An Analysis of Visual and Vibrotactile Cue Interaction in Speech Perception

Arlene Earley Carney, James C. Durkel, and Cindy A. Beachler
Purdue University

Two experiments were designed to answer the questions: (a) is there any additive effect for visual and vibrotactile cues for the perception of vowels and consonants, and (b) do these cues interact in the reception of connected discourse? In the first experiment, subjects were asked to choose the correct vowel or consonant from a closed set of consonant-vowel (CV) syllables under three conditions—visual only, tactile only, and visual plus tactile. Subjects were trained for eight sessions in a live-voice, face-to-face situation with a single-channel vibrotactile device combined with visual cues. Results indicated a significant benefit in phoneme perception in the combined-modality condition for both vowels and consonants. In the second experiment, sentences and paragraphs were presented to subjects under the same three conditions. Half of the subjects tracked the stimuli while the other half received two repetitions for comprehension. In addition, four subjects received visual cues only while the other four subjects received both visual and tactile cues. Results of the latter experiment did not show any benefit for the combined-modality condition. It seems that the addition of vibrotactile cues to the visual stimulus is of limited importance when subjects are asked to perform beyond the sensory-perceptual level.

Discussion of multisensory presentation of speech to hearing-impaired people has generally been confined to an analysis of the interaction of visual and auditory cues. However, a body of research exists which specifically examines the effects of combining visual and vibrotactile cues for speech. Pickett (1963) combined visual and tactile cues in an experiment testing the perception of phonemes and words through a ten-channel vocoder which delivered the signal to ten bone-conduction vibrators placed on subjects' fingers. After training, deaf subjects did receive some small benefit from the combined condition, but this varied according to the subject and speaker. Pickett

Arlene Earley Carney, Ph.D., is an Assistant Professor in the Department of Audiology and Speech Sciences at Purdue University. James C. Durkel, M.S., is a recent graduate from Purdue University and presently employed in a preschool program for the hearing impaired in St. Paul, Minnesota. Cindy A. Beachler, M.S., a recent graduate from Purdue University, is currently employed at the Crossroads Rehabilitation Center in Indianapolis, Indiana.

concluded that the addition of vibratory cues to the visual signal added some new information, particularly about voicing and nasality, and might be useful in making the entire communication less susceptible to interference.

Until the publication of an extensive review of the literature on the tactile reception of speech by Kirman (1973), this area received little further attention. Then, Arnst and Danhauer (1974), using a single-channel vibrator, studied the combination of auditory, visual, and vibrotactile cues with normal and hearing-impaired listeners for the identification of synthetic sentences. Both groups of subjects performed the same on the visual-only and visual-plus-vibrotactile tasks. However, hearing-impaired subjects had maximum performance in the visual-plus-auditory-plus-vibrotactile condition, indicating that some minimal integration of tactile cues occurred when all three modalities were combined. Danhauer and Appel (1976) examined interaction of visual and vibrotactile cues in the perception of CV syllables and concluded that the tactile cues provided no additional information to their normal subjects. These results are somewhat misleading, however, since subjects in these experiments received no more than ten minutes of training with tactile cues.

A more comprehensive study was done by Sparks, Kuhl, Edmonds, and Gray (1978) for three normal subjects. These investigators used a multichannel electrocutaneous device which was worn around the subjects' abdomens. Subjects were trained with both vowel and consonant stimuli in a CVC context for 20 to 35 hours under three conditions—visual only, tactile only, and visual plus tactile. After training, subjects demonstrated nearly perfect recognition performance even in the tactile-only condition for vowel stimuli. For consonants the features of voicing and nasality were transmitted best, with place of articulation considerably poorer in the tactile-only condition. In the combined condition, subjects' performances were better than either of the single-modality conditions. Sparks et al. concluded that subjects were able to combine information from different modalities to make phonemic judgments. Although this study examined visual and vibrotactile interaction in detail, it is difficult to generalize the results. Only three subjects were used, and not all received uniform training. The excellent performance on such features as voicing, nasality, and frication may have been enhanced by the stimulus sets chosen by the experimenters; e.g., one set included six stops plus two nasals and another only fricatives. A stricter analysis seems necessary before extensive work is begun on hearing-impaired individuals at the phonetic level.

Visual and vibrotactile interaction has also been studied in the reception of connected discourse. De Filippo and Scott (1978) described a new technique which they called tracking, in which a subject was asked to repeat every word spoken until the criterion of 100% intelligibility was reached. Each subject's performance was described by the number of words repeated correctly in one minute. When a tactile aid, composed of a vibrotactile and an electrotactile transducer, was used with subjects, their tracking rate increased over the visual-only by approximately 6 to 29%.

A similar study was undertaken by Sparks, Ardell, Bourgeois, Wiedmer, and Kuhl (1979) in which the tracking technique was used with an electrocutaneous instrument. They found that the tracking-rate advantage seen in the combined visual-tactile condition disappeared as training increased; subjects' performances were essentially the same as in the visual-only condition. These two somewhat contradictory sets of results further complicate the issue of visual and tactile interaction and leave several areas open for question.

In both these studies, subjects acted as both speakers and receivers in tracking. Therefore, subjects could not have been tested on the same material, only similar materials. Yet, both sets of investigators stress that vocabulary, literary style, etc., can change tracking rate dramatically. Further, the material used for tracking was almost exclusively literary in nature, rather than informational or communicative and, as such, was far removed from the types of information subjects may be receiving in daily communication. Finally, tactile cues were combined with connected discourse only in the tracking paradigm and not with any other type of procedure.

To attempt to clarify the nature of visual and vibrotactile interaction, two studies were designed, one at the phonetic level and one at the level of connected discourse. The purpose of these studies was to provide more comprehensive information about visual and vibrotactile interaction with very different stimuli by means of a number of methodological changes which allow for a more extensive analysis of subject variability and training techniques.

EXPERIMENT 1

Method

Subjects. The subjects were six females with a mean age of 19.4 years. All six subjects had normal hearing and vision and no previous experience with speechreading and tactile devices.

Apparatus. Tactile stimulation was delivered through a commercially available, single-channel device, the Fonator (Siemens). The vibrotactile transducer is an electro-magnetic disc which can be placed in contact with the subjects' fingers, palms, or wrists. The vibrator oscillates differentially and segmentally in response to variations in the input signal; i.e., as the input signal increases in frequency, the disc vibrates faster; and as the input increases in intensity, the disc vibrates with a greater force. The Fonator has a flat frequency response between 63 and 8000 Hz (±5 dB re 1000 Hz) with the tone control dials set at zero. Subjects in this experiment were instructed to place their fingers on the face of the disc which rested on a table in front of them. All stimuli were delivered to the dynamic microphone of the Fonator.

During each training and testing session, subjects received minimal auditory cues. They wore binaural earplugs (EAR) under headphones (TDH-39 with MX-41/AR cushions) through which was delivered a white noise signal of 98 dB SPL.

Procedures. Each subject attended one pre-training assessment session, eight training sessions, and one post-training assessment session. Each session lasted approximately one and one-half hours. Subjects generally attended two sessions per week for a period of approximately five to six weeks and were tested and trained in pairs.

All training and testing was done by one experimenter in a sound-treated room. Subjects were seated at a table across from the experimenter at a distance of approximately three feet. The experimenter's face was lighted both by overhead lighting and by a high-intensity floor lamp which was focused directly on the experimenter's face.

Each subject was provided with a response board with the consonant-vowel syllable choices printed on it in orthographic symbols. Subjects were instructed to point to the syllable which the experimenter produced. The experimenter recorded each subject's response on a form immediately after each stimulus presentation.

Stimuli used were CV syllables in which either the consonant or vowel member varied. The consonant response set consisted of the following phonemes: /p,b,m,t,d,s,z,l,n,ʃ,ʒ,f,v,θ,ð,w,r,j,k,g/ followed by the vowel /a/. With respect to their feature analysis, seven consonants were voiceless (V^-) and thirteen were voiced (V^+), six consonants were stops, two were nasals, eight were fricatives, and four were semi-vowels or glides. Place of articulation was determined according to the homophenous or viseme categories described by Binnie, Jackson, and Montgomery (1976). The viseme groups were as follows: /p,b,m/; /t,d/; /s,z/; /l,n/; /ʃ,ʒ/; /f,v/; /θ,ð/; /w/; /r/; /j/; and /k,g/. The vowel response set consisted of the following nine phonemes: /i,ɪ,ɛ,æ,a,ʌ,ɔ,ʊ,u/ preceded by the consonant /b/. Consonant response sets and vowel response sets were displayed on two different response boards.

Prior to the initial assessment, each subject was presented with the consonant and vowel stimuli to be used in the experiment. Each subject responded with 100% accuracy to all stimuli in an auditory-visual, face-to-face condition before any assessment was begun.

The initial assessment, as well as all subsequent assessments, was conducted under three conditions: visual-only (V), tactile-only (T), and visual-plus-tactile (VT). Three patterns of stimulus presentation were used, with two subjects receiving each pattern: (a) vowels (VT,T,V), consonants (VT,T,V); (b) vowels (V,T,VT), consonants (V,T,VT); and (c) consonants (V,T,VT), vowels (V,T,VT). These patterns were chosen to reduce any ordering effects in testing.

Each stimulus was presented six times in a random order in each condition. Therefore, the stimulus set for consonants consisted of 120 items (20 consonants × 6 presentations), while the stimulus set for vowels had 54 items (9 vowels × 6 presentations).

During the visual-only condition, the tactile stimulation was disengaged; during the tactile-only condition, the experimenter covered her lower face to eliminate visual cues. The experimenter repeated each stimulus three times, and the subjects then pointed to the perceived stimulus. During assessment, no feedback was given.

The eight training sessions began with one hour of training followed by approximately one-half hour of assessment to determine progress. Training on vowels or consonants first alternated from session to session. The experimenter conducted the training in both tactile-only and visual-plus-tactile modes. During the first two training sessions, stimuli were grouped in various pairs (i.e., /i/ vs. /u/, /t/ vs. /k/) under both training conditions. The subjects were provided with feedback about the correctness of their answers. The experimenter varied the grouping of stimuli to include all possible combinations. As training progressed, the experimenter asked the subjects to choose from larger stimulus sets; i.e., three, four, five, or more stimuli grouped together.

The training session assessments were identical to the initial assessment except that each stimulus was presented only three times per condition in a random order for a total of 60 items for the consonants and 27 items for the vowels. The VT, T, and V conditions were tested for each session; and the order of testing was selected randomly for assessment with vowel and consonant groups. No feedback was given during assessment.

The final assessment was again conducted in the same three ordering patterns as the initial assessment. However, each subject received a different pattern from the first session. In addition, the effect of coarticulatory changes was assessed during the final session. As in the initial assessment, 120 consonant tokens were used (20 consonants × 6 repetitions). For the consonant response set, three vowels /i,a,u,/ were used variably with two repetitions of each vowel paired with each consonant. Similarly, the vowel response set consisted of 54 CV tokens (9 vowels × 6 repetitions) in which three consonants /b,d,g/ were used variably, two repetitions per consonant. For both response sets, the order of consonants or vowels was presented randomly.

Results and Discussion

At the end of the eight training sessions, all six subjects showed evidence of improved phoneme reception under all three experimental conditions: visual-only, vibrotactile-only, and visual-plus-vibrotactile. However, these data do not suggest that subjects have reached asymptotic performance on any of the tasks; rather, they appear to be improving at a slow rate during each training session.

Table 1

Mean Phoneme Reception Scores for Vowels and Consonants in Three Modalities (Percent Correct)

Session	Vowels			Consonants		
	Visual	Tactile	Combined	Visual	Tactile	Combined
Pre	60	23	68	55	12	62
1	73	35	69	60	18	62
2	78	39	79	64	17	72
3	79	34	84	62	22	73
4	81	36	88	67	21	73
5	89	35	90	69	24	66
6	81	38	83	73	31	79
7	86	47	85	70	33	82
8	88	40	89	79	33	83
Post	86	44	87	69	28	75

The data for the vowel stimuli in Table 1 show the large difference in performance between the tactile-only condition and both the visual-only and combined conditions. Throughout pre- and post-testing, as well as throughout training, this difference remains constant at about 37 to 45% poorer for the tactile-only condition. In all three conditions, subjects progressed at about the same rate showing some similarity in the learning across the three modalities. The consonant-environment changes introduced in the post-test do not cause any substantial decrement in performance from the final training session; apparently coarticulation effects are minimal for vowel stimuli under these three conditions.

Although subjects' performances in the tactile-only condition are poor, they are considerably above the chance level of 11% throughout the training. Thus, it appears that subjects can learn to use vibrotactile cues in making choices among vowels. Moreover, the addition of vibrotactile cues to visual cues does appear to enhance performance slightly. Scores for these two conditions are virtually the same throughout the experiment. A t-test for matched pairs was significant (t=-1.83; p<.05).

Results for the consonant stimuli are also shown in Table 1. As in the vowel data, performance in the tactile-only condition is much poorer (approximately 40 to 50%) than in either of the other two experimental conditions. Subjects began with 12% correct on the pre-training test and finished with 33% correct at the final training session. However, as in the vowel data, these scores are considerably above the chance level of 5%.

Subjects again demonstrated a slow but steady increase in performance across training sessions for all three experimental conditions. In contrast to the vowel data, changes in vowel environment introduced during the post-training test caused a substantial decrement in performance—10% for the visual-only, 5% for the tactile-only, and 8% for the combined condition. Regardless of modality, coarticulatory effects appear to influence consonant perception substantially.

As in the vowel data, except for one training session, subjects always received higher scores in the combined condition when compared to the visual-only. This difference was found to be significant by a matched-pair t test (t=-4.33; $p<.001$). After only eight training sessions, subjects were able to receive significant benefits from the addition of vibrotactile cues to visual cues.

A feature analysis of all post-test consonant data is shown in Table 2. These data were based on phoneme-correct performance. It is clear that

Table 2

Mean Performance for Correct Feature in Three Modalities

Feature	Visual	Tactile	Combined
Voicing	84	80	88
Place	90	35	92
Manner	84	61	87

all three features (voicing, place, and manner) are best transmitted in the combined condition, with place of articulation better than the other two. As expected, place of articulation is again best in the visual only condition. However, in the tactile-only condition, voicing is transmitted best, followed by manner of articulation. Place information is considerably poorer than any other feature under any condition.

The results of this experiment are consistent with earlier data of Pickett (1963) and Sparks et al. (1978) despite the large differences in tactile instrumentation and training in this study and those two experiments. At the phonetic level, the addition of tactile to visual cues enhances individual phoneme recognition. More importantly, it contributes feature information about voicing and manner of articulation to the place of articulation from the visual modality. Both Pickett (1963) and Sparks et al. (1978) found the same good transmission of the voicing feature with somewhat poorer transmission of the nasality and frication features. These effects are replicable regardless of methodological differences, suggesting that they are related to the coding characteristics of the skin for the reception of speech stimuli rather than to any particular instrument.

EXPERIMENT 2

Method

Subjects. The subjects were eight females with a mean age of 22.8 years. All eight subjects had normal hearing and vision and had minimal or no previous experience with speechreading and tactile devices.

Apparatus. The apparatus for delivery of tactile stimulation and for masking was the same as in Experiment 1.

Procedure. Each subject attended one pre-training assessment session, five training sessions, and one post-training assessment session. Each session, with the exception of the post-training session, lasted approximately one hour. Subjects generally attended two sessions per week for a period of three to four weeks. Each subject was tested and trained individually.

All testing and training was performed in a sound-treated room by two experimenters, one female and one male. Subjects were seated at a table across from the experimenter at a distance of approximately three feet. The experimenters' faces were lighted both by overhead lighting and by a high-intensity floor lamp which was focused directly on their faces.

All eight subjects received the same pre-training assessment in a visual-only condition. In one part of the assessment, speechreading skill was assessed in a "traditional" fashion. The experimenter read List 1 of the revised CID Everyday Speech Sentences (Hood & Dixon, 1969), two repetitions per sentence, to each subject. The list contained ten sentences with key words indicated for each sentence. There was a total of 50 key words per list. Each subject was instructed to write down what part or parts of the sentence she had comprehended on a response form. Each subject then received a percentage score on the sentence list, based on the number of key words which were reported correctly. In addition, a short paragraph of approximately fifty words was also read twice by the experimenter. Each subject was immediately presented with four multiple-choice questions. Subjects marked their choices on a response form.

The second part of the assessment required the subjects to track the sentences and paragraphs as they were being read by the experimenter. Once again, the stimuli used were contained in a different list, List 2, of ten revised CID Everyday Speech Sentences and a fifty-word paragraph. The same tracking protocol was used by each experimenter. The experimenter read each sentence as a full sentence, and the subject was asked to repeat after the experimenter whatever she had received. If the repetition was not 100% accurate, the experimenter would repeat the sentence, breaking it down into phrases or clauses. The experimenter continued to repeat the phrase or clause until the subject repeated it correctly. However, if the subject could not repeat the message, the experimenter would further break down the clause or phrase into its individual words and occasionally into syllables. If after all the repetition of the material at the sentence, phrase, and word levels the subject failed to repeat a word or words correctly, the experimenter would provide some kind of cue to the subject about the identity of the word. For example, if the problem target word were "tackle," the experimenter might indicate that this word is used in football terminology and it means "to knock down." The target word or words would again be repeated for the subject if necessary. This procedure was continued until all the stimulus words had been correctly repeated by the subject. Four subjects were tested on tracking first, followed by the traditional assessment. Four were tested in the reverse order to minimize order effects.

Following initial assessment, subjects were randomly assigned to one of four groups. Group 1 was trained in the tracking procedure in a visual-only condition, and Group 2 received the same training with the addition of vibrotactile cues. Group 3 was trained in comprehension-oriented techniques in a visual-only condition, and Group 4 received the addition of vibrotactile cues.

The tracking training focused on increasing subjects' tracking rates for sentences; short paragraphs; and short, information-based magazine articles. At the end of each tracking session, ten sentences, similar in length, vocabulary, and syntactic complexity to the CID Everyday Speech Sentences, and a fifty-word paragraph were presented to measure tracking progress.

During the comprehension-oriented training, subjects were instructed in techniques for prediction of words within sentences and for distilling the meaning from communication context, rather than from word-for-word repetition. They were also given information about the visibility and the homophenous nature of phonemes. At the end of each training session, comprehension by the subjects of ten sentences and a short fifty-word paragraph was tested to measure progress.

The four subjects who received vibrotactile cues during training were always assessed in the visual-only condition using the vibrotactile input received during training. After training, all eight subjects were tested first by the experimenter who trained them and then by the other experimenter. These post-tests had the same format as the pre-test. Each subject was tested in both comprehension and tracking skills with Lists 7 and 8 of the revised CID Everyday Speech Sentences and with more complex, two-hundred word passages.

Results and Discussion

The results for the pre-test data for all eight subjects are shown in Table 3. Subjects understood more than half of the CID Everyday Speech Sentences presented to them when key words in the sentence were scored. Their performance on paragraph comprehension is somewhat greater, even compared to the chance score of 25%. All eight subjects have an average tracking rate of

15 wpm for sentences and 11 wpm for paragraphs. Throughout the five training sessions, both in the comprehension group and in the tracking group, those subjects who were trained in the visual-only condition comprehended more and tracked faster than subjects in the visual plus vibrotactile condition.

Table 3
Pre-test Results for All Subjects with Connected Discourse

	Procedures	
Materials	**Comprehension**	**Tracking**
Sentences	52%	15 wpm
Paragraphs	63%	11 wpm

Post-test data for all subjects are shown in Table 4. The data in this table are displayed so that results from the four subjects who received only visual cues, regardless of procedure, are grouped together, as are the four subjects in the combined condition. The results are also separated to allow for a comparison of the first and second post-tests. In all conditions, except for the first post-test sentence comprehension task, subjects who were trained with only visual cues performed better than those who were trained with combined cues.

Table 4
Post-test Results for All Subjects with Connected Discourse

	Post-Test 1			
	Visual		**Visual + Tactile**	
Materials	**Comprehension**	**Tracking**	**Comprehension**	**Tracking**
Sentences	71%	38 wpm	80%	23 wpm
Paragraphs	45%	13 wpm	40%	11 wpm
	Post-Test 2			
	Visual		**Visual + Tactile**	
Materials	**Comprehension**	**Tracking**	**Comprehension**	**Tracking**
Sentences	61%	46 wpm	43%	19 wpm
Paragraphs	75%	20 wpm	60%	13 wpm

It is difficult to find any other consistent trends in the data. Changes in speaker/trainer introduced in post-test two adversely affected sentence comprehension for all subjects, with a greater effect on the subjects who had received the combined training. Conversely, this same experimental change enhanced paragraph comprehension for all subjects, with a greater positive effect on subjects who had been trained with visual cues. Half of the subjects (visual-only training) improved in their tracking performance from the first to the second post-test; the remaining subjects (visual + tactile training) either decreased tracking rate or remained about the same.

When the data were reanalyzed with subjects regrouped according to the procedure with which they had been trained rather than the modality, it was clear that subjects who had been trained with tracking had higher tracking scores than those who had not (38 wpm as opposed to 23 wpm for sentences and 15 wpm as opposed to 10 wpm for paragraphs). This advantage was completely overcome in the second post-test where both groups tracked within 2 wpm of each other across stimulus materials. Of interest is the fact that the tracking-trained subjects had a small advantage over the other subjects even in sentence and paragraph comprehension; e.g., 79% correct in sentences vs. 71% in paragraphs. This advantage, reversed in only one condition, continued to hold in post-test two.

These data are quite different from results reported by De Filippo and Scott (1978) and Sparks et al. (1979), which showed the combined condition to be better than the visual-only condition. However, the present experiment is different from these two studies in one critical methodological aspect. While our subjects were trained with a vibrotactile instrument, the test data obtained to measure progress was always gathered in the visual-only condition. Our concern was not so much whether subjects could perform better in an aided condition but whether the presence of this additional sensory cue during training was helpful to the subject after training ended. This seemed particularly important to us because no wearable tactile aid is yet available.

A second methodological difference in the present study is the testing of two different types of training procedures, a comprehension-oriented technique and a tracking technique. The addition of vibrotactile cues does not appear to affect either one substantially.

A common final result of the present experiments was enormous intersubject variability, regardless of modality or training procedure. Particularly in the tracking procedure, some subjects showed tracking rates two to three times that of other subjects and maintained these scores throughout training.

GENERAL DISCUSSION

Results of the two experiments conducted in this study appear contradictory at first glance; i.e., the addition of vibrotactile cues enhances phoneme perception but has no effect on the reception of connected discourse. However, these results become more reasonable in view of the tasks the subjects were asked to perform. In the first experiment, the subjects were asked to make a nine-alternative, forced choice for vowels and a twenty-alternative, forced choice for consonants. It is essentially an ambiguous pattern-recognition task in the visual-only condition. The addition of vibrotactile cues helps to disambiguate the task and reduce the number of possible alternatives even further. This is most obvious in the selection of a phoneme from within a homophenous group.

In the case of connected discourse, subjects performed an essentially linguistic task. They are either trying to understand a message or repeat a message entirely. For these tasks subjects call on all their semantic and syntactic experience, as well as their perceptual systems, to perform correctly. The addition of a new sensory cue which can add no syntactic or semantic information is not facilitative and, in fact, may be detrimental.

A further argument may be raised regarding the role of experience in these experiments. In the phoneme-choice experiment, subjects are relatively new to the task at a visual level as well as at a tactile level; in the comprehension tracking experiment, subjects may have had quite a bit of exposure to the use of visual cues to assist understanding of messages in noise, etc. The lack of experience in the first experiment may predispose subjects to rely on tactile cues as well as visual cues.

In conclusion, it appears that visual and vibrotactile interaction may be additive at a sensory-perceptual level and inconsequential at a cognitive-linguistic level. The nature of the interaction is not a simple one and is influenced by the experimental situation under which it is studied.

ACKNOWLEDGMENTS

The authors wish to acknowledge Edward J. Carney, Edward M. Burns, and Lawrence Feth for their assistance in the preparation of this manuscript. This research was supported in part by a grant from NINCDS.

REFERENCES

Arnst, D., & Danhauer, J. Improvement in recognition of synthetic sentences with multisensory stimulation. *Journal of Auditory Research,* 1974, *14,* 211-216.

Binnie, C., Jackson, P., & Montgomery, A. Visual intelligibility of consonants: A lipreading screening test with implications for aural rehabilitation. *Journal of Speech and Hearing Disorders,* 1976, *41,* 530-539.

Danhauer, J., & Appel, M. INDSCAL analysis of perceptual judgments for 24 consonants via visual, tactile and visual-tactile inputs. *Journal of Speech and Hearing Research,* 1976, *19,* 68-77.

De Filippo, C., & Scott, P. A method for training and evaluating the reception of ongoing speech. *Journal of the Acoustical Society of America,* 1978, *63,* 1186-1192.

Hood, R.B., & Dixon, R.F. Physical characteristics of speech rhythm of deaf and normal-hearing speakers. *Journal of Communication Disorders,* 1969, *2,* 20-27.

Kirman, J. Tactile communication of speech: A review and an analysis. *Psychological Bulletin,* 1973, *80,* 51-74.

Pickett, J. Tactual communication of speech sounds to the deaf: Comparison with lipreading. *Journal of Speech and Hearing Disorders,* 1963, *28,* 315-330.

Sparks, D., Ardell, L., Bourgeois, M., Wiedmer, B., & Kuhl, P. Investigating the MESA

(Multipoint Electrotactile Speech Aid): The transmission of connected discourse. *Journal of the Acoustical Society of America,* 1979, *65,* 810-815.

Sparks, D., Kuhl, P., Edmonds, A., & Gray, G. Investigating the MESA (Multipoint Electrotactile Speech Aid): The transmission of segmental features of speech. *Journal of the Acoustical Society of America,* 1978, *63,* 246-257.