Prosody and voice characteristics of children with cochlear implants

Jessica M. Lenden*, Peter Flipsen Jr.

Department of Audiology and Speech Pathology, University of Tennessee,
578 South Stadium Hall, Knoxville, TN 37996, USA

Received 9 September 2005; received in revised form 11 April 2006; accepted 25 April 2006

Abstract

This descriptive, longitudinal study involved the analysis of the prosody and voice characteristics of conversational speech produced by six young children with severe to profound hearing impairments who had been fitted with cochlear implants. A total of 40 samples were analyzed using the Prosody–Voice Screening Profile (PVSP; Shriberg, L. D., Kwiatkowski, J., & Rasmussen, C. (1990). Prosody–Voice Screening Profile (PVSP). Tuscon, AZ: Communication Skill Builders). Overall, the children presented with noticeable problems with stress and resonance quality. There were some difficulties noted with rate, loudness, and laryngeal quality, but there were no consistent difficulties with phrasing or pitch. This suggested that prosody and voice characteristics in this population are different from those typically observed in children with severe to profound hearing impairments though some problem areas remain. Some developmental trends were also observed. These findings suggest that cochlear implants offer some significant benefits to children with hearing impairment in terms of prosody and voice outcomes.

Learning outcomes: The reader will be able to: (1) identify aspects of prosody and voice that have been noted as problematic for individuals with severe to profound hearing impairment, and (2) list potential differences between the prosody and voice characteristics of children with cochlear implants and classic descriptions of the speech of individuals with hearing impairment.

© 2006 Elsevier Inc. All rights reserved.
1. Introduction

1.1. Prosody and voice in children with cochlear implants

Since the early 1990s, researchers have documented improvements in speech perception (e.g., Carney, Kienle, & Miyamoto, 1990; Miyamoto, Kirk, Robbins, Todd, & Riley, 1996; Mondain et al., 1997; Tyler, 1990), as well as in speech production in children with prelingual deafness who were fitted with multichannel cochlear implants (Ertmer & Mellon, 2001; Miyamoto et al., 1996; Miyamoto, Kirk, Svirstky, & Sehgal, 1999; Tobey et al., 1991; Tobey, Geers, & Brenner, 1994). Of those studies investigating speech production, most have focused at the segmental level with several studies examining vowels (Ertmer, Kirk, Sehgal, Riley, & Osberger, 1997; Ertmer & Mellon, 2001), while others have focused on consonants (Kirk, Diefendorf, Riley, & Osberger, 1995; Tobey et al., 1991). In these studies, there has been a tendency to use very controlled, previously prepared sentences or word lists. The use of such contexts is understandable given a desire to make careful comparisons across individuals and groups, but it severely limits the ability to examine higher level aspects of production such as the suprasegmentals. It also begs the question of how representative performance in such contexts is of typical speech behavior. A few investigators have made use of conversational speech samples (e.g., Blamey, Barry, & Jacq, 2001; Blamey, Sarant et al., 2001; Bow, Blamey, Paatsch, & Sarant, 2002), though in most cases, those studies have focused on overall intelligibility of the message (Chin, Tsai, & Gao, 2003; Colvard, 2002; Flipsen & Colvard, 2003, 2006; Robbins, Kirk, Osberger, & Ertmer, 1995; Svirstky, Chin, Miyamoto, Sloan, & Caldwell, 2002; Svirstky, Sloan, Caldwell, & Miyamoto, 2000; Tobey et al., 1994). To date, it appears that no study has formally analyzed the suprasegmental aspects of speech in cochlear implant recipients.

1.2. Prosody and voice characteristics in children with hearing impairment

Shriberg, Kwiatkowski, and Rasmussen (1990) identified seven broad dimensions of speech “... that have been observed to contribute to a speaker’s severity of involvement” (p. 3). These include dimensions of prosody (phrasing, rate, stress) and voice (loudness, pitch, laryngeal quality, resonance quality). All of these dimensions have been shown to be at least somewhat problematic for individuals with hearing impairment who have not had the benefit of cochlear implants (Allen & Arndorfer, 2000; Boone, 1966, 1971; Calvert, 1962; Calvert & Silverman, 1983; Finkelstein, Bar-Ziv, Nachmani, Berger, & Ophir, 1993; Fletcher, Mahfuuz, & Hendarmin, 1999; Hargrove, 1997; Kummer & Lee, 1996; Maassen & Povel, 1984; McGarr & Osberger, 1978; Monsen, 1978; Nickerson, 1975; Parkhurst & Levitt, 1978; Smith, 1975; Stathopoulos, Duchan, Sonnenmeier, & Bruce, 1986; Subtelny, Whitehead, & Orlando, 1980; Thomas-Kersting & Casteel, 1989).

Shriberg et al. (1990) define phrasing as “... the flow of speech” (p. 22). Boone (1966) noted that the individuals with hearing impairment in his study tended to produce each word separately as opposed to the continuous, overlapping flow found in normal hearing speakers. This is consistent with studies that have shown lesser amounts of inter-syllabic co-articulation in the speech of the hearing impaired (e.g., Okalidou & Harris, 1999).
Subtelny et al. (1980) reached similar conclusions, reporting poorly blended sounds and words in this population.

Speech rate (i.e., slower speech) has also been shown to be an issue for individuals with hearing impairment and has been characterized by overall slower rate signaled by longer word durations (Hood & Dixon, 1969; Parkhurst & Levitt, 1978). In addition, Boone (1966) reported prolonged vowels and longer interword pauses in this population. Stathopoulos et al. (1986) found that speakers with hearing impairments use longer pauses between sentences and less pausing within sentences.

Another aspect of prosody considered by Shriberg et al. (1990) is stress. In this case, the term is used to denote the use of lexical, phrasal, and emphatic stress as manifested using changes in pitch, loudness, or duration (or some combination thereof). Inappropriate stress patterns have been described as “typical” of individuals with hearing impairment (Hargrove, 1997; Subtelny et al., 1980). Nickerson (1975) reported that individuals with hearing impairment tend to vary pitch less often resulting in excessive stress on all syllables or a flat monotone stress pattern throughout the utterance. Nickerson suggested that it is “almost as if individuals with hearing impairment only produce stressed syllables” (p. 344). Other stress patterns observed by Nickerson included monotonicity or excessive, erratic pitch variation.

Relative to voice parameters, loudness has also been shown to be an issue for those with hearing impairments (Smith, 1975). According to Calvert and Silverman (1983), deaf individuals “…may simply not be able to speak loud enough in any situation” (p. 170).

Another aspect of voice considered by Shriberg et al. (1990) is pitch. The vocal pitch of individuals with hearing impairments has been characterized as including monotonality, excessive changes and diplophonia (Monsen, 1978; Parkhurst & Levitt, 1978; Smith, 1975; Subtelny et al., 1980). Studies using acoustical analysis have shown abnormal pitch as well as restricted pitch range in the speech of the hearing-impaired (Boone, 1966; Stathopoulos et al., 1986). Degree of hearing loss may play a role relative to pitch. McGarr and Osberger (1978) found that no child in their study with a hearing loss greater than 90 dB HL received an appropriate pitch rating.

Relative to laryngeal quality, Shriberg et al. (1990) use this term as one component of overall vocal quality (along with loudness, pitch, and resonance quality). Thus the terms laryngeal quality and voice quality are not necessarily synonymous in the context of the PVSP. Laryngeal quality of individuals with hearing impairment has been described as “tense”, “flat”, “breathy”, “harsh”, and “throaty” (Calvert, 1962). Greater reliance on tactile impressions (as opposed to auditory input) may result in more constriction and tension in hearing impaired speech, which may lead to stridency or harshness. Excessive force on plosives before a vowel, which is common in hearing impaired speech, also may result in a breathy quality (Calvert & Silverman, 1983, p. 169). Subtelny et al. (1980) suggested that a breathy/weak voice could occur due to poor adduction, and extra strain from over adduction could result in an overall tense/strained vocal quality.

Finally, resonance quality differences are a classic part of descriptions of hearing impaired speech. Productions are often characterized by a pharyngeal focused resonance; terms such as “hollow” (Boone, 1971), cul-de-sac resonance or “hot potato voice” (Finkelstein et al., 1993) have been applied. Both hyper- and hyponasality have been
observed in hearing impaired speech. Fletcher et al. (1999) showed that hearing impaired children have significantly more nasalance than normal hearing speakers when nasal consonants are not present and significantly less when an utterance has many nasal consonants.

Relative to the current study, it is important to note that although individuals with hearing impairment do tend to have a very unique vocal quality, this quality may only be audible in continuous speech. Calvert (1962) studied the ability of deaf educators to distinguish the speech of hearing impaired children from that of normal hearing children. He found that they could distinguish between the two groups when they produced sentences but not when they produced isolated vowels.

Although generally discussed in isolation, interactions have also been observed among the suprasegmental aspects of the speech of the hearing impaired. For example, irregularities of rhythm in the speech of individuals with hearing impairment “appear to have their origins in the difficulty of a talker to control the mechanics of varying loudness, pitch and duration” (Calvert & Silverman, 1983, p. 171). Maasen and Povel (1984) note that inappropriate pitch affects the speaker’s ability to correctly produce lexical stress. Stathopoulos et al. (1986) hypothesized that individuals with hearing impairment may rely on the use of long pauses to distinguish sentence boundaries, instead of using falling intonation contours like normal speakers.

In sum, there is much we know about the prosody and voice characteristics of the speech produced by individuals with hearing impairment, but the bulk of the research has focused on individuals who use hearing aids or tactile devices. Relative to users of cochlear implants, the research to date is much more limited. One study of speech perception in children with cochlear implants noted that children implanted with single channel cochlear implants perceived intonation patterns more readily than the number of syllables or syllable stress in a word (Carney et al., 1990). Relative to speech production in children with multichannel cochlear implants, Tobey et al. (1991) noted that after 1 year of cochlear implant experience, speakers were better able to imitate prosodic features of speech as tested by the Phonetic Level Speech Evaluation (Ling, 1976). Imitation of “prosodic characteristics underlying the melody of speech” (p. 170) was noted to significantly improve in 31.1% of the deaf children after implantation. Improvements in the production of non-segmental aspects of speech were greater in the younger children (2;0–9;01) than in the older group (10;0–17;0) with score increases of 33.4 and 17%, respectively. However, neither of the above studies focused directly on prosody and voice characteristics, but rather only briefly mentions findings that are related to speech perception and production. Carter, Dillon, and Pisoni (2002) did examine suprasegmentals aspects of speech in a group of twenty-four 8–10-year-old children fitted with cochlear implants for at least 3.8 years. The analysis focused on imitation of one to five syllable non-word forms. Findings indicated that 64% of the non-words were repeated with the correct number of syllables and 61% were reproduced with the correct placement of primary stress (48% had both the correct number of syllables and correct stress placement). Carter et al. also noted that errors on both parameters increased as the syllable length of the target increased.

1 Expressed in years;months.
1.3. Developmental considerations

As noted by Flipsen and Colvard (2006), the evaluation of intervention outcomes for children with hearing impairments has historically involved consideration of hearing age (HA) or the length of time that the individual has been using amplification. With the advent of Universal Newborn Hearing Screening (UNHS), identification often occurs very early resulting in very little difference between HA and chronological age (CA). With cochlear implants however, the calculation of HA is not so straightforward. Currently children as young as 12 months receive implants, but virtually all children who receive cochlear implants are fitted with hearing aids for some length of time prior to implantation. This pattern arises because a common criterion for implantation is that the potential recipient should have received only “minimal benefit” from a trial with hearing aids. Most studies of outcomes in cochlear implantation have ignored HA and focused on amount of implant use (what might be termed ‘post-implantation age’ or PIA). Flipsen and Colvard suggested that all three age-related variables (CA, HA, PIA) be considered when evaluating outcomes for this population.

1.4. Focus of the current study

The current study examined the prosody and voice characteristics of 40 conversational speech samples obtained from 6 young children with prelingual severe to profound deafness who had been fitted with multichannel cochlear implant devices. The samples were obtained at 3-month intervals over 12–21-month periods and analyzed using the Prosody–Voice Screening Profile (PVSP; Shriberg et al., 1990). Two questions were addressed. First, is there evidence in the conversational speech of children fitted with cochlear implants of difficulty with any of the suprasegmentals measured on the PVSP? Second, are there any developmental trends on any of the suprasegmentals measured on the PVSP relative to CA, HA, or PIA?

2. Method

2.1. Participants

Children were recruited for the current study from Child Hearing Services, a clinical services program at the University of Tennessee, Knoxville. The children were selected based on the following criteria: (a) prelingually deaf (defined as onset of hearing loss before age 3 years), (b) fitted with a multichannel cochlear implant by age 3 years, (c) use of the cochlear implant for at least 18 months at the onset of testing, (d) use of spoken language only as their primary communication mode, and (e) receptive vocabulary performance as measured by the Peabody Picture Vocabulary Test—Third Edition (PPVT-III; Dunn & Dunn, 1997) within two standard deviations of their age group mean (i.e., a standard score of at least 70). The first three criteria increased the likelihood that the children would be using real speech, rather than babbling. The latter criterion was used to ensure that the children would have sufficient language comprehension skill to engage in meaningful conversation.

Six children (five girls and one boy) satisfied the selection criteria and participated in the study. As indicated in Table 1, mean age of identification was 8 months (range birth to 15
Table 1
Participant characteristics

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Age of ID(^a)</th>
<th>Implantation age(^b)</th>
<th>Initial CA(^{a,b})</th>
<th>Initial HA(^{a,c})</th>
<th>Initial PIA(^{a,d})</th>
<th>Implant type</th>
<th>Processing strategy</th>
<th>Number of samples</th>
<th>PPVT-III(^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>0;8</td>
<td>2;4</td>
<td>5;2</td>
<td>4;7</td>
<td>2;11</td>
<td>Clarion</td>
<td>MPS</td>
<td>8</td>
<td>89</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>0;0</td>
<td>2;6</td>
<td>4;5</td>
<td>4;5</td>
<td>1;11</td>
<td>Nucleus-24</td>
<td>ACE</td>
<td>5</td>
<td>99</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>1;0</td>
<td>3;0</td>
<td>6;2</td>
<td>5;3</td>
<td>3;2</td>
<td>Clarion</td>
<td>MPS</td>
<td>8</td>
<td>72</td>
</tr>
<tr>
<td>4</td>
<td>F</td>
<td>0;3</td>
<td>2;0</td>
<td>5;5</td>
<td>5;2</td>
<td>3;6</td>
<td>Nucleus-22</td>
<td>SPEAK</td>
<td>7</td>
<td>77</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>1;3</td>
<td>2;7</td>
<td>4;10</td>
<td>3;7</td>
<td>2;3</td>
<td>Clarion</td>
<td>CIS</td>
<td>6</td>
<td>81</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>0;11</td>
<td>1;8</td>
<td>3;9</td>
<td>2;10</td>
<td>2;1</td>
<td>Nucleus-24</td>
<td>SPEAK/ACE(^e)</td>
<td>6</td>
<td>76</td>
</tr>
</tbody>
</table>

Mean (S.D.)  
0;8 (0;6) 2;4 (0;6) 5;0 (0;10) 4;4 (0;11) 2;8 (0;8) 82.3 (10.0)

\(^a\) Expressed in years;months.
\(^b\) Chronological age at start of current study.
\(^c\) Hearing age (length of total amplification; hearing aid use + cochlear implant use) at start of current study.
\(^d\) Post-implantation age (amount of implant use) at start of current study.
\(^e\) Participant changed from SPEAK to ACE strategy between second and third samples.
\(^f\) Standard score on the Peabody Picture Vocabulary Test—Third Edition (Dunn & Dunn, 1997).
months). None of these children was identified under Universal Newborn Hearing Screening as such a program does not exist in the state of Tennessee. In addition, none of these children was identified via any “at risk” criteria. Mean age of implantation was 2 years, 4 months (range = 1;8–3;0). Mean chronological age at the start of the study was 5 years (range = 3;9–6;2). Mean length of amplification (hearing aid use + implant use) at the start of the study was 4 years, 4 months (range = 2;10–5;3). Mean length of implant experience at the beginning of the study was 2 years, 8 months (range = 1;11–3;6). Mean score of the PPVT-III was 82.3 (range = 72–99). The cause of the hearing loss was unknown for Participants 1–5; Participant 6 (the only boy in the group) had a diagnosis of partial agenesis of the cochlea. None of the children had any other known physical, cognitive, or emotional disability. Parents of the children were all reported to have normal hearing. All six of the children had received individualized intervention prior to the start of the study and continued to do so during the course of the study. In all cases the intervention was a completely oral-based approach (i.e., no signing was involved). None of the children had any special educational placements (i.e., all were in regular education classrooms).

Also shown in Table 1 are implant type and processing strategies used by the participants. This is provided for the information of the reader only. Although it might be tempting to examine outcomes relative to those two variables, the small number of participants in the current study precluded such analyses.

2.2. Conversational speech samples

Conversational speech samples were obtained from each of the children every 3 months for periods ranging from 12 to 21 months. Between 5 and 8 samples were obtained from each child (see Table 1) resulting in a total of 40 samples available for analysis. The samples were evoked as part of a larger protocol lasting 60–90 min per session. The first session for each child included administration of the PPVT-III and thus was slightly longer than the other sessions. Samples were evoked inside a single wall sound-treated booth by one of two trained graduate student clinicians. A parent or the clinician who was providing treatment to the child was also often present and participated during many of the samples.

The samples were recorded on digital audiotape using a Sony PCM-M1 portable digital tape recorder through a Sony TCM-150 microphone held by a puppet. A sampling rate of 48 kHz was used for the recordings. A variety of topics (e.g., favorite movies or cartoons, current activities in therapy) and materials such as age appropriate toys and activity pictures from the Bracken Concept Development Program (Bracken, 1998) were used to evoke the conversational samples. A sample size target of at least 90 different words was selected; samples of this size have been shown to provide a representative sample of English phonemes and canonical forms (Shriberg, 1986). Story telling and other narratives were avoided because of concern that they might evoke the use of atypical prosody in narrative registers (Shriberg et al., 1990). The second author monitored each test session from outside the sound booth through headphones and kept a running tally of the different intelligible words produced by the children. In one case the sample only contained 67 different words, but sampling was terminated after approximately 25 min in order to avoid fatigue effects. All of the remaining samples were allowed to run to at least 15 min of conversation. The resulting samples varied in length from 67 to 199 different
intelligible words (mean = 139.0, S.D. = 26.1) and included between 65 and 216 utterances (mean = 134.3, S.D. = 35.5).

2.3. Prosody–voice coding

The first author conducted all of the prosody and voice coding using procedures from the Prosody–Voice Screening Profile (Shriberg et al., 1990). Before listening to the samples, the training modules developed for the PVSP were completed to ensure proficiency. Training was completed within 2 months. In addition, and in order to obtain practice with coding complete sample sets, the first author carried out PVSP coding on several unrelated conversational speech samples that had previously been coded by a transcriber with extensive experience using the PVSP.

Sample coding was carried out over a 7-week period using transcripts of the samples which had previously been created during the course of narrow phonetic transcription. The phonetic transcripts included the regular spelling gloss. In order to avoid any systematic bias from listening to multiple samples from the same speakers, the samples were coded in random order.

Prior to coding, utterances were evaluated for possible exclusion from the analysis using a set of 32 exclusion codes intended to identify any “...situation or type of utterance that might prohibit, contraindicate, or bias prosody–voice coding” (Shriberg et al., 1990, p. 11). For example, utterances consisting entirely of “I don’t know” responses were excluded, as were utterances involving simultaneous talk by a child and the examiner, or utterances in which the child had assumed a character register. Across the 40 samples, a total of 3205 utterances were excluded; the most common reasons for exclusion were single word utterances (1431 utterances), too many unintelligible words (508), use of narrative register (420), and “I don’t know” responses (140).

Any utterance that was not subject to exclusion was then coded on each of the seven prosody–voice variables as either appropriate or inappropriate. Inappropriate utterances are assigned one or more of 32 prosody–voice codes. A target range of 20–25 coded utterances was set per Shriberg et al. (1990). For the current study, sample coding was terminated once either 25 utterances were coded or the end of the sample was reached. A total of 970 utterances were coded. Of the 40 samples, 35 (88%) included the maximum of 25 coded utterances, and 38 (95%) included at least 20 coded utterances. The remaining two samples (both from Participant 3) included 12 and 16 coded utterances, respectively. Shriberg et al. also recommended that at least 50% of the coded utterances include at least four words. Of the 40 samples, 30 (75%) met this criterion, and 38 (95%) included at least 20% utterances with at least 4 words. The two very short samples mentioned above included 33% (4/12) and 50% (8/16) longer utterances, respectively.

2.4. Data analysis

In order to evaluate the overall pattern of performance, each sample was evaluated on each PVSP variable based on the screening criteria suggested by Shriberg et al. (1990). If 90% of the utterances in a sample were rated as appropriate on a particular variable, the sample was judged to have passed the screening for that variable. Samples with 80–90%
appropriate utterances were judged as borderline, and samples with fewer than 80% appropriate utterances were judged as failing.

In order to examine developmental trends, performance on each of the PVSP variables (i.e., percentage of appropriate utterances) was examined two different ways. First, overall performance by the six children pooled across all of the samples was correlated with HA, CA, and PIA. Given that this study appears to be the first of its kind in this area, an a priori statistical significance value of .05 was set (i.e., Bonferroni corrections were not utilized). The possibility of missing important effects (Type II errors) was considered to be as important as the possibility of mis-identifying spurious effects (Type I errors). The second analysis of developmental trends involved examination of patterns on each variable for each child individually over time.

2.5. Reliability

To evaluate reliability of PVSP coding, a random set of five samples was re-coded by the same coder approximately 2 months after initial coding. The re-coding involved examination of the 125 previously coded utterances in the 5 samples (12.8% of all the coded utterances). Point-to-point exact agreements for coding were 92% (overall), 100% (phrasing), 92% (rate), 83% (stress), 95% (loudness), 94% (pitch), 92% (laryngeal quality), and 85% (resonance quality). These values are consistent with reliability values reported previously for this instrument (e.g., Shriberg, Aram, & Kwiatkowski, 1997).

3. Results

3.1. Overall performance

Findings pooled across the samples for all six participants are summarized in Table 2. As indicated, there were few problems on any of the samples with phrasing or pitch. Some samples yielded non-passing results for rate, loudness, and laryngeal quality. By far, the most problematic aspects of prosody and voice for these children appeared to be the use of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Appropriate utterances (%)</th>
<th>Range (%)</th>
<th>Pass</th>
<th>Borderline</th>
<th>Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrasing</td>
<td>97.1 (4.3)</td>
<td>86–100</td>
<td>36</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Rate</td>
<td>88.4 (10.0)</td>
<td>60–100</td>
<td>22</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Stress</td>
<td>48.2 (26.9)</td>
<td>0–96</td>
<td>2</td>
<td>5</td>
<td>33</td>
</tr>
<tr>
<td>Loudness</td>
<td>92.4 (10.8)</td>
<td>56–100</td>
<td>32</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Pitch</td>
<td>97.6 (4.3)</td>
<td>80–100</td>
<td>38</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Laryngeal quality</td>
<td>87.3 (15.0)</td>
<td>36–100</td>
<td>24</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Resonance quality</td>
<td>10.4 (20.8)</td>
<td>0–84</td>
<td>0</td>
<td>1</td>
<td>39</td>
</tr>
</tbody>
</table>

* Cell entries are means (and standard deviations).
* Number of samples in which at least 90% of coded utterances were judged to be appropriate.
* Number of samples in which 80–90% of coded utterances were judged to be appropriate.
* Number of samples in which less than 80% of coded utterances were judged to be appropriate.
stress (lexical, phrasal, and/or emphatic) and the resonance quality of their voices. For both of these variables, the bulk of the samples resulted in PVSP screening failures. It is noteworthy that, for each of these two variables, a single code accounted for the vast majority of the ratings of inappropriate. For stress, 492/509 (96.7%) of the inappropriate utterances were coded as having ‘excessive/equal/misplaced stress’. For resonance quality, 743/827 (89.8%) of the inappropriate utterances were coded as being ‘nasopharyngeal’ which includes ‘...a percept of sluggish or imprecise tongue movement and using a more “backed” (pharyngeal) rather than forward oral resonance’ (Shriberg et al., 1990, p. 42). This code includes (but is not limited to) the classic cul-de-sac resonance often associated with the speech of the hearing impaired.

3.2. Developmental trends

Findings the correlational analyses are shown in Table 3. Looking at the three left-most columns (which include data from all six participants), the only significant correlations obtained were for laryngeal quality which was significantly correlated \( (p < .05) \) with all three age variables. The magnitudes of the correlations accounted for 14–28% of the variance in the relationships. Overall, laryngeal quality improved over time for these children.

Previous analysis with this same data set (Colvard, 2002; Flipsen & Colvard, 2003, 2006) had suggested that Participant 2 might be an outlier in this particular group. She was the only one identified at birth, she had the highest PPVT-III score, and the intelligibility of her speech was superior to the rest of the group (despite having the least amount of implant experience). To examine the possible influence of this potential outlier, the correlations between the PVSP variables and the 3 age variables were repeated using the 35 samples from the other 5 participants only. Findings for this second set of correlations are shown in the three right-most columns of Table 3. Laryngeal quality was once again significantly correlated \( (p < .05) \) with all three variables; the magnitudes of the correlations increased in all three cases and accounted for 31–33% of the variance in the relationships. In addition,
stress was found to be significantly correlated with both HA and PIA, and resonance quality was found to be significantly correlated with PIA. The magnitudes of these latter correlations accounted for 12–13% of the variance in the relationships. These findings support the probability that Participant 2 was an outlier in this group. These findings also suggest that stress, laryngeal quality, and resonance quality of conversational speech all improved significantly over time.

Performance of individual children on each of the variables over time was then examined. None of the children appeared to have any difficulty with phrasing or pitch during the study period. This is, of course, consistent with the overall findings mentioned previously.

Relative to rate, data for Participant 4 suggested improving performance over time with 76% appropriate utterances on sample 1 which increased to 96% by sample 4 and remained stable at 92% thereafter. In this case, similar proportions of utterances in the early samples had been rated as both inappropriately slow and inappropriately fast. This suggested that this participant was in the process of gaining control over her rate of output during the study period.

Relative to stress, performance was relatively stable (albeit largely inappropriate) for 4/6 participants. By contrast, Participant 1 began at 36% appropriate utterances on sample 1 and increased to 82% by sample 8. This suggested an overall pattern of improvement in the use of stress. A similar pattern was observed for Participant 6 who increased from 10% appropriate utterances on sample 1 to 48% by sample 6. Thus, at least two participants appeared to be in the process of mastering the use of stress during the study period.

Data on loudness were relatively stable (and largely appropriate) for 5/6 participants. Participant 2 increased the percentage of appropriate utterances from 72 to 96% across the five samples. In this case, utterances in the earlier samples were almost exclusively being rated as inappropriate due to soft (or low) volume. Anecdotal reports by the examiners suggested that this participant was initially reluctant to participate in the interactions which may have influenced her voice volume. This appeared to resolve over time.

Relative to laryngeal quality, three participants (2, 3, and 4) evidenced relatively stable (and appropriate) performance across the study period. Data for Participant 1 revealed a large decline in performance from sample 1 (96% appropriate utterances) to sample 2 (36%) followed by a gradual return to passing performance with values of 72, 92, 96, 92, 100, and 95% across the subsequent samples. It was not clear, from available information, what might have precipitated this change in performance. The predominant code being applied to the inappropriate utterances for this participant was ‘rough’ voice quality. Data for Participants 5 and 6 both indicated trends for improvement from failing to passing performance over time. This suggested that both of these participants were gaining control over laryngeal quality during the study period.

Finally, relative to resonance quality, performance was stable (but quite inappropriate) for 4/6 participants. Data for Participant 2 suggested a gradual improvement in performance from 32% appropriate utterances on sample 1 to 84% on sample 4, but a decline to 40% on sample 5. The predominant code for the inappropriate utterances was nasopharyngeal. Information on file (and reports from clinical staff and her family) could not account for this sudden drop in performance. Data for Participant 4 indicated an increasing trend from 0% appropriate utterances on sample 1, to 12% by sample 3, and finally to 60% by sample 7.
4. Discussion

Findings from this study indicated that, unlike similar children with severe to profound hearing impairments who had not been fitted with cochlear implants, phrasing and pitch were not a significant problem for any of these six children. In addition, rate, loudness, and laryngeal quality were only a problem on a small subset of the samples. However, similar to children with hearing impairments without cochlear implants, resonance quality and use of stress were clearly an issue for these children. This is consistent with previous studies of the speech of the hearing-impaired (Boone, 1966; Finkelstein et al., 1993; Fletcher et al., 1999; Nickerson, 1975). Overall, however, these findings suggest that the prosody and voice characteristics of the conversational speech of children with cochlear implants are much less of an issue than has historically been the case with children with hearing impairments. Indeed, general impressions by examiners, as well as coders and transcribers of the samples were that the speech of these children did not sound the same as the classic descriptions of the speech of the hearing-impaired. Thus, the current findings highlight the benefits of cochlear implantation for children with severe to profound hearing impairments. The fact that all of the children in the current study received their implants by age 3 years also provides additional support for both early identification and earlier implantation.

Correlational analysis suggested that laryngeal quality improves over time in children fitted with cochlear implants. When one potential outlier was removed from the analysis, stress and resonance quality were significantly correlated with the amount of auditory experience. The failure to obtain correlations with the other four other PVSP variables (phrasing, rate, loudness, and pitch) likely reflects the overall high level of performance by these children on each of these variables (see Table 2). This suggests that receipt of a cochlear implant (and perhaps specifically receipt at such an early age) allowed these children to master these aspects of conversational speech. This is supported by the fact that at least 56% of the utterances in the samples in the current study were rated as appropriate on these four variables.

Examination of the data for individual speakers on each variable suggested that some developmental patterns might be discerned. Relative to stress, for example, two participants (1 and 6) appeared to be in the process of mastering the use of stress. Mastery of rate was occurring for Participant 4, and learning fine control of laryngeal quality appeared to be ongoing for Participants 5 and 6. Participant 4 appeared to also be in the process of learning to control the resonance quality of her voice.

The size of the current sample precludes broad conclusions about the age of mastery of the prosody and voice variables examined herein by children fitted with cochlear implants. The findings are suggestive however, and additional study with larger samples is clearly warranted. In addition, more study of these aspects of speech with children with much less experience with their cochlear implants (i.e., less than the minimum 23 months for the children in the current study) would appear to be appropriate. Acoustical analysis of various aspects of prosody and voice in this population would also appear warranted to confirm the current findings. It is not clear whether perceptual impressions of speech used to analyze the current samples are sensitive enough to detect changes in the other variables that were not apparent in the current study. Findings from the current study may also have limited generalizability because the children tested may not necessarily represent this
population adequately. The PPVT-III values may be narrower than is typically seen. This is supported by findings from Spencer and Bass-Ringdahl (2004) who studied 19 children fitted with cochlear implants at 12–29 months of age; PPVT-III standard scores ranged from 42–102 after 36 months of implant use.

The current findings do have some potentially important clinical implications. First, it has previously been assumed that natural sounding speech for children with severe to profound hearing impairment was an unrealistic goal. Findings from this study would suggest otherwise for children fitted with cochlear implants. However, assuming that the findings of the current study can be confirmed, clinicians working with this population may still need to focus their efforts on the use of stress and the resonance aspects of speech. The current study also demonstrated that a tool such as the PVSP can be used with this population. For clinicians who wish to track changes in the conversational speech of children with hearing impairments (whether fitted with cochlear implants or not), this tool may be useful. Although narrow phonetic transcriptions were available for the current study, PVSP analysis ultimately only requires samples which are transcribed orthographically. Such samples may already be available in many clinical settings.

Acknowledgments

Many thanks to Julie Beeler, Velvet Buehler, Molly Erickson, Mark Hedrick, Marge Hudson, Gayla Hutsell, Kim Jenkins, Leslie Jolles, and Saneta Thurman for their assistance with this study. Special thanks to Lana Colvard and Michelle LaGesse for testing the participants and to Lisa Moss for transcribing the conversational speech samples. Many thanks also to the children who participated and their parents. The data in the current study were used as part of an unpublished undergraduate honors thesis prepared by the first author at the University of Tennessee, Knoxville. Portions of the findings in this study were presented as a poster at the 2005 Annual Convention of the American Speech-Language-Hearing Association.

Appendix A. Self-study questions

1. Which of the following aspects of prosody and voice had previously been noted as problematic for individuals with severe to profound hearing impairment?
   a. phrasing, rate, stress, resonance;
   b. stress, pitch, loudness, resonance;
   c. pitch, loudness, resonance, laryngeal quality;
   d. all of the above.
2. The Prosody–Voice Screening Profile (PVSP) by Shriberg et al. (1990) is used to analyze all of the following except:
   a. laryngeal quality;
   b. phrasing;
   c. speech intelligibility;
   d. resonance quality.
3. The children in the current study
   a. were all prelingually deaf;
   b. were implanted with a cochlear implant after age 3;
   c. used both spoken language and manual signs (simultaneous communication);
   d. both (a) and (c);
4. The only prosody or voice characteristic that resulted in a significant correlation with hearing age (HA), chronological age (CA), and post-implantation age (PIA) was
   a. resonance quality;
   b. laryngeal quality;
   c. stress;
   d. pitch.
5. Based on the current study, children who have been fitted with cochlear implants may differ from children with severe to profound hearing impairment who have not been fitted with cochlear implants because
   a. Children who have been fitted with cochlear implants demonstrate more problems with rate of speech.
   b. Children who have been fitted with cochlear implants have no problems with resonance quality or loudness.
   c. Children who have been fitted with cochlear implants have little to no problem with phrasing and pitch.
   d. Children who have been fitted with cochlear implants have some difficulty with rate, loudness, and laryngeal quality.

References


