
Speech Production in 12-Month-Old Children With and Without Hearing Loss

Richard S. McGowan
CRESS LLC, Lexington, MA

Susan Nittrouer
The Ohio State University, Columbus

Karen Chenausky
STAR Corporation, Bedford, MA

Purpose: The purpose of this study was to compare speech production at 12 months of age for children with hearing loss (HL) who were identified and received intervention before 6 months of age with those of children with normal hearing (NH).

Method: The speech production of 10 children with NH was compared with that of 10 children with HL whose losses were identified (better ear pure-tone average at 0.5, 1, and 2 kHz poorer than 50 dB HL) and whose intervention started before 6 months of age. These children were recorded at 12 months of age interacting with a parent. Three properties of speech production were analyzed: (a) syllable shape, (b) consonant type, and (c) vowel formant frequencies.

Results: Children with HL had (a) fewer multisyllable utterances with consonants, (b) fewer fricatives and fewer stops with alveolar-velar stop place, and (c) more restricted front-back tongue positions for vowels than did the children with NH.

Conclusion: Even when hearing loss is identified shortly after birth, children with HL do not develop speech production skills as their peers with NH do at 12 months of age. This suggests that researchers need to consider their approaches to early intervention carefully.

KEY WORDS: deaf children, speech production, acoustic analysis

Typically developing children begin to exhibit the influence of the ambient language on their own productions during the first year of life, before they say their first real words. In particular, the effects of the ambient language can be seen in the more global characteristics of babbled productions. For example, Boysson-Bardies, Sagart, Halle, and Durand (1986) computed the long-term spectra of adults and 10-month-old infants whose native languages were English, French, Cantonese, or Algerian Arabic. Long-term spectra derived from speech samples are shaped largely by acoustic characteristics arising from postural settings such as nasalization, pharyngeal constrictions, and general vowel quality. As a result, these spectra vary depending on factors such as how often nasalized segments occur, how common pharyngeal constrictions are, and the size of the vowel space in the language. In the Boysson-Bardies et al. (1986) experiment, long-term spectra of infants' babbling resembled those of speech samples from adults in their respective language communities, indicating that the infants had already started to incorporate language-appropriate postural adjustments into their babbling. Note, however, that some researchers have found the differences between languages in terms of long-term average spectra to be small (Byrne et al., 1994). Further evidence that children are influenced early in life by their linguistic environment is provided by Boysson-Bardies, Halle, Sagart, and Durand (1989), who used discrete formant frequencies from vocalic segments to show that infants produce babbled vowel-like sounds in language-specific ways. These investigators measured first

formant (F1) and second formant (F2) frequencies in the vocal productions of 10-month-olds in the same four language environments that they used in the 1986 study and graphed the results on F1–F2 scatterplots. The infants' plots strongly resembled those of adults in their language communities; in particular, range of F2 (maximum F2 – minimum F2) was similar to that of adult speakers in the language community. Recently, Rvachew, Mattock, Polka, and Ménard (2006) reported a finding for 9- to 18-month-old learners of either Canadian English or Canadian French. The range of F2 variation for vowels in canonical syllables (Oller, 1986) increased in a statistically significant way for the English-learning infants but not for the French-learning infants in this time period. These findings suggest that vowel formant frequencies, particularly F2, can be used to explore the extent to which the language environment has affected very young children presumed to be at risk for delays, such as children with HL.

Formants in vowel-like babbled productions, sometimes called “vocants” (Kent & Murray, 1982), show restricted vowel spaces in early productions. (We use the term *vowel* throughout for both what could be called vocants and vowels in meaningful words.) For example, Kent and Bauer (1985) performed phonetic transcription of vowels for five 1-year-olds with NH and found that these children produced front and neutral vowels when they were produced alone, which was the most common syllable type. In consonant-vowel syllables, mid- and low-back vowels became more numerous, although the neutral and front vowels were still the most common. In general, studies of infants' vowel-like productions show the vowel space (F1–F2 plane) expanding with age through the first couple years of life for typically developing children (e.g., Buhr, 1980; Kent & Murray, 1982). Kent, Osberger, Netsell, and Hustedde (1987) made use of this fact to examine vowel development in twins from ages 8 to 15 months—one twin with HL and one with NH. The expansion in F1–F2 space was found for the twin with NH through 18 months of age but not for his brother with HL (Kent et al., 1987). A study of the speech of English-learning Canadian children compared children with early- (before 6 months) and late-onset (after 6 months or none) otitis media from 6 to 18 months (Rvachew, Slawinski, Williams, & Green, 1996). Rvachew et al. (1996) found that the children in the late-onset group showed a significant expansion in the ranges of their F2s, as measured by within-speaker standard deviations, whereas the children with early-onset otitis media did not. Mean F1, mean F2, and the standard deviation of F1 did not change significantly for either group. So, the children with early-onset otitis media were delayed in acquiring the extent of tongue placement in the front-back dimension, as measured by F2 standard deviation. This finding illustrates the importance of early hearing to vowel acquisition.

In a case study of a child with NH from 14 to 20 months, Davis and MacNeilage (1990) found a discontinuity between vowel production in babbled sequences and early words. The 1 participant in their study favored the neutral and front (mid-to-low) vowels in babbling but used all types of high vowels when word production started. Davis and MacNeilage speculated that babbling is dominated by “jaw wagging,” without active control of the tongue. In this mode, tongue position is high depending on jaw position, and, therefore, F1 exhibits variability. Later, given that the child seemed to be seeking tongue control for words, the authors speculated that high vowels were favored for their greater proprioceptive feedback. Greater tongue control to produce vowels should result in greater front-back movement of the tongue, and, hence, greater F2 variation. The same authors have proposed the “frame-content” hypothesis, in which tongue movement is added to jaw wagging as the infant acquires speech production capability (Davis & MacNeilage, 1995). The jaw oscillations provide the frame for the tongue and lip content as the infant progresses from babbling to speech. Although Davis and MacNeilage based their hypothesis on transcriptional analysis, it seems that vowel production, as indicated by tongue position, and therefore formant frequencies, can be an important indicator of transition from babbling to meaningful speech.

Utterance shape has also been found to differ between children with NH and those with HL. Kent and Bauer (1985) used counts of utterance shapes (V, CV, CVC, VCVC, etc.), both babbling and early word production, and found that their five 1-year-old participants with NH produced 60% “vowels,” or purely vowel-like utterances (Vs). The rest were more complex syllable structures. Stoel-Gammon and Otomo (1986) found that their participants with HL produced fewer multisyllabic utterances than did their counterparts with NH in the 4- to 18-month age range. In their investigation of twins, Kent et al. (1987) found that the syllable types produced by twin brothers were distinctly different from 8 through 20 months, with the brother with HL exhibiting a preponderance of Vs and the brother with NH producing a variety of syllable types and multisyllabic utterances. These observations are consistent with those of Oller, Eilers, Bull, and Carney (1985) and Oller and Eilers (1988), who found that children with HL, even if wearing hearing aids, are delayed in producing canonical babble, which occurs by 8 months of age for children with NH.

Another example of recent research on the differences in utterance shapes between infants with HL and NH comes from children learning Dutch. Koopmans-van Beinum and Doppen (2003), who worked with recordings of 5 infants with HL and 5 infants with NH between 10.5 and 17.5 months, found that the infants with HL produced more multisyllabic utterances than did the

infants with NH. This finding appears to contradict other results described above, but Koopmans-van Beinum and Doppen counted vowels separated by voice breaks as multisyllabic utterances, which is different from the studies cited above. In fact, they attributed the greater quantity of multisyllabic utterances by infants with HL, compared to infants with NH, to various series of vowels with voice breaks. This is consistent with the observation that children with HL are delayed in producing canonical babble, which, by definition, must have consonant-like margins accompanying the vocalic nuclei.

In general, the above review supports a view of early speech development in which infants begin modifying their vocal-tract gestures during their first year of life to resemble those of adult speakers in their language community. Also, the review of the literature illustrates that the speech development of young children with HL diverges from that of children with NH in numerous ways—at least, it has until recently. Most of the work supporting that conclusion was done with children whose hearing losses were identified after the first year of life. Often even those children whose losses could be identified earlier did not receive amplification because there were no devices powerful enough to shift auditory thresholds into a range that would allow children to hear the speech around them. Then in 1988, the U.S. Department of Education and Bureau of Maternal and Child Health convened a group of scientists to advise the government about the feasibility of developing methods for identifying hearing loss at or shortly after birth. This group recommended that demonstration projects be developed to examine the possibility. By 1990 it was clear that methods were available to identify hearing loss at birth. The utility of early identification was demonstrated (e.g., Yoshinago-Itano, Sedey, Coulter, & Mehl, 1998), and that evidence was interpreted by most investigators and clinicians as showing that "... many children with sensorineural hearing loss achieve language abilities similar to hearing peers if comprehensive intervention services are provided by six months of age" (Moeller, 2000, p. 1 of electronic reference). However, many of these studies have indexed the development of speech production using measures that are not particularly sensitive. Frequently, investigators transcribe language samples (e.g., Yoshinago-Itano, Coulter, & Thompson, 2001) and count the number of vowels and consonants produced by young children. But there are inherent problems in using transcription alone to derive dependent measures. Primarily, adult listeners bring to the transcription task their own language-specific perceptual biases. For instance, Kent (1996) wrote about inconsistency among judges, and between judges and instrumental techniques, such as spectrography, in a review of auditory-perceptual means of assessing speech and voice disorders. The disagreements among judges can also occur for normal speech production. If children

are not producing speech in accordance with their native language, then across-category variability will not be accounted for correctly because listeners may not hear acoustic differences associated with nonnative contrasts. In addition, within-category variability in production patterns will not be noted using transcription alone. For these reasons, it is generally preferable to supplement transcribed analysis with instrumental acoustic analysis when studying the development of speech production.

In the present study, we assessed speech production in two ways: with broad segmental transcription, using both perceptual and spectrographic information, and with quantitative analyses of the speech spectra themselves. We used a broad segmental transcription, aided with spectrographic displays to reduce transcription error caused by adult phonemic biases. Here, we used the transcriptions to assess syllable shape and consonant type. We used acoustic measures of formant frequencies as another means to measure and characterize children's speech production. Using these three measures of speech development, syllable shape, consonant type, and vowel formant frequencies, we sought to examine the question of whether, at 12 months of age, children with HL are displaying babbled productions similar to those of children with NH when those children with HL are identified before 6 months of age and given early intervention.

Method

Participants

Speech samples from 10 children with diagnosed HL and 10 children with NH, all 12 months of age, were taken at various test sites across the United States. All children were part of an ongoing study investigating outcomes for children with and without HL between 12 and 48 months of age (Nitttrouer, in press). In that study, children are tested on a variety of measures, including psychosocial development and receptive and expressive language abilities, at each 6-month birthday.

The 10 children with NH whose data were analyzed here had all passed newborn hearing screenings at birth and later passed hearing screenings at 36 months of age consisting of the pure tones 0.5, 1.0, 2.0, and 4.0 kHz presented at 20 dB HL to each ear separately. Table 1 shows better ear pure-tone average hearing thresholds in dB HL for the three frequencies of 0.5, 1.0, and 2.0 kHz for the 10 children with HL. Table 1 also shows age of identification for children with HL. Hearing aids were provided as soon after identification as possible for all these children, and parents of all children reported that hearing aids were worn during all waking hours, except at bath time. At the time of data collection, no child had a cochlear implant. All were in early intervention programs

Table 1. Better ear pure-tone average (BE-PTA) thresholds (at 0.5, 1, and 2 kHz) and age of identification for the children with HL.

Variable	HL1	HL2	HL3	HL4	HL5	HL6	HL7	HL8	HL9	HL10	M
Gender	M	M	F	F	M	F	F	F	F	F	
Mode of communication	Oral	Oral	Oral	Oral	Oral	Oral	Oral	Oral, sign supp.	Oral, sign supp.	Oral, sign supp.	
PTA (dB HL)	120	93	105	95	95	95	120	120	120	73	103.6
Age of ID (months)	1	0	4	1	4	1	1	1	1	0	1.4
Age of amplification (months)	3	1	5	1	4	2	1	2	2	1	2.2

Note. The maximum measurable hearing loss (HL) is 120 dB. M = male; F = female; supp. = supplemental; ID = identification.

in which spoken communication was emphasized, but 3 of these children with HL had spoken language input supplemented with sign language. All families received intervention at least once a week from a provider with a master's degree in deaf education or speech-language pathology. With the exception of 1 child with NH, all children had expressive vocabularies of fewer than 10 recognizable words, as measured by the Language Development Survey, a standardized parent report of vocabulary (Rescorla, 1989). The child who was the exception had 13 words.

Socioeconomic status (SES) was derived using a procedure described elsewhere (e.g., Nittrouer & Burton, 2005). According to this procedure, occupational status and educational level of the primary income earner in the home are used to obtain an SES metric for the household. Two 8-point scales are used, with 8 representing both the highest occupational status and the highest educational level. Derived codes for occupation and education were multiplied to obtain SES metrics, and so scores varied from 1 to 64. The mean SES for the NH group was 36.8 ± 13.2 , and it was 36.2 ± 13.3 for the HL group. This difference is not statistically significant and indicates that children generally came from homes where the primary income earner had a college education.

Procedures

Each child was recorded with one parent sitting on the floor playing with a standard set of toys. The toys were chosen to appeal to the interests of children between 12 and 48 months of age, the range over which the children would be tested for the larger project. Specifically, the toys were a teddy bear; a plastic truck; plastic see-through blocks with moving parts inside; a plastic tea set; a five-member doll family approximately 8 in. (20 cm) tall; a felt board with felt dolls, clothing, and pets; a toy cell phone; and the board book *Goodnight Gorilla* by Peggy Rathmann (1994). Digital videotapes were made using a Sony Digital Handycam. A Sony FM transmitter was used to ensure a high-quality audio signal, and the child wore the transmitter in a vest. Most recording sessions lasted 20 min, and the audio portion of the signal was sampled at 48.1 kHz, with 16-bit

dynamic range. Exceptions to the 20-min duration occurred for 3 children with HL. The durations for their recording sessions were 17.2 min., 15.7 min, and 14.5 min. The videotapes were made at the test site and then sent to the laboratory of the second author. A laboratory assistant separated the audio signal from the video signal. Compact disks containing the audio signals only were sent to the first author for analysis. The first and third authors did all analyses and were blind to the hearing status of the speakers, or whether sign language was used or not, until after analyses were completed.

Children's data were analyzed in two ways, both using the spectral analysis program Speech Station II (Sensimetrics Corporation). First, a broad transcription was made of all utterances produced in the session. Because three of the sessions were shorter than 20 min, syllable types are reported as proportions of total utterances produced in the session. Next, acoustic analyses were performed to derive formant frequencies of a subset of vocalic portions. For these analyses, seven nonoverlapping 1-min intervals of samples were derived from across the sample.

For transcription purposes, the analyst listened to the sound file to identify vocalizations by the child, not including shrieks, cries, and laughter. The time at which each utterance began was recorded, and each utterance was copied into a separate window for closer examination. Discrete utterances were defined by appropriate temporal separations between one vocalization and the next, phrasal intonation, interruptions by the parent, or change of topic focus. At this point, the analyst made a broad phonetic transcription on the basis of both auditory percept and the spectrogram. Using the spectrographic display minimized the extent to which the language proclivities of the analysts biased the transcription, which is a risk to objectivity in all transcriptional analyses.

The transcription of consonants was in terms of manner (stop, fricative, affricate, or trill) and place, but in a way that was coarser than traditional analysis of English phonemes. Stops were identified by abrupt changes in amplitude above 1 kHz, and fricatives had a clear noise component above 1 kHz. There was ambiguity in distinguishing alveolar and velar places of articulation

for stops, and thus these were combined into a single class of stops: AVS. There were also bilabial stops (BLS). There were some classes of segments that do not have counterparts in English, such as bilabial trills that were transcribed as such. Medial aspirants and glottal stops were also transcribed and classified as consonants. Vowels were transcribed according to four categories: voiced (nonnasal and egressive), nasal, ingressive, or whispered. Surrounding glides and liquids were not transcribed separately from the vowel. Apparent liquids were very rare, and glidelike productions were considered part of the vocalic portion.

Utterance shape was derived from the transcriptional data. To characterize syllable shape, vowel-like segments with no constrictions, whether voiced, nasal, ingressive, or whispered, were counted as vocalic nuclei (V). Stop, fricatives, affricates, or trills, as well as medial aspirants and glottal stops, were counted as consonants (C). If a V had an accompanying C, then it was counted as a CV or VC, depending on whether the C was before or after the V. The number of syllables in an utterance corresponded to the number of vocalic nuclei, and it was possible to have neighboring nuclei, Vs, not separated by a C: This was denoted V~V. Neighboring Vs were the result of cessation of voicing without noticeable aspiration or glottal stop.

A detailed set of acoustic measures was made on a subset of the utterances identified in the transcriptional analysis. Specifically, every third minute of the total sample, starting with the first minute, was saved to its own audio file, and the utterances within those minutes were analyzed using detailed acoustic measures. For the three samples that were shorter than 20 min, every third minute starting with the second minute was also analyzed, up to a total of 7 min per speaker. Each utterance to be analyzed was down sampled to a rate of 16 kHz to enhance the details of the spectra in the region of the first three formant frequencies, F1, F2, and F3. For each utterance, a single nonnasal vowel was chosen for analysis. F1, F2, and F3 were estimated at three different times within the vowel portion: within 1/6 total duration of vowel beginning, within 1/4 total duration of the vowel center, and within 1/6 total duration of the end of the vowel. The formant frequency data are reported as averages without regard to where in the vowel the values were estimated. Formant measurements were taken at three different times in a vowel to obtain as complete a picture as possible regarding the extent of formant frequency values during transition as well as the vowel's center.

Both a spectrogram and a spectral cross-section were used to estimate F1, F2, and F3. The spectrogram was plotted from a discrete Fourier transform (DFT) analysis with a 128-point Hanning window with high-frequency preemphasis. The spectral cross-sections were plotted from DFT analysis with a 512-point Hamming window with

preemphasis off. The analysis windows for the spectrograms were usually not long enough to resolve voice harmonics, whereas the voice harmonics were resolved in the spectral cross-sections. There are pitfalls encountered when estimating the formant frequencies for 12-month-old children, which include highly variable fundamental frequencies and high voice fundamental frequencies when the voice harmonics are widely spaced. Furthermore, breathy voices are often found in children, where subglottal resonance frequencies can confuse the measurement of vowel formant frequencies. Nasalized vowels are also problematic with extra formants and zeros added to the spectrum, so these were not considered for measurement at all.

All analyses were done by the first author. However, reliability of both the transcriptional analyses and of derived formant frequencies was checked by having the third author do independent analyses for 3 children with HL and 3 with NH.

Results

Reliability

We analyzed a total of 324 utterances in the reliability study. Of that number, there were 241 agreements in how the speech was transcribed, and 83 disagreements between the first and third authors. The fact that there was 74.4% agreement was considered acceptable, given the great degree of difficulty in analyzing the speech of children so young. Furthermore, none of the disagreements were of a nature that would invalidate the acoustic measures made or the conclusions reached on the basis of those measures. The greatest amount of disagreement occurred over the nasalization of vowels: 33 vowels were judged as nasalized by both authors. The third author judged 33 more vowels as nasalized than the first author had not judged as nasalized. The first author judged 14 still different vowels to be nasalized than the third author did not judge to be nasalized. These differences produced a net difference of 19 more nasalized vowels for the third author. Furthermore, 31 syllables that the first author judged to be isolated vowels only (20 nasalized and 11 nonnasal) were judged to be nasal consonants by the third author. There were no utterances judged to be only nonnasal or nasalized vowels by the third author that were judged to contain consonants by the first author. In total, the first author counted 80 nasal segments, and the third author counted 112. Nasality is recognized as a very difficult quality to judge in speech, particularly in the samples of young children in which other factors, such as subglottal resonances, may create similar spectrographic impressions. Consequently, this amount of disagreement is to be expected and is considered acceptable. When there was disagreement, the transcriptions of the first author were retained.

We performed a separate check of reliability for the first two formant frequencies. Formant frequencies had been measured by the first author as described previously, and the third author independently measured formant frequencies for samples from the 3 children with HL and 3 children with NH whose tokens were checked for reliability in transcriptions. A total of 143 F1 values and 141 F2 values were measured in this way. The average difference between the analysts' measurements was $14\% \pm 14\%$ for F1 and $6\% \pm 7\%$ for F2. This amount of difference between analysts was considered acceptable.

Syllable Shape

There was great variability among children within both the NH and HL groups, which is typical of 12-month-olds. For instance, the numbers of utterances per session (adjusted for the participants recorded for less than 20 min) ranged between 29 and 206 for the NH group and between 21 and 97 for the HL group (see Table 2).

Children in the HL group produced a higher percentage of their utterances as multisyllable utterances (two or more syllables) than children in the NH group. Children with HL produced $40.8\% \pm 25.9\%$ of their total utterances as multisyllable utterances compared with

$20.2\% \pm 13.9\%$ for the children to NH, and these are significantly different, $t(1, 13.778) = 2.224, p < .05$, for unequal variances. The predominance of multisyllables for the children with HL compared to the children with NH was largely due to utterances of the form $V \sim V \sim \dots V$. Children with HL had an average of $79.9\% \pm 27.8\%$ of their multisyllabic utterances that contained only vowel-like segments, compared with $30.6\% \pm 36.5\%$ for children with NH, which is a significant difference, $t(1, 18) = 3.401, p < .01$. In fact, utterances of the form $V \sim V$ accounted for an average of $80.4\% \pm 27.4\%$ of the HL group's two-syllable utterances, compared with $33.8\% \pm 36.4\%$ for the NH group, $t(1, 12.599) = 3.286, p < .01$, for unequal variances. In summary, children with HL produced more multisyllabic utterances than children with NH, but those utterances often contained only vowels.

Utterances that consisted of single syllables were dominated by vowel-only productions for both groups: $85.0\% \pm 18.1\%$ for the children with HL, and $66.2\% \pm 22.8\%$ for children with NH (see Table 2). This difference did not attain statistical significance. The only significant difference between the HL and NH groups for specific single-syllable shapes was for the VC shape. The children with NH had a significantly higher percentage ($8.1\% \pm 5.9\%$) of VC-single syllables than did the children with HL ($2.3\% \pm 4.9\%$), $t(1, 18) = 2.516, p < .05$.

Table 2. Adjusted number of utterances (adjusted inversely proportional to the duration of recording) and syllable statistics.

Hearing status, participant	Adjusted number of utterances	% of utterances that are multisyllabic	% of multisyllabic utterances containing only vowels	% of single syllable utterances containing only a vowel
HL1	63	25.4	86.7	97.6
HL2	40	75.0	83.3	80.0
HL3	35	5.7	100.0	89.3
HL4	36	16.7	100.0	95.8
HL5	97	33.0	93.3	93.9
HL6	81	78.0	13.3	38.5
HL7	21	23.8	100.0	92.3
HL8	44	27.3	100.0	100.0
HL9	28	64.3	61.1	75.0
HL10	26	59.1	61.5	87.5
HL M \pm SD	47.1 \pm 25.2	40.8 \pm 25.9	79.9 \pm 27.8	85.0 \pm 18.1
NH1	60	13.3	0.0	45.2
NH2	30	6.7	0.0	33.3
NH3	49	24.5	75.0	66.7
NH4	30	33.3	90.0	100.0
NH5	48	14.6	72.7	93.6
NH6	206	12.6	34.6	78.4
NH7	96	20.8	0.0	42.3
NH8	29	10.3	0.0	50.0
NH9	55	12.7	0.0	82.6
NH10	70	52.9	33.3	70.0
NH M \pm SD	67.3 \pm 53.0	20.2 \pm 13.9	30.6 \pm 36.5	66.2 \pm 22.8

Note. NH = normal hearing.

For two-syllable utterances, the only statistically significant difference between the two groups, besides the V~V utterances discussed above, was for the syllable shape VCV, $t(1, 9.765) = 2.979, p < .05$, for unequal variances. For children with HL, these utterances constituted $9.1\% \pm 14\%$ of their two-syllable utterances, and for children with NH, these were $36.7\% \pm 28.5\%$ of their two-syllable utterances. Although there were no other statistically significant differences between the groups for other syllable shapes with one- and two-syllable utterances with medial and final consonants (i.e., CVC, CVCV, VCVC, or CVCVC), in all cases the mean percentage for the NH group was greater than for the HL group. Thus, there was a greater tendency for children with NH to produce medial and final consonants in one- or two-syllable utterances than for the children with HL.

Consonant Types

The numbers and types of consonants produced by children with HL and children with NH vary widely (see Table 3). Despite this variation, there were two statistically significant differences between the groups in the types of consonants produced. Children with NH were more likely to produce fricatives than children with HL ($14.3\% \pm 17.5\%$ for NH vs. $1.2\% \pm 2.6\%$ for HL), $t(1, 18) = 2.346, p < .05$ (with unequal variance). There were two classes of stop

consonants that were transcribed: those that were AVS and those that were BLS. Children with NH produced $69.6\% \pm 25.8\%$ of their stops as AVS, whereas children with HL produced only $34.6\% \pm 33.5\%$ of their stops as AVS, which is significant, $t(1, 15) = 2.329, p < .05$, with unequal variance. Finally, there was a group difference in the average percentage of nasal consonants produced, 57.0% for children with HL versus 32.1% for children with NH. However, statistical significance was not attained because of large variation among individual percentages in nasal consonant production.

Vowel Formant Frequencies

The average mean F1 was 906 ± 208 Hz for the children with HL and 915 ± 241 Hz for the children with NH. The average mean F2 was 2305 ± 237 Hz and 2423 ± 678 Hz for the children with HL and the children with NH, respectively. The differences in means of the first two formant frequencies between the groups were not significant.

There were differences between the HL and NH groups in the ranges of formant frequencies. A scatterplot of F1 and F2 measured from vowels for both groups of speakers is shown in Figure 1. (Note that one data point can represent multiple tokens with the same F1 and F2 values.) The data points were from the non-nasal vowels that were a part of the detailed acoustic analysis, and,

Table 3. Adjusted number of consonants (adjusted inversely proportional to the duration of recording) and consonant statistics.

Hearing status, participant	Adjusted number of consonants	% of consonants that are fricatives	% of stop consonants that have a velar-alveolar place	% of consonants that are nasals
HL1	8	0.0	50.0	12.5
HL2	47	2.1	0.0	95.7
HL3	6	0.0	—	100.0
HL4	2	0.0	—	100.0
HL5	10	0.0	0.0	60.0
HL6	287	1.9	72.0	11.1
HL7	2	0.0	50.0	0.0
HL8	2	0.0	—	100.0
HL9	28	0.0	0.0	82.1
HL10	14	8.3	70.0	8.3
HL M ± SD	32.4 ± 63.4	1.2 ± 2.6	34.6 ± 33.5	57.0 ± 44.0
NH1	60	3.3	30.4	71.7
NH2	25	0.0	59.1	12.0
NH3	29	0.0	25.0	58.6
NH4	1	0.0	100.0	0.0
NH5	7	14.3	60.0	28.6
NH6	77	24.7	94.6	26.0
NH7	82	45.1	83.3	0.0
NH8	16	0.0	100.0	81.3
NH9	12	41.7	66.7	25.0
NH10	56	14.3	71.1	17.9
NH M ± SD	36.5 ± 29.8	14.3 ± 17.5	69.6 ± 25.8	32.1 ± 28.8

Note. Em dashes indicate “not applicable.”

thus, there are usually three data points for each vowel analyzed. (In exceptional cases, when a formant frequency could not be determined, it was not recorded.) Figure 1 shows that there is more variation in F2 for the NH group than for the HL group. Most of the extra variation in F2 for NH children occurs for the high vowels, or for low F1.

Children with NH consistently have more variability in F2 than children with HL on the basis of the difference in the group means of the individual standard deviations in F2. Figure 2 shows these data in box-and-whisker format. (The boxes indicate the 25th–75th percentiles and the whiskers the extreme values.) Also, Table 4 shows that only 2 children with HL had F2 standard deviations that exceeded the minimum NH standard deviation of 300.9 Hz. The mean F2 standard deviation for the NH group (i.e., the mean of the 10 individual standard deviations) was 524 ± 111 Hz, and it was 236 ± 81 Hz for the HL group, a difference that was highly significant, $t(1, 18) = 6.309, p < .001$. However, the means in the standard deviations for F1 for the two groups were very similar. The mean F1 standard deviation for the NH group was 196 ± 81 Hz, and it was 174 ± 86 Hz for the HL group.

Discussion

We conducted the analyses in this study to examine whether early identification of hearing loss, early amplification, and early intervention facilitate the speech development of children with HL at 12 months of age. Measures in three areas—syllable shape, consonant type, and formant variability—consistently showed that children with HL behave differently than age-matched children with NH.

The results comparing syllable shape counts between children with HL and children with NH largely conformed to the results of previous research. The children

Figure 1. F1–F2 scatterplot for vowels produced by children with HL and children with NH.

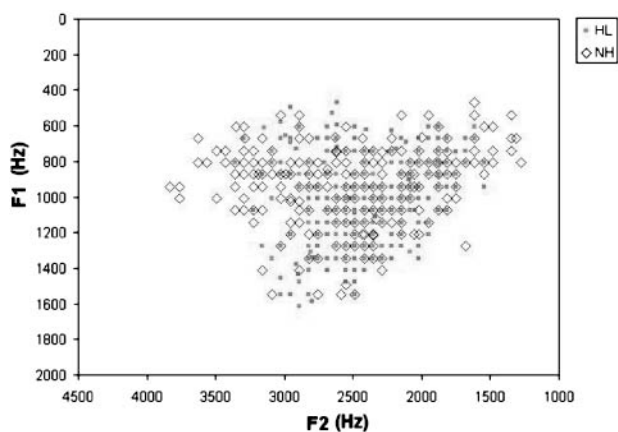
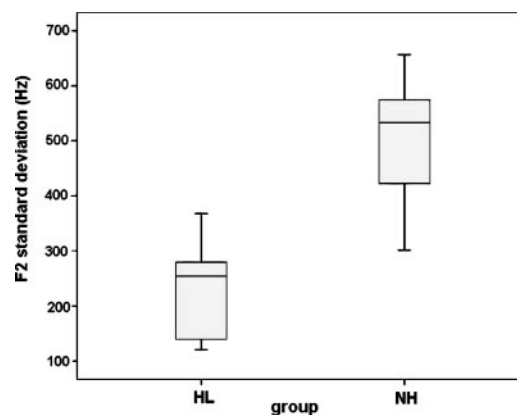


Figure 2. Box-and-whisker plot for F2 standard deviation for HL and NH groups.



with NH produced certain consonants more regularly than did children with HL. Acoustic analyses, in this and previous studies, are providing useful information in understanding the differences in children with HL and children with NH more thoroughly than transcriptional analysis alone, even at the age of 12 months.

Results for syllable shape compare well with previous findings. For instance, the 60% of total utterances that were vowels for five 1-year-old children with NH reported by Kent and Bauer (1985) is within 1 standard deviation of the percentage of single-vowel utterances for children with NH, $45\% \pm 25\%$, and children with HL, $48\% \pm 18\%$. Overall, Kent and Bauer found that children with HL produced more multisyllable utterances than did the children with NH, but most often the children with HL produced multisyllables of the form $V \sim V \sim \dots V$. This is also in agreement with Koopmans-van Beinum and Doppen (2003). The latter researchers found that children with HL produced more multisyllabic utterances, in which they included multisyllabic utterances with only voice breaks between the vowels. The children with NH produced more multisyllable utterances with at least one obstruent or nasal consonant than did the children with HL. These results are in agreement with Oller et al. (1985) and Oller and Eilers (1988), who found that canonical babbling, which involves the production of CVs, was delayed for children with HL.

We also discovered differences between the group of children with HL and the group of children with NH in the types of consonants produced. In particular, children with NH were more likely to produce stops with AVS places of articulation and to produce fricatives than children with HL. These results, although significant statistically, include data from some individuals who produced few consonants. The importance of consonant type will most likely be revealed in future studies as these children progress.

Table 4. The number of measurements and standard deviations of the first two formant frequencies, F1 and F2.

Hearing status, participant	Number of F1 measurements	F1 individual SD (Hz)	Number of F2 measurements	F2 individual SD (Hz)
HL1	33	76.9	33	120.6
HL2	12	261.9	12	279.2
HL3	6	60.7	6	139.4
HL4	10	156.1	10	245.8
HL5	64	260.4	68	263.0
HL6	78	164.8	76	263.1
HL7	9	134.1	9	134.7
HL8	38	236.9	38	236.2
HL9	5	87.9	6	308.5
HL10	26	301.9	27	367.5
HL M ± SD	28.1 ± 25.6	174 ± 86	28.5 ± 25.7	236 ± 81
NH1	50	161.9	50	492.8
NH2	27	192.2	27	385.9
NH3	12	310.3	12	574.1
NH4	21	112.5	21	574.1
NH5	12	135.8	12	618.4
NH6	166	328.8	166	300.9
NH7	71	242.1	72	573.5
NH8	39	189.5	39	518.3
NH9	25	74.3	31	656.6
NH10	72	209.2	71	547.37
NH M ± SD	49.5 ± 46.4	196 ± 81	50.1 ± 46.1	509 ± 111

The acoustic analyses reveals that both the HL and NH groups have attained a normal variation in F1 for their 12 months of age, but the HL group did not move as far from the mean F2 as often as the NH group did, as measured by average standard deviation. Also, the scatterplot (see Figure 1) reveals that the range of F2 was greater for children with NH than for children with HL. (The expansion of the range in F1, particularly toward low vowels, appears to occur in the first 6 months of life for children with NH; Kent & Murray, 1982.) In articulatory terms, the front-back tongue movement for children with NH was more likely to be more distant from mid-position than for children with HL, particularly for high vowels. There appears to be no differences in the use of tongue height between these groups of children.

These acoustic observations agree with the observations of Kent et al. (1987) for the twin brothers: one with HL, and the other with NH. The brother with HL had centralized vowels in which F2 did not approach the maximum F2 of the brother with NH at 12 and 15 months. Kent et al. (1987) also found a restricted F1 range for the brother with HL, but we did not observe this in our group data. The acoustic measurements reported here parallel those of Rvachew et al. (1996) for children with early- and late-onset otitis media. Rvachew et al. found that children with early-onset otitis media had a restricted F2 range but not F1 range. Furthermore, just as the average F1 and F2 do not statistically differ between the early- and late-onset groups, they do not differ between the NH and HL groups in this study.

In summary, these analyses show that the speech production of children with HL differs from that of children with NH at 12 months of age. The differences are exhibited in the relative numbers of multisyllable utterances with obstruent or nasal consonants. Furthermore, there are significant differences in fricative production and place of articulation for stop consonants. These latter differences will need to be explored longitudinally, as the children acquire speech.

Acoustic analyses also provide insight into the articulatory behavior of the children. For instance, the shape of the F1–F2 scatterplot (see Figure 1) during vowel production indicates that the children with HL and the children with NH use the tongue height dimension to about the same extent but that the children with NH use the front-back dimension more than the children with HL. In terms of the frame-content hypothesis of Davis and MacNeilage (1995), the children with HL are not filling the height variation frames with front-back content as much as the children with NH at the age of 12 months. This kind of acoustic analysis can be done without reference to vowel category, which is problematic at this age, and it complements the information gained from a transcription-based analysis.

In future work, speech samples from these children will be examined with the same acoustic parameters and refined transcription as they grow older. Furthermore, it will become possible to assign phonemic identity to the segments that are analyzed, thus providing a finer grained picture of speech production in these two groups

of children. In any event, it is clear that young children must learn about the speech gestures in their language community (e.g., Boysson-Bardies et al., 1989). These gestures include such things as tongue fronting and backing. The children with HL had not accomplished these goals as well as their peers with NH by 12 months of age, in spite of early intervention. Researchers need to ensure that early intervention consists of ample opportunity to hear the speech of others, with an emphasis on providing examples of complete language samples so that children with HL have access to a range of production.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grant NIDCD-06237, awarded to the second author. We thank Chris Chapman for helping to maintain some of the statistics shown here.

References

- Boysson-Bardies, B., Sagart, L., Halle, P., & Durand, C.** (1986). Acoustic investigations of cross-linguistic variability in babbling. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of early speech* (pp. 113–126). New York: Stockton Press.
- Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C.** (1989). A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language*, *16*, 1–17.
- Buhr, R. D.** (1980). The emergence of vowels in an infant. *Journal of Speech and Hearing Research*, *23*, 73–94.
- Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., et al.** (1994). An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America*, *96*, 2108–2120.
- Commission on Education of the Deaf.** (1988). *Toward equality: Education of the deaf. A report to the President and Congress of the United States*. Washington, DC: U.S. Department of Education and Bureau of Maternal and Child Health.
- Davis, B. L., & MacNeilage, P. F.** (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research*, *33*, 16–27.
- Davis, B. L., & MacNeilage, P. F.** (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, *38*, 1199–1211.
- Hollingshead, A.** (1965). *Two factor index of social position*. Cambridge, MA: Harvard University Press.
- Kent, R. D.** (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, *5*, 7–23.
- Kent, R. D., & Bauer, H. R.** (1985). Vocalizations of one-year-olds. *Journal of Child Language*, *12*, 491–526.
- Kent, R. D., & Murray, A. D.** (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, *72*, 353–365.
- Kent, R. D., Osberger, M. J., Netsell, R., & Hustedde, C. G.** (1987). Phonetic development of identical twins differing in auditory function. *Journal of Speech and Hearing Disorders*, *52*, 64–75.
- Koopmans-van Beinum, F. J., & Doppen, L.** (2003). Development of deaf and hearing infants. *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 1033–1066). Barcelona, Spain.
- Moeller, M. P.** (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, *106*, E43.
- Nittrouer, S.** (in press). Development outcomes in the first two years of life: Effects of early identified hearing loss and sign language. *Ear and Hearing*.
- Nittrouer, S., & Burton, L. T.** (2005). The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of Communication Disorders*, *38*, 29–63.
- 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.
- 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). Preschool vocalizations of a deaf infant: A comparison with normal metaphonological development. *Journal of Speech and Hearing Research*, *28*, 47–63.
- Rathmann, P.** (1994). *Good night, gorilla*. New York: Putnams.
- Rescorla, L.** (1989). The language development survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders*, *54*, 587–599.
- Rvachew, S., Mattock, K., Polka, L., & Ménard, L.** (2006). Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *Journal of the Acoustical Society of America*, *120*, 2250–2259.
- Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. L.** (1996). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics*, *24*, 19–28.
- Stoel-Gammon, C., & Otomo, K.** (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders*, *51*, 33–41.
- Yoshinago-Itano, C., Coulter, D., & Thompson, V.** (2001). Developmental outcomes of children with hearing loss born in Colorado hospitals with and without universal newborn hearing screening programs. *Seminars in Neonatology*, *6*, 521–529.
- Yoshinago-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L.** (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, *102*, 1161–1171.

Received February 8, 2007

Accepted October 17, 2007

DOI: 10.1044/1092-4388(2008/064)

Contact author: Richard S. McGowan, CReSS LLC,
1 Seaborn Place, Lexington, MA 02420.

E-mail: rsmcgowan@earthlink.net.