Journal of Speech and Hearing Research, Volume 37, 347–357, April 1994

Chapman F27

Speech Perception Training Can Facilitate Sound Production Learning

Susan Rvachew
Alberta Children's Hospital
Calgary, Alberta, Canada

This study examined the role of speech perception training in the correction of phonological errors. Twenty-seven preschoolers with phonological impairment who misarticulated /ʃ/ were randomly assigned to one of three groups: Group 1 children listened to a variety of correctly and incorrectly produced versions of the word "shoe"; Group 2 children listened to the words "shoe" and "moo"; Group 3 children listened to the words "cat" and "Pete." A computer game was used to provide reinforcement for correct identification of the words. All children received the same traditional sound production training program for correction of their /ʃ/ error, concurrently with speech perception training, during six weekly treatment sessions. On post-testing, Group 1 and 2 children demonstrated a superior ability to articulate the target sound in comparison to Group 3 children. The results are interpreted in relation to previous research on this topic.

KEY WORDS: speech perception, phonological disorders, treatment

---

The role of speech perception training in the treatment of phonological disorders has been a controversial issue throughout the history of the practice of speech-language pathology. Early approaches to articulation treatment involved an extended period of "ear training" (Van Riper, 1963) that preceded the introduction of sound production activities. Ear training was comprised of a number of loosely described activities that included asking the child to identify correct and incorrect versions of the target sound, as produced by the clinician. These activities were directed at helping the child recognize the distinctive elements that defined the contrast between the target sound and its substitution. In addition, it was expected that these procedures would lead to an internalized target that would guide the child during the production phase of treatment.

Ear training separated from sound production practice has been criticized by proponents of perceptual-motor theories of phonetic skill learning. These models focus on the integration of sensory and perceptual processes during sound production learning. The influence of these theories led many authors to advise speech-language pathologists to avoid ear training in the form of interpersonal speech discrimination training, and focus instead on intrapersonal speech discrimination training (i.e., self-monitoring) during sound production activities (Aungst & Frick, 1964; McReynolds, Kohn, & Williams, 1975; Seymour, Baran, & Peaper, 1981; Shelton, Johnson, Ruscello, & Arndt, 1978; Shelton & McReynolds, 1979; Williams & McReynolds, 1975).

More recent linguistic treatment approaches are directed at restructuring the child's system of underlying phonological contrasts and phonological rules (Winitz, 1985). These approaches often, although not necessarily, involve some form of speech perception training (e.g., some minimal pairs training procedures). Hodson and Paden (1983) recommend "auditory bombardment," in which the child listens to a list of words containing the target phoneme, presented by the clinician using an auditory trainer to provide amplification.

0022-4685/94/3702-0347$01.00/0

Despite the long history of this controversy, there is surprisingly little research that directly examines the efficacy of interpersonal speech perception training in the treatment of phonological errors. Winitz (1969, 1975) advocated an approach to sound discrimination training that involved presenting children with recorded, natural speech words that contrasted with a given distinctive feature, followed by reinforcement for correct identifications of the stimuli. This approach was used to modify normal children's perception and production of unusual consonant clusters in a series of studies. However, an effort to show that this approach could lead to improved /r/ articulation by children with delayed phonological skills was unsuccessful (Winitz & Bellerose, 1967). Winitz and Bellerose (1967) concluded that their training procedure was not helpful during the response acquisition phase of treatment, but suggested that speech discrimination training may be useful during the response association phase, if the motor response is available to the child.

Williams and McReynolds (1975) stated that production training affects both perception and production performance, while speech discrimination training affects perception ability with no accompanying improvement in production ability. However, in their study, the training contrasted sounds that did not reflect the child's actual substitution errors. Shelton et al. (1978) compared parent-administered ear training to a reading-talking treatment and a no-treatment control. No differences between the groups were observed for any measures of pre-post gain in auditory or articulatory ability.

More recently, Jamieson and Rvachew (1992) showed that speech sound identification training can facilitate sound production learning for some children with delayed phonological skills. This study is similar to the studies mentioned above in that an identification procedure was employed for speech perception training, and no explicit sound production training was provided. However, there are some significant procedural differences: (a) the training contrasted the target phoneme with the child's substitution for the target; (b) the stimuli were synthesized to ensure that the child's identification performance was based on the critical acoustic cues that define the contrast, rather than on noncritical acoustic or visual information that may be found in natural speech presentations of training stimuli; (c) a single-subject design was used so that performance could be compared for children who did and did not demonstrate pretreatment difficulty with perception of the target contrast; and (d) all subjects were able to imitate the target sound correctly some of the time.

The treatment was shown to be effective in facilitating sound production learning in three cases where the child demonstrated pretreatment difficulties with both the perception and production of the target sound, and where the child successfully learned the perceptual task. In one case the sound identification training procedure did not specifically target the child's error sound (the child lateralized /ʃ/ but received training for the /ʃ/-/s/ contrast for which he demonstrated no pretreatment perceptual difficulties). As expected, there was no observed effect of the treatment on sound production learning in this case. Another child demonstrated pretreatment difficulties with both the perception and production of /s/ but did not learn to identify /s/ and /ʃ/ reliably, and again, no treatment effect was observed.

Although Jamieson and Rvachew (1992) successfully demonstrated that speech perception training can facilitate sound production learning for some children, the procedures used had limited clinical utility, largely because the use of synthetic speech made it difficult to expand the program to cover all of the speech errors typically produced by children with phonological impairment. In addition, it was difficult to maintain the children's attention to the training task, as implemented in this study.

The purpose of the study reported here is to further demonstrate that speech perception training can facilitate sound production learning by children who misarticulate the /ʃ/ sound. In this study a computer game format was used to teach a larger sample of children to identify natural speech tokens as belonging to either the "sh" or not "sh" categories.

Method

Subjects

The subjects were 21 boys and 6 girls ranging in age from 42 to 66 months. All children were referred by a speech-language pathologist who had previously assessed the child and diagnosed a significant phonological delay requiring treatment. The referring speech-language pathologist also determined that the child had normal oral structure and function and that there were no known etiological conditions such as sensory-neural hearing impairment, cleft palate, mental retardation, or neurological impairment. Fourteen of the children did have significant histories of otitis media, however (i.e., the parent reported that the child had had at least one ear infection in every year since birth, or three or more ear infections in any 1 year). Each child was unstimulable for the /ʃ/ sound during pretesting. Stimulability was assessed by asking the child to imitate /ʃ/ in isolation five times. An auditory/visual model was provided along with verbal instructions about tongue retraction and lip protrusion.

The children were randomly assigned to groups.[1] The final group sizes were 10, 9, and 8 children for Group 1, 2, and 3 respectively.[2] Two children had mildly delayed phonological skills, 18 children were moderately delayed, and 7 children had severe deficits, as determined by the Computer Analysis of Phonological Processes (Hodson, 1985). Table 1 shows

[1]Where possible, the children were matched with respect to CA, RLA, ELA, and severity of phonological impairment. Four triads of children who matched on these variables were found. It proved extremely difficult to match children closely on all four variables, however, and consequently this approach was abandoned in favor of an independent groups random assignment procedure. This change in procedure does not threaten the validity of the research design because group assignment was randomly determined for every child and because there were no significant differences between groups with respect to these subject characteristics overall. It did reduce the power of the statistical tests considerably, however (see Note 6).

[2]Eight children were referred but either did not start or did not complete the six treatment sessions. Three children achieved stimulability for /ʃ/ at the word level before beginning treatment, one child was judged to have abnormal oral structure by the author (a marked open-bite), two children refused to complete the program, one child got an ear infection, and another was withdrawn when his mother was hospitalized.

TABLE 1. Group size (N) and mean chronological age (CA), receptive language age (RLA), and expressive language age (ELA), by group.

Subject variable	Group 1 shoe/Xshoe	Group 2 shoe/moo	Group 3 cat/Pete
N	10	9	8
CA	53.4	53.6	51.5
RLA	59.9	66.3	63.7
ELA	50.6	50.4	47.8

Note. All ages are expressed in months.

the mean chronological age (CA), receptive language age (RLA) and expressive language age (ELA) for each of the three groups. There were no significant differences between groups with respect to CA (p = .4), RLA (p = .2), or ELA (p = .8), as determined by randomization tests.

Stimuli

Six sets of stimuli were developed, two practice sets to teach the children the task, one testing set, and three sets for the three different training conditions. The word pairs represented by each of the six sets are as follows: (a) *sheet-cat* (Practice Set 1); (b) *sheet-meat* (Practice Set 2); (c) *sheet-Xsheet* (Word Identification Test, where *Xsheet* refers to misarticulated versions of the word *sheet*); (d) *shoe-Xshoe* (Group 1 Training Set, where *Xshoe* refers to misarticulated versions of the word *shoe*); (e) *shoe-moo* (Group 2 Training Set); and, (f) *cat-Pete* (Group 3 Training Set). The procedures used in the creation of these stimulus sets are described below.

The auditory stimuli used for speech perception testing and training were recorded and digitized at a sampling frequency of 20 kHz, using the Canadian Speech Research Environment (CSRE; Jamieson, Nearey, and Ramji, 1989) and the following hardware: an AST Premium 386C computer, DT2821 D/A, A/D board, TTE 411AFS amplifier and antialiasing filter, and a Sony F-V5T dynamic microphone. The recordings were made in a quiet room. The microphone was suspended approximately 2 to 3 inches in front of each talker's mouth, although the children tended to move closer to the microphone when speaking.

Children with delayed phonological skills and normal speaking adults and children were recruited as talkers. Each talker was shown a picture of a *shoe* and then asked "What is this?" If the talker responded with an incorrect object name, he or she was told, "Say *shoe*." This same procedure was used to obtain a second recording of the word *shoe* and two recordings of the word *sheet* from each talker. Talkers were encouraged to use normal voice quality and loudness, and the amplifier gain control was adjusted in an effort to obtain recordings which were characterized by the highest possible amplitude without peak clipping.

All of the recorded stimuli were edited to remove any extraneous speech, noise, or silence occurring before and after the recorded word. Three speech-language pathologists listened to and transcribed each stimulus. In addition, an acoustic description of each stimulus was obtained, using procedures which are described in the Appendix.

Stimuli were removed from the pool under the following conditions: (a) no agreement among three transcriptions of the word (disregarding disagreements about the voicing status of initial stop segments); (b) signal to noise ratio too low on the basis of subjective perceptual judgment; (c) extraneous clicks or pops in the signal; (d) nonstandard or incorrect vowel; and (e) unusual voice quality. In order to ensure that no more than one version of each word from any one talker remained in the pool, the best version of each pair was selected on the basis of subjective perceptual judgment, and the remaining member of the pair was removed from the stimulus pool. Finally, 40 versions of the word *shoe* were selected for use in the Group 1 speech perception training program, and a parallel set of 40 versions of the word *sheet* were selected for use in the Word Identification Test.

The Appendix contains both a phonetic and an acoustic characterization of each stimulus included in the Group 1 stimulus set. Those stimuli which were judged to contain a misarticulated /ʃ/ sound were grouped according to the manner of the substituted sound (i.e., stop, affricate, fricative, distortion). To some extent the phonetic transcription provides an inadequate characterization of these stimuli. The majority of the incorrect productions were obtained from children who were intending to produce the word *shoe*. Most of the incorrect productions could be transcribed in terms of conventional English phoneme classes (e.g., [tu], [su]). However, the children's productions were not necessarily good quality exemplars of either the intended category (i.e., /ʃ/) or the transcribed category (i.e., /t/, /s/). The entire stimulus set (including correct and incorrect tokens produced by children or adults) is characterized by a large degree of variability in the frequency, duration, and amplitude characteristics of the consonant and vowel portions of the word. It was expected that this degree of variability in stimulus parameters would serve to: (a) present each subject with a contrast between correctly produced /ʃ/ sounds and the subject's typical substitutions for this sound; and (b) focus the child's attention on the critical, stable characteristics of the /ʃ/ sound, and away from irrelevant variation in the acoustic properties of the stimuli.

Four additional sets of auditory stimuli were developed using the procedures described above. The Group 2 stimulus set consisted of one correct version each of the words *shoe* and *moo* (produced by an adult). Unlike the Group 1 stimulus set, this word pair does not represent a contrast which parallels the subjects' errors, as no subject substituted [m] for any fricative sound. However, the single token of *shoe* contained in this set does represent a clear, prototypical exemplar of the /ʃ/ category.

The Group 3 stimulus set consisted of one correct version each of the words *cat* and *Pete* (also produced by an adult talker). This set was used to control for any possible beneficial effects of the computer game itself, while not providing actual speech perception training. This word pair does not contain any fricative sound, and correct responding does not require discrimination of a particular phoneme contrast.

Two pairs of stimuli for use in practice trials were also recorded (*sheet* vs. *cat* and *sheet* vs. *meat*).

TABLE 2. Stimulus words and number of unique tokens presented during each practice, testing, and training block.

Block	Target stimuli Word	Tokens[a]	Foil stimuli Word	Tokens
Practice				
1	sheet	1	cat	1
2	sheet	1	meat	1
Word identification test				
1–4	sheet	5	Xsheet (all)	5
Group 1 training program				
1	shoe	5	Xshoe (stop)	5
2	shoe	5	Xshoe (affricate)	5
3	shoe	5	Xshoe (fricative)	5
4	shoe	5	Xshoe (distortion)	5
5–8	shoe	5	Xshoe (all)	5
Group 2 training program				
1	shoe	1	moo	1
Group 3 training program				
1	cat	1	Pete	1

Note. [a]Misarticulated versions of the words *sheet* and *shoe* are denoted Xsheet and Xshoe, respectively. Each block contains 10 trials with the number of repetitions of each unique token adjusted to ensure five target and five foil stimuli per block.

The experiment generator module of CSRE was used to organize these stimuli into blocks of 10 stimuli each. The number of presentations of each stimulus within a block was adjusted according to the number of unique tokens of target and foil stimuli to ensure five target stimuli and five foil stimuli per block. Table 2 shows the number of blocks and the stimuli presented per block for each of the various programs (i.e., practice, testing, and training blocks). Order of stimulus presentation was random within each block.

In addition to the auditory stimuli a number of graphic stimuli were developed to provide feedback, or to represent response alternatives. The experiment controller module of CSRE was used to coordinate presentation of auditory stimuli, response alternative pictures, and feedback pictures, and to record responses to each auditory stimulus presentation.

Procedure

Each child attended a pretest session, followed by six treatment sessions, and a post-test session. Each session lasted approximately 45 minutes. These sessions were generally scheduled weekly but occasionally the pretest session and the first treatment session, or the last treatment session and the post-test session, occurred during the same week. On average, the children required 8.7 weeks to complete the eight sessions (range = 6 to 11 weeks). There were no significant differences between groups with respect to time required to complete the program ($p = .4$, by randomization test).

Pretest session. During the pretest session each child received the following tests: (a) Reynell Developmental Language Scales (Reynell, 1977), (b) Computerized Analysis of Phonological Processes (Hodson, 1985), (c) a hearing screening, and (d) the *sheet-Xsheet* Word Identification Test.

TABLE 3. Pretreatment substitutions for /ʃ/ during spontaneous production, and pretreatment percent correct scores on each subtest of the *Sheet-XSheet* Identification Test.

Subject	Production test[a]	Pretreatment identification test[b] 1	2	3	4	5
1	θ	90	100	70	20	80
2	h,-	95	100	20	20	20
3	g^{ne}	85	80	80	40	40
4	t	98	100	70	50	40
5	-,θ	58	50	70	40	60
6	t,d	90	20	0	0	0
7	θ	85	100	90	90	90
8	d	75	60	50	60	70
9	s^D,h,-	83	100	40	80	40
10	s	100	100	30	20	20
11	θ	100	90	30	20	10
12	θ,ts	98	100	60	70	30
13	s^D	85	100	90	80	80
14	s,ts	73	80	60	80	60
15	θ	100	100	90	80	70
16	t	100	100	40	90	50
17	s,t,θ	98	100	50	40	60
18	θ	85	70	80	50	60
19	θ	100	100	80	70	70
20	d,t	98	100	50	70	40
21	d,d^{ne}	98	30	20	10	20
22	s	100	100	80	50	60
23	-,h	100	100	80	50	60
24	θ,s	100	100	70	90	50
25[c]	s					
26	dj,j	100	100	70	40	40
27	t,n	100	100	90	100	60

Note. [a]Superscript "ne" refers to nasal emission. Superscript "D" refers to distortion. Dash "-" refers to omission of the initial consonant.
[b]Subtest 1 contains 20 correct versions of the word *sheet.* The remaining subtests contain five incorrectly produced versions of the word *sheet* representing the following error types: subtest 2 = stop substitutions, subtest 3 = affricate substitutions, subtest 4 = fricative substitutions, and subtest 5 = distortions. Each stimulus was presented twice and all 40 stimuli were presented in random order.
[c]This subject's total score was 59%. The raw data were lost before the subtest scores could be recorded.

The *sheet-Xsheet* Word Identification Test was administered as described below. The auditory stimuli and the response alternatives (a picture of a sheet and a picture of a large X) were presented using the hardware and software described in the stimulus section. The auditory stimuli were presented over AKG-K240 headphones. The children were instructed to point to the picture of the sheet when they heard the word *sheet* and to point to the X if the word was "not *sheet*" or "said wrong." The examiner then used the mouse to "click" the picture selected by the child. The computer recorded the child's response and then presented the next auditory stimulus. The two practice blocks (see Table 2) were presented with feedback until the child achieved scores of 90% correct, and then the *sheet-Xsheet* stimuli were presented. No feedback about correctness of response was provided during administration of the *sheet-Xsheet* test. Each of the blocks 1 through 4 (see Table 2) was presented twice, with a different random ordering of stimuli for each block. The children's pretraining performance on this test is shown in Table 3.

In addition to the tests listed above, each child provided a short speech sample by naming the objects shoe, shampoo, shirt, shovel, and shell. These single-word productions were digitized at a sampling frequency of 20 kHz. The author scored each word for correctness of the /ʃ/ production while the child was naming the object. All five words were misarticulated by every subject during the pretest session. The children's pretest substitutions for the /ʃ/ sound are shown in Table 3.

Treatment sessions. During each treatment session the child received 60 perception training trials, using the same procedure for all children regardless of group assignment. The child was placed in front of a monitor which displayed a feedback picture (for example, a duck pond) and two response alternatives (a picture of the target word and a large X). The child was instructed to listen to the words presented over the headphones (the average presentation level was 77 dB SPL), and to point to the target picture if the target word was heard, and to point to the X if the word was not the target or if it was "said wrong." The experimenter then "clicked" the picture selected by the child. Correct responses were rewarded by a change in the feedback picture (e.g., a new duck might be added to the pond). Incorrect responses were recognized by the word *wrong* presented over the headphones. Perception training was administered by a research assistant for most children, and by the author for those children who were scheduled when the assistant was unavailable.

Group 1 children were presented with blocks 1 through 8 (see Table 2) in order. Each of the blocks 1 through 4 was repeated until the child reached a criterion of 80% correct responding over three consecutive blocks, or until the block had been repeated six times. At this time, the next block in the series was presented. When presentation of the first four blocks was complete, blocks 5 through 8 were presented in order during all remaining treatment sessions.

Group 2 children were presented with six blocks of *shoe-moo* stimuli, and Group 3 children responded to six blocks of *cat-Pete* stimuli during each treatment session.

Every child also received 60 production training trials during each treatment session. Production training was administered following completion of 60 perception training trials. The author administered the same production training program to all children, regardless of group placement.[3] A traditional approach was employed,[4] working through the levels: (a) isolation, (b) imitated syllables, (c) imitated words, (d) spontaneous words, (e) imitated phrases, (f) spontaneous phrases, (g) imitated sentences, (h) spontaneous sentences, and (i) carryover to conversational speech.[5] A new level was introduced when the child achieved 90% correct responding over three consecutive blocks of 10 trials, or 100% correct over one block of 10 trials. Phonetic placement was used if necessary to help the children learn correct /ʃ/ production in isolation. However, the child did not reach criterion at Level 1 unless he or she was able to imitate /ʃ/ with 90% accuracy without phonetic placement.

Post-test session. During post-testing the *sheet-Xsheet* Word Identification Test was repeated and a second object-naming speech sample was obtained. In addition, each child produced a single isolated /ʃ/ which was digitized and submitted to centroid analysis (see Appendix).

Reliability

A speech pathology student who had completed course work in phonetic transcription judged the correctness of each stimulus in the Group 1 (*shoe-Xshoe*) training program. In addition, all object-naming speech samples were rescored for correctness of the intended /ʃ/ segment by the student or by a speech-language pathologist. Interjudge agreement for the Group 1 training stimuli was 97.5%, and agreement for the speech sample words was 94%. Correctness of the isolated /ʃ/ samples proved somewhat more difficult to determine and consequently consensus judgments were obtained by using the judgment provided by two of the three speech-language pathologists who listened to these samples. All three listeners agreed for 81% of these stimuli, while pair-wise agreement between the author and the other two listeners was 85 and 88%. All judges except the author were blind to the child's group placement or the time of recording (i.e., pretest or post-test).

Results

Word Identification Test

As noted above, each stimulus on the *sheet-Xsheet* Word Identification Test was presented twice to each subject, both during pretesting and during post-testing. Agreement between responses to pairs of stimuli was calculated for each pretest and post-test. In general, the children were highly reliable in their responses to these stimuli. The mean agreement was 80.92% for the pretreatment test and 83.06% for the post-treatment test.

Overall, performance on both the pretreatment and post-treatment Word Identification Tests was good. The mean score for all children on the pretreatment test was 78% (range 51% to 94% correct), and the mean score on the post-treatment test was 84% (range 53% to 95% correct). Most children had little or no difficulty with the correctly produced stimuli (92% of these stimuli were identified correctly during the pretreatment test).

[3]An effort was made to ensure that I was blind to the child's group placement during production training but this proved impossible for every child. In some cases I was required to administer the perception training due to the unavailability of the research assistant. In other cases, the child revealed his or her group placement when talking about the perception training procedure.

[4]I am not advocating the clinical application of the production training procedures used in this study. The traditional approach was chosen because it allowed me to provide the same treatment program to all children. In normal clinical practice, production training would more likely target multiple phonemes, and target selection would be individualized based on the child's overall phonological pattern. The speech perception training procedures that are evaluated here could be integrated with these more typical production training procedures (see Discussion section for further explanation).

[5]The syllables targeted during production training all contained a prevocalic /ʃ/ followed by a variety of monopthongal and dipthongal vowels. The words targeted were shirt, ship, shopping, sheep, shelf, shower, shell, shovel, shadow, and shooting.

TABLE 4. Mean post-treatment performance data by group, as measured by the *Sheet-XSheet* Identification Test and three indicators of /ʃ/ production ability.

Measure	Group 2 shoe/moo	Group 1 shoe/Xshoe	Group 3 cat/Pete
Identification test[a]	9.0	5.7	−2.4
Production level[b]	3.5	1.9	0.6
Speech sample[c]	2.0	1.0	0.1
Centroid[d]	3763	3719	4829

Note. See text for full description of the dependent variables.
[a]Difference score for pre- and post-treatment identification tests.
[b]Highest level attained during production training.
[c]Number of correctly produced [ʃ] sounds out of 5.
[d]The centroid of the [ʃ] produced in isolation (Hz).

The difference between post-treatment test raw score and pretreatment test raw score was calculated for each child. The mean difference scores on the perception tests are 9.0, 5.7, and −2.4 for Groups 1, 2, and 3 respectively, as shown in Table 4. These differences between groups with respect to pre-post improvement on the sheet-Xsheet test were found to be significantly different (p = .02, by randomization test[6]).

It must be noted that there is an outlier score of −25 among the matched Group 1 scores. It is clear that this child was not attending to task. The percentage agreement within pairs of stimuli associated with this score was only 55%, and performance for the correctly produced stimuli was only 50%. The presence of this outlier obscures the dramatic gains made by some of the Group 1 children (the remaining difference scores in this group range from −4 to +29).

The Group 2 difference scores range from +2 to +16. The Group 3 scores also contain an outlier score (−23), but this score is associated with 90% reliability. The remaining difference scores in this group range from −3 to +6. In order to exclude the Group 1 outlier from the analysis it was necessary also to exclude the scores obtained from the other children in this matched triad. Removing the triad of scores containing this outlier yields significantly different group means of 12.8, 5.6, and 0.57 for Groups 1, 2, and 3, respectively (pooled p = .02, by randomization test[6]).

Production Performance Data

Each child was assigned a number representing the highest level mastered during production training, where Level 1 corresponds to mastery at the isolated production level and Level 9 indicates that carryover to conversation is occurring (see the procedures described above). Many children struggled to succeed even at the isolation level. Twenty children reached criterion at Level 2 or less while only two children progressed to Level 9. On average, however, children in Groups 1 and 2 progressed further during production training than children in Group 3. As shown in Table 4, the mean level mastered during production training was 3.5 (range 0 to 9) for Group 1, 1.9 (range 0 to 9) for Group 2, and 0.6 (range 0 to 2) for Group 3.

As noted above, few children progressed beyond the imitated words phase of production training, and consequently, performance on the post-test object-naming speech sample was poor overall. Nineteen children produced all five words incorrectly with respect to the initial phoneme. As shown in Table 4, the mean number of correctly produced words was 2.0 (range 0 to 5) for Group 1, 1.0 (range 0 to 5) for Group 2, and 0.1 (range 0 to 1) for Group 3.

The centroid was determined for each isolated /ʃ/ production. The children's errors typically involved fronting (e.g., [t], [ts], [s], or [θ] substitutions). In contrast to the children's substitutions, correctly produced /ʃ/ sounds are generally associated with relatively low centroids (see Nittrouer, Studdert-Kennedy, & McGowan, 1989, for a discussion of centroid values for /s/ and /ʃ/). Consequently, it was expected that children in Groups 1 and 2 would produce isolated /ʃ/ with relatively low centroids, when compared to /ʃ/ produced by children in Group 3. This hypothesis is supported by the results as shown in Table 4. The mean centroids for Group 1, 2, and 3, are 3763 Hz, 3719 Hz, and 4829 Hz, respectively. Only one child in Group 3 produced a perceptually correct /ʃ/ sound while six children in each of groups 1 and 2 achieved a correct /ʃ/ in isolation.

The three sets of production data described above were transformed into z scores, summed, and then analyzed by randomization test.[6] The analysis of variance yielded a pooled p = .05. Multiple comparisons yielded the following pooled p values: Group 1 versus Group 3, p = .03; Group 2 versus Group 3, p = .01; and Group 1 versus Group 2, p = .52.

Discussion

The results of this study show that a computer-driven speech perception training program, provided concurrently with sound production training, can facilitate sound production learning by some children with delayed phonological development. This result is consistent with Jamieson and Rvachew (1992) but is discrepant with several other studies (Shelton et al., 1978; Williams & McReynolds, 1975; and Winitz & Bellerose, 1967). These studies differ with respect to a number of specific aspects of the perception training programs, each of which will be discussed individually, below.

Identification of Appropriate Subjects/Clients

Recent studies have shown that although some children who misarticulate have difficulty with the perception of spe-

[6]Randomization tests were selected for analysis of this data because these highly flexible, nonparametric tests allow for analysis of very small samples of data without conversion of scores to ranks. The matched data were analyzed using a randomization test which is analogous to a repeated measures ANOVA. The unmatched data were analyzed using a randomization test analogous to an independent groups one-way ANOVA. The p values obtained for the matched and unmatched samples were then pooled using the additive method of probability pooling to yield the final p value.

The Word Identification Test data were analyzed by determining the post-treatment–pretreatment difference score for each subject and then submitting these difference scores to the analysis described above. The speech production data were submitted to a multivariate analysis, as follows: A composite score was obtained by converting the "level mastered in treatment," "speech sample score," and the centroids to z scores and then summing the three z scores for each subject; the summed z scores were submitted to the analysis described above (c.f. Edgington, 1987).

cific speech sound contrasts, other children exhibit good perceptual ability for sounds that they misarticulate (Broen, Strange, Doyle, & Heller, 1983; Hoffman, Daniloff, Bengoa, & Schuckers, 1985; Rvachew & Jamieson, 1989). Jamieson and Rvachew (1992) demonstrated a positive effect of speech perception training on sound production ability only for those children who experienced difficulties with both the perception and production of the targeted sound contrast. Williams and McReynolds (1975), Shelton et al. (1978), and Winitz and Bellerose (1967) did not assess their subjects' pretreatment perception abilities with respect to any of the sounds targeted for perception training. Certainly the subjects in Winitz and Bellerose (1967) demonstrated marked intersubject variability in their perceptual responses during training. It is possible that many of the subjects in these studies were inappropriate clients for speech perception training.

Unfortunately, there are no clinical assessment tools available for adequately identifying those children who will benefit from speech perception training (see Locke, 1980a, for a discussion of the inadequacies of published speech perception tests). Synthetic speech continua have proven useful for research purposes and have potential for use in the clinical setting. The natural speech word identification test used in this study also has clinical potential. However, these stimulus sets are not currently applicable to clinical populations because they do not cover a broad enough range of speech sound contrasts, they are unavailable to clinicians, and they are not normed. In the meantime, as an alternative to indiscriminate application of this approach, clinicians must rely on intuitive clinical judgment or nonstandardized procedures (see Locke, 1980b) when attempting to select those children who are the most appropriate candidates for speech perception training.

Production Training

Another factor that will determine the child's degree of success with this treatment approach is the availability of the motor response required to produce the targeted sound. In this study all children received traditional production training utilizing imitative models, verbal instruction, and phonetic placement to stimulate production of the target in isolation. Those children who became stimulable for /ʃ/ production and who received one of the experimental speech perception treatments progressed further in treatment than control group children. However, some children in all three groups had difficulty achieving correct production of /ʃ/ in isolation and did not benefit from speech perception training.

It is not surprising that many of the children who participated in this study failed to make significant gains in production ability. On average they were younger than is typical for targeting of this sound, and none of the subjects were stimulable for /ʃ/. The results of Jamieson and Rvachew (1992) also suggest that speech perception training is most effective with children who are able to produce the necessary motor response. Winitz and Bellerose (1967) concluded that their speech perception training failed to improve their subjects' production accuracy because the required motor response was unavailable to the subjects, and the children did not receive any instruction to help them acquire this response. More recently, Saben and Costello Ingham (1991) found that a linguistically based minimal pairs treatment was ineffective without the addition of models and phonetic placement cues to help the child acquire the motor behavior associated with the production of targeted sounds.

Therefore, this approach to speech perception training seems most likely to succeed when traditional target selection criteria are applied, and when explicit production training is provided, either concurrently with or shortly following speech perception training.

To date, proponents of concurrent speech perception and production treatment have focused on intrapersonal speech perception training, and have explicitly proscribed interpersonal procedures (e.g., Seymour, Baran, & Peaper, 1981; Shelton & McReynolds, 1979). Although self-monitoring is undeniably critical to any child's success in treatment, interpersonal speech perception training is not inconsistent with effective self-monitoring by the child. Self-monitoring implies that the child needs to compare his or her own performance with an internal standard. Many children who are phonologically impaired do not have an internalized standard that is adequate to this task, and it seems unlikely that their own poor quality productions provide a good foundation for the development of such a standard.

The speech perception training procedure evaluated in this study does seem to provide the children with a basis for effective self-monitoring. The children in this study who successfully learned to produce the /ʃ/ sound tended to acquire the target in isolation between sessions, rather than during a treatment session. Although all parents were asked to refrain from engaging in home practice with their children, parents and children reported independent practice by the children at home. Those children who learned to produce /ʃ/ were very certain of and proud of their success, and it seems reasonable to assume that effective self-monitoring contributed to their acquisition of correct /ʃ/ production.

Perception Training Contrast

Locke (1980a) has criticized many speech perception assessment tools because they target contrasts irrelevant to children's actual speech errors, and recommends that assessment involve stimuli representing the error sound and the child's substitution for this sound. Jamieson and Rvachew (1992) criticized speech perception training procedures such as that employed by Williams and McReynolds (1975) on similar grounds, and demonstrated that training with more appropriate stimuli can be more effective. In a single-subject study, Jamieson and Rvachew (1992) found that training with stimuli that contrasted the target sound with the child's substitution was effective in facilitating both speech perception and sound production learning. However, a child whose error (lateralization of /ʃ/) was not directly paralleled by the training stimuli (/ʃ/ vs. /s/) did not experience a similar benefit from the training program.

The importance of training with stimuli that contrast the target sound with its substitution also receives tangential support from studies of cross-language speech perception.

For example, in Jamieson and Moore (1991), training subjects to identify /ð/-/θ/ stimuli did not facilitate perceptual performance with /ð/-/d/ stimuli, a contrast more closely related to the production errors typical of the francophone subjects who participated in this study.

The *shoe-Xshoe* training procedure administered to Group 1 subjects also supports this hypothesis, as every child was exposed to contrasts related to his or her error. However, the equivalent success of the *shoe-moo* training procedure does not support the importance of exposing the child to a contrast between the error sound and the substituted sound, as no child in this study substituted /m/ for /ʃ/.

The success of the *shoe-moo* training procedure is consistent with the maximal opposition approach proposed by Gierut (1989, 1990a, 1990b). In these studies sound production training was based upon a procedure involving a maximal or a minimal opposition. Production training with the maximal opposition was shown to be at least as effective as the more commonly used minimal opposition procedure in altering children's productive phonology.

In these papers "maximal opposition" is defined variously as one involving many distinctive feature differences (Gierut, 1989), a distinction across a major sound class (Gierut, 1990a), or a contrasting sound which is always used correctly by the child (Gierut, 1990b). Gierut (1990b) suggests that the child finds it easier to discover a unique combination of multiple features than a single feature distributed repeatedly throughout the language. She recommends that treatment should begin with maximal oppositions, and progress to contrasts involving finer distinctions (Gierut, 1990a). She also suggests that the maximal opposition treatment will be most helpful with children who produce many errors, while the minimal opposition approach will be most appropriate for children who have few errors (Gierut, 1989).

The *shoe-moo* training procedure administered to the Group 2 children in this study is clearly consistent with Gierut's (1989, 1990a, 1990b) approach as it involves a contrast across a major sound class. The *shoe-Xshoe* training procedure is also consistent with her recommendations, as it begins with relatively gross distinctions (e.g., /ʃ/ vs. /b/) and ends with finer distinctions (e.g., /ʃ/ vs. /s/).

The synthetic /ʃ/-/s/ and /s/-/θ/ training stimuli used by Jamieson and Rvachew (1992) do not represent contrasts across a major sound class. However, all of the children who benefited from the training program demonstrated consistently correct production of the contrasting sound, and the children who did not benefit from the training program had difficulty with the production of both the target and its contrast.

Williams and McReynolds (1975) targeted contrasts involving many distinctive feature differences, but in all cases the children misarticulated both members of the stimulus pair. Overall, these studies suggest that the selection of contrasts for speech perception training should parallel the selection of contrasts for production training.

Stimulus Variability

Jamieson and Rvachew (1992) proposed that the primary purpose of speech perception training is to teach the child to attend to the relevant between-category acoustic variability that characterizes the targeted contrast, and to ignore irrelevant within-category acoustic variation. This approach requires that the child be presented with a range of stimuli in which the critical acoustic dimension is varied in a systematic manner with the phonemic category. Acoustic information which does not define the distinction between the target phonemes may be held constant, as in Jamieson and Rvachew (1992), or it may be varied unsystematically, as in the *shoe-Xshoe* training program described above.

Stimulus variability has been shown to be important when teaching second language learners to perceive non-native speech contrasts. Although the teaching of non-native speech contrasts is generally a difficult task (cf. Strange & Dittmann, 1984; Werker & Logan, 1985; Werker & Tees, 1984), Jamieson and Morosan (1986) successfully taught adult Canadian francophone subjects to distinguish the English consonants /ð/ and /θ/, using a synthetic continuum of stimuli in which voice-onset-time was varied systematically among four /ð/ exemplars and four /θ/ exemplars. Logan, Lively, and Pisoni (1991) credit their success in teaching perception of the /r/-/l/ contrast to Japanese listeners to their highly variable natural speech stimuli, recorded from a number of different talkers.

The importance of stimulus variability in speech perception training also receives tangential support from studies which examine the correlation between speech perception and speech production ability. Differences between normal and impaired speakers are less marked when testing for speech perception ability with optimal stimuli (e.g., McReynolds, Kohn, & Williams, 1975), in comparison to less than optimal stimuli such as electronically altered speech (Monnin & Huntington, 1974), synthetic speech stimuli (Broen, Strange, Doyle, & Heller, 1983; Hoffman, Daniloff, Bengoa, & Schuckers, 1985; Rvachew & Jamieson, 1989), and speech recorded from children (Chaney, 1988; Hoffman, Stager, & Daniloff, 1983). The results of these studies suggest that both testing and training should involve a range of stimuli, including some that are clear exemplars of the target sounds, as well as others that are less clear or more ambiguous members of targeted phonemic categories.

The *shoe-Xshoe* training stimuli used in this study were recorded from a number of different talkers, both male and female, child and adult, and thus were characterized by considerable variability in a number of acoustic dimensions. This program did facilitate both perception and production learning and thus might be said to support the importance of stimulus variability for successful speech perception training. However, the *shoe-moo* training procedure, which contained no within-category stimulus variation, was equally successful in enhancing production performance, at least during the early stages of articulation treatment. This apparent contradiction may have occurred because these stimulus sets confound the "variability" dimension with another variable, namely "prototypicality." The *shoe-Xshoe* stimulus set contains a range of stimuli representing both poor and good quality exemplars of the correct category, while the *shoe-moo* stimulus set contains one clearly correct, "prototypical" exemplar of the word shoe.

Kuhl (1991) has shown that adults can reliably rate members of a vowel category with respect to "category goodness," and that these ratings are based upon an internalized standard (i.e., prototype) for the vowel which is common to speakers of a given dialect. She has also shown that the prototype acts as a perceptual magnet, in that it perceptually assimilates similar members of the vowel category. Conversely, the perceptual distance between a nonprototype and other members of the category is increased relative to what one might expect on the basis of acoustic distance alone. Infants also demonstrate knowledge of vowel prototypes, and their responses to vowel stimuli are similarly affected by the perceptual magnet effect (Grieser & Kuhl, 1989; Kuhl, 1991).

The development of these prototypes in infancy is dependent upon experience with the language being acquired by the child (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992); however, details about the kind of experience that optimally promotes prototype development are not yet available. It seems likely that the relationship between stimulus variability and prototypicality will be important to these investigations. These variables should also be systematically manipulated when attempting to develop the most effective speech perception training program for use with children who are phonologically impaired. It is possible that presenting varied stimuli that are all relatively prototypical would be more effective than either the *shoe-Xshoe* or *shoe-moo* training programs evaluated in this study. It is also possible that the timing of different types of experience is important.

One might suggest beginning with highly prototypical stimuli; as the child moves from the response acquisition stage of production training to sentence and conversation level tasks, speech perception training could progress to stimulus sets which contain an increasing proportion of nonprototypical stimuli. Such a program would serve two different but complementary purposes: (a) helping the child to develop an internalized standard against which to judge his own and others' productions of a given sound; and (b) helping the child to discover the critical acoustic differences that define the boundaries between one sound and another.

Amplification

Hodson and Paden (1983) strongly advocate the use of amplification when presenting stimuli to children for the purpose of "auditory bombardment." Amplified training stimuli were used in both the current study and Jamieson and Rvachew (1992). Stimuli were presented live-voice without amplification in the Shelton et al. (1978) study. The presentation level of the recorded stimuli used by Winitz and Bellerose (1967) and by Williams and McReynolds (1975) is unknown or unreported. This is another variable worthy of empirical study.

Computer-Driven Training

Shelton et al. (1978) note that the parents in their study found the procedure to be frustrating. Winitz and Bellerose (1967) made numerous adjustments to their reinforcement system throughout their study in an effort to maintain the children's interest in the task. My experience with the training procedure employed in Jamieson and Rvachew (1992) also corroborates the view that it is extremely difficult for children to comply with speech perception training activities for any length of time.

The computer-driven feedback procedure developed for this study admirably overcame this difficulty. All of the children enjoyed the task, and often asked to continue with it after the 60 trials for a given session were completed. It was much more difficult to engage the children's willing participation in the pre- and post-tests, during which visual feedback was not provided. It was often necessary to allow the children play breaks between stimulus blocks during testing, but not during training. In fact, the lack of visual feedback during post-testing may account for at least one child's markedly negative difference score on the post-treatment word identification test.

Summary

In summary, this study demonstrates that interpersonal speech perception training can facilitate sound production learning for some children who are phonologically impaired. Further research is required to determine the type of stimuli that optimally promote perception and production learning. In addition, the role of stimulus amplification in speech perception training needs to be explored, clinically useful speech perception assessment tools need to be developed, and the proper timing of speech perception and speech production training needs to be examined. Although further research is required, some suggestions for clinical application of this approach can be made on the basis of the currently available information.

Speech perception training should probably be provided concurrently with speech production training. The training should involve asking the child to identify words that contrast sounds that belong to the targeted phoneme category with sounds that do not belong to this phoneme category. Initially, stimuli that represent the targeted category would be good quality exemplars but more ambiguous exemplars of the target sound should be presented with increasing frequency as the child's production abilities improve. Similarly, the contrasting word should involve a maximal opposition in the early phases of production training. Finer distinctions should be introduced gradually over time.

Acknowledgments

This research was supported by a grant to the author from the M.S.I. Foundation. I am also grateful to Dr. E. S. Edgington for his advice regarding the statistical analysis, to Dr. E. Slawinski for her advice regarding the acoustic analyses, to Dr. D. G. Jamieson and Peter Bangarth (Hearing Health Care Research Unit, University of Western Ontario) for software support, to Tom Milutinovic (Pennan Inc.) for his delightful art work, to Michelle Gaffney for her excellent services as a research assistant, and to Norma Wood, director, and staff at the Alberta Children's Hospital Speech-Language Department for numerous forms of support throughout the duration of this project.

References

Aungst, L. F., & Frick, J. V. (1964). Auditory discrimination ability and consistency of articulation of /r/. *Journal of Speech and Hearing Disorders, 29,* 76–85.

Broen, P., Strange, W., Doyle, S., & Heller, J. H. (1983). Perception and production of approximant consonants by normal and articulation delayed preschool children. *Journal of Speech and Hearing Research, 26,* 601–608.

Chaney, C. (1988). Identification of correct and misarticulated semivowels. *Journal of Speech and Hearing Disorders, 53,* 252–261.

Edgington, E. S. (1987). *Randomization Tests* (2nd ed.). New York: Marcel Dekker, Inc.

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

Gierut, J. A. (1989). Maximal opposition approach to phonological treatment. *Journal of Speech and Hearing Disorders, 54,* 9–19.

Gierut, J. A. (1990a). Differential learning of phonological oppositions. *Journal of Speech and Hearing Research, 33,* 540–549.

Gierut, J. A. (1990b). Linguistic foundations of language teaching: Phonology. *Journal of Speech-Language Pathology and Audiology, 14,* 5–16.

Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology, 25,* 577–588.

Hodson, B. W. (1985). *Computer analysis of phonological processes.* Stonington, Illinois: Phonocomp.

Hodson, B. W., & Paden, E. P. (1983). Targeting intelligible speech: A phonological approach to *remediation.* Boston: College-Hill Press.

Hoffman, P. R., Daniloff, R. G., Bengoa, D., & Schuckers, G. H. (1985). Misarticulating and normally articulating children's identification and discrimination of synthetic [r] and [w]. *Journal of Speech and Hearing Disorders, 50,* 46–53.

Hoffman, P. R., Stager, S., & Daniloff, R. G. (1983). Perception and production of misarticulated /r/. *Journal of Speech and Hearing Disorders, 48,* 210–215.

Jamieson, D. G., & Moore, A. E. (1991). Generalization of new speech contrasts trained using the fading technique. *Proceedings of the XIIth International Congress of Phonetic Sciences, 5,* 286–289.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception and Psychophysics, 40,* 205–215.

Jamieson, D. G., Nearey, T. M., & Ramji, K. (1989). CSRE: A speech research environment. *Canadian Acoustics, 17,* 23–35.

Jamieson, D. G., & Rvachew, S. (1992). Remediation of speech production errors with sound identification training. *Journal of Speech-Language Pathology and Audiology, 16,* 519–521.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50,* 93–107.

Locke, J. L. (1980a). The inference of speech perception in the phonologically disordered child. Part I: A rationale, some criteria, the conventional tests. *Journal of Speech and Hearing Disorders, 45,* 431–444.

Locke, J. L. (1980b). The inference of speech perception in the phonologically disordered child. Part II: Some clinically novel procedures, their use, some findings. *Journal of Speech and Hearing Disorders, 45,* 445–468.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America, 89,* 874–886.

McReynolds, L. V., Kohn, J., & Williams, G. C. (1975). Articulatory-defective children's discrimination of their production errors. *Journal of Speech and Hearing Disorders, 40,* 327–337.

Monnin, L. M., & Huntington, D. A. (1974). Relationship of articulatory defects to speech-sound identification. *Journal of Speech and Hearing Research, 17,* 352–366.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research, 32,* 120–132.

Reynell, J. (1977). *Reynell Developmental Language Scales* (Revised). Windsor, England: NFER Publishing Co.

Rvachew, S., & Jamieson, D. G. (1989). Perception of voiceless fricatives by children with a functional articulation disorder. *Journal of Speech and Hearing Disorders, 54,* 193–208.

Saben, C. B., & Costello Ingham, J. (1991). The effects of minimal pairs treatment on the speech-sound production of two children with phonologic disorders. *Journal of Speech and Hearing Disorders, 34,* 1023–1040.

Seymour, C. M., Baran, J. A., & Peaper, R. E. (1981). Auditory discrimination: Evaluation and intervention. In N. L. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 6, pp. 1–58). New York: Academic Press.

Shelton, R. L., Johnson, A. F., Ruscello, D. M., & Arndt, W. B. (1978). Assessment of parent-administered listening training for preschool children with articulation deficits. *Journal of Speech and Hearing Disorders, 43,* 242–254.

Shelton, R. L., & McReynolds, L. V. (1979). Functional articulation disorders: Preliminaries to treatment. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 2, pp. 2–112). New York: Academic Press.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics, 36,* 131–145.

Van Riper, C. (1963). *Speech correction: Principles and methods.* Englewood-Cliffs, NJ: Prentice-Hall.

Werker, J., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics, 37,* 35–44.

Werker, J. F., & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, 75,* 1866–1878.

Williams, G. C., & McReynolds, L. V. (1975). The relationship between discrimination and articulation training in children with misarticulations. *Journal of Speech and Hearing Research, 18,* 401–412.

Winitz, H. (1969). *Articulatory acquisition and behavior.* New York: Appleton-Century-Crofts.

Winitz, H. (1975). *From syllable to conversation.* Baltimore: University Park Press.

Winitz, H. (1985). Auditory considerations in articulation treatment. In P. W. Newman, N. A. Creaghead, & W. Secord (Eds.), *Assessment and remediation of articulatory and phonological disorders* (pp. 249–358). Columbus, OH: Charles E. Merril Publishing Co.

Winitz, H., & Bellerose, B. (1967). Relation between sound discrimination and sound learning. *Journal of Communication Disorders, 1,* 215–235.

Received September 16, 1992
Accepted September 17, 1993

Contact author: Susan Rvachew, M.Sc., Child Health Centre, Alberta Children's Hospital, 1820 Richmond Road, SW, Calgary, Alberta, Canada T2T 5C7. Email: srvachew@acs.ucalgary.ca.

Appendix

Each stimulus was submitted to acoustic analysis as follows. The duration and amplitude of the total stimulus, the initial fricative or stop portion, and the vowel portion of each recorded word was determined using CSRE. In addition, autoregression analysis was used to determine the relative amplitude value at each of the 256 equidistant frequency values between 0 and 10,000 Hz, averaged across the entire duration of the initial noise portion of each word. From these values, the centroid was calculated, using the procedure described by Forrest, Weismer, Milenkovic, and Dougall (1988). The centroid is the mean frequency of the noise spectrum, weighted by amplitude. The phonetic and acoustic characteristics of the stimuli contained within the Group 1 stimulus set are described below in Tables 5 and 6.

TABLE 5. Acoustic characteristics of the Group 1 Stimulus Set—correctly produced "shoe."

	Initial consonant			Total stimulus	
Stimulus	Duration	Amplitude ratio[a]	Centroid	Duration	Amplitude
Adult produced					
[ʃu]	361	.67	4245	702	1.36
[ʃu]	284	.69	3629	658	1.31
[ʃu]	209	.43	3988	503	.97
[ʃu]	145	.47	4384	348	1.60
[ʃu]	343	.30	3707	563	1.60
[ʃu]	231	.54	4340	452	2.06
[ʃu]	399	1.05	3804	605	1.03
[ʃu]	163	.32	4230	329	1.27
[ʃu]	168	.39	3568	388	1.79
[ʃu]	240	.20	4800	450	1.68
[ʃu]	267	1.80	3518	434	1.29
[ʃu]	184	.47	3913	435	2.26
Child produced					
[ʃu]	284	.39	3875	523	1.35
[ʃu]	198	.37	4544	446	.95
[ʃu]	185	.56	4065	304	1.89
[ʃu]	181	.62	3976	512	1.08
[ʃu]	271	.74	4806	561	1.22
[ʃu]	278	.54	3600	540	1.77
[ʃu]	196	.28	4173	196	1.31
[ʃu]	192	.43	4938	365	1.87

Note. The durations are expressed in milliseconds, the amplitudes in volts, and the centroids in Hertz.
[a]This is the ratio of the initial consonant amplitude to the vowel amplitude.

TABLE 6. Acoustic characteristics of the Group 1 Stimulus Set—incorrectly produced "shoe."

	Initial consonant			Total stimulus	
Stimulus	Duration	Amplitude ratio[a]	Centroid	Duration	Amplitude
Stop substitutions					
[du]	48	.39	3516	333	1.54
[gu]	40	.29	3711	497	1.13
[gu]	57	.31	4148	510	.68
[bu]	29	.23	2960	635	1.88
[tu]	129	.52	5463	391	1.20
Affricate substitutions					
[dðu]	64	.32	5299	256	2.11
[tʃu]	276	.70	4200	614	2.01
[tsu]	229	.29	6039	592	1.11
[dθu]	148	.32	2831	391	1.30
[tθu]	135	.08	5964	449	2.14
Fricative substitutions					
[su]	48	.39	3516	333	1.54
[su]	40	.29	3711	497	1.13
[θu]	57	.31	4148	510	.68
[hu]	29	.23	2960	635	1.88
[θu]	129	.52	5463	391	1.20
Distortions					
[ʃˡu][b]	64	.32	5299	256	2.11
[ʃˡu][b]	276	.70	4200	614	2.01
[ʃu][c]	229	.29	6039	592	1.11
[hju]	148	.32	2831	391	1.30
[sju]	135	.08	5964	449	2.14

Note. The durations are expressed in milliseconds, the amplitudes in volts, and the centroids in Hertz.
[a]This is the ratio of the initial consonant amplitude to the vowel amplitude.
[b]Lateral distortion.
[c]Produced with a low, flat tongue and some lateral airflow, resulting in a breathy fricative.