Chapter VI. Discussion

VI-1 Summary of Main Results

The potential for communication through the kinesthetic aspect of the tactual sense was examined in a series of preliminary experiments employing Morse Code signals. Experienced and inexperienced Morse Code operators were trained to identify Morse Code signals that were delivered as sequences of motional stimulation through up-down (≈10 *mm*) displacements of the fingertips. Performance on this task was compared to that obtained for both vibrotactile and acoustic presentation of Morse Code using a 200-*Hz* tone delivered either to the fingertip through a minishaker or diotically under headphones. For all three modalities, the ability to receive Morse Code was examined as a function of presentation rate for tasks including identification of single letters, random three-letter sequences, common words, and sentences. Equivalent word-rate measures (i.e., product of percent-correct scores and stimulus presentation rate) were nearly twice as high for auditory presentation as for vibrotactile and motional presentation.

The main body of this work was aimed at developing a tactual display with a high information transfer rate. A multi-finger positional stimulator, called the TACTUATOR, was developed for the thumb, the index finger, and the middle finger. The TACTUATOR has three physically separated and independently controlled channels for the three fingers. Each channel has an excitable bandwidth of over 300 *Hz* and is capable of delivering signals with amplitudes from absolute detection thresholds to roughly 50 dB above detection thresholds across this whole bandwidth. The peak-to-peak range of motion was approximately 25 *mm* at very low frequencies and 90 μm at 300 *Hz*. All channels are reasonably linear, exhibit small harmonic distortion and low background noise (mainly power-line components), and generate little crosstalk between channels. Overall, the TACTUATOR is well suited for studying the tactual continuum.

In exploring the stimulus attributes that are most effective for producing a large set of clearly distinguishable elements with the TACTUATOR, it was found that subjects could naturally categorize motions over a frequency range of 2 to 300 *Hz* into three distinctive groups: slow motion (up to about 6 *Hz*), rough / fluttering motion (10 to 70 *Hz*), and smooth vibration (above about 150 *Hz*). When motions from the different categories were combined, their individual perceptual qualities could still be discerned. It was also found that subjects could exploit variations in site of stimulation, provided that the distinction between single- and multiple-finger stimulation was clear and the same waveforms were used to stimulate multiple fingers. Thus, the set of possible sites chosen consisted of each finger stimulated alone, plus all fingers stimulated together with the same waveform.

Preliminary experiments were performed with roving backgrounds to gain insight into perceptual interaction between various stimulus parameters. We then selected a set of thirty 500-*msec* waveforms to be used in our stimulus set. The choice of an initial stimulus duration of 500-*msec* was based on the trade-off between our desire to ensure that stimulus duration was not the limiting factor in perception, and our desire to keep the value of the lowest frequency sufficiently small (i.e., 2 *Hz*) and yet still be able to deliver one-cycle of a sinusoid. We did not vary stimulus duration within a given stimulus set because using stimulus durations that are easily distinguishable would make some stimuli too long (we considered stimulus duration in excess of 1 *sec* to be too long). For the 30 waveforms of the 500-*msec* stimulus set, eight single-frequency waveforms were selected, including 2 *Hz* at 35 and 44 dB SL, 4 *Hz* at 35 and 44 dB SL, 10 *Hz* at 35 dB SL, 30 *Hz* at 40 dB SL, 150 *Hz* at 44 dB SL, and 300 *Hz* at 47 dB SL. Then sixteen double-frequency waveforms and six triple-frequency waveforms were constructed by combining waveforms from different frequency ranges using the eight single-frequency waveforms. Furthermore, each of these waveforms could be applied to each of the following stimulation sites: the thumb alone, the index finger alone, the middle finger alone, or all three fingers together (stimulated with the same waveform). The largest stimulus set thus contained 120 alternatives (30 waveforms $\times$ 4 finger locations). A response code based on the time-domain sketches of the waveforms were laid out on a digitizing tablet to ease the problem of labeling associated with this relatively large stimulus set.

In order to examine the effect of stimulus duration on information transfer, two additional stimulus sets with durations of 250 *msec* and 125 *msec* were constructed. Both stimulus sets employed the same four stimulation sites as those used in the 500-*msec* stimulus set. The 30 waveforms in the 250-*msec* stimulus set were very similar to those in the 500-*msec* stimulus set except for a few changes to the low-frequency components due to shortened stimulus duration. The 19 waveforms in the 125-*msec* stimulus set were more extensively redesigned so as to keep subject's performance at a high level. There were a total of 120 and 76 alternatives in the 250-*msec* and 125-*msec* stimulus sets, respectively.

Identification experiments were performed while attempting to keep subject's performance at high levels. Three reasons motivated this strategy. First, we have some limited data indicating that it is more efficient to estimate information transfer with a stimulus uncertainty that is close to expected information transfer (thereby resulting in high identification accuracy). Second, in the case where stimulus uncertainty is high, it appears to be too time consuming to collect sufficient trials to obtain an unbiased estimate of information transfer. We have instead used a conservative estimate of IT based on percent-correct scores. According to the empirical data we have, percent-correct scores need to be high in order for this estimate to be a conservative lowerbound estimate. Third, we wanted to minimize training time associated with the three stimulus sets. In general, subjects are more motivated if their performance level is high.

Training was conducted with all three stimulus sets until each subject had completed either one perfect run of 100% correct or 3 runs with percent-correct scores of 95% or higher. With the 500-*msec* stimulus set, subjects were trained incrementally; they had to reach the training criterion on a subset of the stimulus set before new signals were introduced. With the 250- and 125-*msec* stimulus sets, subjects were trained with all alternatives in the stimulus set. One subject was also the experimenter; since she developed all the stimulus sets, it was impossible to estimate the number of hours she had been exposed to the signals. The other two subjects took between 20 to 27 *hours* to reach the training criterion with all three stimulus sets.

All subjects were then tested to estimate IT with all three stimulus sets. The estimated IT values averaged over the three subjects were 6.5 *bits* (i.e., corresponding to perfect identification of 90 *items*) for the 500-*msec* stimulus set, 6.4 bits (i.e., 84 items) for the 250-*msec* stimulus set, and 5.6 *bits* (i.e., 49 *items*) for the 125-*msec* stimulus set. In other words, there was only a 0.9 *bit* loss in IT when signal duration was reduced by a factor of four. These results seemed to suggest that a higher IT *rate* might be achieved by using signals of the shortest duration. It turned out, however, that the IT rate depended mainly on the stimulus presentation rate, not the stimulus duration alone.

The IT rate for the TACTUATOR was estimated indirectly. Using an identification paradigm with both forward and backward masking, it was found that the optimal stimulus presentation rate was approximately 3 *items/sec* independent of stimulus duration (for durations T_1 in the region $125 \leq T_1 \leq 500$ msec). A constant presentation rate suggests that constant processing time was the principal limiting factor. The estimated IT rate averaged over three subjects was approximately 12 *bits/sec*.

In additional to the above work, several important issues that warrant further investigation were identified: selection of stimulus uncertainty to maximize information transfer, definition of stimulus-set dimensionality, and possible relationships between the capability to receive motional input sequences and one's ability to deliver the same motor outputs.

VI-2 Comparison with Previous Work

The air-driven finger stimulator developed by Bliss (1961) was an impressive hardware system capable of delivering three-degrees-of-freedom motion on each of eight finger rests. Each single-finger stimulator was powered by three orthogonally-placed, push-pull bellows assemblies; thus, each finger rest could deliver motional pulses in any combination of the ±x, ±y and ±z directions. Considering the lack of computational power in the early sixties, it was remarkable that test material could be delivered automatically by using paper tapes with holes that acted as air valves. Compared with Bliss's finger stimulator, the TACTUATOR system provides fewer degrees of freedom per finger and stimulates three instead of eight fingers simultaneously. But unlike the Bliss device that could only deliver gross (and non-graded) motional pulses, the TACTUATOR is capable of delivering well-controlled arbitrary waveforms over large amplitude and frequency ranges.

The OMAR system developed by Eberhardt *et al.* (1994) and the TACTUATOR share many features. Both systems provide kinesthetic as well as vibrotactile stimulation to multiple fingers. Performance differences between the two systems have to do with their abilities to deliver motions with intermediate frequencies and amplitudes. According to Eberhardt *et al.* (1994), low-frequency movements (i.e., up to 20 *Hz*) are reproduced with high fidelity (without loading); in the case of high-frequency low-amplitude vibrations (i.e., above 100 *Hz*), the system operates "open-loop" due to "limited dynamic range" of the feedback potentiometer and drive circuit. It is not clear if OMAR can deliver stimulation between 20 and 100 *Hz*. If so, it is not obvious how the transition from closed-loop control at low frequencies to open-loop control at high frequencies can be accomplished. The TACTUATOR has a continuous excitable frequency range of over 300 *Hz* and is closed-loop controlled throughout the whole frequency range. We have shown that motions with intermediate frequencies induce characteristic perceptions that are important contributors to the overall information transmission achievable with the TACTUATOR. Other differences between OMAR and the TACTUATOR are structural. For example, in one configuration, OMAR system delivers 2-dof planar motion to a single finger using two actuators (Bernstein, 1995, personal communication).

Many studies have investigated the information transmission capabilities of the various human sensory systems. Miller (1956) summarized the early experiments involving single stimulus attributes and came to the conclusion that our capacity for processing information along uni-dimensional stimulus sets is limited by the magical number seven, plus or minus two (i.e., 2.3 to 3.2 *bits*). Pollack & Ficks (1954) were able to obtain IT values between 5 to 7 *bits* with elementary auditory displays involving six or eight stimulus aspects. These authors showed that (1) extreme subdivision of each stimulus aspect fails to produce substantial improvement in IT, and (2) similar IT values were obtained with a six-attribute, quinary-coded display and an eight-attribute, binary-coded display. The stimulus sets we used with the TACTUATOR involved many stimulus aspects, with a mainly binary coding scheme. Our IT value of 6.5 *bits* obtained with the 500-*msec* stimulus set appears impressive considering the fact that the tactual system is often thought to have a low channel capacity and, in any case, is not accustomed to receiving motional stimulation (especially at very low frequencies). It is also important to note that it is the highest IT that has been obtained with tactual artificial displays of any kind. For example, the IT obtained from a tactile display involving vibratory intensity, frequency, and contactor area was 4–5 *bits* (Rabinowitz *et al.*, 1987), and the IT obtained from the four movement channels of an artificial facial movement display was 3–4 *bits* (Tan *et al.*, 1989).

The information transfer rates obtained with several tactual communication devices can be compared. Using his air-driven finger stimulator, Bliss (1961) reported an IT rate of 4.5 *bits/sec*[1] for one experienced typist who received letters and a few punctuation symbols (4.9 *bits/symbol*) at a presentation rate of 1.32 *symbols/sec*. In an earlier one-finger experiment (Bliss, 1961), an IT rate of 4.7 *bits/sec*[2] was obtained with six subjects who identified motions in six directions (i.e., ±x, ±y, and ±z, with 2.58 *bits/movement*) at a presentation rate of 2.8 *movements/sec*. It appears that not much was gained in terms of IT rate by stimulating eight fingers instead of one. However, during the single-finger experiment, subjects were presented with three movements at a time; during the multi-finger experiment, the subject received a sequence of 130 symbols at the specified rate and responded orally by naming the symbols as they were received. The other important factor was that the 30 symbols used in the multi-finger experiment were presented in random order to form the 130-symbol sequence. In other words, the subject was not able to take advantage of any contextual cues, although letters and punctuation symbols were used. Using the display for the Vibratese language, Geldard (1957) reported that one subject was able to handle 38 *wpm*, or equivalently, 5.1 *bits/sec*.[3] Using the Optacon device (see Linvill & Bliss, 1966) and English

1. Bliss estimated information transfer as IT = IS×(1–e), but did not explain why. Using our conservative estimate of IT = IS×(1–2e), the IT rate would have been 2.6 *bits/sec* based on an IS of 4.9 *bits* and an error rate of 30%. It is questionable, though, whether IT can be reliably estimated from percent-correct scores with this relatively large error rate, because IT would depend heavily on the distribution pattern of the errors.
2. Information per presentation was computed from the stimulus-response confusion matrix. Had Bliss used IT = IS×(1–e) to compute IT based on an error rate of 23%, the IT rate would have been 3.9 *bits/sec*. This would have been a lowerbound estimate in this case.
3. The information transfer rate was estimated from word rate based on two assumptions. First, according to Shannon (1951, Fig. 4), the uncertainty for strings of eight letters (including the 26 letters of the English alphabet and space) or more has an upper bound of 2 bits / letter. For simplicity, it is assumed that the test material is longer than eight letters. Second, it is assumed that the average word length is 4 letter / word. It follows that the information content in words is 2 bits / letter × 4 letter / word, or 8 bits / word. The information rate is, therefore, 8 bits / word × 38 words / minute, or equivalently, 5.1 bits / sec.

sentences as test material, Cholewiak *et al.* (1993) reported that their best subject was able to reach a word rate of 40 *wpm*, or equivalently, 5.4 *bits/sec* (see Footnote 3). Foulke & Brodbeck (1968) reported that experienced Morse code operators were able to receive the code by electrocutaneous stimulation at a rate of 10 *wpm*, or equivalently, 1.3 *bits/sec* (see Footnote 3). The IT rate obtained from our study on Morse code reception through up-down finger motions using conversational English material was 2.7 *bits/sec* (Appendix A). The relatively lower IT rates obtained by Foulke & Brodbeck (1968) and our study on Morse code reception may be partly due to the inefficiency of the Morse code.

Overall, the IT rates measured with man-made tactual displays are much lower than the rates demonstrated by natural tactual communication methods. Reed, Durlach & Delhorne (1992) estimated that the information rate is about 7.5 *bits/sec* for tactual fingerspelling, 12 *bits/sec* for Tadoma, and 14 *bits/sec* for tactual sign language. These authors noted that whereas the information rate for fingerspelling appears to be limited by the speed at which handshapes can be made, the information rate for Tadoma and sign language appear to reflect limitations of tactual perception. Our estimated IT rate of 12 *bits/sec* appears to be quite promising. To the extent that this IT rate can be substantiated by future research using English material, we will finally be able to communicate through a tactual device at a rate comparable to that achieved by Tadoma users. Furthermore, results obtained on the perception of speech through the TACTUATOR can be used to address the role of the direct tie-in to the articulatory process to the success of Tadoma. Proponents of the motor theory of speech (e.g., Liberman & Mattingly, 1989) would argue that Tadoma is successful because of the tight coupling between the perception of speech and the feedback provided by the production of speech sounds. Thus, if similarly high IT rates for speech can be demonstrated for both the TACTUATOR and Tadoma, then such a finding would suggest that the monitoring of the articulatory process per se is likely not the key component to the success of Tadoma.

VI-3 Future Research

The immediate next step in this research is to use the TACTUATOR with English test material. Of particular interest to us is the development of a tactual automatic cueing system as a supplement to speechreading. Sounds that look alike from mouth movements can be conveyed effectively in Cued Speech (Cornett, 1967), which is a system that combines handshapes (eight for American English) representing groups of consonant phonemes, hand placements (four for American English) denoting groups of vowel phonemes, and mouth movements to present a visually distinct model of the counterpart sound code of a traditionally spoken language. It serves to distinguish visually sounds that are ambiguous through speechreading alone for individuals who are deaf or hard-of-hearing. For example, different handshapes, combined and synchronized with the relevant visible mouth movement, are employed in Cued Speech to convey look-alike consonant sounds /b/, /m/, and /p/. Although Cued Speech is normally received visually, Delhorne, Besing, Reed, & Durlach (1990) have shown that manual cues associated with Cued Speech can be received tactually and combined effectively with visual speechreading. Advances in the development of automatic speech recognition systems make it possible to obtain classes of phonemes that can be displayed through a tactual stimulator to the human hand. Linking the TACTUATOR to such an automatic phoneme-classification system provides an opportunity to study many issues. First, a basic signal set for the 8×4 hand postures used in Manual Cued Speech needs to be devised. These signals could be “natural”, meaning that they mimic the actual handshapes and hand positions used in delivering Cued Speech. They can also be designed exclusively on the basis of the perceptual distinctiveness of the signals. Second, individuals who are skilled at delivering and/or receiving Cued Speech, as well as those who have no prior knowledge of Cued Speech, can be trained on such a tactual cueing system. By using either natural or perception-based signals, we can study whether the experience in outputting Cued Speech, or the knowledge of the code itself, affects one’s ability to receive Cued Speech tactually. Third, because manual cueing has to be synchronized with the visual presentation of mouth movements, the issue of how stimulus duration affects the perception of tactual cueing needs to be addressed. Fourth, subjects can be trained to “chunk” individual cues into meaning signals. Finally, the information-transfer rate achievable with such an automatic tactual cueing system can be assessed using continuous speech material. This rate will be compared with the estimate of 12 *bits/sec* we have obtained in this thesis work.

The three issues summarized in Chap. V are closely related to this thesis work and warrant further investigation.

Finally, the existence of a general "additivity" law needs be investigated. In this thesis, we used an empirically-based conservative estimate based on percent-correct scores to estimate information transfer associated with a relatively large number of stimulus alternatives. In general, it is time-consuming to obtain an unbiased estimate of information transmission with a multi-dimensional stimulus set which usually contains a large number of stimulus alternatives. Due to perceptual interactions among the dimensions, the information-transfer obtained with an *M*-dimensional identification experiment is usually smaller than the sum of the information transfers obtained with the *M* corresponding uni-dimensional experiments. It is a tempting goal to try to estimate multi-dimensional information transfer from unidimensional information transfers, because of the difference in the total number of trials needed for an unbiased estimate of information transfer. Without loss of generality, let us assume that each of the *M* dimensions contains *k* alternatives. The total number of trials needed for the *M*-dimensional experiment would be $5\times(M\times k)^2$. The total number of trials needed for the *M* uni-dimensional experiments is $M\times(5\times k^2)$. When *M* is relatively large, the saving in the total number of trials can be substantial. Durlach, Tan, MacMillan, Rabinowitz, & Braida (1989) proposed the hypothesis that multi-dimensional IT is always equal to the sum of the corresponding unidimensional ITs, independent of whether the variables are independent, provided only that the background parameters are roved over the appropriate ranges in the unidmensional experiments. Using a multi-dimensional tactile display, limited supporting data are available from Tan *et al.* (1989) using a multi-dimensional tactile display. One-dimensional identification experiments with *fixed* and *roving* background as well as four-dimensional identification experiments were performed. Transmitted information, averaged over subjects, was 3.3 *bits* for four-dimensional identification, 6.5 *bits* for the sum of the four uni-dimensional experiments with *fixed* background, and 3.4 *bits* for the sum of the four uni-dimensional tests with *roved* background. More recently, Campbell (1993) provided additional data by studying tonal stimuli. Transmitted information, averaged over subjects, was 3.5 *bits* for three-dimensional identification, 5.0 *bits* for the sum of the three one-dimensional experiments with *fixed* background, and 3.4 *bits* for the sum of the three uni-dimensional experiments with *roved* background. Thus these data support a general additivity law relating multidimensional and unidimensional resolution measures when background variables are roved over the appropriate ranges. This issue is closely related to the definition of "dimensionality."

This page is intentionally left blank.