Acoustic Characteristics of the Speech of Young Cochlear Implant Users: A Comparison with Normal-Hearing Age-Mates

Rosalie M. Uchanski and Ann E. Geers

Objective: **The primary objective of this study was to compare select acoustic characteristics of the speech of deaf children who use cochlear implants (young cochlear implant users) with those of children with normal hearing. A secondary objective of this study was to examine the effect, if any, of the deaf child's education (oral versus total communication) on the similarity of these acoustic characteristics to those of normal-hearing age-mates.**

Design: **Speech was recorded from 181 young cochlear implant users and from 24 children with normal hearing. All speech was produced by imitation, and consisted of complete sentences. Acoustic measures included voice onset time (/t/, /d/), second formant frequency (/i/, /ɔ/), spectral moments (mean, skew and kurtosis of /s/ and /ʃ/), a nasal manner metric, and durations (of vowels, words, and sentences).**

Results and Discussion: **A large percentage (46 to 97%) of the young cochlear implant users produced acoustic characteristics with values within the range found for children with normal hearing. Exceptions were sentence duration and vowel duration in sentence-initial words, for which only 23 and 25%, respectively, of the cochlear implant users had values within the normal range. Additionally, for most of the acoustic measures, significantly more cochlear implant users from oral than from total communication settings had values within the normal range.**

Conclusions: **Compared with deaf children with hearing aids (from previous studies by others), deaf children who use cochlear implants have improved speech production skills, as reflected in the acoustic measures of this study. Placement in an oral communication educational setting is also associated with more speech production improvement than placement in a total communication setting.**

(Ear & Hearing 2003;24;90S–105S)

The speech of deaf children has been studied for many years (e.g., Hudgins & Numbers, 1942; Smith, 1975; Tobey, Geers, & Brenner, 1994), and continues to be studied because deaf speech is often not normal and often not highly intelligible (e.g., Monsen, 1983; Osberger, Maso, & Sam, 1993; Porter & Bradley, 1985). Reports on the speech of deaf children have examined 1) errors in their speech production (Hudgins & Numbers, 1942; Smith, 1975); 2) differences in the speech of deaf children as a function of hearing loss and/or perceptual abilities (Smith, 1975; Tye-Murray 1992; Tye-Murray, Spencer, & Gilbert-Bedia, 1995); 3) differences in the speech of deaf children as a function of hearing device (Osberger, Robbins, Berry, Todd, Hesketh, & Sedey, 1991; Osberger et al., 1993; Tobey et al., 1994); 4) longitudinal changes in the speech of deaf children and the assessment thereof (Fourakis, Geers, & Tobey, 1993; Tobey et al., 1994); or 5) deviations in the speech acoustics of deaf children from those of normal-hearing children (Angelocci, Kopp, & Holbrook, 1964; Monsen, 1976a, 1976b). However, few previous studies have examined the similarity in acoustic characteristics of deaf children's speech with those of normal-hearing children. Until the recent successes of many children with cochlear implants (Tobey, Geers, Brenner, Altuna, & Gabbert, 2003), such similarity comparisons with normal-hearing children have generally not been considered, largely because deaf children's speech was so obviously impaired.

Measurement of acoustic energy present in various frequency regions of the speech signal for different phonetic contrasts can provide an objective index of perceived differences in speech sound production. Acoustic contrasts have been used not only to characterize speech sounds, but to simulate and reproduce human speech and to describe the nature of speech production disorders. With regard to the speech of individuals with hearing loss, certain acoustic characteristics have been associated with the intelligibility of their speech to hearing listeners (Metz, Samar, Schiavetti, Sitler, & Whitehead, 1985; Monsen, 1978). Although there is considerable range in the values obtained from acoustic analysis of speech sounds produced by intelligible speakers, it is generally assumed that deviations outside of this range are associated with reduced intelligibility. The inability to produce specific phonetic contrasts (e.g., to differentiate between high

Central Institute for the Deaf, St. Louis, Missouri.
DOI: 10.1097/01.AUD.0000051744.24290.C1

0196/0202/03/241S-0090S/0 • Ear & Hearing • Copyright © 2003 by Lippincott Williams & Wilkins • Printed in the U.S.A.

and low vowels) may reflect a more general lack of control of the articulators, which results in problems producing a wide variety of phonetic contrasts. The study reported here examines a set of acoustically measured speech features that may be associated with more general aspects of speech production skill in deaf talkers. To the extent that the acoustic measurements resemble those of hearing talkers, these deaf talkers may be considered to have good articulatory control and hence, presumably are more likely to have intelligible speech. Admittedly, though, the relation between acoustic measures and intelligible speech is not simple or well understood.

In this study, two questions are examined. First, do children with cochlear implants produce speech sounds with acoustic values similar to those of normal-hearing children? And, in particular, more similar to normal-hearing children than has been reported in the past for children with profound deafness who used hearing aids? One might expect that the use of cochlear implants would result in improved speech production skills in children with profound hearing impairment, due to the implant's delivery of more auditory information relative to hearing aids. There is good evidence that the more recent processing strategies in cochlear implants provide better perception of consonants than both older implant processing strategies and hearing aids (Geers & Brenner, 1994; Tobey et al., 2003). For example, with a hearing aid, fricative sounds with high-frequency energy such as /s/ and /ʃ/ might be inaudible. Even when these sounds are audible, they might still be indistinguishable (Boothroyd, 1984). However, recent reports on postlingually deafened adults indicate these fricative sounds are often distinct in their perceived manner and place of articulation when perceived through a cochlear implant (Skinner, Fourakis, Holden, Holden, & Demorest, 1999).

The second question addressed in this study is whether the communication mode used in the child's education (oral versus total communication) has an effect on the acoustic characteristics of the speech of deaf children who use cochlear implants. Though cochlear implants may deliver more auditory information than hearing aids for the profoundly deaf, cochlear implants are still deficient in their delivery of auditory information relative to normal hearing. Hence, one might expect the type of special education a child receives to have an effect on speech production skills (Geers & Moog, 1992). Specifically, oral education programs, with their emphasis on speech production and perception, might have a beneficial effect on the speech production skills of deaf children using cochlear implants, as manifest in the acoustic characteristics of their speech.

In this article, we present select acoustic characteristics of the speech of young deaf children who use cochlear implants and we compare these characteristics with those of similarly aged children with normal hearing. In addition, we examine whether the acoustic characteristics of speech from deaf children educated in an oral communication (Oral) setting are closer to those of normal-hearing children than are the acoustic characteristics of speech from deaf children educated in a total communication (TC) setting. Finally, when possible, the acoustic data from our cochlear implant users are compared with previously reported data from deaf talkers without implants who presumably wore hearing aids.

The intelligibility of these children, as assessed by naïve human listeners, is described in detail in an accompanying article by Tobey et al. (2003). Although acoustic characteristics do not necessarily correspond directly to the quality or intelligibility of deaf speech to normal-hearing listeners, they provide an objective means of documenting differences in articulation of specific speech sounds between groups of children. Intelligibility judgments, although presenting greater face validity, contain considerable variability and subjectivity within each rating. These acoustic measurements permit us to examine whether the greater intelligibility of children enrolled in Oral settings when compared with that of children educated in TC settings (see Tobey et al., 2003) is also reflected in more "normal" values on a more objective metric. We hypothesize that the positive effects of a spoken language emphasis in a child's educational program combined with increased auditory information provided by the cochlear implant will be reflected in his or her speech through acoustically measured values that more closely resemble those of their hearing age-mates.

METHODS

Participants

Speech was recorded from 181 8- and 9-yr-old children with prelingual profound deafness who had worn a cochlear implant for at least 4 yr. Detailed information on subject selection criteria and demographics are provided in Geers and Brenner (2003). The educational variable examined in the following analysis was based on the communication mode used in the child's classroom before and since receiving a cochlear implant. Ratings of classroom communication mode that were provided by the child's parents included three levels of TC programs (1-sign emphasis; 2-equal speech and sign emphasis; and 3-speech emphasis) and three levels of Oral programs (4-cued speech; 5-auditory/oral; and 6-audito-

TABLE 1. Table of sentences for acoustic measurement.

	VOT		F2		Moments			Ratio Vow/ Word	
Orthography	t	d	i	ɔ	s	ʃ	Nasal Metric	Init	Final
1. I am **ta**ll.	1			1					
2. H**e** has a blue *pen*.			1						1
3. *Chuck* **see**ms thirsty after the *race*.			1					1	1
4. **She** **nee**ds str**aw**berry jam on her **t**oast.	1		1	1		1	1		
5. **D**id you like the zoo this spring?		1							
6. **D**addy **t**ook his **n**ew ***sh**oes*.	1	1				1	1		1
7. **Sh**ow **me** the little ***d**uck*.		1	1			1	1		1
8. ***S**it* at the **t**able please.	1				1			1	
9. ***S**ue's* friend b**ou**ght a **t**oy ***sh**ip*.	1			1	1	1		1	1
10. *Feel* the **so**ft ***d**og*.		1		1	1			1	1
11. **M**ay I **see** that *rock*?			1		1		1		1
Totals	5	4	5	4	4	4	4	4	7
Use both repetitions	—	—	—	—	×2	×2	×2	—	—
Grand totals	5	4	5	4	8	8	8	4	7

Italics indicate words used for duration measurements. Letters in bold indicate sounds used for VOT (voice-onset-time), F_2 (second formant frequency), moment (spectral moments), and nasal metric measurements.

ry/verbal). Communication mode ratings, averaged across the preimplant year and 4 yr after implantation, were used to summarize this variable. A total of 89 children obtained average mode ratings between 1.0 and 3.9, indicating predominant placement in total communication classrooms. The remaining 92 children obtained average mode scores between 4.0 and 6.0, indicating predominant placement in oral classrooms. Although the average mode rating does not necessarily correspond with consistent use of a particular communication mode, it does reflect the overall degree of emphasis on speech and auditory skill development over a period of years.

Speech was also recorded from 24 8- and 9-yr-old children with normal hearing. The children with normal hearing were recruited from a local elementary school and from another study at Central Institute for the Deaf (CID).

Recording Procedures

For the deaf children, speech was recorded in a quiet hotel room, and for the normal-hearing children, speech was recorded in a quiet observation room at CID. Recording equipment was the same for all children. Each child used an Audio-Technica ATM75 cardioid headworn microphone and the teacher/clinician wore an Audio-Technica AT803B omnidirectional lavalier microphone. Speech was recorded to both a Sony 75ES DAT recorder and a Technics RS-B107 analog cassette recorder. For some speech materials, speech was also recorded directly to a computer, using a sampling frequency of either 20 KHz or 22.05 KHz depending on the recording site and equipment.

Speech was elicited by imitation. A teacher/clinician, fluent in sign and signed English, communicated with each child using his or her preferred mode of communication. The deaf child was instructed to use his or her "best speech," regardless of any simultaneous use of sign, and was instructed to repeat each sentence after the teacher/clinician produced it. The teacher/clinician first said (or said and signed) a sentence. Then, before the child spoke, a printed version of the sentence (either on an index card or computer screen) was made available to the child. For the children with normal hearing, the talker was instructed to repeat each sentence after it was produced by the teacher/clinician. For these children, no special instructions were given regarding speech clarity or enunciation.

Speech Materials

A list of 11 sentences (SAM: Sentences for Acoustic Measurement) was recorded twice from each talker. Table 1 lists these SAM sentences, and indicates the words and sounds chosen for subsequent acoustic analyses. In addition, 36 sentences from the study by McGarr (1983) were recorded to assess the intelligibility of each talker as judged by adult normal-hearing listeners (Tobey et al., 2003) and were used in this analysis to provide a measure of sentence duration.

Acoustic Analyses

Twelve acoustic characteristics were measured from the recorded sentences. They were selected based on three criteria: 1) their previous use in studying the speech of deaf children (Goldhor, 1995;

Monsen, 1974, 1976b, 1976c), 2) their previous association with intelligibility judgments (Metz et al., 1985; Monsen, 1978), or 3) their likelihood of being affected by the high-frequency sensitivity provided by a cochlear implant (Skinner et al., 1999).

The following acoustic measurements were made for each talker (see Table 1):

1. VOT, voice onset time, /t/
2. VOT, voice onset time, /d/
3. F_2, second formant frequency, /i/
4. F_2, second formant frequency, /ɔ/
5. Spectral moments (mean, skew, and kurtosis) /s/
6. Spectral moments (mean, skew, and kurtosis) /ʃ/
7. Nasal manner metric /m,n/
8. Vowel duration in sentence-initial position CVC-type words (/Λ/, /I/, /u/, /i/)
9. Word duration of sentence-initial position CVC-type words (/Chuck/, /Sit/, /Sue's/, /Feel/)
10. Vowel duration in sentence-final position CVC-type words (/ϵ/, /e/, /u/, /Λ/, /I/, /ɔ/, /a/)
11. Word duration of sentence-final position CVC-type words (/pen/, /race/, /shoes/, /duck/, /ship/, /dog/, /rock/)
12. Average sentence duration (McGarr sentences)

Measurements of voice onset time (VOT), F_2, and duration were made by hand. Measurements of spectral moments and the nasal metric were made by computer programs once the sound segments' time boundaries had been marked by hand (cursors placed at appropriate time locations and marked with labels). Details about each measurement are provided below.

Voice Onset Time • Measurement of the VOTs in English plosive sounds provides a useful means of distinguishing acoustically between sounds that lie on a continuum between the fully prevoiced and unaspirated to the completely unvoiced and heavily aspirated stops (Lisker & Abramson, 1965). The ability of a deaf talker to produce plosive sounds with typical VOTs presumably reflects an ability to hear this voicing distinction and may reflect a general skill in producing brief, properly timed speech events. In addition, the difference in VOT for /t/ versus VOT for /d/ has been found to account for a significant amount of the variance in speech intelligibility scores for deaf children's speech (Monsen, 1978). VOT for the English alveolar stops was measured as the time from the plosive burst release to the onset of voicing or phonation. Both waveform and spectrogram displays were used to determine the times associated with these events, namely burst release and onset of voicing. VOT is positive if voicing occurs after the release of the plosive burst. VOTs for /t/ and /d/ were measured only for sound segments in the intended position that were plosive-like in manner. If a talker substituted a sound with a different manner of articulation, e.g., produced a fricative-like /s/ for an intended /t/ or /d/ sound, no VOT measurement was made. If a talker produced a plosive with a different place of articulation for an intended alveolar plosive, the VOT for that produced plosive was measured. As a byproduct of this VOT measurement rule, we tabulated the percent of intended alveolar plosives that were produced as plosive-like sounds. As shown in Table 1, the maximum number of plosives examined is nine (five /t/s and four /d/s) for the SAM sentences. If a talker produced all nine of these /t/s and /d/s as plosive-like sounds, he or she would have produced 100% "good plosives." For each talker, the average VOT of the five possible /t/s, the average VOT of the four possible /d/s, and the difference of the averages (avgVOT /t/ − avgVOT /d/) were computed.

Second Formant Frequency (F_2) • The vowels /i/ and /ɔ/ have the highest and lowest second formant frequencies, respectively, amongst all the vowels of American English (Peterson & Barney, 1952). Production of these vowels with the appropriate F_2 values presumably reflects appropriate articulatory gestures of the tongue for these vowels, and may reflect a general ability to control tongue movement for all vowels sounds (Monsen, 1976b). Similar to the VOT difference, the difference in F_2 for /i/ versus F_2 for /ɔ/ has also been found to account for a significant amount of the variance in speech intelligibility scores for deaf children's speech (Monsen, 1978). F_2 was measured at the approximate midpoint of the steady state region of the intended vowel sound. The spectrogram display was used to determine this approximate midpoint. Then, both FFT and LPC spectra were computed. LPC spectra were computed using the autocorrelation method, a 12 msec Hamming window, pre-emphasis ($a = 0.98$), and 24 coefficients. F_2 was recorded using the LPC estimate, and was confirmed using a wideband spectrogram display. F_2 was measured for the vowel sound produced in the position of the intended /i/ or /ɔ/, regardless of the perceived vowel identity. For each talker, the average F_2 of the five /i/ vowels, the average F_2 of the four /ɔ/ vowels, and the difference of the averages (avgF_2/i/ − avgF_2/ɔ/) were computed.

Spectral Moments • The fricative sounds /s/ and /ʃ/ have prominent spectral peaks at very high frequencies (e.g., 3300 to 7500 Hz; Jongman, Wayland, & Wong, 2000). As such, these fricative sounds might be perceived more easily and distinctly with cochlear implants than with hearing aids. Recently, many researchers have used statistical measures (usually, mean, skewness, and kurtosis) of spectral energy distributions as acoustic descriptors of fricative sounds because these measures summarize nicely their spectral concentration, tilt, and peakedness (e.g., Avery & Liss; 1996; Forrest, Weismer, Milenkovick, & Dougall, 1988; Jongman et al., 2000). Appropriate placement of the articulators (tongue, lips, etc.) in the production of these two fricative sounds is reflected in a greater spectral mean for /s/ than for /ʃ/ (Forrest et al., 1988; Jongman et al., 2000). And, vice versa. That is, a greater spectral mean for /s/ than for /ʃ/ presumably reflects appropriate articulator placement for those fricative sounds. Based on previous research on fricatives, three spectral moments of the sounds /s/ and /ʃ/ were computed (Forrest et al., 1988). Using both waveform and spectrogram displays, the beginning and end of each fricative sound were labeled. A custom-written program using Entropics software performed the spectral moments calculations for each labeled sound segment as follows. First, the speech was pre-emphasized and high-pass filtered (f_c = 80 Hz). Then, over the duration of the fricative sound, FFT spectra were computed every 20 msec. From the real and imaginary components of the Fourier spectra, power spectra were computed and then averaged across the duration of the sound segment, omitting the very first and very last spectra for that segment. The resultant average power spectrum was replaced by a weighted average power spectrum using a Bark scale conversion (Forrest et al., 1988; Syrdal & Gopal, 1986). Then, each component of this weighted average power spectrum was normalized by the total power, yielding component values consistent with a probability distribution (that is, individual component values between 0 and 1 with their sum across Barks equal to 1). This resultant probability distribution represents a normalized average Bark spectrum of a fricative sound. Finally, moments were calculated from this probability distribution using standard statistical formulas for central moments and coefficients of these moments. Based on previous studies, the 1st central moment, and the coefficients of the 3rd and 4th central moments were calculated. These are known as the spectral mean, coefficient of skewness and coefficient of kurtosis, respectively (Forrest et al., 1988). Analogous to the rule used for plosive measurements, spectral moments for /s/ and /ʃ/ were measured only for sound segments in the intended position that were fricative-like in manner. If a talker substituted a sound with a different manner of articulation, e.g., produced a plosive-like /t/ for the intended /s/ or /ʃ/ sound, no spectral moments were measured. If a talker produced a fricative with a different place of articulation for the intended /s/ or /ʃ/, spectral moments were computed as described above. As a consequence of this fricative measurement rule, we tabulated the percent of intended /s/ and /ʃ/ fricatives that were produced as fricative-like sounds. As shown in Table 1, the maximum number of plosives examined is 16 (eight /s/s and eight /ʃ/s) for the SAM sentences. If a talker produced all 16 of these /s/s and /ʃ/s as fricative-like sounds, he or she would have produced 100% "good fricatives." For each talker and each of the three spectral moments, the average of the eight possible /s/s, the average of the eight possible /ʃ/s, and the difference of the averages (avg/s/ − avg/ʃ/) were computed.

Nasal Manner Metric • Previously, Monsen (1978) found that a measure of the correct manner of production of nasal and liquid sounds correlated well with speech intelligibility scores for deaf children's speech. This measure, though numerical, was based on a subjective assessment of the distinctiveness of the boundary between a syllable-initial nasal (or liquid) and the following vowel. More recently, Goldhor (1995) developed a useful acoustic correlate of nasal manner phonetic judgments for the speech of hearing-impaired children. This acoustic correlate, an objective nasal manner metric, is composed of three elements; 1) the duration of the nasal segment (*dur*), 2) the difference between the total log power in the consonant and the log power below 750 Hz in the consonant (*consdiff*), and 3) the difference between the log power below 750 Hz in the consonant and the log power below 750 Hz in the vowel of the same syllable (*reldiff*). The values of each of these three elements was converted to a number ranging from 0 to 2 based on formulas developed by Goldhor (1995). Values of 2 or close to 2 occur when a good or typical nasal is produced, and values near 0 are associated with poor nasal productions. The formulas are:

1. If *dur* is greater than 40 msec, D = 2; if *dur* is less than 10 msec, D = 0; else D = 1 + (*dur* − 25)/30
2. If *consdiff* is 0, L = 2; if *consdiff* is greater than 10 dB, L = 0; else L = 2 − 0.2 × *consdiff*
3. If *reldiff* is greater than −5 dB, R = 2; if *reldiff* is less than −25 dB, R = 0; else R = 2.5 + 0.1 × *reldiff*

The new values, represented by the variables D, L, and R, respectively, reflect a Duration measure, a measure of Low-frequency power, and a measure of Relative low-frequency power. Finally, the combined nasal metric is computed by subtracting 1 from the cuberoot of the product, D × L × R; metric = cuberoot(D × L × R) − 1. This resultant metric can have values between −1 and +1, where −1 corresponds to very poorly produced nasals and +1 corresponds to good nasals. The nasal manner metric was computed for eight nasal sounds in word-initial positions, two /n/s and two /m/s with two repetitions each. For each sound segment, the beginning and end times of the intended nasal segment were determined by hand using both spectrogram and waveform displays. Viewed with a spectrogram display, the boundaries of nasal-like sounds were usually determined by the presence of anti-resonances or a lack of energy in frequencies above about 1 kHz, coincident with significant energy at low frequencies. The waveform would generally exhibit a diminished amplitude and a smoother periodic waveform shape relative to the subsequent vowel sound. Using Entropics software, a custom-written program computed the nasal manner metric as described above. This nasal manner metric was computed for the consonant sound produced in the position of the intended /m/ or /n/, regardless of the manner or place of articulation of the consonant produced. For each talker, the average D value, the average L value, the average R value, and the average nasal manner metric were computed across the eight intended nasal tokens.

Vowel and Word Duration • One prominent characteristic of the speech of deaf talkers is the tendency to elongate the duration of vowel sounds in their speech (Pratt & Tye-Murray, 1997). More typical vowel durations may reflect more normally developed speech motor skills. For these speech materials, durations of the vowel and the entire word were measured for monosyllabic, CVC-type words in sentence-initial and sentence-final positions. As shown in Table 1, four instances of CVC-type words in sentence-initial positions and seven instances of CVC-type words in sentence-final positions are used. Both vowel and word duration were measured by hand using waveform and spectrogram displays to determine the beginning and end times for each event. Typically, vowel boundaries were determined by high-amplitude, periodic signals in the waveform display and strong, relatively steady formants in the spectrogram display. However, portions of a vowel sound produced with weak phonation (formants still apparent but not as dark) would also be included in the vowel boundaries. Word boundaries encompassed all the speech-like segments of the syllable associated with the first (or last) word of the sentence. Then, for each word, the vowel duration, the word duration and the ratio of vowel duration to word duration was computed (i.e., voweldur/worddur). Finally, for each talker, the average vowel duration, the average word duration, and the average ratio were computed for the initial and final words separately. Durations were measured regardless of any sound substitutions and/or omissions in the intended monosyllabic words.

Sentence Duration • Previous researchers have found sentence duration (or equivalently, rate of speech production) to be correlated with speech intelligibility (Monsen, 1978) and to account for a large part of the variance in judgments of suprasegmental aspects of deaf children's speech (Smith, 1975). For our recordings, the beginning and the end of each of the 36 McGarr sentences were marked. The time difference between these time events, or total duration of the sentence, was computed and then an average duration across the 36 sentences was found for each talker.

Acoustic measurements of speech, especially of impaired speech, can be difficult. Though great care was taken when making these measurements, due to the large number of measurements (nearly 20,000, of which most were judgments of time boundaries), some errors are bound to exist. To examine measurement repeatability, over 500 measurements were repeated. For VOTs, 74% of the repeated measurements were within ±5 msec of the first measurement, and 92% were within ±10 msec. For vowel and word duration measurements combined, 74% of the repeated measurements were within ±5 msec and 86% were within ± 10 msec. For the formant measurements, 60% were within ± 50 Hz and 77% were within ±100 Hz. Finally, for the time markers (boundaries) associated with the labeling of fricative and nasal sounds, 62% were within ±5 msec and 82% were within ±10 msec (and 92% within ±20 msec). Though there is not perfect agreement between the first and second measurements (the poorest agreement is found for the formant values), repeatability is generally high.

RESULTS AND DISCUSSION

Individual values for each participant have been computed for each acoustic measure and are presented in Figures 1 to 10. The format is the same for all figures. Participants have been divided into three groups in each display, two groups of children with implants (TC and Oral) and one group of children with normal hearing (NH). The 89 children from TC classrooms are grouped on the far left and the 92 children primarily educated in Oral classrooms are in the middle. At the far right are the values obtained from the 24 talkers with normal hearing. Within each group, individual data are ordered from lowest to highest. Within each of the two groups with implants, TC and Oral, we calculate the percent of those talkers who have data values within normal limits (WNL), defined as that range of values from the minimum to the maximum found across the 24 talkers with normal hearing. Table 2 provides a summary of the number and percent of talkers

TABLE 2. Talkers within normal limits

Measure	Number of Talkers TC	Number of Talkers Oral	Percent of Talkers TC	Percent of Talkers Oral	All CI Talkers Number	All CI Talkers Percent
% good plosives	69	87	78	95*	156	86
% good fricatives	33	56	37	61*	89	49
VOT /t/	55	78	62	85*	133	73
VOT /d/	70	81	79	88	151	83
Δ VOT	56	77	63	84*	133	73
F2 /i/	75	84	84	91	159	88
F2 /ɔ/	80	78	90	85	158	87
Δ F2	62	76	70	83	138	76
Spectral mean /s/	42	72	47	78*	114	63
Spectral mean /ʃ/	72	83	81	90	155	86
Δ Spectral mean	48	80	54	87*	128	71
Spectral skew /s/	73	86	82	93*	159	88
Spectral skew /ʃ/	79	86	89	93	165	91
Δ Spectral skew	62	84	70	91*	146	81
Spectral kurtosis /s/	67	83	75	90*	150	83
Spectral kurtosis /ʃ/	81	83	91	90	164	91
Δ Spectral kurtosis	44	69	49	75*	113	62
Nasal metric	60	80	67	87*	140	77
Nasal R factor	34	63	38	68*	97	54
Nasal L factor	74	87	83	95*	161	89
Nasal D factor	84	91	94	99	175	97
Duration of initial vowel	10	36	11	39*	46	25
Duration of initial word	45	61	51	66*	106	59
Vowel to word duration ratio-initial	32	52	36	57*	84	46
Duration of final vowel	51	73	57	79*	124	69
Duration of final word	40	59	45	64*	99	55
Vowel to word duration ratio-final	52	79	58	86*	131	72
Sentence duration	8	33	9	36*	41	23
Number of talkers	89	92			181	

*Talkers are grouped by classroom communication mode, Total Communication (TC) or Oral Communication (Oral), and all together (CI). Statistically significant differences (at the 0.05 level) in the percent of TC and Oral talkers with acoustic measures WNL are indicated by an * (using Fisher's exact probability test).*

(divided into TC and Oral groups, and overall cochlear implant users) who have acoustic values WNL. Additionally, for each measure, Fisher's exact probability test was used to determine whether the distributions for TC and Oral children (i.e., the percent of TC cochlear implant users and the percent of Oral cochlear implant users with acoustic values WNL) are significantly different at the 0.05 level. Significantly different distributions are indicated in Table 2. Finally, presented in Table 3 are inter-talker statistics for each acoustic measure. Specifically, for each talker and for each measure, an average was computed across the sound tokens produced by that talker (e.g., an average of the five VOTs for /t/). Then, across all the talkers in each of three groups (TC, Oral, NH), the inter-talker mean, standard deviation, minimum and maximum for these "token averages" were found.

Figure 1 shows individual data on percent good plosives and percent good fricatives for the children with cochlear implants and the children with normal hearing. This value represents the percentage of productions that matched the intended sound in manner of articulation (e.g., production of /b/ when the target was /d/ was counted as a "good" plosive). All of the normal-hearing children produced 100% of the fricative targets as fricatives, but one hearing child produced something other than a plosive for one target (produced an affricate). Therefore, when the speech of cochlear implant users was judged with respect to normal production, the criterion was more lax for plosive sounds (89%) than for fricative sounds (100%). More cochlear implant users produced plosive sounds at a level comparable with hearing age-mates compared with fricative sounds.

Greater accuracy in the production of plosives than fricatives has been documented in children who use hearing aids (Geffner, 1980; Smith, 1975) and who use cochlear implants (Blamey, Barry & Jacq, 2001; Tobey, Pancamo, Staller, Brimacombe, & Beiter, 1991). In the study by Smith (1975) the speech of 40 severely to profoundly deaf children, who wore hearing aids (age range 8 to 10 and 13 to 15 yr), was examined for correct phoneme production. Averaged across those 40 talkers, the plosives /d/ and /t/ were produced correctly as /d/ and /t/, 40% and 37% of the time. Analogously, the fricatives /s/ and /ʃ/ were produced correctly as /s/ and /ʃ/, 24%

TABLE 3. Inter-talker data for three talker groups.

	TC (n = 89)		Oral (n = 92)		NH (n = 24)	
Measure	mean (SD)	min:max	mean (SD)	min:max	mean (SD)	min:max
% good plosives	87 (23)	0:100	96 (8)	44:100	99 (3)	88:100
% good fricatives	72 (30)	0:100	91 (19)	6:100	100 (0)	100:100
VOT /t/ (msec)	65 (32)	−32:149 **85**	80 (21)	18:142	72 (15)	40:101
VOT /d/ (msec)	17 (11)	−46:37 **86**	17 (9)	−43:36	17 (5)	1:25
Δ VOT (msec)	49 (32)	−40:133 **84**	63 (22)	−5:132	55 (15)	21:83
F2 /i/ (Hz)	2838 (360)	1437:3557	3003 (237)	2437:3451	2978 (181)	2511:3379
F2 /ɔ/ (Hz)	1430 (152)	1144:2052	1408 (143)	1096:1789	1422 (145)	1211:1678
Δ F2 (Hz)	1408 (400)	13:2239	1596 (277)	748:2162	1555 (194)	1248:1983
Spectral mean /s/ (barks)	18.0 (1.1)	13.9:20.4 **79**	18.7 (0.9)	15.8:20.5 **90**	19.4 (0.6)	18.2:20.5
Spectral mean /ʃ/ (barks)	17.6 (0.8)	14.9:19.7 **85**	17.7 (0.7)	15.4:19.4 **91**	17.6 (0.6)	16.4:18.6
Δ Spectral mean (barks)	0.4 (0.9)	−3.3:2.5 **78**	1.0 (0.7)	−1.1:2.6 **89**	1.8 (0.7)	0.2:3.1
Spectral skew /s/	−2.2 (1.3)	−9.2:0.5 **79**	−2.7 (1.2)	−6.2:0.0 **90**	−2.9 (1.2)	−4.7:−0.3
Spectral skew /ʃ/	−1.0 (0.8)	−3.7:0.6 **85**	−0.9 (0.8)	−3.4:0.4 **91**	−0.2 (0.8)	−2.4:0.8
Δ Spectral skew	−1.2 (1.3)	−7.9:0.9 **78**	−1.9 (1.1)	−5.0:0.7 **89**	−2.7 (1.3)	−5.1:−0.2
Spectral kurtosis /s/	16.9 (20.0)	0.4:168 **79**	23.8 (15.0)	0.7:102.6 **90**	23.5 (10.4)	5.7:49
Spectral kurtosis /ʃ/	8.3 (5.4)	1.2:29.9 **85**	8.6 (7.6)	0.7:45.0 **91**	5.4 (4.6)	0.2:19
Δ Spectral kurtosis	9.0 (19.9)	−19.6:159 **78**	15.3 (13.1)	−6.0:96.5 **89**	18.1 (10.4)	3.3:40
Nasal metric	0.71 (0.3)	−1:1.0	0.86 (0.2)	−0.4:1.0	0.94 (0.1)	0.70:1.0
Nasal R factor	1.58 (0.4)	0.0:2.0	1.78 (0.3)	0.4:2.0	1.95 (0.1)	1.80:2.0
Nasal L factor	1.95 (0.2)	0.2:2.0	1.98 (0.1)	1.6:2.0	2.0 (0.01)	1.97:2.0
Nasal D factor	1.91 (0.2)	1.1:2.0	1.95 (0.1)	1.2:2.0	1.91 (0.1)	1.40:2.0
Duration of initial vowel (msec)	270 (103)	140:701	210 (85)	108:604	138 (19)	106:176
Duration of initial word (msec)	509 (119)	255:832	480 (125)	303:1006	382 (49)	318:512
Vowel to word duration ratio-final	0.54 (0.16)	0.32:1.00	0.44 (0.12)	0.25:0.78	0.36 (0.04)	0.30:0.44
Duration of final vowel (msec)	290 (136)	153:1307	242 (62)	150:464	210 (29)	161:277
Duration of final word (msec)	566 (132)	297:1355	546 (79)	407:797	498 (44)	414:563
Vowel to word duration ratio-initial	0.52 (0.14)	0.28:0.96	0.44 (0.09)	0.31:0.89	0.41 (0.05)	0.31:0.51
Sentence duration (sec)	2.94 (1.11)	1.58:7.78	2.22 (.68)	1.30:4.51	1.32 (.15)	1.14:1.83

For each talker and for each acoustic measure, an average was computed across the sound tokens. In this table are the means, standard deviations, minima, and maxima of those data, i.e., inter-talker data, for each of three talker groups. Groups are deaf talkers from Total Communication settings (TC), deaf talkers from Oral Communication settings (Oral), and normal-hearing talkers (NH). The number of talkers included in each set of statistics is given below the group name (e.g., there are 89 TC talkers), except where indicated by bold numbers after the max value. In these cases, the number of talkers is less due to a few talkers not producing a plosive-like or fricative-like sound for any of the intended plosives or fricatives in our sentences.

and 22% of the time. From the error analyses in Smith's report we can compute the overall percent of /d/ and /t/ that were produced as "good" plosives (by keeping manner-associated errors as errors, but changing voicing and place errors to nonerrors), and can perform a similar calculation for the fricatives /s/ and /ʃ/. For the 40 talkers in Smith's study who used hearing aids, 57% of the /d/s and /t/s were produced as "good" plosives. For the 181 cochlear implant users in this study, 92% of the intended /d/s and /t/s were produced as "good" plosives. Analogously, 36% and 82% of the intended fricative sounds (/s/ and /ʃ/) were produced as "good" fricatives, respectively, by the 40 deaf talkers in Smith's study and the 181 cochlear implant users in this study. Hence, for the manner of production of these two types of sounds, the cochlear implant users were much more accurate than those in the Smith study. (Note: in Smith's article, data from both severely and profoundly hearing-impaired children were combined. One would expect the percentages computed from Smith's data to be even lower if productions from only the profoundly hearing-impaired children had been used.)

Our cochlear implant users' performances for these sounds are comparable with those reported by others for cochlear implant users. In a study by Tobey et al. (1991), 80% of the cochlear implant

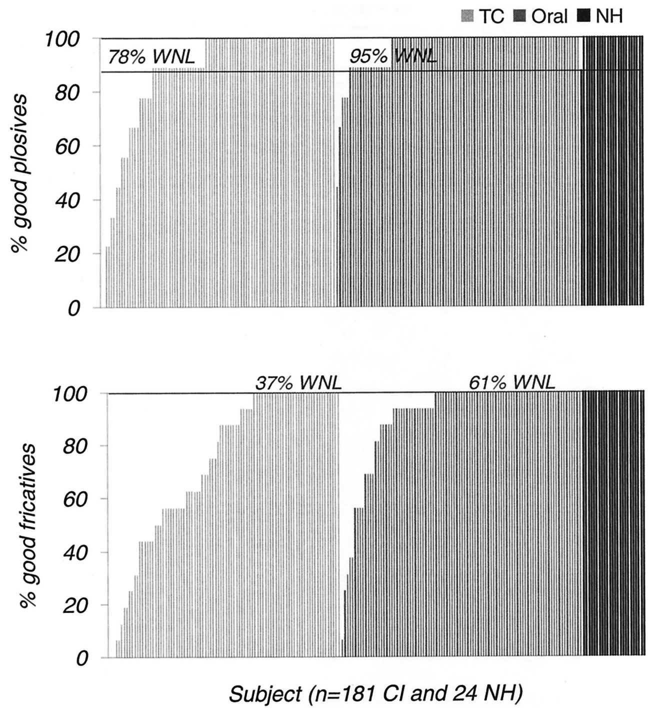

Figure 1. Percent good plosives (top) and fricatives (bottom) produced by each talker. Talkers are separated into three groups, from left to right, deaf children in Total communication settings (TC), deaf children in Oral communication settings (Oral), and children with normal hearing (NH). Within each group, data are ordered by their ordinal value. For the TC and Oral groups, the percent of those talkers who have values within normal limits (WNL) is indicated along the top.

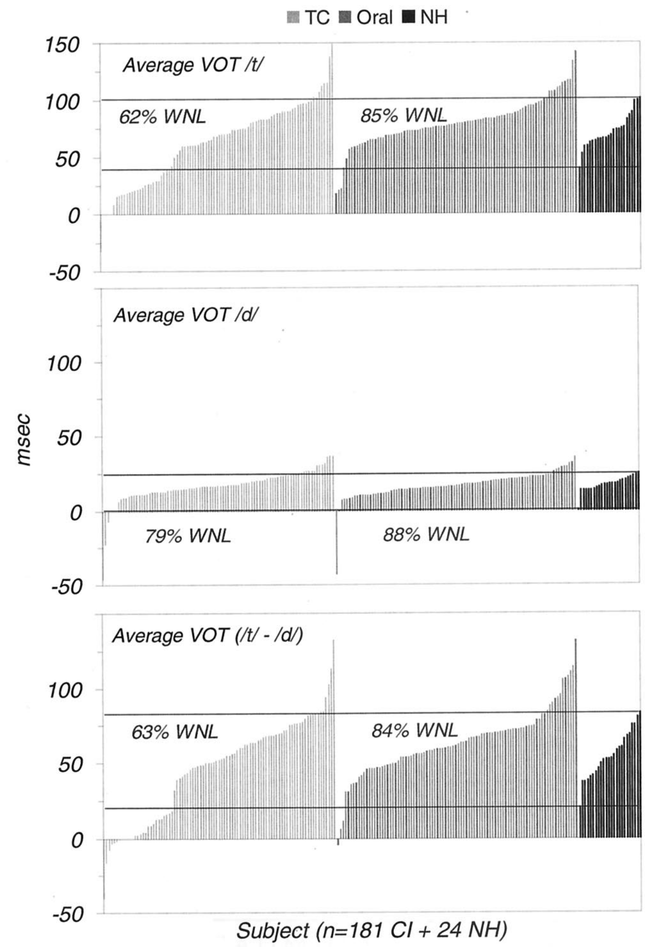

Figure 2. Average voice onset time (VOT) for intended /t/s (top), intended /d/s (middle), and their difference, VOT /t/ − VOT /d/ (bottom), produced by each talker. Talker groups and WNL are described in the Figure 1 caption. A handful of TC talkers did not produce a plosive-like sound for any of the intended /t/s and/or /d/s, and these talkers are represented by zero values in the appropriate panels.

users could produce plosive sounds 1 yr postimplant and 55% could produce fricative sounds. Blamey et al. (2001) reported about 75% of plosive sounds produced by cochlear implant users reached a target criterion, and roughly 50% of fricative sounds reached this same criterion. In our data, the ability to produce plosive and fricative sounds also appears to be associated with the emphasis on spoken language provided in the child's educational program. A significantly greater percent of children enrolled in Oral settings (95 and 61%) produced plosives and fricatives, when called for, compared with those in TC settings (78 and 37%).

Figure 2 shows VOT data. Note, there were a few TC talkers who did not produce a plosive-type sound for any of the intended /t/s or /d/s. These few talkers are represented by zeros in the graphs of Figure 2, and they are omitted from the inter-talker statistics provided in Table 3 (see Table 3 for the exact numbers of talkers with measurable VOTs). For our normal-hearing talkers, the measured VOTs for /t/ and /d/ agree fairly well with those reported previously for both adults and children. For adults producing stressed syllables in sentences, Lisker and Abramson (1965) reported a range of VOT /t/ values from 15 to 90 msec with a mean of 48 msec. The range of VOT /t/ values for our NH young talkers is 22 to 145 msec with a mean of 72 msec (note: this range of values considers the VOTs from each of the five /t/ tokens separately, whereas data in Figure 2 and Table 3 reflect the average VOT /t/ for each talker across the five /t/ tokens). Eguchi and Hirsh (1969) reported an average VOT for /t/ from 8- and 9-yr-old talkers of 69 msec. For /d/s produced by adults, Lisker and Abramson (1965) reported a range of VOT /d/ values from −80 to 25 msec with a mean close to 0 msec. Our NH young talkers exhibit a range of VOT /d/ values from −45 to 38 msec with a mean of 17 msec (similarly, this range considers the VOTs from each of the four /d/ tokens separately, whereas data in Figure 2 and Table 3 reflect the average VOT /d/ for each talker across the four /d/ tokens).

VOT data from deaf talkers, both children and adults are much more variable. In Figure 2, it is evident that the inter-subject ranges in average VOT /t/, VOT /d/, and ΔVOT for the deaf talkers are much greater than those seen for the normal-hearing talkers. Yet, over half the deaf talkers (both TC and Oral) produce VOTs and ΔVOTs WNL. McGarr and Lövquist (1982) measured average VOTs for /t/ and /d/ from one normal-hearing adult and three profoundly deaf adults. The deaf talkers' average VOT values differed from those of the normal-hearing talker (NH adult talker: mean VOT /t/ = 121 msec, mean VOT /d/ = 23 msec), and generally exhibited a small or insignificant time difference (ΔVOT) between the unvoiced and voiced plosives, namely 83 versus 47 msec, 20 versus 21 msec, and 69 versus 59 msec, respectively, for the three deaf talkers. Monsen (1976c) measured VOTs from six normal-hearing children and 30 profoundly hearing-impaired children age 11 to 16 yr old who used hearing aids. (In this study by Monsen, data from a total of 37 deaf children are reported. However, for our analysis, the seven severely hearing-impaired talkers were omitted and only the data from the remaining 30 profoundly hearing-impaired children were used.) Using the WNL range shown in Figure 2, 63% of the hearing aid talkers in Monsen's study produced VOTs for /t/ WNL. Analogous percents for VOTs /d/ and ΔVOT are 40% and 27%. Compared with these hearing aid talkers, cochlear implant users are producing much more normal VOTs (73%, 83%, and 73% of the cochlear implant users are WNL for VOT /t/, VOT /d/, and ΔVOT, respectively). The proportions of cochlear implant talkers producing VOTs for /d/ and ΔVOTs WNL are significantly greater than those proportions for hearing aid talkers, as found using a Pearson χ^2 test ($p < 0.01$). However, the proportion of talkers with VOTs for /t/ WNL is not statistically different for the cochlear implant and hearing aid users. Finally, a significantly greater percent of children in Oral settings (85 and 84%) produced VOTs for /t/ and ΔVOTs WNL than did children in TC settings (62 and 63%). TC and Oral children did not differ significantly in their ability to produce VOTs for /d/ WNL.

Figure 3 shows second formant frequency (F_2) data. Our normal-hearing children produced vowels with F_2s similar to those reported previously for children with normal hearing, especially of similar ages. Our average NH F_2 values for /i/ and /ɔ/ are 2978 Hz and 1422 Hz. For /i/ and /ɔ/, others report average second formant frequencies of 3105 Hz and 1355 Hz (8- and 9-yr-old talkers in Eguchi & Hirsh, 1969), 2776 Hz and 1067 Hz (11- to 14-yr-old talkers in Angelocci et al., 1964), 2414 Hz and 1146 Hz (12- to 16-yr-old talkers in Monsen, 1976b), and 3200 Hz and 1060 Hz (children of unknown ages in Peterson & Barney, 1952).

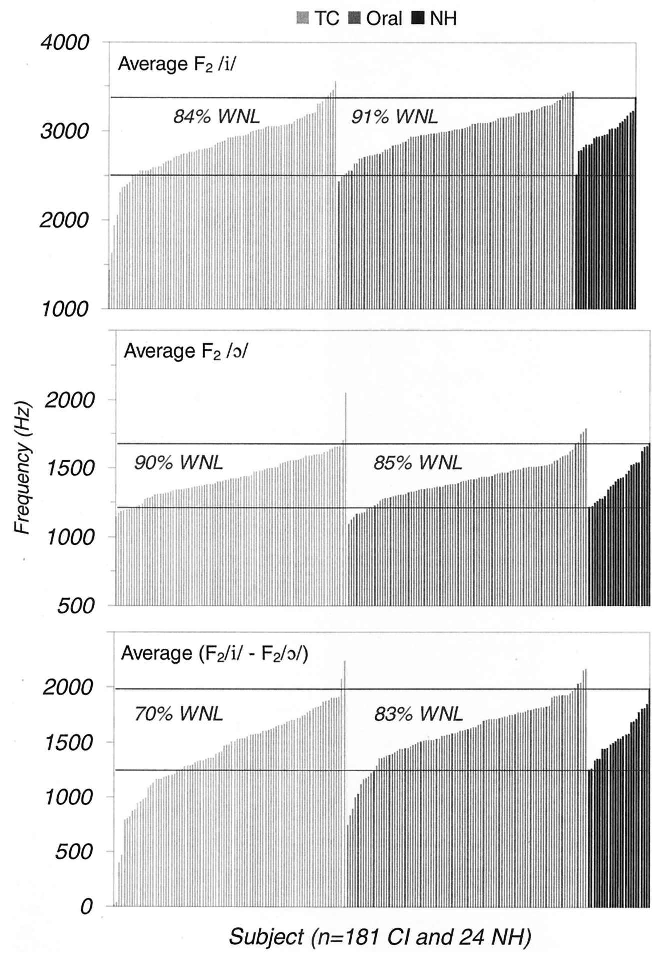

Figure 3. Average second formant frequency for intended /i/s (top), intended /ɔ/s (middle), and their difference, F_2 /i/ − F_2 /ɔ/ (bottom), produced by each talker. Talker groups and WNL are described in the Figure 1 caption.

Deaf talkers' formant frequencies often deviate from those of normal-hearing talkers (Angelocci et al., 1964; Monsen, 1976b). Here, nearly all the cochlear implant users produced these two vowels with F_2s WNL (88% and 87%, respectively, for /i/ and /ɔ/). A somewhat smaller percent of the cochlear implant users produced a difference in F_2 that was WNL (76%). In a study by Angelocci et al. (1964), average F_2 values from 18 deaf boys (ages 11 to 14 with both severe and profound hearing impairment) were 2325 Hz and 1177 Hz, respectively, for /i/ and /ɔ/ as compared with 2922 Hz and 1419 Hz for our cochlear implant users. In particular, F_2 for /i/ is much higher for the cochlear implant users than found for the hearing aid users. (The F_2 /i/ datum from Angelocci et al. is relatively low even when compared with similarly aged boys. For 11 to 13 yr old boys with normal hearing, using the mean ± 1 SD from Eguchi and Hirsh's data, one would expect the F_2 for /i/ to range from 2450 to 3010 Hz.) The speech of 29 profoundly hearing-impaired children (age 12 to 16 yr old) was measured by Monsen (1976b). (Monsen reports data from a total of 36 talkers, of whom seven did not have a profound hearing loss. Again, to make comparisons as similar as possible with respect to hearing loss, the seven talkers with severe hearing loss were omitted from our analysis.) Using the ranges from normal-hearing talkers in this study, 21 (72%), 7 (24%), and 5 (17%) of those hearing aid users had F_2 values WNL for /i/, /ɔ/, and /i/ − /ɔ/, respectively. For all three measures (F_2 /i/, F_2 /ɔ/, and ΔF_2), the proportions of cochlear implant users WNL are significantly greater than those found for the hearing aid users in Monsen's study. Thus, again, the cochlear implant users produced acoustic values much more like those of normal-hearing talkers than did hearing aid users. In this case, though hearing loss is similar for these comparisons, age is not. However, if one uses the range of normal values from the four similarly aged talkers in Monsen's study, the percentage of profoundly deaf talkers with values WNL are the same or even lower than those found when we use the range of normal values from our 24 (albeit younger) normal-hearing talkers.

Figures 4, 5, and 6, show data for the spectral moments mean, coefficient of skewness, and coefficient of kurtosis, respectively. (Note, as with the plosive sounds, there were a few talkers who did not produce a fricative-type sound for any of the intended /s/s or /ʃ/s. These talkers are represented by zeros in the appropriate graphs of Figures 4, 5, and 6, and they are omitted from the inter-talker statistics provided in Table 3.) Our NH data agree with data from others in that /s/ spectra have a greater mean than /ʃ/ spectra, /s/ spectra are more negatively skewed than /ʃ/ spectra, and /s/ spectra are more peaked (more positive in kurtosis) than /ʃ/ spectra (Forrest et al. 1988; Jongman et al., 2000; Matthies, Svirsky, Perkell, & Lane, 1996; Nittrouer, 1995). Also, in agreement with Nittrouer (1995), the overall skewness of our young talkers (8 and 9 yr olds) is much less positive than the skewness found for adult talkers' fricative spectra (Jongman et al., 2000). The NH talkers are consistent in these trends. That is, every one of our NH talkers exhibits a greater spectral mean for /s/, more negative spectral skew for /s/, and more peaked spectra (more positive kurtosis) for /s/ than for /ʃ/. Although most cochlear implant users also exhibit these trends (71%, 81%, and 62% of the cochlear implant users are WNL for the difference in spectral mean, skewness, and kurtosis, respectively), many cochlear implant users do not. For example, the spectral means (centroids) of the /s/ spectra, when not WNL, are lower in Bark (or frequency value) than those found for NH talkers, and often are not very different from the centroids for /ʃ/. In such cases, these two fricative sounds are probably not produced distinctively. Because these centroid values presumably reflect front cavity resonances and, hence, indicate the position of the vocal tract constriction, it is possible that such cochlear implant users do not place their tongue as far forward for /s/ as do NH talkers. Finally, there also seems to be an effect of educational setting on the ability of the cochlear implant users to produce fricatives with acoustic measures WNL. For all three spectral moments measures (mean, coefficient of skewness and coefficient of kurtosis), significantly more Oral talkers have values WNL than do TC talkers for both the /s/ and

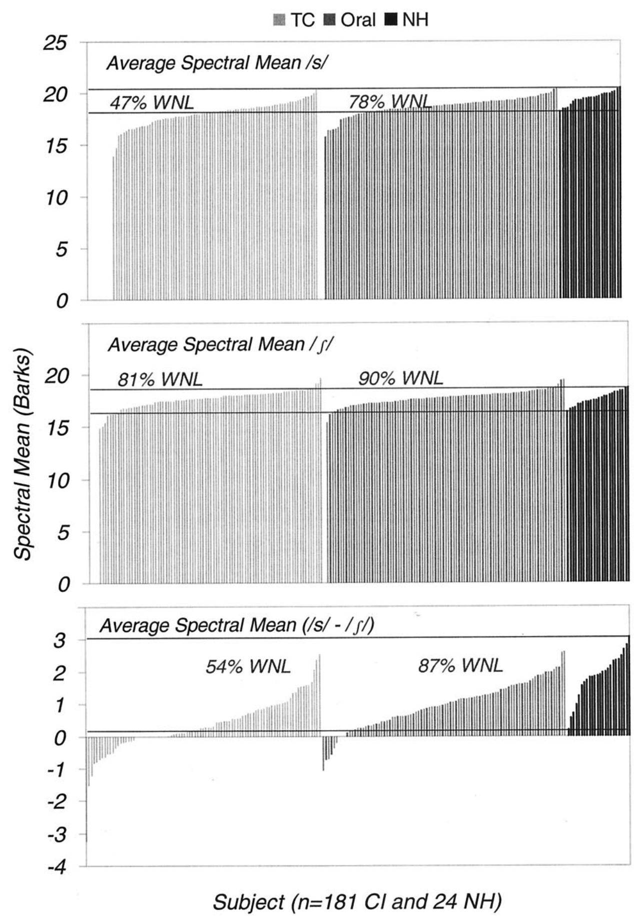

Figure 4. Average spectral mean for intended /s/s (top), intended /ʃ/s (middle), and their difference, /s/ − /ʃ/ (bottom), produced by each talker. Talker groups and WNL are described in the Figure 1 caption. A handful of talkers did not produce a fricative-like sound for any of the intended /s/s and/or /ʃ/s, and these talkers are represented by zero values in the appropriate panels.

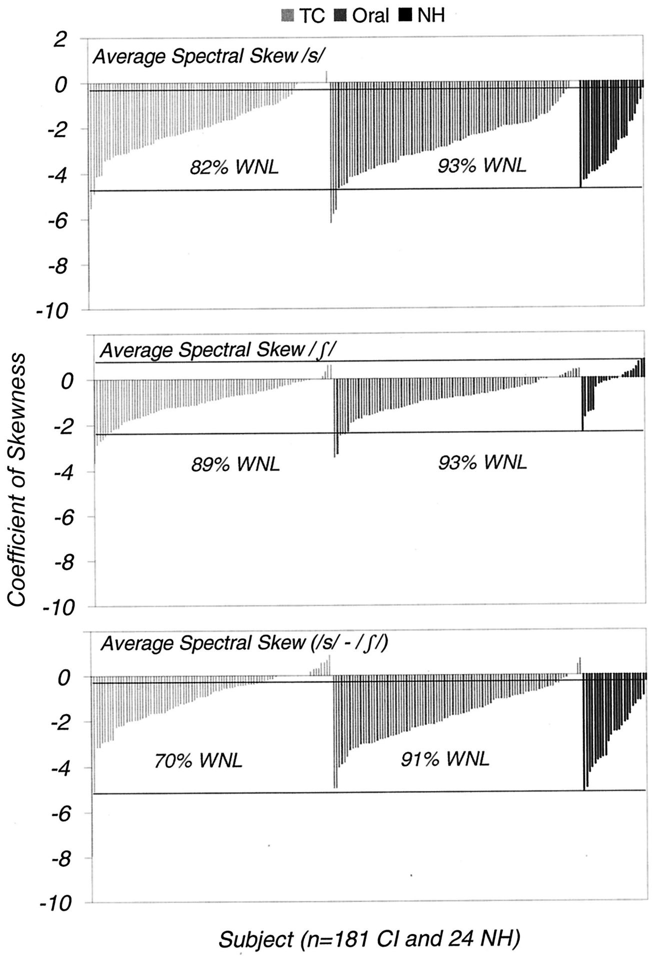

Figure 5. Average coefficient of spectral skewness for intended /s/s (top), intended /ʃ/s (middle), and their difference, /s/ − /ʃ/ (bottom), produced by each talker. Talker groups and WNL are described in the Figure 1 caption. A handful of talkers did not produce a fricative-like sound for any of the intended /s/s and/or /ʃ/s, and these talkers are represented by zero values in the appropriate panels.

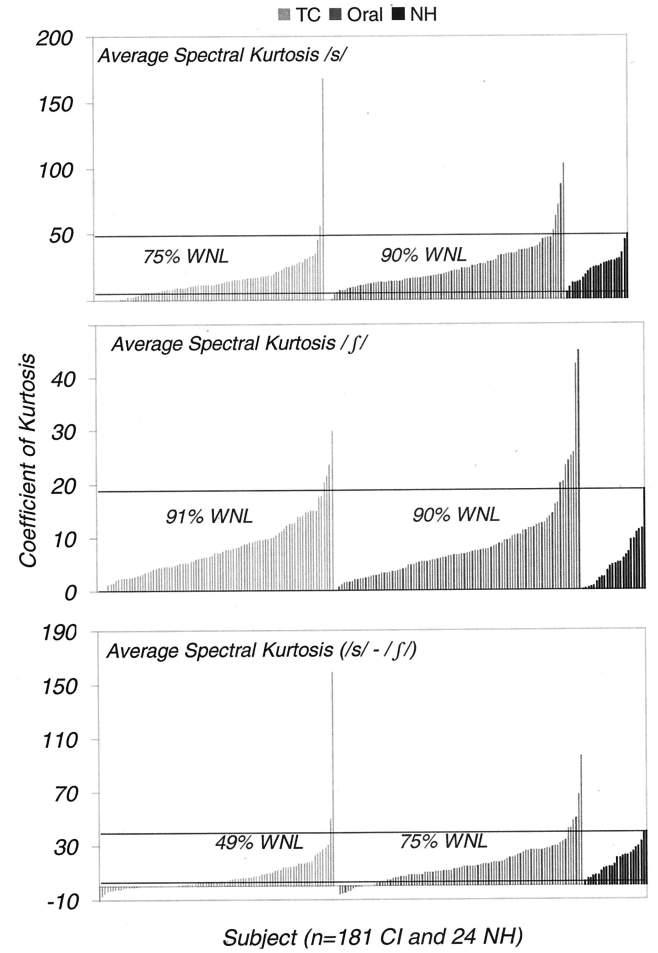

Figure 6. Average coefficient of spectral kurtosis for intended /s/s (top), intended /ʃ/s (middle), and their difference, /s/ − /ʃ/ (bottom), produced by each talker. Talker groups and WNL are described in the Figure 1 caption. A handful of talkers did not produce a fricative-like sound for any of the intended /s/s and/or /ʃ/s, and these talkers are represented by zero values in the appropriate panels.

difference measures, though not for the /ʃ/ measures.

Figure 7 shows data for the nasal manner metric and two of its three separate components. By design (Goldhor, 1995), values near 2 for each of the three component factors (D, L, and R), and values near 1 for the composite nasal manner metric should represent good nasal productions. The factor D is not plotted as 94% of the TC talkers and 99% of the Oral talkers produced their intended nasal sounds with a duration factor (D) WNL. The variability in the cochlear implant users' values for the nasal manner metric is primarily due to variability in the R factor, which represents the difference between the log power below 750 Hz in the consonant and the log power below 750 Hz in the subsequent vowel. If the log power difference is quite negative (<−5 dB), R has a low value, whereas if the log power difference is slightly negative or positive (>−5 dB), R has a value near 2. Thus, when a talker's nasal sound has low-frequency energy similar to that found in the subsequent vowel, the R factor contributes positively to the nasal manner metric. Many of the cochlear implant users did not have an R factor WNL (only 54% WNL). As such, these low values for R presumably reflect inadequate low-frequency energy in the intended nasal sound, possibly due to insufficient opening of the velopharyngeal port. It is difficult to compare our data with Goldhor's because data from both normal-hearing talkers and hearing-impaired talkers are combined in his report. Nevertheless, a significantly greater percent of children enrolled in Oral settings had R factor values, L factor values, and nasal manner metric values WNL compared with children in TC settings.

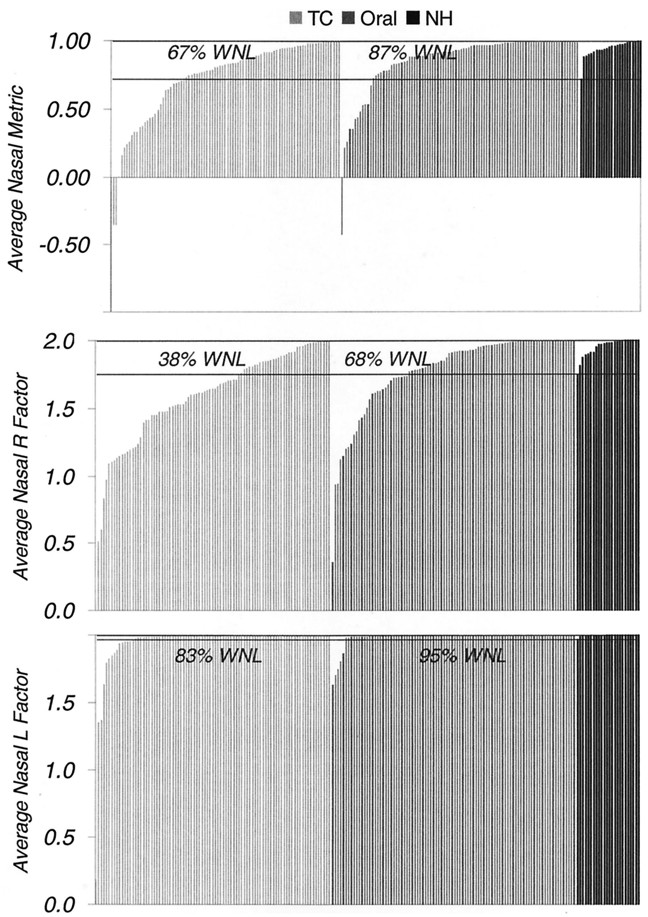

Figure 7. Average nasal manner metric (top), R factor (middle) and L factor (bottom) calculated from the intended nasals /m/ and /n/ produced by each talker. Talker groups and WNL are described in the Figure 1 caption.

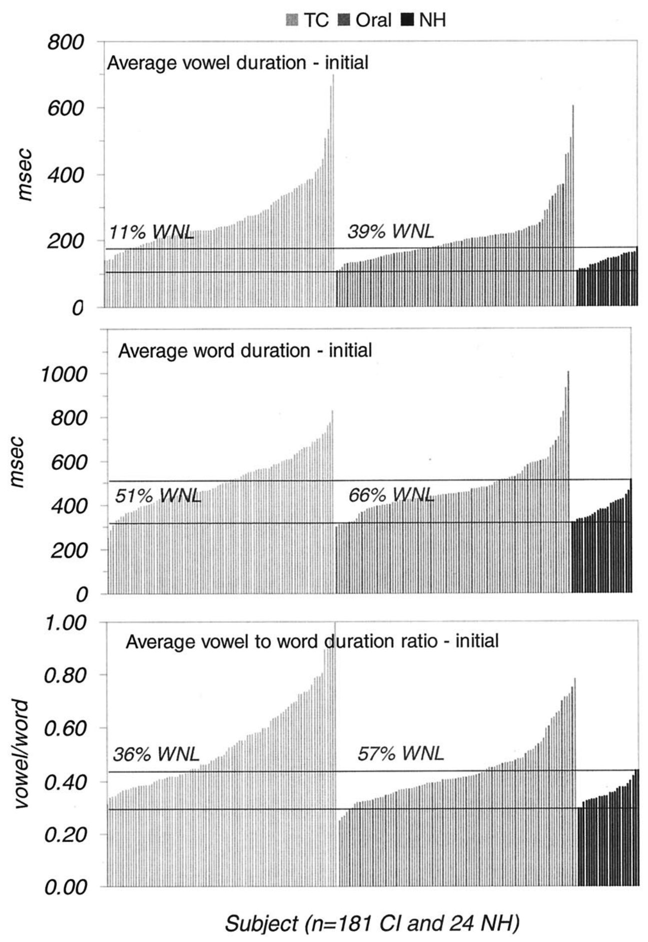

Figure 8. Average vowel duration (top), word duration (middle) and ratio of vowel to word duration (bottom) for sentence-initial words, produced by each talker. Talker groups and WNL are described in the Figure 1 caption.

Figures 8 and 9 show duration data from words in sentence-initial and sentence-final positions, respectively. In general, cochlear implant users produced vowels and words with similar or longer durations than did children with normal hearing. In addition, for the cochlear implant users, the vowel sounds in these CVC words account for a larger percent of the word duration than found for NH talkers (see bottom panels of Figs. 8 and 9). However, all talkers did not produce the same sequence of phonetic segments even though the intended words and sounds were the same across all talkers. Hence, some of the large duration ratio values found for the cochlear implant users reflect sound substitutions and omissions. That is, if a talker substituted a voiced plosive for a fricative, one would expect the vowel to constitute a larger proportion of the word duration (a talker says /dIp/ instead of /ʃIp/). Or, if a talker omitted a consonant sound, again, one would expect the vowel sound to constitute a larger proportion of the word duration. For the NH children, there is also an increase in the absolute and relative durations in sentence-final words compared with those in sentence-initial words (e.g., average NH ratio of 0.36 for sentence-initial words versus 0.42 for sentence-final words). By contrast, the cochlear implant users exhibit quite similar absolute and relative durations for both sentence-initial and sentence-final words (e.g., average cochlear implant ratios of 0.49 and 0.48, respectively, for initial and final positions). This may reflect the absence of a prepausal-lengthening vowel duration rule for the cochlear implant users (Klatt, 1976). Based on relative measures of phrase-level events, Robb and Pang-Ching (1992) found that although hearing-impaired talkers produced speech with absolute greater durations, their relative timing for phrase-level events was no different from that of talkers with normal hearing. Our data agree with Robb and Pang-Ching regarding absolute greater durations. However, because the relative measures in our study and in theirs are based on different length structures (words versus phrases), it is difficult to compare our duration ratio

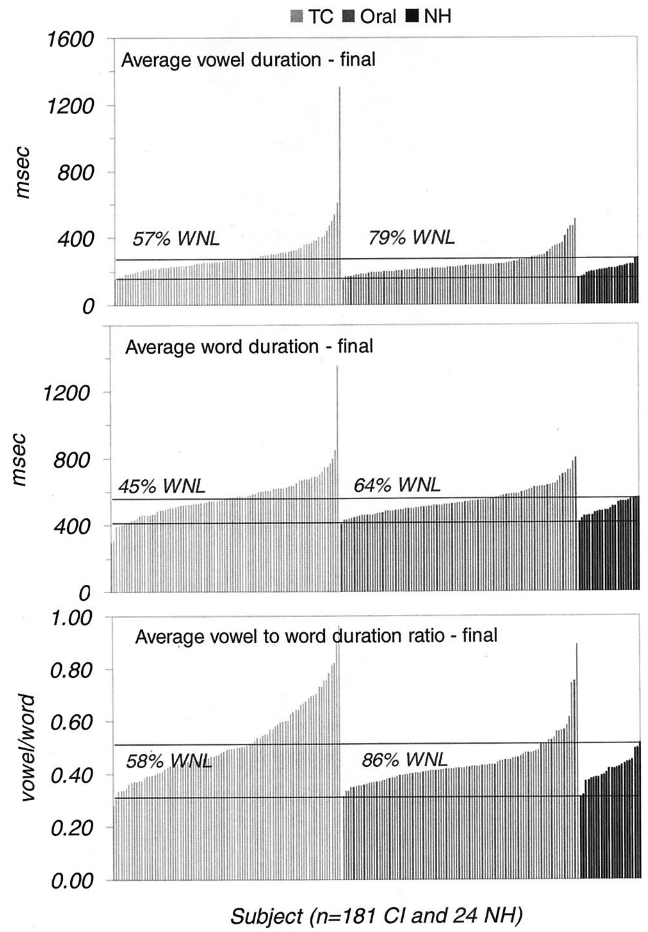

Figure 9. Average vowel duration (top), word duration (middle) and ratio of vowel to word duration (bottom) for sentence-final words, produced by each talker. Talker groups and WNL are described in the Figure 1 caption.

results with their data. In general, our cochlear implant users spend proportionately greater time, in the CVC word, on the vowel sound than do NH talkers. Overall, only 46% and 72% of the cochlear implant users exhibited vowel-to-word duration ratios WNL for sentence-initial and sentence-final words, respectively. Again, there is also a significantly greater percentage of Oral than TC talkers with values WNL. This result was found for all three duration measures in both sentence-initial and sentence-final words.

Figure 10 displays average sentence duration data. These duration data are similar to those of Figures 8 and 9 in that absolute duration is much longer for the cochlear implant users than for the NH talkers. Only 9% of the TC and 36% of the Oral talkers produced average sentence durations WNL (a significant difference between these two subject groups); overall, 23% of the cochlear implant users did. For a different set of sentences, Monsen (1978) reported a sentence duration range from 1.2 to 3.9 sec for 69 severe-to-profound hearing-impaired children (age 11 to 16 yr). He also reported that sentence duration was negatively correlated with intelligibility with coefficients of −0.47 and −0.33 for two subsets of these talkers. Although a slower duration per se does not necessarily imply poor speech, for these talkers a slower duration may reflect poorly developed speech skills in general.

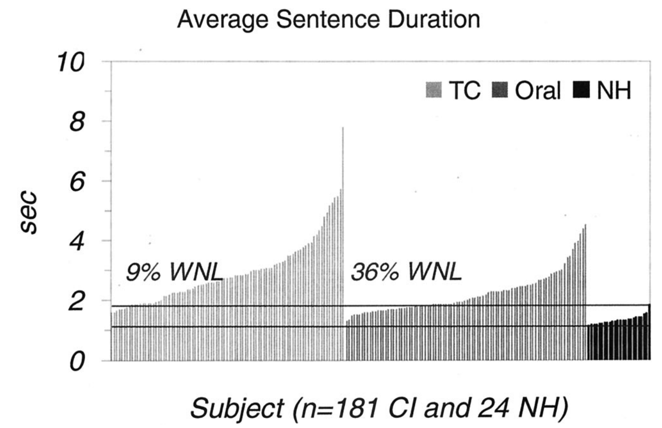

Figure 10. Average sentence duration produced by each talker. Talker groups and WNL are described in the Figure 1 caption.

Conclusions

Improvement in the production of manner features (plosive and fricative) after cochlear implantation has been shown by others (Brown & McDowall, 1999; Kirk, Diefendorf, Riley, & Osberger, 1995; Matthies et al., 1996; Sehgal, Kirk, Svirsky, Ertmer, & Osberger, 1998; Tobey et al., 1991, 1994). However, none of these studies compared the acoustic characteristics of the speech of children with cochlear implants with similarly aged normal-hearing talkers. Our definition of normal values (WNL), though reasonable, is possibly generous. Despite this qualification, our results are unlike those found in the past. For a large percentage of the profoundly deaf children in this study, many acoustic measures of their speech reach or closely approach normal values. This large proportion of speech sounds with normal acoustic values possibly reflects a combined effect of improved speech perception from cochlear implantation and auditory-oral instruction. For many of our acoustic measures, a significantly larger proportion of children from oral education settings, who depended on listening to and producing speech for communication on a daily basis, produced speech that resembled that of normal-hearing talkers. This result confirms the intelligibility findings reported by Tobey et al. (2003) and further defines the nature of articulatory control exhibited by orally educated children. They are better able to produce clear voicing contrasts, to differentiate fricative sounds and to produce nasal sounds properly, and they exhibit more normal vowel and sentence durations than children whose communication mode includes sign.

ACKNOWLEDGMENTS:

This study was supported by Grant No. DC03100 from the National Institute on Deafness and other Communication Disorders (NIDCD) of the National Institutes of Health to Central Institute for the Deaf (CID). The authors are grateful to all the children and their families who participated in this study. We also thank Chris Brenner, Gina Torretta, Debra Sawyer, Jennifer Micheletto and Melissa Scheuermann for their invaluable assistance in data collection, data management, and programming. We thank Mary Gamache from the Special School District of St. Louis County who elicited the speech samples from children with cochlear implants.

Address for correspondence: Rosalie M. Uchanski, Ph.D., Central Institute for the Deaf, 4560 Clayton Avenue, Saint Louis, Missouri 63110. E-mail: ruchanski@cid.wustl.edu.

Accepted September 12, 2002

REFERENCES

Angelocci, A. A., Kopp, G. A., & Holbrook, A. (1964). The vowel formants of deaf and normal-hearing eleven- to fourteen-year-old boys. *Journal of Speech and Hearing Disorders*, *29*, 156–170.

Avery, J. D., & Liss, J. M. (1996). Acoustic characteristics of less-masculine-sounding male speech. *The Journal of the Acoustical Society of America*, *99*, 3738–48.

Blamey, P. J., Barry, J. G., & Jacq, P. (2001). Phonetic inventory development in young cochlear implant users 6 years postoperation. *Journal of Speech, Language, and Hearing Research*, *44*, 73–79.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech and Hearing Research*, *27*, 134–144.

Brown, C., & McDowall, D. W. (1999). Speech production results in children implanted with the Clarion® implant. *Annals of Otology, Rhinology and Laryngology*, *108*, 110–112.

Eguchi, S., & Hirsh, I. J. (1969). Development of speech sounds in children. *Acta Oto-Laryngologica*, *257* (Suppl.), 1–51.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America*, *84*, 115–123.

Fourakis, M., Geers, A., & Tobey, E. (1993). An acoustic metric for assessing change in vowel production by profoundly hearing-impaired children. *The Journal of the Acoustical Society of America*, *94*, 2544–2552.

Geers, A., & Brenner, C. (1994). Speech perception results: Audition and lipreading enhancement. *The Volta Review*, *96*, 97–108.

Geers, A., & Brenner, C. (2003). Background and educational characteristics of prelingually deaf children implanted before five years of age. *Ear and Hearing*, 24 (Suppl.), 2S–14S.

Geers, A. E., & Moog, J. S. (1992). Speech perception and production skills of students with impaired hearing from oral and total communication education settings. *Journal of Speech and Hearing Research*, *35*, 1384–1393.

Geffner, D. (1980). Feature characteristics of spontaneous speech production in young deaf children. *Journal of Communication Disorders*, *13*, 443–454.

Goldhor, R. (1995). The perceptual and acoustic assessment of the speech of hearing-impaired talkers. In A. Syrdal, R. Bennett, & S. Greenspan (Eds.), *Applied Speech Technology* (pp. 521–545). Boca Raton, FL: CRC Press.

Hudgins, C., & Numbers, F. (1942). An investigation of the intelligibility of the speech of deaf and normal subjects. *Genetic Psychology Monographs*, *25*, 289–392.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, *108*, 1252–1263.

Kirk, K. I., Diefendorf, E., Riley, A., & Osberger, M. J. (1995). Consonant production by children with multichannel cochlear implants or hearing aids. In A. S. Uziel & M. Mondain (Eds.), *Cochlear Implants In Children*. *Advances In Oto-Rhino-Laryngology*, *50*, 154–159. Basel: Karger.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Lisker, L., & Abramson, A. S. (1965). Voice onset time in the production and perception of English stops. *Speech Research, Haskins Laboratories, SR-1*, 3.1–3.15.

Matthies, M. L., Svirsky, M., Perkell, J., & Lane, H. (1996). Acoustic and articulatory measures of sibilant production with and without auditory feedback from a cochlear implant. *Journal of Speech and Hearing Research*, *39*, 936–946.

McGarr, N. S. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research*, *26*, 451–458.

McGarr, N. S., & Lövquist, A. (1982). Obstruent production by hearing-impaired speakers: Interarticulator timing and acoustics. *The Journal of the Acoustical Society of America*, *72*, 34–42.

Metz, D. E., Samar, V. J., Schiavetti, N., Sitler, R. W., & Whitehead, R. L. (1985). Acoustic dimensions of hearing-impaired speakers' intelligibility. *Journal of Speech and Hearing Research*, *28*, 345–355.

Monsen, R. (1974). Durational aspects of vowel production in the speech of deaf children. *Journal of Speech and Hearing Research*, *17*, 386–398.

Monsen, R. (1976a). Second formant transitions of selected consonant-vowel combinations in the speech of deaf and normal-hearing children. *Journal of Speech and Hearing Research*, *19*, 279–289.

Monsen, R. (1976b). Normal and reduced phonological space: The production of English vowels by deaf adolescents. *Journal of Phonetics*, *4*, 189–198.

Monsen, R. (1976c). The production of English stop consonants in the speech of deaf children. *Journal of Phonetics*, *4*, 29–41.

Monsen, R. B. (1978). Toward measuring how well hearing-impaired children speak. *Journal of Speech and Hearing Research*, *21*, 197–219.

Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders*, *48*, 286–296.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *The Journal of the Acoustical Society of America*, *97*, 520–530.

Osberger, M. J., Maso, M., & Sam, L. (1993). Speech intelligibility of children with cochlear implants, tactile aids, or hearing aids. *Journal of Speech and Hearing Research*, *56*, 186–203.

Osberger, M. J., Robbins, A. M., Berry, S. W., Todd, S. L., Hesketh, L. J., & Sedey, A. (1991). Analysis of the spontaneous speech samples of children with cochlear implants or tactile aids. *The American Journal of Otology*, *12* (Suppl.), 151–164.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.

Porter, K. F., & Bradley, S. (1985). A comparison of three speech intelligibility measures for deaf students. *American Annals of the Deaf*, *130*, 514–525.

Pratt, S., & Tye-Murray, N. (1997). Speech impairment secondary to hearing loss. In M. McNeil (Ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 345–387). New York: Thieme Medical Publishers, Inc.

Robb, M. P., & Pang-Ching, G. K. (1992). Relative timing characteristics of hearing-impaired speakers. *Journal of the Acoustical Society of America*, *91*, 2954–2960.

Sehgal, S. T., Kirk, K. I., Svirsky, M., Ertmer, D. J., & Osberger, M. J. (1998). Imitative consonant feature production by children with multisensory aids. *Ear and Hearing*, *19*, 72–84.

Skinner, M. W., Fourakis, M. S., Holden, T. A., Holden, L. K., & Demorest, M. E. (1999). Identification of speech by cochlear implant recipients with the multipeak (MPEAK) and spectral peak (SPEAK) speech coding strategies II. Consonants. *Ear and Hearing*, *20*, 443–460.

Smith, C. R. (1975). Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research*, *18*, 795–811.

Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on auditory representation of American English Vowels. *Journal of the Acoustical Society of America*, *79*, 1086–1100.

Tobey, E., Geers, A., & Brenner, C. (1994). Speech production results: speech feature acquisition. *Volta Review*, *96*, 109–129.

Tobey, E. A., Geers, A. E., Brenner, C. A., Altuna, D., & Gabbert, G. (2003). Factors associated with development of speech production skills in children implanted by age five. *Ear and Hearing*, 24 (Suppl.), 36S–45S.

Tobey, E. A., Pancamo, S., Staller, S. J., Brimacombe, J. A., & Beiter, A. L., (1991). Consonant production in children receiving a multichannel cochlear implant. *Ear and Hearing*, *12*, 23–31.

Tye-Murray, N. (1992). Articulatory organizational strategies and the roles of audition. *The Volta Review*, *94*, 243–259.

Tye-Murray, N., Spencer, L., & Gilbert-Bedia, E. (1995). Relationships between speech production and speech perception skills in young cochlear-implant users. *The Journal of the Acoustical Society of America*, *98*, 2454–2460.