Reducing Consonant Voicing Inconsistency in a Child With a Severe-to-Profound Hearing Loss With Computer-Based Visual Feedback

Sheila R. Pratt
University of Pittsburgh

Using a computer-based system that provided visual feedback, a 13-year-old child with a severe-to-profound hearing loss was treated for a pattern of inconsistent consonant devoicing within an ABAB single-subject treatment design with multiple baselines. During a 5-week treatment phase his consonant devoicing was treated to criterion at the consonant-vowel level. A treatment effect was demonstrated and generalization to other phonemes at the CV-syllable level was noted. The child's intelligibility also improved at the word level but remained low in continuous speech indicating that generalization was restricted to isolated productions. These results suggested that despite a reduction in consonant devoicing, additional treatment sessions and possible changes in the treatment design were needed to stabilize the skill and promote generalization to more complex contexts.

Speech errors and reduced speech intelligibility are common sequelae of prelingual hearing loss (Hudgins & Numbers, 1942; Smith, 1975). The nature and magnitude of the impact of hearing loss on speech production varies across children, but many children with hearing loss require some type of intervention in order to speak in a manner acceptable and intelligible to most listeners. Children with severe or profound hearing loss usually require extensive long-term treatment if they are to become effective oral-aural communicators. Even children with mild-to-moderate hearing loss are at risk for resonance and segmental

---

Correspondence concerning this article should be addressed to Sheila R. Pratt, Ph.D., Department of Communication Science and Disorders, University of Pittsburgh, 4033 Forbes Tower, Pittsburgh, Pennsylvania 15260. Voice: 412-383-6537; Fax: 412-383-6555. E-mail: spratt@pitt.edu

speech problems and may benefit from speech therapy (Elfenbein, Hardin-Jones, & Davis, 1994; Oller & Kelly, 1974; West & Weber, 1973). Despite this need for therapy, only a limited number of well-controlled studies exists that document the effectiveness of speech treatment for children with hearing loss. Moreover, little is known about what treatment approaches and characteristics work best with the various types of speech impairments that occur with hearing loss in children.

Although limited to a single subject, the following treatment study tested the effectiveness of online visual feedback for consonant voicing with a commonly used computer-based feedback system (Voicing Awareness module of the International Business Machines [IBM] SpeechViewer II; IBM Corporation, 1992). The effectiveness of using visual feedback was assessed by treating a child who had a severe-to-profound hearing loss and exhibited a pattern of inconsistent devoicing across his consonant inventory.

Visual Feedback

The debate over the use of visual and other non-auditory forms of feedback for treating the speech of children with hearing loss has a long history and distinguishes the major therapeutic approaches that have been proposed for treating speech impairment in children with hearing loss. The issues are whether non-auditory (particularly artificial) forms of feedback are relevant to the speech sensory-motor system; whether they interfere with the development of the auditory system and its linkage with speech production; and if artificial forms of feedback are used in therapy, how they can be most effectively employed when training speech skill in children. These issues are relevant to most computer-based feedback systems because most depend on artificial visual feedback. Moreover, the visual feedback is often transformed and focuses on one acoustic feature for ease of interpretation (Bernstein, 1989; Pratt & Hricisak, 1994).

Most clinicians who use a multisensory treatment approach are very willing to incorporate non-auditory types of feedback. Carhart (1947, 1963) and later Silverman (1971) argued for using various forms of feedback to supplement auditory feedback when training speech skills in children with hearing loss. In describing a speech training approach used at the Lexington School for the Deaf, Magner (1979) endorsed using artificial feedback and cueing systems, even with infants. Youdelman, MacEachron, and Behrman (1988) used a combination of visual and tactile sensory aids to promote more normal pitch in a group of adolescent children with hearing loss, and Ertmer and Maki (2000) used various auditory, visual, and tactile forms of feedback in their noninstrumental therapy with adolescent deaf children. The perspective of these traditional treatment approaches is that it is acceptable to use whatever modalities and means available in order to compensate for limited auditory input and stimulate speech production. It therefore seems appropriate to use artificial visual feedback to train speech skills.

Ling (1976; 2002), however, has been more cautious. He acknowledged that visual and other forms of feedback may be needed for some children when initially learning to produce a particular sound, but he warned that augmented or artificial forms of feedback should be withdrawn as quickly as possible so that a dependency on the feedback is not established. His concern has support in the motor-learning literature in that artificial or augmented feedback, such as computer-based feedback, can hinder skill learning of motor tasks when intrinsic feedback is minimal or difficult to interpret (Lintern, Roscoe, & Sivier, 1990). Developing a dependency results from the feedback becoming a part of the motor memory during practice and, when it is withdrawn, the weaker intrinsic feedback is not sufficient to support the skill (Proteau & Cournoyer, 1990; Proteau, Marteniuk, Girouard, & Dugas, 1987). Dependence on augmented feedback is particularly problematic when the feedback schedule is 100%, and the task and feedback are simple (Weinstein & Schmidt, 1990; Wulf, Shea, & Matschiner, 1998). However, complex behaviors with multiple characteristics, such as speech, are less affected by the frequency of the feedback, so that the common use of 100% feedback by clinicians and computer-feedback systems may not be detrimental to learning.

Some proponents of auditory-verbal approaches hold an even more restrictive view on artificial feedback (Estabrooks, 1994; Pollack, 1985). They have argued that the therapeutic focus should be the improved use of residual hearing, and that other forms of feedback impede the development of auditory skills that are needed to support speech production. As a form of constraint-based therapy, auditory-verbal therapeutic approaches restrict visual feedback, with some clinicians restricting even naturally occurring visual feedback, in an effort to strengthen the auditory system or force strategy formations to better utilize the speech-acoustic information that is perceived. Constraint-based therapeutic approaches have been found to be effective for other types of conditions with other clinical populations (Freeman, 2001; Schaechter et al., 2002), but empirical evidence is lacking for (and against) its application in the treatment of speech in children with hearing loss. It also should be noted, relative to the following study, that improvements in speech with artificial feedback do not discount the effectiveness of treatment approaches that advocate an auditory focus.

Effectiveness of Computer-Based Feedback

Computer-based speech feedback systems have been proposed for over two decades as a means of altering speech production in children with hearing loss because of the visual nature of the feedback (Bernstein, 1989; Watson & Kewley-Port, 1989). One goal for using computer-based feedback systems is to provide feedback that might facilitate learning new speech behaviors or the modification of existing speech behaviors. Another goal is that children will be able to use computer-based feedback systems independently so that therapy time can be augmented without substantial increase in clinician contact and cost (Osberger, Moeller, Kroese, & Lippmann, 1981; Watson & Kewley-Port, 1989). Independent use is a critical matter with children with hearing loss who tend to require long-term treatment and with whom many clinicians have limited experience. To that end, the visual feedback provided by computers must be comprehensible to children and useful in effecting change (Bernstein, 1989; Pratt & Hricisak, 1994; Watson & Kewley-Port, 1989). Most of the commercially available systems are multi-modular and target various properties of speech such as voicing, vocal pitch and intensity, and spectral adequacy (Pratt & Hricisak, 1994). The modules in most systems vary substantially in the type and complexity of feedback provided. As such, they need to be tested with various types of children who appear to match the particular focus of the modules, cognitive requirements of the feedback, and interest level.

Only a limited number of studies have documented the effectiveness of the systems available to most clinicians. The results of early studies that used laboratory prototypes were mixed and difficult to interpret because they tended to lack adequate controls (Fletcher & Hasegawa, 1983; Fletcher & Higgins, 1980; Osberger et al., 1981). The results from more recent studies that have used commercially available systems, or systems that are similar to those on the market, have been more promising. In a group study with school-aged children with profound hearing loss, Dagenais, Critz-Crosby, Fletcher, and McCutcheon (1994) compared Ling's (1976) approach for treating consonant production to training consonant production using palatometric feedback. They found that palatometry was more effective than Ling's approach when the target productions were compared, but intelligibility at the word level was similar for the two treatment groups. In a study using single-subject design with replication across 5 children, Pratt, Heintzelman, and Deming (1993) found that the Vowel Accuracy module of the IBM SpeechViewer effectively promoted correct vowel production in some preschool-aged children with hearing loss but that clinician involvement was a critical consideration when using the module with young children. In another study using single-subject-design, Ertmer, Stark, and Karlan (1996) found that spectrographic feedback targeting the first and second formants was effective with two 9-year-old deaf children in improving vowel production, although generalization to untreated vowels was limited and clinician support and instruction was an integral component of the treatment protocol. More recently, Ertmer and Maki (2000) compared the efficacy of noninstrumental instruction against noninstrumental instruction plus visual feedback from a spectrographic display. The productions of /m/ and /t/ in monosyllabic words were trained with 4 teenaged deaf children who did not wear sensory aids. Both approaches resulted in improved performance but the relative differences between approaches were child and phoneme dependent making it difficult to determine the influence of the spectrographic display.

It is not unreasonable that the efficacy of visual feedback via computer would be dependent on the speech characteristic being treated. Speech features that are particularly susceptible to hearing loss, and speech gestures associated with limited visibility and proprioceptive feedback may be particularly responsive to visual feedback. Consonant voicing is a speech feature that, if not well perceived through audition, may be effectively treated with computer-based visual feedback because it is not easily viewed, is associated with limited proprioceptive feedback, and requires strict temporal coupling between articulators.

Consonant Voicing Errors

Speakers with severe-to-profound loss typically produce multiple types of speech errors, and exhibit compromised function at various levels of the speech production system. One of the more common types of error is improper control of consonant voicing (McGarr & Lofquist, 1982; Monsen, 1976; Smith, 1975). Problems with consonant voicing have been observed in individuals with post-lingual as well as pre-lingual hearing impairment. It also has been a characteristic that changes with modifications in auditory feedback in adults with adventitious deafness (Kishon-Rabin, Taitelbaum, Tobin, & Hildesheimer, 1999; Lane, Wozniak, Matthies, Svirsky, & Perkell, 1995). Consonant voicing errors produced by children with hearing loss have been reported by a number of researchers and clinicians although there is inconsistency in the patterns described. Hudgins and Numbers (1942) and Markides (1970) reported that children with hearing loss tend to devoice voiced consonants. Similarly, Nober (1967) observed that voiceless consonants are more often correct than are voiced consonants. In contrast, Smith (1975) found voiced production of voiceless consonants to be more common. Heider, Heider, and Sykes (1940) as well as Carr (1953) observed that children with hearing loss are more likely to use voiced than voiceless cognates, suggesting that voiced consonants may be easier to produce. In addition, Millin (1971) reported that some speakers with hearing loss produce continuous voicing throughout an utterance as a way to compensate for inadequate voicing control. Predominant voicing patterns also may be dependent on sound class and influenced by general developmental patterns. For example, Tobey, Geers, and Brenner (1994) found that children with profound hearing loss tend to develop voiceless fricatives and affricates prior to the voiced cognates, but show the reverse pattern with stop consonants.

It also has been reported that both types of consonant voicing errors can be observed in some children and often are inconsistently exhibited (Pratt & Tye-Murray, 1997). It is atypical for children with hearing loss to use only voiced or only unvoiced cognates, and when consonant voicing errors do occur their expression often varies over time and across and within contexts (Smith, 1975). The differences in the patterns observed, the inconsistency of their expression, as well as the variability in production over time likely reflect the influence that audition has on the development of speech motor-control.

Pratt and Tye-Murray (1997) concluded that inappropriate consonant voicing and devoicing are probably due to reduced temporal coordination of the larynx with the upper-airway articulators and, as such, represent phonetic rather than phonologic differences in most children with severe to profound hearing loss. Similarly, Millin (1971) proposed that voiced-for-voiceless errors are probably due to a mistiming of initiation and termination of phonation. It also has been suggested that some speakers with hearing loss miscue the perception of consonant-voicing distinctions by improperly timing the durations of the preceding vowels or the durations of the target segments (Gold, 1980; Osberger & McGarr, 1982; Tobey, Pancamo, Staller, Brimacombe, & Beiter, 1991). The implication is that consonant voicing errors may be cued by mistiming at various points in production, or possibly the production of reduced temporal co-ordination across multiple segments and not just mistiming of the individual segments. Because listeners are sensitive to phonetic context effects, incompatible contextual information, such as incongruent voicing cues, can result in misperceptions and reduced intelligibility (Fowler, Brown, & Mann, 2000; Holt & Lotto, 2002; Holt, Lotto, & Kluender, 2000).

Treatment Rationale

The following treatment focused only on voice onset as a cue for consonant voicing. The child who was treated was unable to maintain consistent control of voicing with audition alone, so the rationale for using a computer-based visual feedback system was to provide an online visual analogue of his acoustic speech. The assumption was that the child would use the visual feedback to compensate for poor auditory feedback, which would allow for the establishment of a better linkage between his oromotor and orosensory systems, and more accurate correspondence with his internal auditory representation or model of voiced speech sounds (Perkell et al., 1997; Perkell, Matthies, Svirsky, & Jordan, 1995).

METHODS

Subject

The subject was a 13-year-old boy with a congenital severe-to-profound bilateral sensorineural hearing loss (see Figure 1). His hearing loss was identified at approximately 10 months of age and he was fitted with binaural hearing aids at 12 months. At the time of the study he wore an FM system with button receivers (Phonic Ear 471) in school and linear behind-the-ear hearing aids when not in school. Reports from teachers, parents, and the child indicated that he consistently wore his sensory aids. A listening check of his FM system and hearing aids was completed at the beginning of every school day by the classroom teacher, and his FM system was checked by the treating clinician (the author) prior to each treatment session. In addition, the school audiologist periodically evaluated the child's FM system and hearing aids electroacoustically.

The child's auditory performance was restricted even while wearing his sensory aids, although it was consistent with his age and hearing-loss severity. On the *CID Early Speech Perception Test* (Moog & Geers, 1990) he performed at a level consistent with Category 4 in that he demonstrated pattern perception, some ability to recognize a limited set of spondees and over 50% correct recognition (14 out of 24) of a limited set of monosyllabic words. On the *Test of Auditory Comprehension* (Office of the Los Angeles County Superintendent of Schools, 1980) he passed only the first two subtests (out of 10) and on the four subtests completed, his *t* scores ranged from 55 to 59, indicating that his performance was

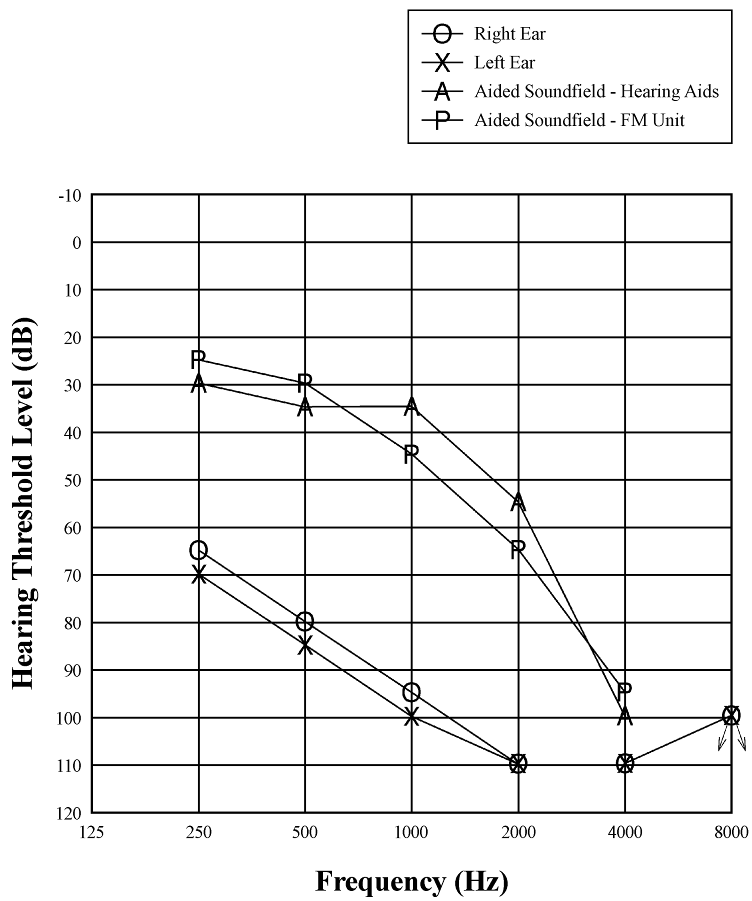

Figure 1. The child's puretone audiogram and aided sound field thresholds.

at expected levels when compared to his age and hearing-loss-severity group. Due to his restricted auditory skills, he relied heavily on speechreading and other visual cues in conversation. His visual skills appeared to support a reliance on visual cues in that his performance on the *Motor-Free Visual Perception Test* (Colarusso & Hammill, 1972) was at age level.

The child's auditory-oral vocabulary skills and his speech skills were substantially depressed when compared to normal-hearing children his age. When assessed with the *Comprehensive Receptive and Expressive Vocabulary Test* (Wallace & Hammill, 1994) and *Peabody Picture Vocabulary Test – Revised* (Dunn & Dunn, 1981) he scored at the first percentile on both tests compared to children with normal hearing. The child also exhibited a substantial number of speech errors typical of children with severe-to-profound hearing loss. On the *Goldman-Fristoe Test of Articulation* (Goldman & Fristoe, 1986) he produced 35 errors out of 73 target sounds and his speech intelligibility at the single-word level was 26% as tested with the *CID Picture SPINE* (Monsen, Moog, & Geers, 1988). Although some pitch and resonance differences were observed, his segmental errors appeared to reduce his intelligibility more significantly than his suprasegmental differences. This observation was confirmed by his performance on the *Fundamental Speech Skills Test* (Levitt, Youdelman, & Head, 1990) on which he performed at or above the eighth decile for all subskills except Elementary Articulation when compared to his age and hearing-loss-severity group. He exhibited glottal replacement, stopping, and deaffrication but his most notable segmental errors were omissions and consonant devoicing. These two types of errors were particularly evident in conversational speech. Despite the severity of the child's speech impairment, an oral mechanism examination revealed an intact oral-speech mechanism and oral-motor function typical for his age (Robbins & Klee, 1987).

The child was enrolled in an oral-aural school for the deaf at the time of the study. He consistently had received speech and language intervention and auditory-habilitation services since shortly after he first received amplification although he had not received any computer-based speech training. Just prior to the initiation of this study he was receiving speech, language, and auditory therapy from his school speech-language pathologist. He also was receiving supplementary speech treatment from a private speech-language pathologist. For the duration of the study he continued to receive therapeutic services to improve his language skills but speech production was not targeted other than in the study.

Treatment Design

An ABAB single-subject treatment design with multiple baselines was employed to document the effectiveness of computer-based visual feedback for voicing and to determine if generalization occurred. Voicing for the production of /b/ in consonant-vowel (CV) syllables was targeted for treatment because the child could produce the speech gestures for /b/ except for the correct timing of voicing. Voicing during the production of /d/ + V was monitored for generalization for the same reason and because it would reflect generalization of appropriate consonant voicing across place of articulation. At the eighth session of the pre-treatment baseline it was decided to also monitor generalization across voicing and manner of production, so the assessment of /v/ + V production was included. The delay in monitoring the /v/ was not considered a substantive weakness because treatment was not initiated until the /v/ pre-treatment performance characteristics were established.

In addition to measuring generalization at the CV-syllable level, it also was measured at the passage and single-word level. The *Zoo Passage* (Fletcher, 1978) and the *Rainbow Passage* (Fairbanks, 1960) were read at the beginning of each session to assess the treatment impact on intelligibility in continuous speech. As supplementary data, the *CID Picture SPINE* was administered before and after the treatment to assess intelligibility at the single-word level.

Baselines also were obtained relative to the spectral adequacy of /s/ and /S/ and the place of articulation of /k/ and /g/. These baseline data were not expected to change in response to the treatment because the child's /s/ and /S/ errors were spectral and not temporal in nature, and he was unable to produce the speech gestures associated with /k/ and /g/. The purpose of collecting these baselines was to monitor for non-treatment related influences that could account for changes in the treatment baseline such as general development, classroom activities, or changes in hearing aid performance.

Treatment

The Voicing Awareness module of the IBM SpeechViewer II residing on a desktop computer (Gateway 2000, 486) with a 15-in. monitor (Dell Ultrascan P1528U) was used to provide the visual feedback for consonant voicing. A hand-held microphone (Shure SM58), attached to a microphone stand approximately 12 in. from the child, was used to input the speech signal, and the visual feedback was presented via computer monitor. The child's productions also were audio recorded on cassette tape for later analysis. The visual display consisted of an image (e.g., clown, trees, dinosaur), which included colored portions when the child spoke. One color was displayed when unvoiced speech was present and another color was displayed when voiced speech was present. The feedback was based on the software detecting periodicity consistent with voicing in the speech signal. The color combinations were image specific. The child was allowed to select the image used in each session and he tended to vary his selection across sessions although not in any systematic fashion.

The treatment occurred individually in a quiet therapy room at the child's school (<50 dB SPL-A) and the software sensitivity was calibrated to the level of the background noise each session. The treatment sessions were scheduled twice weekly and lasted approximately 30 min. As with most treatment studies using a single-subject design, the duration of the pre-treatment and treatment phases of the study were determined by the performance of the child and the characteristics of the baselines. The data for the withdrawal phase (second A-phase) of the treatment baseline were collected after a 1-month withdrawal of treatment that occurred at a natural break in the child's academic calendar.

During treatment the child was asked to produce /b/ in CV syllables using three different vowel contexts, /A/, /i/ and /u/. The /b/ was produced 10 times in each of the vowel contexts with the order of all 30 cards entirely randomized for each session. Cards (3" × 5") with the individual CV syllables written orthographically on them were presented singularly and the child tried to say the syllable with the initial consonant voiced. He was not provided with an acoustic or spoken model to imitate and, therefore, had to rely on his own internal representation of the syllables. This was done to simulate a computer-based treatment approach in which a clinician would not be present to provide acoustic or visual-speech models. This method also was used with the assumption that without an external model the treatment would force the child to alter his internal representation. The task was to produce the syllables without any indication of consonant devoicing from the computer. The child was instructed during each treatment session to maintain a constant color on the video image throughout the production of the /b/-syllables and that after saying each syllable he was to indicate to the clinician whether the /b/ was produced correctly based on color characteristics of the visual feedback. The criterion for a correctly voiced production was conservative in that it only allowed voice-onset-times (VOT) of 0 ms or less (verified spectrographically on 30 treatment samples, allowing for 5 ms measurement error). Using such a restricted voicing performance goal made the feedback more categorical for the child and, as a result, easier for him to judge whether his productions were voiced. It also provided a target that would be less likely to shift to VOTs more consistent with consonant devoicing in continuous speech.

Requiring the child to indicate whether each /b/-syllable was produced correctly based on the visual feedback was included in order to monitor his knowledge of the results. Monitoring the child's judgments guaranteed that he was attending to the visual feedback and that his interpretation of the feedback was accurate. The child had little difficulty evaluating the visual feedback and indicating whether his productions were correct based on that feedback. Differences between the child and the clinician relative to these judgments were rare with percent agreement being 96% as measured over four sessions. The clinician did not correct the child if judgment errors were made although if the child had produced judgment errors in excess of 20% he would have received more extensive instruction on task requirements. It also should be acknowledged that by requiring an overt judgment from the child, an element was added to the treatment that could have affected the outcome beyond the visual feedback.

Assessment

The assessment of the treated and untreated target sounds occurred at the beginning of each subsequent session. This delay in the assessment eliminated the immediate effects of the treatment and allowed examination of learning and retention across sessions. All of the sounds were elicited in CV syllables with /A/, /i/, and /u/ vowels. The /b/ trials were completed first, followed by trials for /d/, /v/, /k/, /g/, /s/, and then /S/. Each consonant was produced five times per vowel context with the vowel order completely randomized per consonant. As was done during the treatment, cards depicting the syllables were presented and the child would attempt to say them. For the /b/, /d/, and /v/ syllable productions, consonant voicing was evaluated by the clinician using the same computer module and criterion for consonant voicing that were used during the training, but the computer monitor was turned away from the child so he could not view the visual feedback. Preventing him from viewing the computer monitor reduced the likelihood of additional learning during the assessment tasks. The clinician recorded whether the child produced correct voicing as indicated by no color change in the visual image on the computer monitor for each production. The training criterion was set at 80% correct consonant voicing for three consecutive sessions.

Performance with the /k/ and /g/ syllables was not based on voicing but overall accuracy as judged perceptually by the clinician. Voicing was not the sole parameter assessed because the child also was unable to coordinate the place and manner of production for /k/ and /g/, which was the reason they were monitored for changes independent of the treatment. They would be unlikely to change in response to the treatment but could possibly change in response to uncontrolled external or developmental factors. To establish the reliability of the clinician's judgments of the child's /k/ and /g/ productions, the clinician and three judges independently assessed 60 randomly selected productions of the child's /k/ + V and /g/ + V productions. The agreement between the clinician and the three judges was 100%.

The adequacy or goodness of /s/ and /S/ were analyzed by the Phoneme Accuracy Skill Building module in the SpeechViewer II software package. The software compared the productions of these sounds in CV contexts against a spectral model derived from previously obtained speech samples of 16 school-aged children with severe-to-profound hearing loss who had high speech intelligibility. The software produced a goodness score indicating the extent of spectral match between the child's production and the computer-derived model. The /s/ and /S/ productions also were assessed with the computer monitor turned away from the child so that he could not use the visual display to modify his productions. An audio signal reflecting goodness of the syllables also was produced by the software, so it too was turned off during the assessment of /s/ and /S/.

Following the assessment of the consonants at the CV syllable level, the intelligibility of the *Zoo Passage* and the *Rainbow Passage* was evaluated and was defined as the percentage of words identified by the clinician for each passage. To assess the reliability of the clinician's judgment of productions at the passage level, three independent listeners (audiology graduate students with at least 2 years of experience with deaf children) listened to 5 samples of both passages. Percent intelligibility for the two passages did not vary between the clinician and the three judges by more than 15%. The good inter-judge agreement may be due, in part, to the child's poor intelligibility in continuous speech.

Analysis

The treatment data were graphed and reviewed independently by three judges familiar with single-subject design and who had previously published studies that employed single-subject design (Jayaratne, Tripodi, & Talsma, 1988). They were asked to determine if a treatment effect was evident for /b/ and if generalization occurred for /d/ and /v/. They also were asked to indicate if substantive change occurred over time in the control baselines for /g/, /k/, /s/, and /S/ productions. A treatment effect; generalization; and changes in the /g/, /k/, /s/, and /S/ baselines were considered present only if all three judges agreed. Exact nonparametric statistics also were applied to the data to support the results of the judges. Exact nonparametric statistics were used because they are associated with fewer assumptions and are more appropriate when evaluating restricted data sets (Bakeman & Robinson, 1997; Cytel Software Corporation, 2001).

RESULTS

Treatment Effect

The treatment baseline and the baselines for the untreated sounds are illustrated in Figures 2 through 4 and represent 18 weeks of enrollment in the study. The pretreatment phase (initial A-phase) of the treatment baseline illustrates data collected over a 6-week interval, while the first treatment phase (initial B-phase) occurred over a 5-week interval. After a 4-week withdrawal that occurred at a natural break in the child's school program, the second A-phase was instituted over a 2-week period, and then the second B-phase was completed within a single week. Despite sessions being planned twice weekly, scheduling conflicts and illness resulted in some missed sessions, although at least one scheduled session occurred each week except during the withdrawal.

The three judges viewed the figures, and all three judges determined that a treatment effect was evidenced. Although the treatment baseline (i.e., production of /b/) was quite variable during the pretreatment phase (see Figure 2), treatment was evidenced by a shift between the pretreatment and treatment phases with performance in the treatment phase reaching the treatment criterion. The drop in performance on /b/ after withdrawal of treatment, and then a return to above criterion performance with the resumption of treatment provided further evidence that the improved performance in the treatment phase was due to the treatment. The clinician perceived all of the /b/ + vowel productions from the sessions with

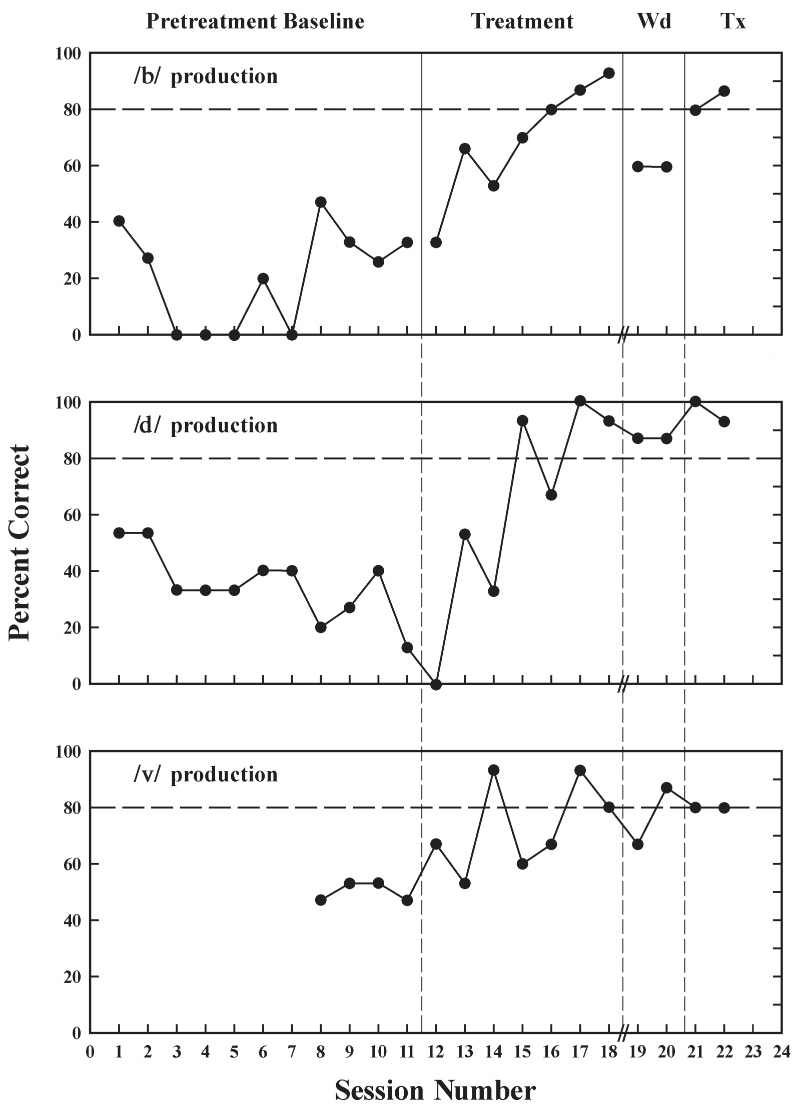

Figure 2. Percentage of voiced /b/, /d/, and /v/ productions in CV syllables as a function of treating voicing in /b/ at the CV-syllable level. The dropped dashed line represents the 80% criterion level.

above criterion-level performance as voiced, providing support for a treatment effect. Three independent listeners evaluated 30 of these productions and also perceived them all as voiced. These productions also were measured spectrographically and all had VOTs of < 5 ms, which is consistent with voiced productions.

To corroborate the judges' results, the phases of the treatment baseline (the child's production of /b/) were compared with exact permutation tests. The pre-treatment and the treatment phases were significantly different (p = .0002, one-tailed). However, the treatment and the withdrawal phases were not significantly different (p = .1667, one-tailed) nor were the withdrawal and final treatment phases (p = .1000, one-tailed) although these two comparisons were based on a very limited number of data points and should be viewed with caution.

Generalization

Beyond finding a treatment effect, the judges also unanimously agreed that

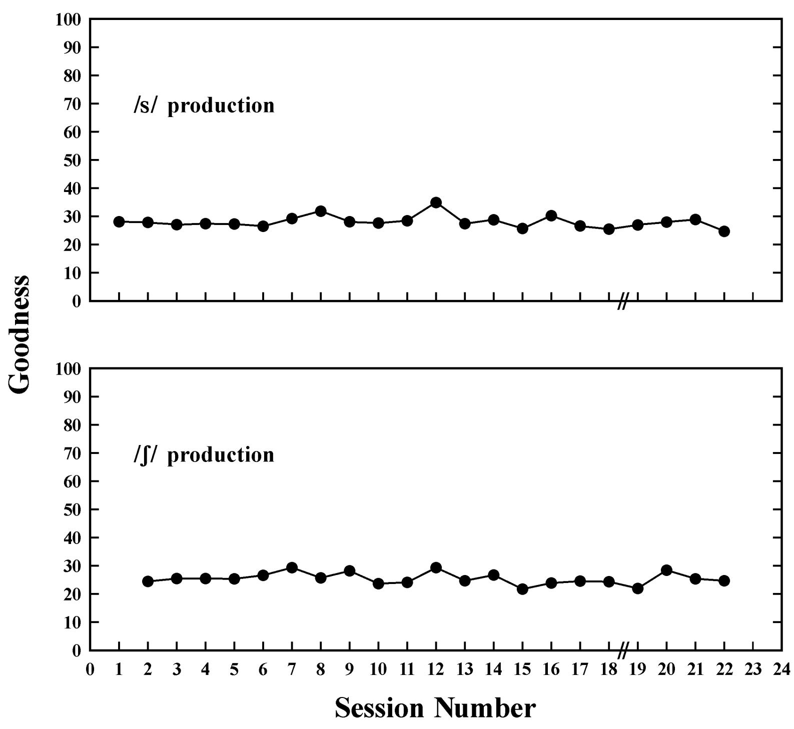

Figure 3. Spectral goodness of /s/ and /S/ productions in CV syllables with treatment of voicing in /b / at the CV-syllable level. Goodness was a metric of spectral match between the child's productions and previous samples collected from highly intelligible children with hearing loss.

generalization occurred for voicing of /d/ and /v/. These two untreated baselines, shown in the middle and lower panels of Figure 2, shifted up in correspondence with the changes in the treatment baseline, the upper panel of Figure 2. Exact Spearman's correlation coefficients for the /b/ and the /d/ and /v/ baselines were .71 (p = .00009) and .72 (p = .00008) respectively indicating a significant correspondence in the baseline changes for /d/ and /v/ with the treatment of /b/. In contrast, the judges found no changes in the /s/ or /S/ baselines (see Figure 3), or in the /k/ and /g/ baselines, suggesting that the treatment and generalization effects for consonant voicing were due to the treatment rather than other potential factors such as general developmental gains.

Single-word speech intelligibility, as measured by the *CID Picture SPINE,* was consistent with the treatment and generalization effects, with the child's single-word speech intelligibility increasing from 28.5% to 53.5% over the course of the study. However, intelligibility at the connected-speech level was largely unaffected (see Figure 4). Intelligibility of the read passages ranged from 23% to 31% (M = 27.6%) for the *Zoo Passage* and 28% to 38% (M = 31.8%) for the

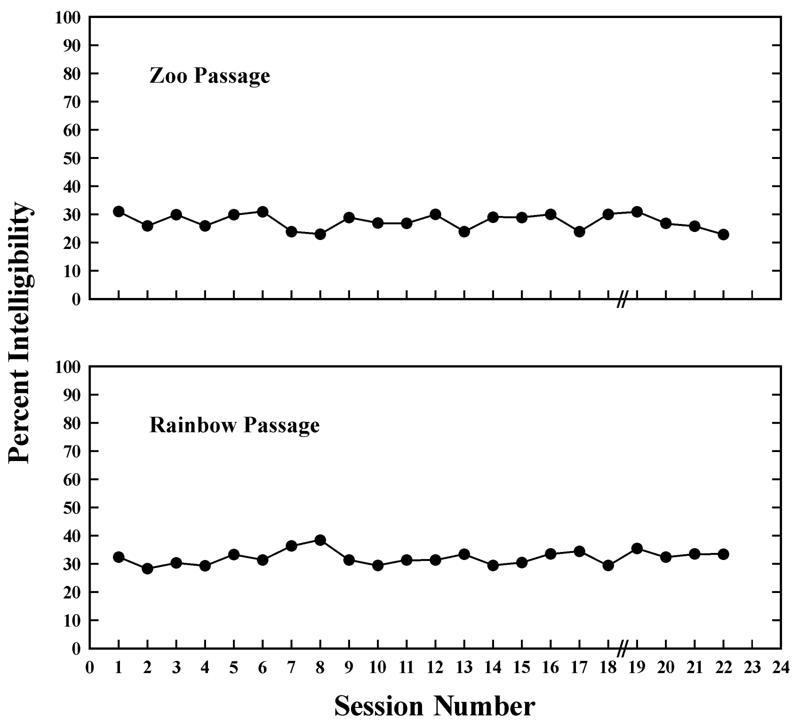

Figure 4. Changes in intelligibility of the *Zoo Passage* and *Rainbow Passage* with treatment of voicing in /b/ at the CV syllable level.

Rainbow Passage. The mean percent intelligibility across both passages was 29.7%. The limited impact on speech intelligibility was not entirely unexpected. The child was treated only at the CV-syllable level and no effort was made to generalize or stabilize his newly trained behaviors. The child likely needed to remain above criterion level for a longer period of time so that consolidation of the skills could occur. In addition, the introduction of variability in the feedback schedule and contexts of the training, and systematically varying the difficulty of the task might have resulted in more consolidation of the behaviors and greater generalization.

DISCUSSION

This single-subject treatment study showed that using computer-based visual feedback was a reasonable approach to training consistency in consonant voicing in this child, at least at the initial skill acquisition level. Moreover, this treatment was implemented without clinician support or augmentation. Prior to the treatment, the child occasionally produced voiced consonants correctly but not uniformly, even at the single syllable level. With repeated visual feedback from the computer he produced the target sound with appropriate voicing with at least 80% accuracy at the CV-syllable level, and began showing higher levels of consistency across and within sessions. He also improved his voicing on the non-treated /d/ and /v/ sounds indicating that he was learning a coordinative sensory-motor skill independent of the individual target sound and its particular combination of features. That is, the child likely was developing a more consistent internal representation of voicing, at least in CV syllables. Although not proof, the results are concordant with inconsistent consonant voicing being a phonetic-level (motor process) problem with children with hearing loss since the child was only given phonetic-level feedback (Van der Merwe, 1997).

The results further suggest that the child attended to the categorical characteristics of the visual feedback (color or no color change) as well as its timing characteristics (timing of the color change onset if it occurred) and made appropriate adjustments in his vocal coordination with the upper-airway articulators. He accurately evaluated the correctness of his productions independent of the clinician and benefited from knowledge of performance derived from prior visual feedback. This is in contrast to the study by Pratt et al. (1993) who found that young children (5 years of age), particularly those with poor speech skills, experienced difficulty using computer-based visual feedback independently and needed a clinician to facilitate their use of the fedback information. It is likely that children need to reach certain cognitive and speech levels before they can independently make adjustments in their speech in response to computer-based feedback. It also is likely that some forms of computer-based feedback are more conducive to independent corrections than others. The feedback in the present study was used in a binary fashion, making it easier for the child to judge the correctness of his productions. More graded feedback may have been less interpretable, although a more accurate reflection of consonant voicing in ongoing speech.

Despite evidence of a treatment effect and generalization at the CV-syllable level (and possibly single words), generalization to connected speech was not evidenced. A number of factors could have influenced this outcome. The child demonstrated skill learning with a 100% feedback schedule; however, it is possible that he would have improved with fewer sessions and generalized to more complex levels of production if the feedback schedule had been more variable. Another consideration is the simplicity of the feedback. If the feedback had been judged on a more continuous basis or included instructions for correction, the task demands would have increased in difficulty, which likely would have been associated with a longer duration for skill learning but greater generalization. Difficulty level also could have been modified by varying the syllable structure or linguistic level at which training occurred. Restricting training to just CV syllables greatly simplified the learning task, whereas introducing the target sound in numerous contexts in irregular schedules would likely have resulted in more applicability and, therefore, greater generalization of the skill. However, as indicated previously, little is known about what therapeutic parameters best promote speech production skills in children with hearing loss. Moreover, information on the factors that influence the development of audition and speech in infants and young children with hearing loss is very limited and unable to guide our treatment approaches.

Alternatively, models of normal speech development may be useful when considering current treatment approaches. Recent auditory models of speech development and production provide theoretical support for auditory-based treatment approaches although they do not discount the impact of other sensory modalities (Callan, Kent, Guenther, & Vorperian, 2000; Guenther, 1995; Perkell et al., 1997; Perkell et al., 1995). These theoretical models contend that the formation of a multi-dimensional auditory space provides the basis for the development of speech in infancy, and that speech production is learned by mapping the orosensory information to acoustic-based targets or regions within the acoustic space. Given this perspective, speech treatment would be most effective through audition if the auditory system can provide sufficient sensory input and feedback. However, if hearing is limited to the extent that children cannot form a complete internal auditory space, they may need to use other types of sensory information to develop an analogous internal representation. It is not likely that current computer-based feedback systems can provide sufficient information to stimulate the development of such a representation, but they might provide complementary or supplementary information to augment the compromised auditory sensory input. Future systems may provide more complete cues, as well as information that can be mapped to speech acoustics transparently. Future systems also may become sufficiently flexible so that speech information can be tailored to the perceptual and speech production needs of individual children. Tailoring speech treatment approaches is consistent with the fact that children with hearing loss are heterogeneous in their auditory and speech characteristics and are unlikely to respond best to any one treatment approach. As such, efforts should be made to match child characteristics with optimal treatment characteristics.

Another issue associated with computer-based training methods and the need to individualize treatment is whether they motivate children sufficiently to promote active involvement in the treatment over multiple sessions. This is a particularly relevant issue for children with hearing loss because they often require long-term treatment programs. The child in the present study did not appear to be more or less motivated by the computer-based treatment compared to his previous clinician-based treatment approach. His motivation may have been affected by the presence of the clinician during the training or the fact that he missed some sessions due to scheduling conflicts and illness. He did not complain about the training and assessment tasks, despite their being repetitive and simple. Cooperation did not diminish substantively over the course of the study, although the child did not express interest in the tasks or a desire to work on them independently or more frequently. It should be acknowledged that other children may have needed more interesting tasks, greater variation in the task characteristics, or supplementary reinforcement to remain motivated over multiple sessions. Finding the right combination of feedback, training methods, and interest for individual children with hearing loss is a challenge for the developers of computer-based speech training programs and clinicians alike.

REFERENCES

Bakeman, R., & Robinson, B.F. (1997). When Ns do not justify the means: Small samples and scientific conclusions. In L.B. Adamson & M.A. Romski (Eds.), *Communication and language acquisition: Discoveries from atypical development* (pp. 50-72). Baltimore: Paul H. Brookes Publishing.

Bernstein, L. (1989). Computer-based speech training for profoundly hearing-impaired children: Some design considerations [Monograph]. *The Volta Review, 19,* 19-28.

Callan, D.E., Kent, R.D., Guenther, F.H., & Vorperian, H.K. (2000). An auditory-feedback-based neural network model of speech production that is robust to developmental changes in the size and shape of the articulatory system. *Journal of Speech, Language, and Hearing Research, 43,* 721-736.

Carhart, R. (1947). Conservation of speech. In H. Davis (Ed.), *Hearing and deafness, a guide for the laymen* (pp. 300-317). New York: Murray Hill Books.

Carhart, R. (1963). Conservation of speech. In H. Davis & S.R. Silverman (Eds.), *Hearing and deafness* (pp. 387-402). New York: Holt, Rinehart & Winston.

Carr, J. (1953). An investigation of the spontaneous speech sounds of five-year-old deaf-born children. *Journal of Speech and Hearing Disorders, 18,* 22-29.

Colarusso, R.P., & Hammill, D.D. (1972). *Motor-Free Visual Perception Test.* Novato, CA: Academic Therapy Publications.

Cytel Software Corporation. (2001). *StatXact* (Version 5) [Computer software]. Cambridge, MA.

Dagenais, P., Critz-Crosby, P., Fletcher, S., & McCutcheon, M. (1994). Comparing abilities of children with profound hearing impairments to learn consonants using electropalatography or traditional aural-oral techniques. *Journal of Speech and Hearing Research, 37,* 687-699.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test – Revised.* Circle Pines, MN: American Guidance Service.

Elfenbein, J., Hardin-Jones, M., & Davis, J. (1994). Oral communication skills of children who are hard of hearing. *Journal of Speech and Hearing Research, 37,* 216-226.

Ertmer, D.J., Stark, R.E., & Karlan, G.R. (1996). Eliciting prespeech vocalizations in a young child with profound hearing impairment: Usefulness of real-time spectrographic speech displays. *American Journal of Speech-Language Pathology, 4,* 33-38.

Ertmer, D.J., & Maki, J.E. (2000). A comparison of speech training methods with deaf adolescents: Spectrographic versus noninstrumental instruction. *Journal of Speech, Language, and Hearing Research, 43,* 1509-1523.

Estabrooks, W. (1994). *Auditory-verbal therapy.* Washington, DC: Alexander Graham Bell Association for the Deaf.

Fairbanks, G. (1960). *Voice and articulation drillbook.* New York: Harper & Brothers.

Fletcher, S., & Hasegawa, A. (1983). Speech modification by a deaf child through dynamic orometric modeling and feedback. *Journal of Speech and Hearing Disorders, 48,* 179-185.

Fletcher, S., & Higgins, J. (1980). Performance of children with severe to profound auditory impairment in instrumentally guided reduction of nasal resonance. *Journal of Speech and Hearing Disorders, 45,* 181-194.

Fletcher, S.G. (1978). *Diagnosing speech disorders from cleft palate.* New York: Grune & Stratton.

Fowler, C.A., Brown, J.M., & Mann, V.A. (2000). Contrast effects do not underlie effects of preceding liquids on stop-consonant identification by humans. *Journal of Experimental Psychology: Human Perception & Performance, 26* (3), 877-888.

Freeman, E. (2001). Unilateral spatial neglect: New treatment approaches with potential application to occupational therapy. *American Journal of Occupational Therapy, 55,* 401-408.

Gold, T. (1980). Speech production in hearing-impaired children. *Journal of Communication Disorders, 49,* 58-64.

Goldman, R. & Fristoe, M. (1986). *Goldman-Fristoe Test of Articulation.* Circle Pines, MN: American Guidance Service.

Guenther, F.H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102,* 594-621.

Heider, F., Heider, G., & Sykes, J. (1940). A study of the spontaneous vocalizations of fourteen deaf children. *The Volta Review, 43,* 10-14.

Holt, L.L., & Lotto, A. (2002). Behavioral examinations of the level of auditory processing of speech context effects. *Hearing Research, 167,* 156-169.

Holt, L.L., Lotto, A.J., & Kluender, K.R. (2000). Neighboring spectral content influences vowel identification. *Journal of the Acoustical Society of America, 108,* 710-722.

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

International Business Machines Corporation. (1992). *IBM SpeechViewer Series II* [Computer Software]. New York.

Jayaratne, S., Tripodi, T., & Talsma, E. (1988). The comparative analysis and aggregation of single-case data. *Journal of Applied Behavioral Science, 24,* 119-128.

Kishon-Rabin, L., Taitelbaum, R., Tobin, Y., & Hildesheimer, M. (1999). The effect of partially restored hearing on speech production of postlingually deafened adults with multichannel cochlear implants. *Journal of the Acoustical Society of America, 106,* 2843-2857.

Lane, H., Wozniak, J., Matthies, M., Svirsky, M., & Perkell, J. (1995). Phonemic resetting versus postural adjustments in the speech of cochlear implant users: An exploration of voice-onset-time. *Journal of the Acoustical Society of America, 98,* 3096-3106.

Levitt, H., Youdelman, K., & Head, J. (1990). *Fundamental Speech Skills Test.* Englewood, CO:

Resource Point.
Ling, D. (1976). *Speech and the hearing-impaired child: Theory and practice.* Washington, DC: A.G. Bell Association for the Deaf.
Ling, D. (2002). *Speech and the hearing-impaired child: Theory and practice* (2nd ed.). Washington, DC: A.G. Bell Association for the Deaf.
Lintern, G., Roscoe, S.N., & Sivier, J. (1990). Display principles, control dynamics, and environmental factors in pilot training and transfer. *Human Factors, 32,* 299-317.
Magner, M. (1979). Techniques of teaching. In L.E. Connor (Ed.), *Speech for the deaf child: Knowledge and use* (pp. 254-264). Washington, DC: A.G. Bell Association for the Deaf.
Markides, A. (1970). The speech of deaf and partially hearing children with special reference to factors affecting intelligibility. *British Journal of Disorders of Communication, 5,* 126-140.
McGarr, N., & Lofquist, A. (1982). Obstruent production by hearing-impaired children speakers: Interarticulator timing and acoustics. *Journal of the Acoustical Society of America, 72,* 34-42.
Millin, J. (1971). Therapy for reduction of continuous phonation in the hard-of-hearing population. *Journal of Speech and Hearing Disorders, 36,* 496-498.
Monsen, R. (1976). The production of English stop consonants in the speech of deaf children. *Journal of Phonetics, 4,* 29-42.
Monsen, R.B., Moog, J.S., & Geers, A.E. (1988). *CID Picture SPINE.* St. Louis: Central Institute for the Deaf.
Moog, J.S., & Geers, A.E. (1990). *Early speech perception test for profoundly hearing-impaired children.* St. Louis: Central Institute for the Deaf.
Nober, E.H. (1967). Articulation of the deaf. *Exceptional Child, 33,* 611-621.
Office of the Los Angeles County Superintendent of Schools. (1980). *Test of Auditory Comprehension.* North Hollywood, CA: Foreworks.
Oller, D.K., & Kelly, C.A. (1974). Phonological substitution processes of a hard-of-hearing child. *Journal of Speech and Hearing Disorders, 39,* 221-283.
Osberger, M.J., & McGarr, N.S. (1982). Speech production characteristics of the hearing impaired. *Speech and Language, 8,* 221-283.
Osberger, M.J., Moeller, M., Kroese, J., & Lippmann, R. (1981). Computer-assisted speech training for the hearing impaired. *Journal of the Academy of Rehabilitative Audiology, 14,* 145-158.
Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22,* 227-250.
Perkell, J.S., Matthies, M.L., Svirsky, M.A., & Jordan, M.I. (1995). Goal-based speech motor control: A theoretical framework and some preliminary data. *Journal of Phonetics, 23,* 23-35.
Pollack, D. (1985). *Educational audiology for the limited-hearing infant and preschooler* (2nd ed.). Springfield, IL: Charles C. Thomas.
Pratt, S.R., Heintzelman, A., & Deming, S. (1993). The efficacy of using the IBM SpeechViewer Vowel Accuracy Module to treat young children with hearing impairment. *Journal of Speech and Hearing Research, 36,* 1063-1074.
Pratt, S.R., & Hricisak, I. (1994). Commercially available computer-based speech feedback systems. *Journal of the Academy of Rehabilitative Audiology, 27,* 89-106.
Pratt, S.R., & Tye-Murray, N. (1997). Speech impairment secondary to hearing loss. In M.R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 345-387). New York: Thieme.
Proteau, L., & Cournoyer, L. (1990). Vision of the stylus in a manual aiming task: The effects of practice. *Quarterly Journal of Experimental Psychology, 42B,* 811-828.
Proteau, L., Marteniuk, R.G., Girouard, Y., & Dugas, C. (1987). On the type of information used to control and learn an aiming movement after moderate and extensive training. *Human Movement Science, 6*, 181-199.
Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young

children. *Journal of Speech and Hearing Disorders, 52,* 271-277.

Schaechter, J.D., Kraft, E., Hilliard, T.S., Dijkhuizen, R.M., Benner, T., Finklestein, S.P., Rosen, B.R., & Cramer, S.C. (2002). Motor recovery and cortical reorganization after constraint-induced movement therapy in stroke patients: A preliminary study. *Neurorehabilitation and Neural Repair, 16,* 326-338.

Silverman, S. (1971). The education of deaf children. In L.E. Travis (Ed.), *Handbook of speech and language pathology* (pp. 399-430). Englewood Cliffs, NJ: Prentice-Hall.

Smith, C.R. (1975). Residual hearing and speech production in the deaf. *Journal of Speech and Hearing Research, 19,* 795-881.

Tobey, E., Geers, A., & Brenner, C. (1994). Speech production results: Speech feature acquisition [Monograph]. *The Volta Review, 96,* 109-129.

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

Van der Merwe, A. (1997). A theoretical framework for the characterization of pathological speech sensorimotor control. In M.R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 1-25). New York: Thieme.

Wallace, G., & Hammill, D. (1994). *Comprehensive Receptive and Expressive Vocabulary Test.* Dallas, TX: PRO-ED.

Watson, C., & Kewley-Port, D. (1989). Advances in computer-based speech training: Aids for the profoundly hearing impaired [Monograph]. *The Volta Review, 91,* 29-46.

Weinstein, C.J., & Schmidt, R.A. (1990). Reducing frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16,* 677-691.

West, J., & Weber, J. (1973). A phonological analysis of the spontaneous language of a four-year-old hard-of-hearing child. *Journal of Speech and Hearing Disorders, 38,* 25-35.

Wulf, G., Shea, C.H., & Matschiner, S. (1998). Frequent feedback enhances complex skill learning. *Journal of Motor Behavior, 30,* 180-192.

Youdelman, K., MacEachron, M., & Behrman, A.M. (1988). Visual and tactile sensory aids: Integration into an ongoing speech training program. *The Volta Review, 90,* 197-207.