

REFERENCES

ATKINSON, KAY; BRIAN MACWHINNEY; and CAROLINE STOEL. 1970. An experiment on the recognition of babbling. Papers and Research in Child Language Development 1.71–6.

BATES, ELIZABETH. 1976. Language and context: The acquisition of pragmatics. New York: Academic Press.

——. 1979. The emergence of symbols: Cognition and communication in infancy. New York: Academic Press.

BLOOM, LOIS. 1973. One word at a time: The use of single-word utterances before syntax. The Hague: Mouton.

BRANIGAN, GEORGE. 1979. Some reasons why successive single word utterances are not. J. Ch. Lg. 6.411–22.

BUSH, CLARA N., et al. 1973. On specifying a system for transcribing consonants in child language: A working paper with examples from American English and Mexican Spanish. Stanford University: Department of Linguistics.

CLARK, EVE V. 1973. What's in a word? On the child's acquisition of semantics in his first language. Cognitive development and the acquisition of language, ed. by Timothy E. Moore, 65–110. New York: Academic Press.

CRUTTENDEN, ALAN. 1970. A phonetic study of babbling. British Journal of Disorders of Communication 5.110–17.

DE BOYSSON-BARDIES, BÉNÉDICTE; LAURENT SAGART; and NICOLE BACRI. 1981. Phonetic analysis of late babbling: A case study of a French child. J. Ch. Lg. 8.511–24.

DE BOYSSON-BARDIES, BÉNÉDICTE; LAURENT SAGARD; and CATHERINE DURAND. 1984. Discernible differences in the babbling of infants according to target language. J. Ch. Lg. 11.1–15.

DORE, JOHN. 1978. Cognition and communication in language acquisition and development. Paper presented at the Boston University Conference on Language Development.

——; MARGERY B. FRANKLIN; ROBERT T. MILLER; and ANDRYA L. H. RAMER. 1976. Transitional phenomena in early language acquisition. J. Ch. Lg. 3.13–28.

ELBERS, LOEKIE. 1982. Operating principles in repetitive babbling: A cognitive continuity approach. Cognition 12.45–63.

ENGEL, WALBURGA VON RAFFLER. 1973. The development from sound to phoneme in child language. Studies of child language development, ed. by Charles A. Ferguson & Dan I. Slobin, 155–8. New York: Holt, Rinehart & Winston.

FERGUSON, CHARLES A. 1978. Learning to pronounce: The earliest stages of phonological development in the child. Communicative and cognitive abilities: Early behavioral assessment, ed. by Fred D. Minifie & Lyle L. Lloyd, 273–97. Baltimore: University Park Press.

——, and CAROL B. FARWELL. 1975. Words and sounds in early language acquisition. Lg. 51.419–39.

——, and MARLYS A. MACKEN. 1983. The role of play in phonological development. Child language IV, ed. by Keith Nelson, 231–54. Hillsdale, NJ: Erlbaum.

[13] R = the sum of the signed ranks; N = number of subjects. 1/2 is added to R if R is less than zero, or subtracted if R is greater than zero. This statistic follows the standard normal distribution quite well for N of at least 8.

HALLIDAY, MICHAEL A. K. 1975. Learning to mean: Explorations in the development of language. New York: Elsevier North-Holland.
INGRAM, DAVID. 1974. Phonological rules in young children. J. Ch. Lg. 1.49–64.
——. 1979. Phonological patterns in the speech of young children. Language acquisition, ed. by Paul Fletcher & Michael Garman, 133–48. Cambridge: University Press.
IRWIN, ORVIS C. 1947a. Infant speech: Consonant sounds according to place of articulation. Journal of Speech Disorders 12.397–401.
——. 1947b. Infant speech: Consonant sounds according to manner of articulation. Journal of Speech Disorders 12.402–4.
——, and HAN PIAO CHEN. 1946. Infant speech: Vowel and consonant frequency. Journal of Speech Disorders 11.123–5.
JAKOBSON, ROMAN. 1941. Kindersprache, Aphasie und allgemeine Lautgesetze. Uppsala. [Translated by Allan R. Keiler: Child language, aphasia, and phonological universals. The Hague: Mouton, 1968.]
——, and LINDA WAUGH. 1979. The sound shape of language. Bloomington: Indiana University Press.
JENG, HENG-HSIUNG. 1981. From babbling to speech: A case study of the acquisition of Chinese phonology. MS.
KENT, RAYMOND D., and ANN D. MURRAY. 1982. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. JASA 72.353–65.
KIPARSKY, PAUL, and LISE MENN. 1977. On the acquisition of phonology. Language learning and thought, ed. by John Macnamara, 47–78. New York: Academic Press.
LABOV, WILLIAM, and TERESA LABOV. 1978. The phonetics of *cat* and *mama*. Lg. 54.816–52.
LEOPOLD, WERNER F. 1939. Speech development of a bilingual child: A linguist's record. Vol. I, Vocabulary growth in the first two years. Evanston: Northwestern University Press.
LOCKE, JOHN. 1983. Phonological acquisition and change. New York: Academic Press.
McCUNE-NICOLICH, LORRAINE. 1981. Toward symbolic functioning: Structure of early pretend games and potential parallels with language. Ch. Dev. 52.785–97.
MACKEN, MARLYS A. 1980. Aspects of the acquisition of stop systems: A cross-linguistic perspective. In Yeni-Komshian et al., 143–68.
——, and CHARLES A. FERGUSON. 1983. Cognitive aspects of phonological development: Model, evidence, and issues. Child language IV, ed. by Keith Nelson, 255–82. Hillsdale, NJ: Erlbaum.
MACNEILAGE, PETER F. 1979. Speech production. Report to the 9th International Congress of Phonetic Sciences, Copenhagen.
McNEILL, DAVID. 1970. The acquisition of language: The study of developmental psycholinguistics. New York: Harper & Row.
McSHANE, JOHN. 1980. Learning to talk. Cambridge: University Press.
MENN, LISE. 1976. Pattern, control and contrast in beginning speech: A case study in the development of word form and word function. Urbana: University of Illinois dissertation.
MENYUK, PAULA, and LISE MENN. 1979. Early strategies for the perception and production of words and sounds. Language acquisition, ed. by Paul Fletcher & Michael Garman, 49–70. Cambridge: University Press.
MILLS, ANNE E. 1984. The development of phonology in the blind child. To appear in Hearing by eye, ed. by Barbara Dodd & Ruth Campbell. Hillsdale, NJ: Erlbaum.
OLLER, D. KIMBROUGH. 1980. The emergence of the sounds of speech in infancy. In Yeni-Komshian et al., 93–112.
——; LESLIE A. WIEMAN; WILLIAM J. DOYLE; and CAROL ROSS. 1976. Infant babbling and speech. J. Ch. Lg. 3.1–11.
RAMER, ANDRYA L. H. 1976. Syntactic styles in emerging language. J. Ch. Lg. 3.49–62.

SCOLLON, RONALD. 1976. Conversations with a one-year-old: A case study of the developmental foundation of syntax. Honolulu: University of Hawaii Press.

STARK, RACHEL E. 1980. Stages of speech development in the first year of life. In Yeni-Komshian et al., 73–92.

STOCKMAN, IDA J.; DAVID R. WOODS; and ABRAHAM TISHMAN. 1981. Listener agreement on phonetic segments in early infant vocalizations. J. Psycholing. Res. 10.593–617.

STOEL-GAMMON, CAROLINE, and JUDITH A. COOPER. 1984. Patterns of early lexical and phonological development. J. Ch. Lg. 11.247–71.

UŽGIRIS, INA C., and J. MCV. HUNT. 1975. Assessment in infancy: Ordinal scales of psychological development. Urbana: University of Illinois Press.

VANVIK, ARNE. 1971. The phonetic-phonemic development of a Norwegian child. Norsk Tidsskrift for Sprogvidenskap 24.269–325. Oslo: Universitetsforlaget.

VIHMAN, MARILYN MAY. 1976. From pre-speech to speech: On early phonology. Papers and Research in Child Language Development 12.230–43.

——. 1981. Phonology and the development of the lexicon: Evidence from children's errors. J. Ch. Lg. 8.239–64.

WINITZ, HARRIS. 1969. Articulatory acquisition and behavior. New York: Appleton-Century-Crofts.

WUNDT, WILHELM. 1912. Völkerpsychologie. Stuttgart.

YENI-KOMSHIAN, GRACE H., et al. 1980 (eds.) Child phonology I: Production. New York: Academic Press.

[Received 17 June 1983;
revision received 14 May 1984;
accepted 3 December 1984.]