Vowel Production in a Prelinguistic Child Following Cochlear Implantation

Walter H. Manning*
Jeffrey N. Moore*
Marilyn J. Dunham*
F. Ling Lu†
Elizabeth Domico‡

Abstract

Formant frequencies were determined from vocalic utterances of a prelinguistically deafened child implanted with a Nucleus 22-channel device at age 5 years. Speech samples were obtained from recordings of speech made 5 and 2 months prior to implantation and at 6, 12, 18, 24, and 36 months post implantation. Prior to implantation, first formant values showed a greater range than those of normal hearing children of comparable age, and second formant values showed a greatly reduced range and clustered around a mean of 1800 Hz. By 36 months post implantation, first formant values approximated age-matched norms. By 6 months post implantation, higher second formant values were apparent. A progressive shift in second formant values was noted through 36 months post implantation when the vocalic space most closely corresponded to age-matched normative data.

Key Words: Child, vowel, formant, cochlear implant, longitudinal

Since the introduction of implantable cochlear prostheses in the 1960s, a number of reports have appeared describing the perceptual capabilities of patients using these devices. A smaller number of investigations have considered the changes in speech production associated with implant use (Kirk and Edgerton, 1983; Kirk and Hill-Brown, 1985; Eisenberg et al, 1986; Osberger, 1989; Tartter et al, 1989; Tobey et al, 1991). Data describing longitudinal change, particularly for children with prelinguistic onset of deafness using the Nucleus 22-channel device are lacking.

The purpose of this investigation was to provide data concerning longitudinal formant changes in the vocalic segments of one prelinguistic child following implantation of a multichannel cochlear prosthesis. This child was selected for study because she has responded well to all aspects of intervention associated with cochlear implantation. Documentation of systematic changes in the formant frequencies allows inferences regarding characteristics of vocal tract shaping. Methods that rely only on listener judgments may not be as sensitive to such changes, particularly when the vowels produced are in distorted or non-standard phonologic form.

METHOD

Subject

The subject was a profoundly deaf girl with hearing loss of probable congenital onset and unknown origin. The child had worn binaural, high-power hearing aids from the time of the initial diagnosis of hearing loss at approximately 1 year of age until fitting of a Nucleus 22-channel wearable speech processor at age 5.0 years.

This child's history of aural rehabilitation involved speech and language training using high-power binaural aids up to the initial pre-stimulation session. Communication with the child throughout toddler and preschool years
involved the use of amplified residual hearing, speech reading, and sign language. The child was evaluated for possible cochlear implantation at the Shea Clinic in Memphis and at the Memphis Speech and Hearing Center, Memphis State University.

The child was implanted at the Shea Clinic in July, 1987. The operation was unremarkable with all active rings and all stiffening rings inserted. The coding scheme for the speech processor employed fundamental frequency as well as first and second formant information (WSP III with information through 3500 Hz). Following implantation, the patient received new maps as appropriate throughout poststimulation follow-up testing.

Procedure

Tape recordings of the child's interactions with a parent and one or two investigators were made at 6-month intervals. Two recordings were made prior to implantation at 4 years, 8 months and 4 years, 11 months. Additional recordings were made at 6, 12, 18, 24, and 36 months post implantation.

Following the initial preimplant session the patient received auditory training emphasizing discrimination of word and syllable differences. She also received intervention speech training following the Ling (1976) protocol. This training was provided on the average of four times a week prior to and following implantation.

Vocalic utterances were sampled at random intervals from the recordings in a manner similar to that described by Kent and Murray (1982) and Kent et al (1987). Digitization of the tape-recorded analog signal and subsequent acoustic analysis was conducted using the MacSpeech Lab software/hardware system. Because the video recordings allowed more precise identification of the subject's participation in the conversations, acoustic data from these recordings were used for analysis. The number of vocalic utterances sampled ranged from 52 to 99 and averaged 81.6 utterances (Table 1). Steady-state vocalic portions of the subject's speech with clearly observable formant structure were used for spectrographic analysis. Formant frequencies were estimated on separate occasions by two persons making use of wide-band (300 Hz) spectrograms as well as discrete spectral (FFT) analysis of corresponding vocalic portions. Intrajudge and interjudge reliability of the formant estimates was similar to that of previously published reports on the estimation of formant frequencies of children's vowels (Kent and Murray, 1982; Kent et al, 1987). Interjudge agreements for first formant determination averaged 35 Hz (r = 0.91). Interjudge agreements for second formant determination averaged 45 Hz (r = 0.98). Intrajudge agreements for first and second formant determination averaged 24 Hz (r = 0.98).

As a measure of speech intelligibility, the child produced the 36 McGarr (1983) sentences during each recording session. These three-to-seven syllable sentences are divided into 18 high and 18 low context items with a single target word within each sentence. The child's tape recorded sentences were transcribed by two listeners, one who was familiar and one listener who was unfamiliar with the speech of hearing-impaired subjects. One of the authors then determined the ability of the listeners to

<table>
<thead>
<tr>
<th>Session</th>
<th>Pre A</th>
<th>Pre B</th>
<th>Post 6</th>
<th>Post 12</th>
<th>Post 18</th>
<th>Post 24</th>
<th>Post 36</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>4-8</td>
<td>4-11</td>
<td>5-7</td>
<td>6-1</td>
<td>6-7</td>
<td>7-1</td>
<td>8-1</td>
</tr>
<tr>
<td>First formant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>99</td>
<td>72</td>
<td>52</td>
<td>91</td>
<td>88</td>
<td>81</td>
<td>88</td>
</tr>
<tr>
<td>Mean</td>
<td>752</td>
<td>892</td>
<td>696</td>
<td>797</td>
<td>767</td>
<td>706</td>
<td>748</td>
</tr>
<tr>
<td>SD</td>
<td>189</td>
<td>153</td>
<td>184</td>
<td>203</td>
<td>212</td>
<td>199</td>
<td>148</td>
</tr>
<tr>
<td>Range</td>
<td>316-1164</td>
<td>388-1123</td>
<td>336-1153</td>
<td>338-1184</td>
<td>377-1184</td>
<td>316-1082</td>
<td>494-1050</td>
</tr>
<tr>
<td>Second formant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1800</td>
<td>1998</td>
<td>2159</td>
<td>2189</td>
<td>2057</td>
<td>1963</td>
<td>2087</td>
</tr>
<tr>
<td>SD</td>
<td>211</td>
<td>246</td>
<td>435</td>
<td>546</td>
<td>489</td>
<td>480</td>
<td>445</td>
</tr>
<tr>
<td>Range</td>
<td>1205-2256</td>
<td>1164-2379</td>
<td>1153-2941</td>
<td>1153-3196</td>
<td>1215-3196</td>
<td>1122-3186</td>
<td>1321-3212</td>
</tr>
</tbody>
</table>
identify the target word yielding a percent correct score for all 36 target words.

RESULTS

Table 1 provides the number of samples, mean, standard deviation, and range for the first and second formant frequencies across two prestimulation and five poststimulation conditions. Changes in the speaker's $F_1/F_2$ vowel space during the two preimplantation and the five postimplantation recording sessions are shown by the seven plates contained in Figure 1. Age-matched normative data obtained by Eguchi and Hirsh (1969) for children (boys and girls combined) are provided for comparison.

First formant values were somewhat higher than the normative data and were characterized by a greater than normal range. The first formant values begin to approximate those of normal hearing children by 24 months postimplantation (CA = 7-1 for this subject). By the 36-month postimplantation recording session first formant values showed the closest correspondence to the age-matched norms.

In comparison to normative data, the second formant values show a reduced range (approximately 100 Hz). Preimplantation values averaged 1800 Hz, identical to the second formant frequency noted by Monsen (1976) as characteristic of young hearing-impaired speakers. By 6 months post implantation second formant values were markedly higher, resulting in an increased range of values. A progressive shift in second formant values continued through 36 months post implantation when they showed the closest approximation to the age-matched norms.

Table 2 indicates the changes in intelligibility scores during the subject's production of the McGarr sentences. Percent correct scores are shown for low and high context stimuli using both familiar and unfamiliar listeners during the two preimplantation and the five postimplantation recording sessions. The key words from these sentences remained unintelligible from the initial preimplantation session (age 4-8) through the 6-month postimplantation session (age 5-7). Improvement in intelligibility was first reflected by these scores at the 12-month post implantation session (age 6-1). Subsequent 6-month recordings showed continued improvement in intelligibility through the 36-month postimplantation recording session with scores from the high context sentences as judged by familiar listeners yielding the highest ratings.

DISCUSSION

It is, of course, not reasonable to generalize results obtained from a single subject. As Osberger (1989) pointed out, subjects show a highly individualized response to implantation. For this young female speaker, selected speech production characteristics were monitored over a period of approximately three and one-half years — from age 4 years, 8 months until just beyond 8 years. Following implantation at age 5 years, this child demonstrated consistent and obvious improvement in her speech intelligibility in the form of the McGarr sentence scores. Beginning at 12 months post implantation, these scores began to improve in an obvious and consistent fashion. Because there was a minimum of 6 months between all post-implantation administrations of the McGarr sentence, practice effects for the subject were considered to be minimal.

Figure 1 A and B, Changes in $F_1/F_2$ vowel space during the two preimplantation recording sessions (CA = 4-8 and 4-11). The area bounded by the solid lines represents age-matched normative data for children from Eguchi and Hirsh (1969).
The formant changes noted during vocalic productions over a 3-year period indicated progressive changes in first and second formant values in the direction of age-matched normative values. Obvious changes in formant values were seen by the 6-month postimplantation session. However, considerable additional change took place during the following 2.5 years. By the 36-month postimplantation recording session these formant values provided a close, but not identical, correspondence to the $F_1/F_2$ vowel space characteristic of normal hearing children.

Although no single vocal tract gesture can be associated with the production of a particular vowel, it is typically the case that the second
Table 2 Changes in Intelligibility Ratings (in percent) for the Subject’s Production of the McGarr Sentences for both Low- and High-Context Stimuli and Familiar and Unfamiliar Listeners during Preimplantation and the Five Postimplantation Recording Sessions*

<table>
<thead>
<tr>
<th>Session</th>
<th>Pre A to Post 6</th>
<th>Post 12</th>
<th>Post 18</th>
<th>Post 24</th>
<th>Post 36</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA Context</td>
<td>4-8/5-7</td>
<td>6-1</td>
<td>6-7</td>
<td>7-1</td>
<td>8-1</td>
</tr>
<tr>
<td>Low High</td>
<td>0 0</td>
<td>17 20</td>
<td>11 22</td>
<td>39 56</td>
<td>50 83</td>
</tr>
<tr>
<td>Listener</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar</td>
<td>0 0</td>
<td>0 5</td>
<td>6 28</td>
<td>33 56</td>
<td>39 72</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>0 0</td>
<td>0 5</td>
<td>6 28</td>
<td>33 56</td>
<td>39 72</td>
</tr>
</tbody>
</table>

*Scores from preimplantation sessions A and B and the postimplantation sessions are combined due to zero intelligibility scores during each of these recording sessions.

Formant increases as the tongue is placed in a more anterior and elevated position. Correspondingly, the second formant tends to be lowered with a back tongue constriction. In addition, the first formant tends to increase with increased constriction of the pharynx (Fant, 1960; Lindblom and Sundberg, 1971; Pickett, 1980). Although both formants tend to decrease with lip constriction and both formants tend to increase with lip flaring (or jaw lowering), these articulatory gestures tend to be produced correctly by a hard-of-hearing or deaf child (Monsen, 1976).

During the 3-year period following cochlear implantation, the young speaker studied in this investigation demonstrated F1/F2 changes suggesting a greater degree of differential shaping of the oral cavity, particularly in terms of higher and more anterior positioning of the tongue. That is, second formant values associated with high, front vowels began to assume normal limits by the 6-month postimplantation session. However, both first and second formant values continued to show instances of higher than normal values for low, back vowels. This is most obvious in the data from the 6-, 12-, and 18-month postimplantation recording sessions and continues to be apparent in the 24- and 36-month sessions. These results provide indirect evidence of vocal tract shaping during production of low back vowels that is characterized by a greater than normal constriction of the tongue in the pharyngeal portion of the vocal tract (resulting in a higher than normal first formant) combined with a lack of tongue constriction in the posterior portion of the oral cavity (resulting in a higher than normal second formant). While there may be other explanations for these acoustic characteristics, these results coincide with reports that deaf speakers tend to have an abnormally posterior positioning for the tongue root/body during speech production (Stevens et al, 1983).

The improvement in the speech production of the subject of this investigation is likely the result of several factors. Information provided by the cochlear device appeared to be one important factor. Although a small shift in vowel formant relationships were observed from the first to the second preimplantation recording session, this may have been the result of the prestimulation training. Using the Ling approach (1976) this training focused on vowel production and vocal tract shaping. Following implantation, formant relationships began to show an obvious shift toward the normal F1/F2 vocalic space. Moreover, McGarr scores showed little or no change during the two preimplantation sessions but began to show consistent improvement by 12 months post implantation.

A second factor is the intensive aural rehabilitation program including speech and language treatment and auditory training in which the subject participated from the time implantation was first considered. Preoperative speech perception testing in the aided, auditory-only condition indicated minimal pattern perception skills and no ability to recognize vowel and consonant sounds. Following implantation, the subject demonstrated statistically significant (p < .05) improvement in closed-set identification of words based on suprasegmental and segmental speech features. Moreover, after 18 months of experience with the implant, the child was able to recognize a limited number of familiar words and phrases in the open-set condition. This subject’s improvement in auditory speech sound recognition via formant information provided by the cochlear implant system may have facilitated concurrent changes in the formant characteristics of her speech via the development of an internal feedback loop.
The emphasis placed on oral communication and on the importance of listening by the subject's family members, particularly her parents, also played a significant role. Her parents consistently requested good speech efforts from their daughter and were cognizant of the need to always use speech when communicating with her. The child's own motivation to wear the cochlear implant system and her willingness to undergo intensive training comprise a fourth factor. Finally, the commitment of the professionals working with this child at school and privately to optimize the benefits of the cochlear implantation should also be considered an important influence.

Acknowledgment. This research was supported in part by a State of Tennessee Higher Education Commission Center of Excellence Grant awarded to the Department of Audiology and Speech Pathology, Memphis State University. Appreciation is expressed to the staff of the Shea Clinic, Memphis, Tennessee for their support and cooperation during this investigation.

Portions of this manuscript were presented to the Annual Meeting of the American Speech-Language-Hearing Association, November, 1990, Seattle, WA

REFERENCES


