Comparison of Perceptual and Acoustic Analyses of Two Infants' Phonetic Ability

Prédiction des retards du langage: Le rôle éventuel de l'analyse acoustique

Susan Rvachew, MSc
Alberta Children's Hospital
University of Calgary

Elzbieta B. Slawinski, PhD
University of Calgary

Megan Williams, MSc, Carol L. Green, MSc
Alberta Children's Hospital

Key words: prelinguistic speech development, babbling, infants, phonetic development, speech analysis, language delay

Abstract

This paper describes the early speech development of two normally developing infants during the period six through 18 months of age, one of whom suffered chronic otitis media with effusion during the first year of life. Four different methods of phonetic and acoustic analysis of speech were used. The infraphonological (Oller, 1986) and vowel formant analyses proved most sensitive to differences in speech development for the two boys. The babbling level analysis (Stoel-Gammon, 1989) was quite sensitive to differences in phonetic development when the subjects were 15 and 18 months old, but may not be appropriate for younger infants. The phonetic contrast estimators (Bauer, 1988) appear to be a useful research tool but did not yield results that varied with the subjects' developing language abilities in this study.

Abrégé

Le présent article décrit les débuts du langage chez deux nourrissons au développement normal entre l'âge de 6 et de 18 mois. Un des deux enfants avait souffert d'otite moyenne chronique avec épanchement au cours des première année de vie. On s'est servi de quatre méthodes distinctes d'analyse phonétique et acoustique du langage. L'analyse infraphonologique (Oller, 1986) et l'analyse de la formation des voyelles se sont avérées les plus sensibles pour déceler les variations du développement du langage chez les deux garçonnets. L'analyse du degré de lallation (Stoel-Gammon, 1989) est très sensible aux fluctuations du développement phonétique quand les sujets atteignent 15 et 18 mois, mais ne convient pas aux enfants plus jeunes. Enfin, l'évaluation des phonèmes par contraste (Bauer, 1988) semble également un outil de recherche utile, mais les résultats ont fluctué avec l'aptitude de développement du langage chez les deux sujets dans le cadre de cette étude.

Prelinguistic speech (i.e., babble) has been viewed previously as a largely reflexive behaviour of little if any linguistic significance. For example, Locke (1983) reviewed a large number of studies which found no evidence for individual differences in the development of babbling, even as a consequence of variations in the infants' auditory environments or hearing abilities. During the past decade this view has been altered dramatically with the development of new theoretical approaches along with more sophisticated techniques for describing infant speech.

Recent research confirms the earlier view that the course of infant speech development is dependent upon anatomical and neuromotor maturation of the articulatory system (Kent, 1992). However, it has also been shown that prelinguistic speech is influenced by the infant's auditory environment. The nature of the auditory environment is determined by a number of infant, caregiver, and environmental characteristics, including the infant's hearing status and the linguistic behaviour of its caregivers. For example, de Boysson-Bardies, Halle, Sagart, and Durand (1989) found that the acoustic characteristics of vowels produced by 10 month-old infants are significantly different across groups of infants exposed to English, French, Arabic, or Cantonese (see de Boysson-Bardies, Sagart, Halle, and Durand, 1986, and de Boysson-Bardies, Vihman, Roug- Hellichius, Durand, Landberg, and Arao, 1992, for further evidence of "babbling drift").

The role of the auditory environment in prelinguistic speech development is further illuminated by studies of infants with profound hearing impairment. Oller and Eilers (1988) found that the onset of canonical babbling is delayed significantly among these infants in relation to their normal hearing peers. Stoel-Gammon and Otomo (1986) collected samples longitudinally from normal hearing and hearing-impaired infants during the period four through 15 months of age and calculated the number of consonantal types per sample. They found that infants with normal hearing had larger phonetic repertoires and produced a higher proportion of true consonants than did hearing-impaired infants throughout this period. Similar findings are reported by Kent, Osberger, Netsell, and Hustedde (1987) in their longitudinal study of twins, one of whom was profoundly hearing-impaired. They also found that the vowel space expanded over time in the speech of the normal hearing twin, while the vowel space became increasingly restricted in the speech of the hearing-impaired twin. Thus, current research contradicts previous opinions and demonstrates that infants differ in the development of babbling. These individual differences can be due to such factors as the infant's hearing status and the auditory environment.

The linguistic significance of babble has been highlighted in studies which suggest that poor quality babble may predict delayed speech development in young children. Whitehurst, Smith, Fischel, Arnold, and Lonigan (1991) analyzed the phonetic characteristics of babble produced by two year-old children with specific expressive language delay, and then used this information to predict individual differences in the children's language abilities 5 months later. They reported that the "single strongest correlate of language outcome was the proportion of consonantal to vowel babble" (p. 1121). Furthermore, Stoel-Gammon (1989) observed that two infants who displayed atypical patterns of prelinguistic speech development later demonstrated delayed phonological and language abilities at age 24 months.

Paul and Jennings (1992) studied the phonetic abilities of toddlers with slow expressive language development, and found that they were more likely to produce a restricted range of consonants and less complex syllable types than were their normally developing peers.

The relationship between prelinguistic phonetic abilities and later speech and language development requires further examination before clinicians can attempt to predict the risk of speech delay from the characteristics of babble. In particular we must determine the degree of overlap, as well as the differences between groups of children with normally developing and those with delayed speech and language abilities. The necessary studies will require the use of analysis techniques that are sensitive to individual differences in the babble produced by children both within and between these groups.

This article describes several methods of analysis that were applied to samples of babble obtained from two children during the period six through 18 months of age. Specifically, the following types of analysis were used: (a) infraphonological analysis[1] (Oller, 1986), (b) vowel formant analysis (de Boysson-Bardies et al., 1989; Kent et al., 1987), (c) babbling level analysis (Stoel-Gammon, 1989), and (d) phonetic contrast estimators (Bauer, 1988).

As will be further described in this paper, both children[2] were developing normally in all respects, although one child's rate of speech development was markedly more rapid than is typical, while the other child's speech abilities were in the average range at 18 months of age. The purpose of this article is not to speculate about the reasons for the difference in rate of speech development observed for these two children. Rather, the purpose is to examine the relative sensitivity of the various analysis methods to the observed differences. It is hoped that these case studies will encourage clinicians to provide systematic descriptions of their young clients' phonetic skills more routinely, even during the prelinguistic phase of development. It is also hoped that researchers will be prompted to examine the predictive validity of these measures of early phonetic ability.

Method

Subjects

Both children, J. and T., were born healthy following uneventful pregnancies. The parents and caregivers of both children were all native speakers of English. J. has no siblings. His parents are both elementary school teachers who worked outside the home throughout the study period. J. was found to be allergic to milk, but remained healthy throughout the period from birth to 18 months of age. According to parent reports, there was no history of speech, language or learning difficulties in the immediate or extended family at the beginning of the study. However, J.'s cousin experiences recurring otitis media with effusion (OME) and was referred for assessment of his language development at the age of 18 months.

T. has a brother who was aged five years at the completion of the study. His mother was employed as a legal secretary and his father completed his engineering degree during the course of the study. T. suffered his first bout of OME at age two months and experienced recurring infections until age 13 months. At age 14 months he was hospitalized with croup, and at 15 months ventilating tubes were inserted in both ears. While there is no history of speech, language, or learning difficulty in the family, T.'s brother was assessed by a speech-language pathologist at age two and a half and again at age four years due to parental concerns about stuttering and slow speech development. However, at the age of four, his language skills were found to be above average. T.'s brother often accompanied him to assessments and the investigator did not observe any noticeable degree of difficulty with his articulation, language, or fluency.

Equipment

Sound field threshold testing was accomplished in an Ekoustic double-walled sound chamber with the following equipment: Interacoustics Clinical Audiometer (model AC 30) and DALI speakers. Middle ear impedance measures were obtained with a GSI-33 Middle Ear Analyzer. Speech samples were tape-recorded with a Sony Walkman Professional tape recorder and a Crown PZM-6D microphone. The speech samples were digitized using the Computerized Speech Research Environment (CSRE; Jamieson, Nearey, and Ramji, 1989) and the following hardware: an AST Premium 386C computer, DT2821 D/A, A/D board (12 bit), and a TTE 411AFS amplifier and antialiasing filter.

Procedure

Both children visited the audiology department at a regional paediatric hospital for approximately one hour once every three months. All assessments were conducted within two weeks of the subjects exact chronological age of six, nine, 12, 15, and 18 months, except that T's first assessment was scheduled for three weeks after his six-month birthday. In most cases the hearing and impedance measures were obtained first, and a taped speech sample was obtained immediately thereafter. In one case the speech sample was obtained one week following the audiology assessment. At each assessment the parents were asked to complete a word-production checklist as a means to estimate expressive vocabulary size. Formal assessments of expressive and receptive language development were conducted during the visit at 18 months.

The audiology assessments were conducted by one of two certified paediatric audiologists. The speech and language assessments and analysis of the speech samples were completed by the first author, a CASLPA certified speech-language pathologist. The procedures used to complete the audiology and speech-language assessments and to analyze the recorded speech samples are described below.

Audiology Assessment. Audiometric testing was performed in a double-walled Ekoustic sound chamber. Subjects were seated between two loudspeakers positioned at 90 degrees relative to midline. Visual reinforcement audiometry (VRA) was used to assess auditory sensitivity to live voice and to warbled tones presented at 250, 500, 1000, 2000, and 4000 Hz. Lighted toys with animation were used as visual reinforcers. A response was considered positive when the subject localized toward the signal source upon presentation of the stimulus. Criterion for a threshold was met at the lowest intensity level which elicited two or more reliable head turning responses. Tympanometry was performed using a 226 Hz probe tone. Two values produced by the instrument were recorded: peak pressure in decaPascals and tympanic membrane compliance in millilitres. At each visit ipsilateral reflexes were attempted at 500, 1000, 2000, and 4000 Hz at the previously recorded peak pressure value.

Language Assessments. During each assessment, the parent was asked to complete one of the five forms of the word production checklist described by Reznick and Goldsmith (1989)[4]. These checklists are equivalent forms, each containing a list of 123 words covering 20 different categories including nouns, verbs, prepositions, pronouns, and modifiers. A score on any one checklist represents a reasonable estimate of one- fifth of the total expressive vocabulary size.

At the age of 18 months, the *Receptive-Expressive Emergent Language Test* (REEL-2; Bzoch & League, 1991) was administered. Although this is an interview scale, most receptive items were administered directly to the child to confirm the parent's report. The parent's report of the child's expressive language abilities were also verified through direct observation where possible. The results of this test are expressed as a standard score with a mean of 100 and a standard deviation of 15.

The *Child Development Inventory* (Ireton, 1992) was given to the parent for completion at the 18-month assessment. This parent report scale provides information about social, self-help, gross motor, fine motor, expressive language, and language comprehension abilities.

Speech Sample Collection. The speech samples were recorded in an Ekoustic sound treated chamber. The mother was instructed to interact with her child in the usual manner. The mother and child were provided with the same set of quiet toys during each assessment (e.g., soft blocks, cloth books, pop beads, stuffed toys, a ball, and puppets). No effort was made to restrict the child's movements during recording sessions; rather, the microphone was moved when necessary so that it was within 1 to 2 feet of the child, preferably positioned with the child facing the microphone (the pressure zone microphone used was capable of capturing almost all speech produced within the sound chamber, even when whispered). The recording session continued until the child produced 60 utterances, which generally took between 10 and 30 minutes.

Speech Sample Analysis Procedures

Acoustic Analyses. Three of the four speech analysis methods to be described below involve acoustic analysis of at least some of the infants' utterances. These utterances were digitized at a sampling frequency of 20 kHz and low pass filtered at 10 kHz. These utterances were then submitted to autoregression analyses using a 128 millisecond (ms) analysis window, 128 Hz frequency bands, preemphasis, and a Hanning window.

Infraphonological Analysis. This form of analysis has been described in detail elsewhere (e.g., Oller, 1986; Oller, Eilers, Bull, and Carney, 1985; Oller, Eilers, Steffens, Lynch, & Urbano, 1994). In this application, 50 consecutive utterances were selected from the tape, each meeting the following criteria: the utterance was bounded by one-second of silence, an audible inspiration, or adult speech; the utterance was perceived to have a *unifying pitch contour* (a criterion which overrode the first criterion in some instances so that two utterances were coded with less than a one-second separation); the utterance was produced and recorded with sufficient loudness for coding; and the utterance was not so obscured by adult speech or other noise as to prevent accurate coding. These utterances comprised both babble and words but no effort was made to distinguish meaningful and nonmeaningful utterances for any of the analyses. Nonspeech sounds such as crying, laughing, burping, grunting, etc., were excluded.

Each utterance was coded using the criteria described in detail by Oller (1986) and in brief as follows:

1. Other: this category includes squeals, raspberries, growls, and yells.
2. Quasiresonant nucleus (QRN): these are syllabic nasals or nasalized vowels that contain little energy over 1200 Hz.
3. Fully resonant nucleus (FRN): these are vowel-like utterances with at least two measurable formants and with resonances above 1200 Hz in addition to resonances in the lower frequency range.
4. Marginal babble (MB): these utterances are transcribed as consisting of consonant-vowel (CV) or vowel-consonant (VC) syllables, but do not meet the criteria for canonical babble; some of the more common reasons for coding an utterance as MB included formant transitions longer than 120 ms in duration, absent or unmeasurable formant transitions, absent upper formants beyond F1, or abnormal phonation (i.e., CV and VC syllables with a squealed, yelled, or whispered vowel were coded as MB rather than Other).
5. Canonical babble (CB): these utterances also contain at least one consonant (excluding glottal stops) combined with at least one fully resonant vowel, but additionally fit the acoustic definition of a canonical syllable; the criterion that was considered most carefully in this study was the requirement for a smoothly changing formant transition between 25 and 120 ms duration[3].

Utterances coded as Other, QRN, or FRN could generally be identified by listening to the tape. Acoustic analysis was sometimes required to distinguish QRN and FRN utterances. Acoustic analysis was also required to distinguish utterances transcribed as /hV/ because it was often difficult to determine if these were Other, FRN, MB, or CB. All utterances that were transcribed as containing a C and V were digitized and submitted to acoustic analysis in order to differentiate MB and CB utterances and syllables. It should be noted that even with acoustic analysis we found that these categorizations involved a great deal of subjective judgment. In many cases, the spectrographic analysis confirmed the perceptually ambiguous character of the utterance. This was particularly true for the differentiation of FRN and QRN type utterances. When these were difficult to classify perceptually, the spectrograms would indicate strong low frequency energy with a faint second formant (F2); as the definition of an FRN includes "substantial energy above 1200 Hz", it was necessary to make a subjective judgment about whether the F2 contained enough energy to be included in this category. The differentiation of MB and CB utterances also involved some subjective judgments regarding the "smoothness" of the formants. Despite these difficulties, reliability of these judgments was reasonably good (see section on Reliability, page 00).

This analysis yielded two outcomes for each sample: (a) a frequency count for each of the five utterance types, and (b) a canonical babble ratio. The canonical babble ratio was calculated by dividing the number of canonical syllables contained in the sample by the number of utterances (i.e., 50). Note that any utterance containing at least one canonical syllable was coded as a CB utterance and thus multisyllable canonical utterances might contain both canonical and noncanonical syllables. These two measures are related in that the canonical babble ratio will increase with both the number and length of canonical utterances contained in the sample. Spectrograms of some of these utterance types are shown in Figure 1.

Vowel Formant Analysis. The vowel of each canonical syllable was submitted to autoregression analysis in order to determine the frequency of the first and second formants (i.e., F1 and F2). When the nucleus of a syllable was a diphthong, the frequency spectrum was obtained for the first vowel only. Spectra were calculated for 10 millisecond segments located at the juncture of the first and second thirds of the steady state portion of the syllable, when appropriate. In many cases the segment most likely to yield a valid result was selected by eye. Often the formants contained gaps caused by intermittent breathiness or harshness in the infants' voices and it was necessary to avoid such gaps. This was especially true for the second formant which is often quite low in energy in infant vowels. This problem is reflected by better reliability for F1 than F2 frequency analyses. For example, the difference between intercoder judgments was less than 10 Hz for 78% of F1 judgments, but were this close for only 55% of F2 judgments (see reliability section below).

Note that this analysis could have been applied to FRN utterances. In this application the vowel formant analysis was restricted to canonical utterances, because the examiners are interested in applying this analysis as described by de Boysson-Bardies et al. (1989) to the larger sample of infants from which these case studies were drawn. Although de Boysson-Bardies et al. (1989) do not justify their utterance selection criteria, we found that some practical problems were solved by excluding vowel-only utterances. For

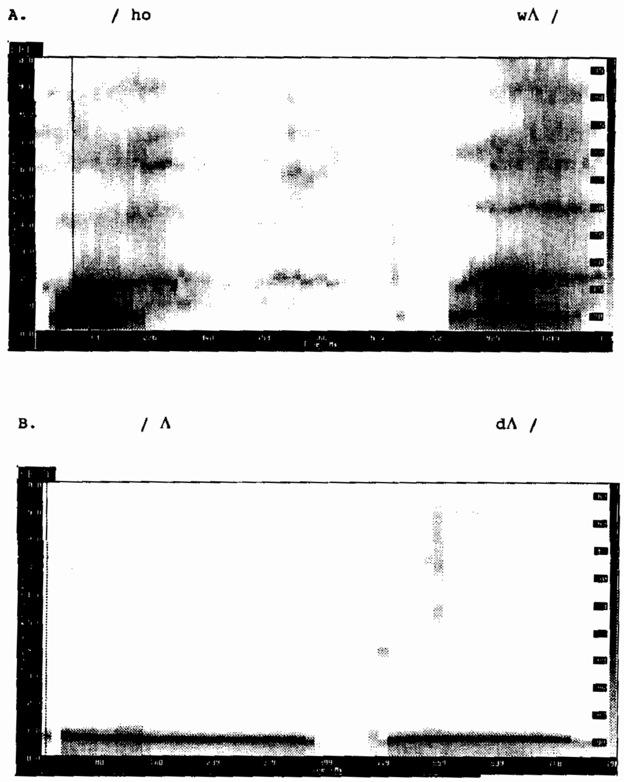


Figure 1. Spectrograms of utterances produced by infant "J" at 9 months of age. Panel A (top) shows a canonical utterance transcribed as /ho.w*/. The first syllable is not canonical because there is no observable transition of the formants from the consonantal /h/ portion (shown before the cursor) into the vocalic /o/ portion of the syllable (shown after the cursor). In contrast, the second syllable shows a clear, smoothly changing second formant transition that is 64 milliseconds in duration. Panel B (bottom) was readily transcribed as /*d*/ but can be seen to contain quasiresonant syllable nuclei, as indicated by the absence of resonances above 1200 Hertz, and was thus coded as marginal babble.

example, it is not unusual for FRN utterances to be very long and to contain considerable variation in vowel identity throughout the utterance. In this case it is not possible to characterize the utterance by a single set of vowel formants. Canonical syllables by definition are restricted in duration to the range between 50 and 500 milliseconds and tend to represent a single monopthongal or dipthongal vowel.

Babbling Level Analysis. Each of the 50 utterances was transcribed phonetically and then coded using the criteria described by Stoel-Gamon (1989). Level I utterances consist of a vowel, a syllabic consonant, or a CV or VC syllable in which the C is non-true consonant (i.e., glottal consonant or glide). Relative to the infraphonological analysis, all OTHER, QRN, and FRN type utterances and some MB and CB utterances would be coded at this level. Level II utterances consist of CV or VC syllables containing true consonants, but with no variation in place or manner across these consonants in multisyllable utterances. Level III utterances contain at least two true consonants that contrast place or manner of articulation. Following this coding a mean babbling level was calculated for the sample by summing the levels assigned to each utterance and dividing by the total number of utterances (i.e., 50). Phonetic description of infant speech is difficult, especially without the support of acoustic analysis. One advantage of this method of coding is that a high degree of accuracy for the transcription of each segment is not required. Rather, the primary differentiation is between true and non-true consonants. This judgment is relatively easy to make and is reflected in the high degree of intercoder reliability observed for this analysis.

Phonetic Contrast Estimators. Phonetic contrast estimators reflect the number and variety of articulatory movements produced by the child. There are two estimators which assess a child's ability to produce consonants with different manners and places of articulation: the phonetic product - P/P and the phonetic product - M/P. The phonetic product - P/P codes each closant[5] in an utterance according to five place categories (bilabial, apical, palatal, velar, and glottal) and each vocant according to three place categories (front, central, and back). The phonetic product - M/P codes each closant by manner of articulation (stop, fricative, glide, and nasal) and each vowel by place of articulation (front, central, and back). Coding was accomplished using both auditory and acoustic analyses. The number of elements in each category was tallied and then the phonetic products were calculated for each utterance using the formulae shown below (where Cb refers to the number of bilabial closants contained within the utterance, and so on for the remaining terms). These formulae incorporate weights that reflect the frequency of occurrence of these place and manner categories in American English:

P/P: $[(Cb)(0.1658)+1][(Ca)(0.3149)+1][(Cp)(0.01129)+1][(Cv)(0.04945)+1][(Cg)(0.04945)+1][(Vf)(0.18)+1][(Vc)(0.1431)+1][(Vb)(0.0709)+1]$

M/P: $[(Cs)(0.1981)+1][(Cf)(0.1617)+1][(Cg)(0.1496)+1][(Cn)(0.1108)+1][(Vf)(0.18)+1][(Vc)(0.1431)+1][(Vb)(0.0709)+1]$

For example, consider an utterance transcribed as [bawi]. The accuracy of this transcription would be confirmed by spectrographic information indicating longer transition durations and a higher frequency second vowel formant for the last syllable in comparison with the first. Subsequent to these phonetic and acoustic analyses, the utterance would be coded as containing one bilabial stop, one bilabial glide, one back vowel and one front vowel. The resulting phonetic products would be P/P = 1.68 and M/P = 1.74. The spectrographic analysis did not always resolve confusion about the place or manner of articulation however, especially for vowels. The difficulty inherent in making these judgments is reflected by the relatively poor intercoder agreement observed for this analysis in comparison with the other methods described in this paper. Bauer (1988) reports much higher reliability, a discrepancy which may be due to higher levels of knowledge about the acoustic characteristics of speech sounds among the reliability coders who participated in his studies. In this study however, the reliability coders had as much training and experience with acoustic analysis as could be expected among speech-language pathologists, and thus the results here are likely more valid for the clinical context.

Bauer (1988) applied this method to all speech-like utterances in samples collected from 13- and 24-month old infants. However, in this application the estimators were calculated for each canonical utterance, as it was not clear from Bauer's published description how the coding could be applied to some of the less speech-like utterances produced frequently by younger infants (e.g., QRNs and GROWLS). It was also thought that it would be interesting to examine whether this method provides information beyond that provided by infraphonological analysis by differentiating among canonical utterances of greater and lesser complexity.

Reliability. Two speech-language pathologists (SLPs) were trained by the first author to transcribe and code infant speech samples on the basis of both perceptual and acoustic analyses. One of these SLPs analyzed a six month sample produced by one infant, while the other SLP analyzed an 18 month sample produced by another infant. The reliability coders were unaware of the first author's judgments regarding the utterances and were blind to the subject's history of middle ear function; however, they were aware of the infant's age at time of recording. One hundred utterances were transcribed and then coded using the infraphonological analysis procedures. The transcriptions were used to code each utterance using the babbling level and phonetic complexity analysis procedures. Point-by-point agreement between the first author and the SLP who conducted the reliability analysis was calculated for a number of different judgments: (a) the infraphonological classification of each utterance using all possible categories, including those grouped as OTHER (in other words, the judges agreed if both judges coded an utterance as GROWL; they disagreed if one judge coded an utterance as GROWL and the other coded the same utterance as RASPBERRY, even though both judgments would place the utterance in the OTHER category); (b) the number of CB syllables when coding all MB and CB utterances (for example, if both judges found two CB syllables within the utterance [baba], then two agreements were counted; if one coder judged only the first syllable to be CB, while the second found two CB syllables, then one agreement and one disagreement were counted); (c) the first and second formant frequency values for vowels in all CB syllables (the percentage of judgment pairs within 100 Hz of each other was calculated); (d) coding of each utterance as Level I, II, or III using the babbling level analysis; and (e) the coding of each segment contained within MB and CB utterances by place and manner of articulation using the procedures described for determining phonetic contrast estimators. Table 1 shows that agreement ranged from 82 to 90 percent for the five types of coding.

Table 1. Reliability of Coding for Utterances, Syllables, Segments, and Vowel Formant Frequencies

Type of Judgment	% Agreement
Infraphonological Analysis	
(1) classification of each utterance	87
(2) number of CB syllables per MB/CB utterance	89
Vowel Formant Analysis	
(3) within 100 Hz	85
Babbling Level Analysis	
(4) classification of each utterance	90
Phonetic Contrast Estimators	
(5) place and manner classification of each segment	82

Note: See text for further descrption of the procedure used to calculate agreement for each type of analysis.

Results

Audiology Assessments. Table 2 shows the Pure Tone Average (PTA; for the frequencies 500, 1000, and 2000 Hz) and Speech Detection Thresholds (SDT) and describes the results of the tympanometry assessment for each child for all assessments. The tympanometric results were judged to reflect normal middle ear function using criteria suggested by Silman and Silverman (1991): middle ear pressure greater than -100 decaPascals and acoustic reflex present at 1 kHz. The results shown in Table 1 indicate normal hearing and middle ear function for J. throughout the study, while T. demonstrated OME with associated mild hearing loss during three of the five assessments (six, 12, and 15 months).

Table 2. Results of the Audiometric Assessments

	Age Months	PTA[a] (dB HL)	SDT[b] (dB HL)	Tympanometry (Tympanograms, Reflexes)
J.				
	6	18.33	20.00	Normal, Normal
	9	20.00[c]	10.00	Normal, Normal
	12	18.33	5.00	Normal, Normal
	15	10.00	10.00	Normal, Could Not Test
	18	15.00	5.00	Could Not Test
T.				
	6	41.67	35.00	Flat, absent
	9	30.00	20.00	Normal, Normal
	12	40.00	30.00	Flat, absent
	15	41.67	35.00	Flat, Could Not Test
	18	20.00	5.00	Patent PE tube (left), negative middle ear pressure (right)

Note. Procedures for obtaining puretone and sound detection thresholds and the criteria for determining normal tympanometric results are outlined in the text. Tympanometry could not be completed if the infant was crying vigorously, which occurred on three occasions.

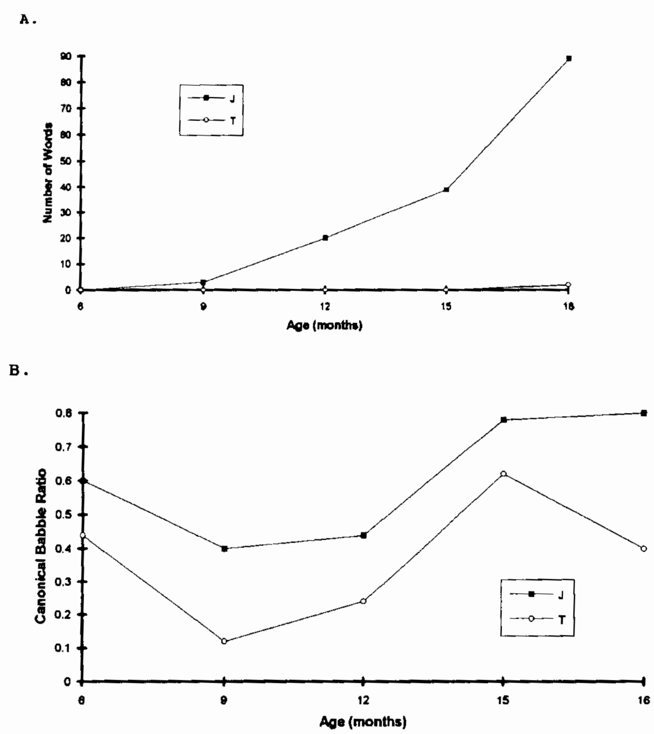
[a]Pure Tone Average (average threshold for the frequencies 500, 1000, and 2000 Hz). [b]Sound Detection Threshold. [c]Screened at 20 dB, thresholds were not obtained.

Language Assessments. The results of the Child Development Inventory, which covers social, motor, and language development, indicate that both children's overall development was within normal limits. However, the assessments specific to language skills indicate a different course of development for the two boys. The results for the Word Production Checklist are shown in Panel A of Figure 2 as a function of age for the two children under examination. J. began producing single words before the age of nine months. Expressive vocabulary growth was extremely rapid. At the age of 18 months he obtained a score of 89 on the Word Production Checklist. Two-word utterances were first observed at 15 months and were relatively frequent at 18 months. In contrast, T. scored zero on the Word Production Checklist until age 18 months, when his score increased to two words. T. was reported to produce "dada" at nine months, and the words "dada", "mama", and "bottle" (/baba/) at 12 months. However, at 15 months of age only one word was in consistent use ("bottle"). By 18 months of age his father reported that he was using between 10 and 15 words in total (which is consistent with his score on the Word Production Checklist, since this measure estimates approximately one-fifth of total expressive vocabulary size).

On the REEL-R, J. obtained standard scores of 133 and 167 for receptive and expressive language skills, while T. obtained standard scores of 111 and 89 for receptive and expressive language skills respectively. (For the larger sample from which these two children were drawn, REEL-R scores are high on average, probably due to the parents' relatively high levels of education; thus far 10 children have completed the study and have obtained mean receptive and expressive language scores of 113 and 116 respectively).

Infraphonological Analysis. The Panel B of Figure 2 shows the canonical babble ratios for J. and T. at each of the five assessments. This figure indicates that both boys showed a similar pattern of development, although J. consistently produced a higher ratio of canonical syllables relative to T. However, an examination of Figure 3 shows that their pattern of babbling does differ qualitatively as well as quantitatively. For example, both children show a decline in the canonical babble ratio at nine months relative to the six-month sample. For J. this decline in the ratio occurs despite an increase in the number of canonical utterances because his utterance length decreased as his babble took on a more "word-like" character. In contrast, T. tended to produce more multisyllable babbles at both ages. However, at six months many of these utterances involved true consonants (e.g., /dadada/) while at 9 months /h/ was the

Figure 2.

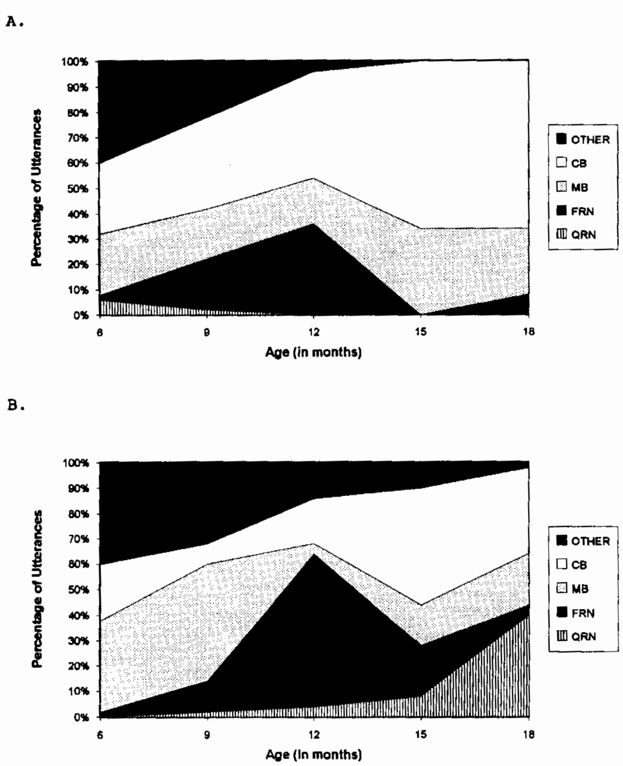


Panel A (top) shows each infant's score on the Word Production Checklist as a function of age. Panel B (bottom) shows the Canonical Babble Ratio for the speech samples recorded from each infant at the ages six, nine, 12, 15, and 18 months.

preferred consonant. As a consequence the proportion of MB utterances rose and the proportion of CB utterances declined. Furthermore, Figure 3 shows different patterns of development for both OTHER and QRN type utterances. For J. the number of OTHER type utterances is negligible at 12 months, while this level is not achieved by T. until 18 months of age. The number of QRNs decreases to zero at 12 months for J., but actually increases with age for T. so that they comprise 40% of his sample at age 18 months.

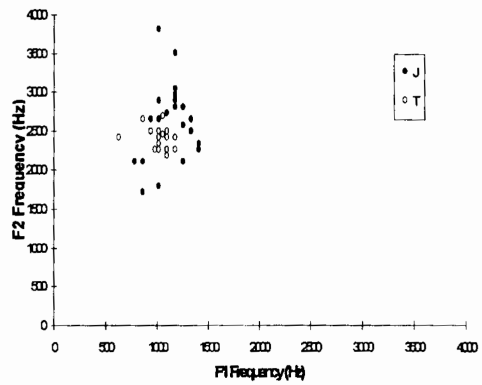
Vowel Formant Analysis. Figure 4 plots the F1 and F2 frequencies of each vowel produced within a canonical syllable by both J. and T. at six months of age. It can be seen that T. is producing a more restricted set of vowels in comparison with J. In fact nearly all vowels produced by T. at age six months were transcribed as /*/. The F1 and F2 frequencies arising from the vowel formant analysis of a given speech sample were reduced to a single measure, the

Figure 3.



Frequency of each of the five utterance types contained in a sample of fifty utterances as a function of age. The utterances were coded as OTHER (raspberries, squeals, growls, and yells), CB (canonical babble), MB (marginal babble), FRN (fully resonant nucleus), and QRN (quasiresonant nucleus). These utterance types are defined further in the text. Panel A (top) shows the frequency of the utterance types produced by J while Panel B (bottom) shows the frequency of the utterance types produced by T.

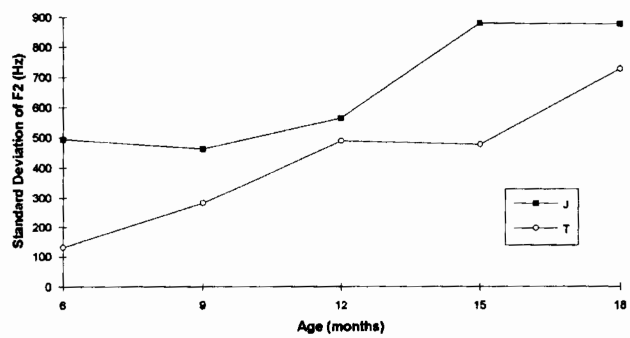
Figure 4.



F1-F2 plot of all vowels produced within canonical syllables by J (filled circles) and T (open circles) at six months of age.

standard deviation of the F2 values. This measure was chosen because the data from the larger sample from which these two children were drawn indicates that the range of F2 values generally increases with age. Thus far, in the larger study, no consistent pattern of change in the mean formant frequencies or in the standard deviation of the first formant frequency has been observed. Therefore, the standard deviation of F2 for each sample is shown in Figure 5. The variability of F2 values produced by T. remains less than that for J. at all age levels, although the restriction in T.'s vowel space is most obvious at ages six and nine months.

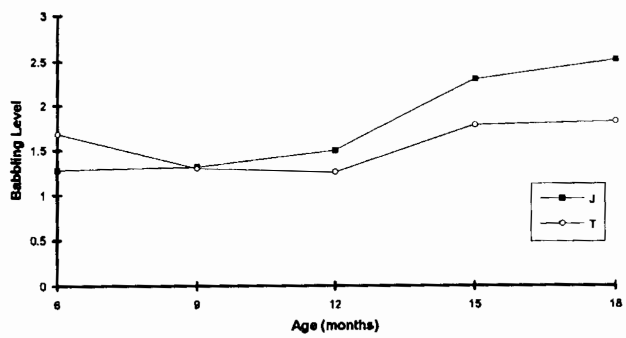
Figure 5.



Standard deviation of F2 values for all vowels produced within canonical syllables as a function of age for infants J (closed squares) and T (open circles).

Babbling Level Analysis. The mean babbling level for each child at each age is shown in Figure 6. With this analysis T. obtained a higher score than J. at six months because T. produced many strings of syllables containing true consonants, while J. tended to produce strings of syllables containing glides. At nine months of age both infants obtained the same mean babbling level because this analysis does not distinguish between CB and MB type utterances. For example, the utterance shown as spectrogram B in Figure 2 would be coded as babbling level II even though this utterance is obviously marginal due to the lack of higher frequency information. J.'s babbling level exceeds T.'s at ages 15 and 18 months. This occurred because nearly all of J.'s utterances contained a consonant and a vowel at these ages, while T. continued to produce a relatively large proportion of QRN, FRN, and OTHER type utterances that would be coded as babbling level I.

Figure 6.



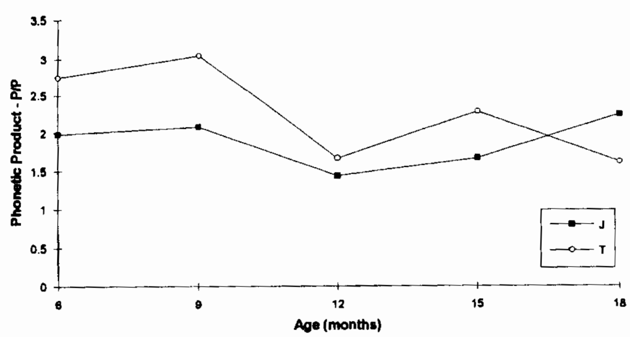
Mean Babbling Level for the samples recorded from J (closed squares) and T (open circles) at the ages six, nine, 12, 15, and 18 months.

Phonetic Contrast Estimators. As noted, the phonetic contrast estimators were calculated for and averaged across canonical utterances only. The results would likely have been very different if this analysis had been applied to all types of utterances in each sample, because the inclusion of FRN and QRN utterances would lower the mean complexity scores for both children.

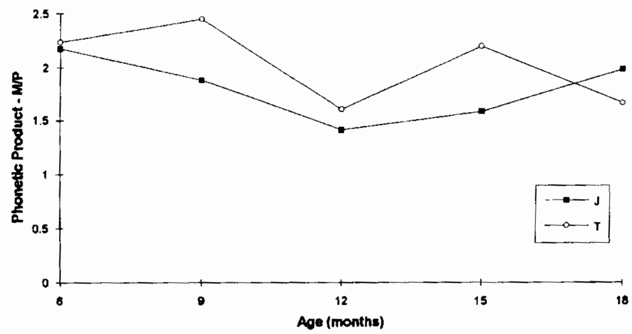
The Phonetic Product scores P/P and M/P are shown in Panels A and B of Figure 7 and indicate that T.'s canonical utterances are actually more complex than J.'s at all ages, except 18 months. Again, this reflects T.'s tendency to produce long strings of syllables and J.'s tendency to produce shorter, more word-like utterances (in fact, at 15 and 18 months J.'s utterances were almost exclusively real words). For example, at 15 months T.'s /d*s*d*d.d*/ would be more complex in comparison with J.'s /ba/ ("ball") or /m*mi/ ("mummy").

Figure 7.

A.



B.



Phonetic contrast estimators for all canonical utterances produced by J (closed squares) and T. (open circles) as a function of age. Panel A shows the Phonetic Product-P/P which focuses on variation in place of consonant articulation while Panel B shows the Phonetic Product-M/P which focuses on variation in manner of consonant articulation.

Discussion

This paper describes the phonetic characteristics of the speech-like utterances produced by two boys during the period six through 18 months of age. One child was healthy throughout the period of observation, while the other suffered chronic OME during the first year of life. Both children attained the canonical babbling stage relatively early and both demonstrated above average receptive language skills at 18 months. Nonetheless these children showed different patterns of phonetic and expressive vocabulary development during infancy.

J. demonstrated precocious development from at least the age of six months, when his babble consisted of multisyllabic strings of canonical syllables involving some consonants and a variety of vowels. By the age of nine months, he began producing a few true words and his babble took on the character of short, word-like utterances containing true consonants. At the age of 18 months his expressive vocabulary size was extraordinarily large and two-word utterances were appearing in his speech.

In contrast, T.'s vowel space was restricted at six months, despite a high proportion of canonical utterances, and his vowel repertoire remained smaller than J.'s throughout the period of study. His canonical babble ratio dropped below .2 at nine months of age, a startling observation given that Oller and Eilers (1988) observed such a regression only among their profoundly hearing impaired sample. Although the canonical babble ratio recovered at the age when ventilating tubes were inserted, he continued to produce fewer canonical syllables than J., and demonstrated a second regression in the canonical babble ratio at 18 months of age. Despite above average receptive language skills, his expressive language abilities at age 18 months were slightly below average, and his expressive vocabulary did not exceed 15 words.

The four methods of phonetic analysis that were examined here varied in their sensitivity to the different course of phonetic and language development observed for these two boys. The infraphonological and vowel formant analyses revealed superior performance for J. relative to T. at each observation interval. The babbling level analysis also indicated a higher level of babbling for J. than for T., but only when they were older than 12 months of age. The boys performed similarly on this measure when younger because this analysis does not discriminate between marginal and canonical CV and VC syllables, as does infraphonological analysis. Rather, the babbling level analysis distinguishes between syllables containing "true" and "non-true" consonants. Studies that examine the predictive validity of these two measures of prelinguistic phonetic ability could help to determine the functional significance of these distinctions.

Although T. produced fewer canonical utterances than J. at all ages, the phonetic contrast estimators suggest that T.'s canonical utterances were more complex than J.'s for most of the samples. The utility of these measures as predictors of language delay depends upon the assumption that phonetic complexity increases with both age and level of language ability. These assumptions have not received unequivocal support in the literature.

Oller (1980), Roug, Landberg, and Lundberg (1989), and Stark (1980) all proposed that babbling becomes increasing complex during the second half of the first year, as indicated by the progression from reduplicated to variegated babble. Elbers (1982) documents a very clear pattern of increasing utterance complexity during this period for a single child. Stoel-Gammon (1989) also provides evidence of increasing frequency of variegated babbling among 32 children during the period nine through 18 months.

In contrast, Mitchell and Kent (1990) did not find a clear pattern of increased frequency of multiple syllable utterances or of utterances containing variations in place or manner of articulation among eight children who were recorded at seven, nine, and 12 months of age. They propose that reduplicated and variegated babbling does not occur during separate stages in the first year.

Similarly, there is no support for a positive association between linguistic skill and phonetic complexity. In fact, Nelson and Bauer (1991) found that phonetic complexity decreased with increasing utterance length and linguistic complexity.

The two case studies presented in this paper suggest that some aspects of prelinguistic speech development may be related to early language development. However, this paper cannot properly address this issue because only two subjects were studied, both children were functioning within the normal range, and the children were not followed long enough to describe outcomes relating to syntax, phonology, or other later developing language skills. It is necessary to conduct further longitudinal studies to determine if prelinguistic phonetic skill might predict language ability in older children. The phonological skills of preschool age children would be of particular interest.

In summary, the infraphonological and vowel formant analyses proved most sensitive to differences in phonetic development for two boys throughout the period six to 18 months of age. Unfortunately, these analyses are time consuming and require specialized equipment and software that is not routinely available in clinical settings. In comparison, babbling level analysis (Stoel-Gammon, 1989) is elegant in its simplicity and ease of use. This method proved to be quite sensitive to differences in phonetic development at the 15 and 18 month observation intervals, but may not be appropriate for younger infants. The phonetic contrast estimators appear to be a good research tool for further examination of the complex relationship between phonetic and linguistic complexity. However, it is as costly to use as infraphonological analysis, but in this study did not provide additional useful information.

As noted, some of these analyses require specialized equipment and skills. The necessary equipment has become relatively inexpensive in recent years. The cost efficiency of computers is further enhanced by the increasing variety of software programs that have become available for the assessment and treatment of the full spectrum of speech and language disorders. Speech-language pathologists are themselves gaining in technological sophistication. Given these developments, there are few practical impediments to the clinical application of the analysis tools that have been outlined in this paper. At this time, the primary need is for the collection and organization of normative data along with a thorough examination of the predictive validity of these measures.

Please address all correspondence to: Susan Rvachew, Speech-Language Section, Alberta Children's Hospital, 1820 Richmond Rd. SW, Calgary, Alberta, T2T 5C7. E-mail: srvachew@acs.ucalgary.ca

Footnotes

[1]Until recently this form of analysis was called metaphonological analysis. Consistent with Oller et al. (1994), the term infraphonological will be used in this paper in order to avoid confusion with the other meaning of the term metaphonological.

[2]These two children were drawn from a larger study in which 18 children received hearing and prelinguistic speech assessments at six, nine, 12, 15, and 18 months of age in order to examine the impact of otitis media with effusion on infant speech development.

[3]Formant transition durations were measured by visually placing the cursors at transition onsets and offsets. The acoustic cues that are associated with transition onsets vary with the type of consonant and the position of the consonant within the syllable; for a discussion of these issues see Kent and Read (1992). We also used auditory feedback to ensure that the cursors marked the consonant transition, rather than some other segment such as the transitional portion of a dipthongized vowel. If any one of F1, F2, or F3 was observed to have a transition duration within 25 to 120 ms the syllable was judged to be canonical (provided that the other relevant criteria were met). Absolute accuracy for the measurement of transition durations was not considered to be of critical importance; rather, the reliability assessment focuses on the final judgment about the category membership of each syllable.

[4]Although this was not the case when this study began, these checklists are now available as part of The MacArthur Communicative Development Inventory (1993) Singular Publishing Group: San Diego, CA. However, the checklists are not divided into equivalent forms and the instructions and procedures are somewhat different for this test in comparison with the procedures employed in this study and cited in the Reznick & Goldsmith (1989) paper.

[5]When describing infant vocalizations, Bauer (1988) prefers to use the terms closant and vocant to refer to consonant-like and vowel-like speech sounds respectively, in recognition of the acoustic and articulatory differences between infant and adult produced vocalizations

Author Notes

This research is supported by a studentship granted by the Medical Research Council of Canada to the first author who is currently completing her doctoral dissertation. We are grateful to the infants and their families for their participation in this study. We are also thankful to the audiology and speech-language sections of the Alberta Children's Hospital for their support of this project. In particular, we appreciate the assistance of Naida Feldman and Linda Hornsberger, who conducted the reliability analyses. We also appreciate the careful and thoughtful comments provided by two anonymous reviewers, one of whom supplied the title for this paper.

References

Bauer, H.R. (1988). The ethological model of phonetic development: I. Phonetic contrast estimators. *Clinical Linguistics & Phonetics, 2*, 347-380.

Bzoch, K.R., & League, R. (1991). *Receptive-Expressive Emergent Language Test* (2nd Ed.). Austin, Texas: Pro-Ed.

de Boysson-Bardies, Halle, P., Sagart, L., & Durand, C. (1989). A cross linguistic investigation of vowel formants in babbling. *Journal of Child Language, 16*, 1-17.

de Boysson-Bardies, Sagart, L., Halle, P., & Durand, C. (1986). Cross-linguistic variability in babbling. In B. Lindblom and R. Zetterstron (Eds.), *Precursors of early speech* (pp.113-126). New York: Stokton Press.

de Boysson-Bardies, Vihman, M.M., Roug-Hellichius, L. Durand, C., Landberg, I., & Arao, F. (1992). Material evidence of infant selection from target language: A cross-linguistic phonetic study. In C.A. Ferguson, L.Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 369-392). Timonium, Maryland: York Press.

Computerized Speech Research Environment (CSRE 4.2). London, ON: AVAAZ Innovations Inc.

Elbers, L. (1982). Operating principles in repetitive babbling: A cognitive continuity approach. *Cognition*, 45-63.

Ireton, H. (1992). *Child Development Inventory.* Minneapolis, Minnesota: Behavioral Science Systems, Inc.

Jamieson, D.G., Nearey, T.M., & Ramji, K. (1989). CSRE: A speech research environment. *Canadian Acoustics, 17*, 23-35.

Kent, R.D. (1992). The biology of phonological development. In C.A. Ferguson, L.Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 65-90). Timonium, Maryland: York Press.

Kent, R.D., Osberger, M.J., Netsell, R., & Hustedde, C.G. (1987). Phonetic development in identical twins differing in auditory function. *Journal of Speech and Hearing Disorders, 52*, 64-75.

Kent, R.D., & Read, C. (1992). *The acoustic analysis of speech.* Singlular Publishing Group: San Diego, CA. Locke, J.L. (1983). *Phonological Acquisition and Change.* New York: Academic Press, Inc.

Mitchell, P.R., & Kent, R.D. (1990). Phonetic variation in multisyllable babbling. *Journal of Child Language, 17*, 247-265.

Nelson, L.K., & Bauer, H.R. (1991). Speech and language production at age two: Evidence for tradeoffs between linguistic and phonetic processing. *Journal of Speech and Hearing Research, 34*, 879-892.

Oller, D.K. (1980). The emergence of the sounds of speech in infancy. In G.H. Yeni-Komshian, J.F. Kavanaugh, & C.A. Ferguson (Eds.), *Child Phonology* (Vol. 1, pp. 93-112). New York: Academic Press.

Oller, D.K. (1986). Metaphonology and infant vocalizations. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 21-36). New York: Stockton Press, Inc.

Oller, D.K., & Eilers, R.E. (1988). The role of audition in infant babbling. *Child Development, 59*, 441-449.

Oller, D.K., Eilers, R.E., Bull, D.H., & Carney, A.E.(1985). Prespeech vocalizations of a deaf infant: A comparison with normal infraphonological development. *Journal of Speech and Hearing Research, 28*, 47-63.

Oller, D.K., Eilers, R.E., Steffens, M.L., Lynch, M.P., & Urbano, R. (1994). Speech-like vocalizations in infancy: An evaluation of potential risk factors. *Journal of Child Language, 21*, 33-58.

Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech and Hearing Research, 35*, 99-107.

Reznick, J.S., & Goldsmith, L. (1989). A multiple form word production checklist for assessing early language. *Journal of Child Language, 16*, 91-100.

Roug, L., Landberg, I., & Lundberg, L.-J. (1989). Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. *Journal of Child Language, 16,* 19-40.

Silman, S., & Silverman, C.A. (1991). Auditory Diagnosis: Principles and Applications. San Diego,CA: Academic Press, Inc.

Stark, R. (1980). Stages of speech development in the first year of life. In G.H. Yeni-Komshian, J.F. Kavanaugh, & C.A. Ferguson (Eds.), *Child Phonology* (Vol 2, pp. 73-92). New York: Academic Press.

Stoel-Gammon, C. (1989). Prespeech and early speech development of two late talkers. *First Language, 9*, 207-224.

Stoel-Gammon, C., & Otomo, K. (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders, 51*, 33-41.

Whitehurst, G.J., Smith, M., Fischel, J.E., Arnold, D.S., & Lonigan, C.J. (1991). The continuity of babble and speech in children with specific expressive language delay. *Journal of Speech and Hearing Research, 34*, 1121-1129.