Training severely hearing-impaired children in vowel imitation

DANIEL LING, *PhD*
CLINTON W. BENNETT, *PhD*

This study is concerned with the production of vowel sounds by hearing-impaired children who have little or no auditory sensitivity above 1,000 Hz. Vowel sounds are vitally important in communication by speech, for they not only serve to differentiate between words in their own right (for example, bill, bell, ball), but carry information about adjacent consonants (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967). Consonant-to-vowel and vowel-to-consonant transitions bear so much information that it is essential that hearing-impaired children learn to use all vowels in their language as early in life as possible.

However, the task proves to be particularly difficult for pre-schoolers who have limited hearing and restricted auditory-vocal experience. Such children tend to produce a limited range of speech sounds (Sykes, 1940; Stark, 1967; Lach, Ling, Ling, and Ship, 1970). Prior to formal speech training, their most commonly used vowel is the schwa /ə/, as in away, which accounts for about 40 percent of vowels used. The schwa is produced when the tongue is neutrally placed, i.e. assumes a position of rest in the mouth. The central vowels, such as, /ae/ as in cat; /ʌ/ as in cut; and /a/ as in father, are produced by moving the tongue a short distance from the neutral position. These central vowels are the next most frequently occuring vowels in the speech of young hearing-impaired children. The vowels /i/ as in heed

Dr. Ling is associated with the program in Human Communication disorders at McGill University. He has done extensive research in the area of low frequency hearing aids and diagnostic audiology with infants.
Dr. Bennett is supervisor of speech pathology, and coordinator of research at the Willie Ross School for the Deaf in Longmeadow, Massachusetts.

Authors' address: *D. Ling, School of Human Communication Disorders, 1266 Pine Avenue West, Montreal 112, P.Q.*

Abstract

Severely hearing-impaired children have difficulty learning to produce natural vowel sounds because they have insufficient feedback cues for tongue placement. Analysis of this problem suggests that additional tactile-kinesthetic cues in the early stages of training might facilitate the acquisition of vowels. To test this hypothesis, operant conditioning techniques were employed to train four hearing-impaired children to imitate various vowels. In the first experiment, two children were trained using an auditory-visual approach, and in the second, another two children were trained using additional tactile-kinesthetic cues. The addition of tactile-kinesthetic cues led to superior performance. Clinical implications are discussed.

and /u/ as in sh<u>oe</u>, which require placement of the tongue in high-front and high-back positions respectively, rarely occur in the early spontaneous speech of children whose hearing impairment is severe.

These findings suggest that the major obstacle to vowel production by hearing-impaired infants relates to tongue placement. As this problem persists among children who may be quite adept at imitating visible movements of the lips and tongue, the problem appears to stem not only from reduced auditory-vocal cues, but from inadequate visual information, i.e. masking of the tongue's position to some extent by the teeth, but mainly by the lips which are rounded for the vowel /u/ and spread for the vowel /i/. These vowels can be lipread with a fair degree of accuracy by subjects who have already acquired the English language in which tongue position and lip position are systematically related. As the tongue moves forward in the mouth and spreads, the lips also move from a rounded to a spread position.

As the tongue and lips assume the positions required to produce particular vowels, cavities and apertures of different shapes and sizes are created. These give rise to the peaks of resonance (formants) which characterize the acoustic pattern of each vowel. Denes and Pinson (1963) provide average formant frequency values for several vowels: for an adult male speaker, formants of the vowel /u/ occur at about 300, 870 and 2,240 Hz. Those for the vowel /a/ occur around 730, 1,090 and 2,440 Hz. Formants of /i/, at 270, 2,290 and 3,010 Hz, are the most widespread of any vowel. Formants produced by a woman tend to be slightly higher than the values given above and those produced by a child may be higher again, perhaps by as much as 40 percent. Consistently accurate identification of a vowel usually requires that two or more of its formants are audible. If only the lowest formant (F_1) is audible, then confusions may arise because the F_1 frequency of two different vowels may be similar. Thus when frequencies above 1,000 Hz are eliminated, either by electronic filtering or by hearing loss, the vowel /i/ sounds rather like /u/, and /ɛ/ as in be<u>d</u>, closely resembles /o/ as in s<u>o</u>.

Thus, neither by vision nor by aided audition alone can a child with little or no hearing above 1,000 Hz learn to identify all vowels.

However, by using both the auditory and visual cues available to him he should have sufficient information for their discrimination. Using the two modalities together, he might be expected to process cues in the following manner: low formant with lips widely spread = /i/; low formant with lips rounded = /u/; medium-low formant with moderately spread lips = /ε/; medium-low format with moderately rounded lips = /o/; strong, medium frequency formants, lips maximally parted = /a/. For a young hearing-impaired child to learn to produce these vowels correctly requires further complex skills since he must not only remember lip patterns and be able to imitate them without direct visual feedback, but he must also derive information on his tongue position by producing formants which, while closely related to the model provided by an adult, may be 40 percent higher in frequency.

Despite the evident difficulty of the task, an auditory-visual approach to teaching speech to the young hearing-impaired children has long been widespread. Recommended and appraised by such authorities as Urbantschitsch (1895), Goldstein (1939), Ewing and Ewing (1954) and Whetnal and Fry (1964), it has been applied successfully to many, but not all, hearing-impaired children. Zaliouk (1954), whose methods have received little attention, suggested that a greater proportion of hearing-impaired children would achieve better standards of speech if tactile and kinesthetic cues were more fully exploited in teaching them to talk. Accordingly, he recommended that any such child should be allowed to feel positions of the tongue with a finger of one hand placed in his mouth and the corresponding finger of the other hand placed in his teacher's mouth. Zaliouk was apparently unaware that this procedure had been traditionally employed in England during the 17th and 18th centuries (Arrowsmith, 1819). In theory, such a technique provides direct and immediate sensory feedback on the tongue position for each vowel, the production of which could then be sustained through auditory, visual and kinesthetic cues.

The purpose of this study was two-fold. In the first experiment we sought to determine whether vowel production by severely hearing-

Résumé

Plusieurs enfants déficients auditifs ont des difficultés d'apprentissage à produire des sons clairs pour les voyelles. Ils n'ont pas assez de données sensorielles sur lesquelles se baser pour un bon placement de la langue. Une analyse du problème suggère qu'on donne dès les premières phases de la formation des données tactile-kinesthétiques additionnelles. Ces données faciliteraient l'acquisition des voyelles. Pour éprouver cette hypothèse on employa des techniques opérantes de conditionnement chez quatre enfants déficients auditifs afin qu'ils imitent diverses voyelles. Lors de la première expérience on utilisa une approche auditive-visuelle pour deux enfants et, lors de la deuxième expérience, deux autres enfants reçurent des données tactile-kinesthétiques additionnelles. Ainsi, ce supplément de données tactile-kinesthétiques conduit à une supériorité dans l'exécution. Nous discuterons aussi des implications cliniques.

impaired children could be effectively achieved through intensive auditory-visual training and in the second whether the diphthong /au/ as in the word cow, and high-front vowel /i/ as in the word see, could be more effectively taught by employing the Arrowsmith-Zaliouk procedure than by auditory-visual procedures.

EXPERIMENT I

Method

Tom and Dean, aged four and five years respectively, served as subjects. Both had been profoundly hearing-impaired since birth. Their average hearing loss was greater than 90 dB ISO at 1,000 Hz. Tom had no measurable hearing at 2,000 Hz, whereas Dean's threshold for this frequency was 100 dB. Each had worn two high-powered, body-type hearing aids since their admission to school two years previously. Without hearing aids neither could hear speech, but with aids the F_1 of all vowels was audible if spoken within a distance of six feet. All unvoiced consonants were inaudible to these children and their vocalizations were limited to central vowels and bilabial consonants. Although each of the boys was of normal intelligence and could imitate visible lip and tongue movements accurately, neither had acquired more than a limited understanding of English since the language spoken in Tom's home was Chinese, and in Dean's home, French.

An electric train on an oval track, which could be activated by means of a push-button switch, was used to provide consequent reinforcement for correct responses. A Uher 4000 Reporter tape recorder was used to provide binaural amplification through TDH 39 earphones and simultaneously record each training session. The same tape recorder was used to obtain samples of speech elicited both through the presentation of a variety of colourful pictures of people engaged in a range of everyday activities and through imitation of the experimenter (CB), a male, who presented all vowel stimuli by live voice. These stimuli were the sounds /u/, /o/, /a/, /ae/ and /i/ as in the words shoe, toe, father, man and heed, respectively.

Before training began, recordings of Tom's and Dean's vocalizations were made. One recording was of each child's "spontaneous" speech elicited by their teacher who questioned the child about the content of the various pictures. The other was of each child's imitation of the five vowels specified above as spoken by the experimenter (CB). The first 30 vocal responses in the "spontaneous" speech sample were transcribed with the help of an independent judge and used as baseline measures. This number of vocal respirations was selected because we wished to compare our results with those obtained on normally hearing infants by Irwin (1947) who considered that a sample of this size provided a reliable representation of a child's phonology. A further sample of "spontaneous" speech was obtained in the same manner at the conclusion of the experiment.

All testing and training were carried out in a quiet, distraction-free room. The children were each trained individually by means of operant conditioning procedures. One child and the experimenter sat together at a small table with the train situated on the floor by the child. The microphone of the tape recorder was held by the examiner who spoke into it at a distance of about six inches and who held it at a similar distance from the child when a response was required.

A baseline measure (pre-test) was administered to determine the child's level of performance on each vowel before training was begun, and probe measures were administered at intervals throughout training to determine what progress had been made and whether training on one vowel or another had improved the child's ability to produce vowels on which no training had yet been given.

To establish each child's baseline for vowel imitation, the five stimuli were each presented ten times in the order /a/, /ae/, /o/, /i/ and /u/. To ensure reliability this procedure was repeated the next day. Recordings of each child's responses were again transcribed with the help of an independent judge and only those agreed to be accurately imitated and produced without nasalization were regarded as correct. These procedures were also followed in the training and generalization probes described below.

Reinforcement of each correct response was provided only in the course of training — not during baseline measures or probes. It consisted of activating the electric train for about five seconds. In order to maintain the child's interest, its direction was sometimes reversed and carriages added or changed. Incorrect responses were ignored.

Each subject was seen individually for training four times a week, and each training session lasted for about ten minutes, sufficient time for two sets of 25 trials to be administered. Subjects were trained until they could imitate each vowel correctly on 15 trials in a set of 25 presentations. When this criterion had been reached for each vowel, a generalization probe was administered. This probe, like the baseline measures of vowel imitation, consisted of ten presentations of each of the five vowels distributed throughout the test in random order. We began training with /a/, proceeded to /ae/, then /o/, /i/ and finally /u/. The order in which the vowels were presented demanded placement of the tongue towards the front and back of the mouth at successively increasing distances from the neutral position as the training proceeded.

Results

Results of this first experiment are summarized in Table 1, which shows the number of correct imitations obtained from each child at baseline and for each probe. This table also shows the number of training trials required by each child to reach the criterion of 15 imitations within a set of 25 presentations.

The main findings in this experiment were that the subjects' ability to imitate all vowels improved dramatically with training; that the vowel /i/ proved to be the most difficult of the five vowels to teach; that there was only one result (on Probe 5) which suggested that generalization might have affected the production of a vowel on which no training had been given; and that there was no measurable carry-over from this training into the spontaneous speech of the subjects.

Baseline measures indicated that both Tom and Dean could occasionally imitate /a/ and /u/, Dean more successfully than Tom. Indeed, Dean could imitate /a/ with better than a 60 percent, and /u/ with at least a 40 percent level of success prior to training. Even so, both children were on their third set of 25 training trials before they were able to imitate 15 presentations of the vowel /a/ correctly. Tom required 150 training trials to reach criterion for the vowel /ae/, imitated only 3/10 presentations of this vowel correctly on Probe 2 and dropped back in his ability to repeat /a/ on this probe. In contrast, Dean succeeded in reaching criterion on /ae/ during the first set of 25 presentations, imitated 6/10 presentations of this vowel correctly on Probe 2, and maintained his ability to imitate the previously taught vowel /a/. Following training on the vowel /o/, on which criterion was achieved by both children within 50 trials, Tom began to achieve better results on vowels previously taught. Thus, by Probe 3, Tom's ability to imitate /a/ and /ae/ was superior to that found on Probes 1 and 2. From this point in training, both boys either maintained or improved on imiation of vowels taught at an earlier stage. The number

TABLE 1

Number of correct responses in each set of ten items made on the two baseline measures (BL) and the five generalization probes (GP) and the number of training trials required to reach criterion (T-C) for each vowel. (Asterisks indicate which set of ten items in the 50 item probe immediately followed training on a particular vowel.)

Measure	Tom Vowel Employed					Dean Vowel Employed				
	a	ae	o	i	u	a	ae	o	i	u
BL 1	2	1	0	0	2	7	0	1	0	4
BL 2	1	0	0	0	1	6	2	1	0	4
GP 1	6*	1	1	0	2	9*	1	1	0	3
GP 2	1	3*	2	0	0	9	6*	1	1	3
GP 3	8	6	7*	4*	1	10	8	8*	2	3
GP 4	10	9	8	4*	3	10	7	10	9*	8
GP 5	9	8	8	5	7*	10	8	10	9	10*
T-C	75	150	50	500	25	75	25	50	225	25

of training trials required to reach criterion on the vowel /i/ was greater than that required for all four other vowels combined. Probe 4, which followed this extensive training, yielded the only result suggesting generalization of training. On this probe, Dean's imitation of /u/, on which no training had yet been given, had improved substantially. Results of Probe 5 showed that training had yielded a dramatic increase from baseline for both subjects and all vowels.

Analysis of the 30 "spontaneous" vocalizations elicited from the two subjects before and after training yielded no interpretable results. Tom produced only six vocalizations which utilized a target vowel before training. Following training he produced ten. For Dean, the corresponding totals were 15 and 18. This level of responding was so subnormal (Irwin, 1947) that no pre- and post-training comparisons were attempted.

EXPERIMENT II

The most salient finding in Experiment I was the enormous difficulty experienced by both subjects in learning to produce the high-front vowel /i/ through imitation based on auditory-visual input. Though Dean achieved and maintained a high level of performance following extensive training, Tom's performance did not exceed 50 percent correct for this vowel even after 500 training trials. Dean's relatively better performance was probably related to his having learned to attend to the high F_2 of this vowel which falls just within the frequency range of his residual audition. Tom's poor performance might have been due to his ability to hear only the first formant of /i/. If this is so, then auditory-visual presentation of this vowel might prove to be equally ineffective as an initial training procedure with other children having hearing-impairment similar to Tom's. Less difficulty might have been experienced by both boys had a tactile cue been introduced. This probability cannot be explored directly since it is impossible to undo the effects of the specific training subjects have already received. However, should two similar naive subjects trained with tactile cues achieve better levels of performance in imitating this vowel in less trials, it would seem reasonable to infer that training with a tactile cue is advantageous. This is the hypothesis underlying the second experiment, reported below.

Method

Sam and Nell, both 4½ years old, served as subjects. Both had profound hearing loss since birth which averaged more than 100 dB ISO at 500 and 1,000 Hz. Nell had no measurable hearing at or above 2,000 Hz while Sam had a 90 dB loss at 2,000 and 4,000 Hz. Each wore two body-type hearing aids and had been in school for about one year. Their vocalizations were generally limited to bilabial consonants and central vowels. Both were able to imitate visible movements of the lips and tongue. English was the maternal language of both children.

A Uher 4000 tape recorder was used to provide binaural amplification through TDH 39 earphones and simultaneously record both subjects' responses to baseline and generalization probe measures. Tokens, small plastic chips, were awarded to the subjects for each correct response during training, and various trinkets — plastic men, toy animals, picture cards, etc. — were available for "purchase" with these tokens.

The experiment was conducted in a quiet room and the child and the experimenter were seated facing each other next to a small table. The microphone of the tape recorder was held by the experimenter who spoke into it at a distance of about six inches and similarly held it for the child when a response was required.

Prior to training, we assessed each child's ability to imitate the vowels /i/ and /o/ and the diphthong /au/ presented with tactile cues. To establish baselines for imitation, the three stimuli were each presented ten times in the order /i/, /au/ and /o/. To ensure reliability, baseline measures were repeated and recordings of each child's responses were made for transcription by an independent judge. A correct response was defined as any utterance agreed by the experimenter and the independent judge to be imitated accurately, without nasalization, on the first attempt. Generalization probes were similarly administered and scored during training.

After baseline measures were completed, we ensured that both children understood that tokens would be awarded for correct responses and that these could be exchanged for trinkets at the end of each session. This was done by having each child imitate 25 presentations of the syllable /pa/ which was already within his repertoire. As in subsequent training trials a token was awarded for each correct response and incorrect responses were ignored. We introduced the phoneme /p/ to release the desired vowel or diphthong. This bilabial step was already in each subject's repertoire, and its presence provided the experimenter with consonant-vowel transition cues which helped him assess the integrity of the child's imitation of the target phoneme.

During training, the syllables were presented with tactile-kinesthetic (T-K) cues (the Arrowsmith-Zaliouk approach) and also without T-K cues in which condition the child had only auditory-visual (A-V) information about the stimuli. The target syllables /pi/ and /pau/ were presented under each condition in one order to Sam and the reverse order to Nell, so that the effects of previous experience on acquisition could be observed. Neither was trained to imitate /po/, which was included in the baseline and probe measures to determine whether generalization occurred. The training involved four steps for each subject. Each step was terminated when the criterion for acquisition of 15/25 correct responses was reached. Training sessions, which were of about 30 minutes duration for steps 1 and 2 when the two series of target syllables were both presented, lasted about 15 minutes for steps 3 and 4. One training session was provided each school day for a period of three weeks. At the end of each session tokens were exchanged for a trinket.

Sam was first trained on /pi/ using the T-K procedure and /pau/ using the A-V approach. This step was repeated on subsequent days until Sam had reached the target criterion of acquisition, namely 15 correct in a set of 25 trials for one or the other phoneme. On reaching this criterion, Probe 1 was presented. The purpose of this probe was to determine the relative effectiveness of the two methods of training in relation to base-line and to determine the extent to which initial

T-K approach facilitated imitation from a model later presented under A-V conditions. In the second training step both target syllables were taught through the A-V approach. Training sessions were again repeated until criterion was reached on one or other of the target syllables. The second probe was then administerd. In the thrid step, /pau/, the syllable previously presented only through the A-V approach was presented with tactile-kinesthetic cues. On reaching criterion, Probe 3 was administered. Following this, the fourth training step, further presentation of /pau/ through the A-V approach was initiated. The final probe, Probe 4, was presented when Sam reached criterion. The same paradigm was followed with the other subject, Nell, except that the order in which target syllables were presented was reversed.

Results

The number of correct responses obtained at baseline under each condition during training and for each probe is shown in Table 2. Results for the two children are similar. Both acquired the target syllable more rapidly through the T-K than through the A-V approach, and they required many fewer trials to reach criterion on /i/ than did the subjects in the first (A-V) experiment. Tom, who scored 2/10 at baseline, learned to imitate /i/ after only 50 T-K trials and Nell, who scored zero on this vowel at baseline and a maximum of 4/25 after 75 A-V trials, reached criterion within four T-K and three further A-V training sessions. A similar result was obtained for the target syllable /pau/. In contrast to results obtained in Experiment I, both children's imitation of the syllable on which no training had been given improved as the experiment proceeded.

Discussion

Results obtained support the hypothesis that initial tactile-kinesthetic (T-K) training of profoundly hearing-impaired children yields more rapid acquisition of ability to imitate vowels which require tongue placement in high-front or high-back positions than auditory-visual (A-V) training. The results on both pairs of children suggests that the acquisition of this skill is enhanced if there is measurable hearing above 2,000 Hz. Thus, initial T-K training seems advantageous whether or not hearing above 2,000 Hz is present.

Results of Experiment II also indicate that initial T-K training readily generalized to imitation of A-V patterns, the condition under which these children have to function in everyday life. The fact that the syllable initially used in the T-K mode yielded slightly better scores on Probe 4 than the syllable first presented in the A-V mode suggests that there may be an interference effect from initial training in the less effective modality, or conversely, that there is a facilitative effect with initial training in the more effective modality. Whether or not this is truly so would require further research. Whether such an effect might persist for a significant period would require study involving considerably more training sessions than were given in Experiment II.

There was almost no generalization of training noted in Experiment I, yet improved imitation of the syllable on which no training had been given was noted in Experiment II. There are several possible reasons for this difference. In Experiment II, we used syllables rather than isolated vowels, but since the releasing consonant was already

TABLE 2

Number of correct responses to each syllable made by Sam and by Nell at baseline (BL) during each tactile-kinesthetic (T-K) or auditory-visual (AV) training session and at each probe (P). Baseline and probe scores were obtained on duplicated sets of ten presentations of each syllable. Training sessions each contained 25 items. Criterion for acquisition was 15/25 correct imitations, following which probes were administered.

	Syllable	BL	BL	TK	AV	TK	AV	P_1	P_1	AV	AV	P_2	P_2	TK	TK	TK	P_3	P_3	AV	AV	AV	AV	AV	P_4	P_4
Sam	pi	2	2	14		22		5	4	19		8	8				10	9						10	6
	pau	1	0		6		6	1	2		5	2	1	11	12	15	3	3	10	11	10	10	15	8	8
	po	1	0					5	3			3	4				5	4						8	7

	Syllable	BL	BL	TK	AV	P_1	P_1	AV	AV	AV	AV	P_2	P_2	TK	TK	TK	P_3	P_3	AV	AV	AV	P_4	P_4
Nell	pi	0	0		4	1	1	3		4	1	0	10	10	10	16	5	7	11	10	16	6	7
	pau	0	0	24		5	5	13		18		5	7				9	8				9	9
	po	1	0			3	2					3	1				4	4				4	5

in each child's repertoire, such generalizations could not readily be explained in these terms. Nor is it likely that the second pair of subjects would have greater ability to generalize than the first pair. All subjects came from similar backgrounds and had similar histories of training and performance in school. The most likely explanation is that the better generalization in the second experiment was due to the type of training provided. Schroeder and Baer (1972) found that probe accuracy following concurrent training was consisently greater than that following serial training, and in the second experiment, the first and second steps provided concurrent training of the two target syllables. Additionally, the T-K training in more accurately identifying tongue placement as the critical variable for the child might have facilitated generalization. Also, an association between lip rounding and placing the tongue towards the back of the mouth could have been established through training with the syllable /pau/. If this were so, then training diphthongs first, and vowels next, might prove to be a tactic superior to that of training tongue placement primarily in isolated vowels.

Although the teaching of speech to hearing-impaired children has been practiced for several centuries, little experimental work has been undertaken to determine how best to achieve natural sounding patterns. The increasing use of hearing aids over the past 20 years has coincided with a general trend in the education of the hearing-impaired towards a so-called "natural language" approach (Groht, 1958) in which systematic, structured teaching has been largely neglected in favour of simply talking to the child in the expectation that speech perception and production will be acquired through intensive exposure to auditory-visual patterns. While this approach may be effective with children who have ample residual hearing, children such as those who served as subjects in the present study may be unable to benefit from such an approach. The tactile-kinesthetic procedures reported here appear to be advantageous in initiating certain speech skills. The operant procedures employed could also be adopted in training hearing-impaired children in other aspects of verbal expression. (Bennett and Ling, 1972; Bennett, 1973).

Acknowledgements

Thanks are due to Doris Leckie, Principal, Montreal Oral School for the Deaf, and her staff for their collaboration in this study, which was supported by Canadian National Health Grant 604-7-725.

References

Arrowsmith, J. P. 1819. *The art of instructing the infant deaf and dumb.* London: Taylor and Hessey.

Denes, P. B., and Pinson, E. N. 1963. *The speech chain.* Murray Hill: Bell Telephone Laboratories.

Ewing, I. R., and Ewing, A. W. G. 1954. *Speech and the deaf child.* Manchester: Manchester University Press.

Goldstein, M. A. 1939. *The acoustic method.* St. Louis: Laryngoscope Press.

Groht, M. A. 1958. *Natural language for deaf children.* Washington, D. C.: Alexander Graham Bell Association for the Deaf, Inc.