Information transmission with a multifinger tactual display

HONG Z. TAN, NATHANIEL I. DURLACH, CHARLOTTE M. REED, and WILLIAM M. RABINOWITZ
Massachusetts Institute of Technology, Cambridge, Massachusetts

In this work, the tactual information transmission capabilities of a tactual display designed to provide stimulation along a continuum from kinesthetic movements to cutaneous vibrations are assessed. The display is capable of delivering arbitrary waveforms to three digits (thumb, index, and middle finger) within an amplitude range from absolute detection threshold to about 50 dB sensation level and a frequency range from dc to above 300 Hz. Stimulus sets were designed at each of three signal durations (125, 250, and 500 msec) by combining salient attributes, such as frequency (further divided into low, middle, and high regions), amplitude, direction of motion, and finger location. Estimated static information transfer (IT) was 6.5 bits at 500 msec, 6.4 bits at 250 msec, and 5.6 bits at 125 msec. Estimates of IT rate were derived from identification experiments in which the subject's task was to identify the middle stimulus in a sequence of three stimuli randomly selected from a given stimulus set. On the basis of the extrapolations from these IT measurements to continuous streams, the IT rate was estimated to be about 12 bits/sec, which is roughly the same as that achieved by Tadoma users in tactual speech communication.

This work was motivated by our interest in using the sense of touch as an alternative communication channel. The potential to receive information tactually is well illustrated by some natural (i.e., nondevice-related) methods of tactual speech communication. Particularly noteworthy is the so-called Tadoma method, which is employed by some individuals who are both deaf and blind. In Tadoma, one places a hand on the face and neck of a talker and monitors a variety of actions associated with speech production. Our previous research has documented the remarkable abilities of experienced Tadoma users (see, e.g., Reed et al., 1985); these individuals can understand everyday speech at high performance levels, allowing rich two-way conversation with both familiar and novel talkers. In contrast, attempts to develop artificial tactual speech communication devices have had only limited success, with none achieving performance comparable with that demonstrated by Tadoma (see, e.g., Reed, Durlach, Delhorne, Rabinowitz, & Grant, 1989).

Such performance differences may be partly attributed to the fact that Tadoma users have had more experience with the method than any laboratory subjects have had with artificial tactual aids. However, several studies show

that (1) long-term users of artificial tactile aids demonstrate limited improvements in their speechreading abilities and (2) normal subjects with simulated deafness and blindness can achieve a certain level of skill in discriminating and identifying speech through the Tadoma method after a modest amount of training.

Reed and Delhorne (1995) tracked the performance of long-term (i.e., up to 4 years and 2 months) users of artificial tactile aids and found that the relative gain (the difference between aided and unaided speechreading scores normalized by the maximum possible improvement) observed for the reception of words in isolated sentences averaged roughly 25%. No demonstrable speech-reception ability through the tactile aids alone was observed with these subjects. Therefore, the benefits of current wearable tactual devices to speechreading appear to be limited, even after subjects have worn these aids for several years.

Studies with inexperienced users of the Tadoma method suggest that some ability to receive speech through the tactual sense alone can be acquired with limited amounts of training. In an analytic study of the Tadoma method, Reed, Rubin, Braida, and Durlach (1978) reported that 2 normal subjects with simulated deafness and blindness performed at least as well as an experienced Tadoma user on the discrimination of words, vowels, nonsense syllables, and consonants. In another study, Reed, Doherty, Braida, and Durlach (1982) reported that, with modest training of 100 h each, two inexperienced Tadoma users performed in a way comparable with that of an experienced Tadoma user on consonant and vowel identification tasks. Two inexperienced Tadoma users also demonstrated some ability to receive sentences constructed from a limited vocabulary. Therefore, the exceptional ability of the experienced Tadoma user does not seem to be due to the ability to process

segmental speech units, but rather to the ability to comprehend a stream of connected speech and the knowledge of a large *tactile vocabulary*. The differences in performance between Tadoma users and users of artificial aids may reflect differences in the amount of information available in a talking face, as compared with that available from current artificial tactual aids.

One characteristic common to most previous tactual devices, such as the vocoder (J. M. Pickett & P. H. Pickett, 1963), the Optacon (Linvill & Bliss, 1966), the Tickle-Talker (Blamey & Clark, 1985, 1987; Cowan, Blamey, Sarant, Galvan, & Clark, 1991), and Tactaid II and Tactile VII (Reed & Delhorne, 1995), concerns the nature of the output display. These displays are composed of multiple stimulators that deliver a relatively homogeneous cutaneous stimulation, in the sense of exhibiting few distinctive perceptual qualities. In contrast, Tadoma is perceptually rich through its simultaneous display of various stimulation qualities that engage the kinesthetic as well as the cutaneous sensory systems (see Reed et al., 1985).

Recognition of the need for richer tactual displays is now prevalent. An artificial mechanical-face display built around a model plastic skull has shown promise in conveying information important in Tadoma (Leotta, Rabinowitz, Reed, & Durlach, 1988; Rabinowitz, Henderson, Reed, Delhorne, & Durlach, 1990; Reed et al., 1985). A more general display for studying haptic perception by the hand (the OMAR system, developed by Eberhardt, Bernstein, Barac-Cikoja, Coulter, & Jordan, 1994), like the display discussed in this report, was designed to deliver kinesthetic as well as cutaneous stimulation to one or more fingers.

The present research is concerned with exploring the capabilities of the kinesthetic and cutaneous sensory systems for information transmission. Using the Tactuator, a multifinger display (Tan & Rabinowitz, 1996), stimuli can be presented on a continuum from low-frequency large-amplitude motions to high-frequency small-amplitude vibrations. The goal of stimulus design was to achieve as high an information transfer (IT) as possible with minimal training of the human observers. It is well known from Miller's (1956) classical paper that, given a one-attribute[1] stimulus set, IT for human observers is limited to roughly 2.3–3.2 bits (corresponding roughly to 7±2 perfectly identifiable stimuli). It is also well known that this limit can be overcome by employing multi-attribute stimulus sets (see, e.g., Pollack & Ficks, 1954). The IT that can be ultimately achieved depends on the perceptual interactions among the various attributes, with greater independence being among the attributes leading to higher IT. Thus, a key principle in designing stimulus sets for high information transmission is to recruit as many stimulus attributes as possible with as little perceptual interaction among them as possible. Since any movement signal can be viewed in the frequency domain as the sum of sinusoidals through Fourier transform, sinusoidal waveforms were used as the basic building blocks. Thus, possible attributes for stimulus construction were amplitude and frequency of waveforms, site of stimulation, and onset direction of motion.

In general, identification performance may depend not only on the characteristics of the stimulus set, but also on the extent to which the response set and the mapping between stimuli and responses are *natural* (i.e., on the degree of *stimulus–response compatibility*; see, e.g., Alluisi, Muller, & Fitts, 1957; and Proctor & Reeve, 1990, for a review). Although much work has been done in the area of stimulus–response compatibility, such studies have been mainly concerned with spatial congruency between stimuli and responses or with verbal versus motor responses for highly familiar stimuli, such as digits. Thus, we had no general guidelines to follow in designing our response codes. After some informal experimentation, graphic-based response icons were developed.

Measurements of information transmission through the Tactuator device were obtained in two sets of experiments. In Experiment 1, three stimulus sets with stimulus durations of 500, 250, and 125 msec, respectively, were designed. Estimates of IT per stimulus (i.e., *static* IT) were then obtained, using absolute identification paradigms.

In Experiment 2, estimates of IT rate (i.e., *dynamic* IT) were assessed, using the same three stimulus sets and a masking paradigm. Ideally, IT rate should be assessed by presenting continuous streams of signals with varying amounts of IT per signal and varying presentation rates. The subjects' task would be to report back the signals received in the proper order. Very often, owing to limitations on motor output speed, it is not possible for a subject to maintain an output rate consistent with the input rate. Therefore, a representation of the input signal stream needs to be stored in short-term memory until the subject is able to record the response. The difficulty with this approach, however, is the length of time required to train a subject. The first step in such training is for the subject to learn to recognize the individual signals that make up the continuous presentation stream. This process takes relatively little time, provided that the stimulus and response sets are well designed. The next step is to become highly proficient in processing the basic signals, so that recognition is almost automatic (i.e., recognition time is minimized). This prepares the subject for the next stage, that of organizing basic signals into meaningful *chunks* that can be stored in short-term memory and later retrieved. According to Miller (1956), the span of immediate memory, or the number of chunks people can recall correctly, is about seven items in length. However, there is no limit on the amount of information each chunk can contain. Therefore, the goal of the third step is to maximize the chunk size in bits/chunk. This process can take many years. There is also evidence that reaching a temporary plateau in performance does not necessarily imply completion of the chunking process. For example, Bryan and Harter (1899) showed that students of Morse code reached several plateaus with regard to their ability to receive the code. The plateaus in the code reception curves were interpreted as evidence that a student of telegraphy first learned to receive indi-

vidual letters, then developed the skills to receive common words as the basic units, and eventually, after many years of full-time practice on the railway, learned to receive short phrases.

Our subjects were not trained extensively on chunking, for two reasons. First, our signals were abstract and, therefore, were not particularly well suited for chunking.[2] Second, it is impossible to predict the length of time that would be required to complete the training process (allowing for the possibility of several performance plateaus). In view of these problems, our strategy in assessing IT rate was to measure identification performance in the context of other signals, using an identification paradigm with both forward and backward masking. Estimates of IT rates were obtained by (1) sequencing three random stimuli, (2) having the subject identify only the middle stimulus, and (3) extrapolating this IT to that for continuous streams.

GENERAL METHOD

Apparatus

The experimental apparatus (the Tactuator) consists of three independent, single, point-contact, one-degree-of-freedom actuators interfaced individually with the fingerpads of the thumb, the index finger, and the middle finger (see Figure 1). The motion trajectory for the thumb is perpendicular to that for the index and middle fingers, thereby maintaining an approximately natural hand configuration. The range of motion provided by the display for each digit is about 26 mm. All the motions begin and end with the three digits at the middle of their respective range of motion. Each digit can thus be moved either outward (i.e., extension, defined as the "+" position) or inward (i.e., flexion, or the "−" position).

Each actuator utilizes a disk-drive head-positioning motor augmented with angular position feedback from a precision rotary variable differential transformer that has a response bandwidth (-3 dB) of 1 kHz and essentially infinite resolution, owing to its electromagnetic coupling. A floating-point DSP system with 16-bit analog-to-digital and digital-to-analog converters provides real-time positional control, using a digital PID controller. The overall system performance is well suited for psychophysical studies of the tactual sensory system for several reasons. First, each movement channel has a *continuous* frequency response from dc to 400 Hz (beyond which the disk-drive motor hardly moves). The magnitude and phase delay can be well modeled by a second-order system with no zeros. This means that the Tactuator can deliver stimulation in the kinesthetic (i.e., low-frequency) and cutaneous (i.e., high-frequency) ranges, as well as in the mid-frequency range. Second, each channel has an overall dynamic range of 96 dB. The maximum achievable
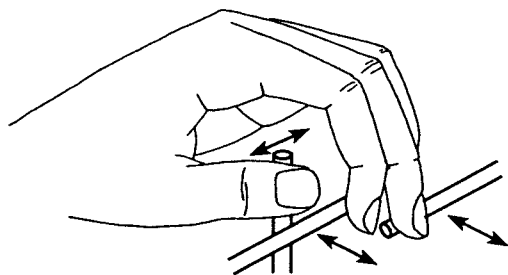

Figure 1. Schematic drawing illustrating finger placement on the Tactuator and motion trajectories.

displacement (limited by the built-in mechanical stops in the motor assembly) is 82 dB μm peak (i.e., 26 mm peak to peak) at low frequencies (e.g., several hertz). The minimum perceivable displacement (limited by the human detection threshold) is -14 dB μm peak (i.e., 0.4 μm peak to peak) at around 250 Hz (Tan & Rabinowitz, 1996, Figure 11). It thus follows that the overall system dynamic range (across frequencies) is 82 dB μm peak $- (-14$ dB μm peak), or equivalently, 96 dB. Third, across the frequencies of dc to 300 Hz, an amplitude range of at least 47 dB *per frequency* can be achieved, thereby encompassing the perceptual range from gross motion to vibration. This follows from noting that (1) measured detection threshold is about 35 dB μm peak at 2 Hz (Tan & Rabinowitz, 1996, Figure 11), resulting in a 47-dB range in sensation level (SL; i.e., 82 dB μm peak $- 35$ dB μm peak) at low frequencies, and (2) the system saturates at 39 dB μm peak at around 250 Hz, providing a stimulation range of threshold to 53 dB SL [i.e., 39 dB μm peak $- (-14$ dB μm peak)] at high frequencies. This range is adequate for psychophysical studies, because stimulation levels exceeding 50–55 dB SL can induce discomfort and fatigue (Verrillo & Gescheider, 1992). Fourth, measurements with single and multiple frequency inputs at various levels indicate that each channel is highly linear, distortion (in terms of harmonics) is low, and interchannel cross talk (in terms of induced response on a channel owing to stimulation of another channel) is small. This allows high-fidelity delivery of waveforms of arbitrary frequency content and stimulation level (e.g., 26-mm low-frequency motion with superimposed 30-μm high-frequency vibration).

The effect of *loading* (i.e., resting the finger lightly on the actuator's moving bar) was measured for 1 subject. It was found that loading reduced the magnitude of stimulation by an average of 1.5 dB at 2 Hz, 2.7 dB at 20 Hz, and 0.1 dB at 200 Hz. This did not pose a significant problem for the experiments reported here, because the stimuli are generally strong (i.e., at least 17 dB SL) and are sparsely spaced along the amplitude axis (i.e., the amplitudes of two stimuli of the same frequency differ by at least 9 dB). Loading also increased harmonic distortions by a modest amount, especially for stimulation at around 30 Hz, which is roughly the system's resonant frequency. More detailed description of the Tactuator and its performance characteristics can be found in Tan and Rabinowitz (1996).

Subjects

Three subjects (S_1, S_2, and S_3) participated in the experiments. At the time of this study, S_1 (the leading author) was a 30-year-old female graduate student at MIT, S_2 was a 42-year-old male, and S_3 was a 20-year-old male undergraduate student at MIT. All 3 subjects had participated in a previous study of the reception of Morse code through tactual and auditory stimulation (Tan, Durlach, Rabinowitz, Reed, & Santos, 1997). S_1 had previously designed and performed other tactual psychophysical experiments and was highly familiar with the stimulus sets used in this study, owing to her role in their development and design. S_2 and S_3 had not participated in any other psychophysical experiments prior to the Morse code study. All the subjects were right-handed, with no known tactual impairments of their hands.

Basic Procedure

During all the experiments, the Tactuator was visually hidden from the subjects. The subjects wore earplugs and earphones, with pink noise to eliminate auditory cues. (The Tactuator produces little audible noise, except in the neighborhood of 300 Hz.)

Both Experiment 1 and Experiment 2 made use of a standard absolute identification (AI) paradigm involving a set of k stimuli S_i ($1 \leq i \leq k$), a set of k responses R_j ($1 \leq j \leq k$), and a one-to-one mapping between the stimuli and the responses. The stimuli were presented one at a time in random order with equal a priori probabilities, and the subject was instructed to respond to each stimulus presentation with the response defined by the one-to-one mapping.

Without loss of generality, we can assume that the stimuli and responses are labeled in such a way that R_i is the correct response to S_i. In some experiments, the subject was provided with trial-by-trial correct-answer feedback.

Data Analysis

Results of AI experiments are typically summarized in terms of IT. This quantity measures the increase in information about the signal transmitted that results from knowledge of the received signal. For a particular stimulus–response pair (S_i, R_j), it is given by $\log_2[P(S_i/R_j)/P(S_i)]$, where $P(S_i/R_j)$ is the probability of S_i given R_j, and $P(S_i)$ is the a priori probability of S_i. The average IT is given by the weighted sum of $\log_2[P(S_i/R_j)/P(S_i)]$:

$$IT = \sum_{j=1}^{k} \sum_{i=1}^{k} P(S_i, R_j) \log_2\left(\frac{P(S_i/R_j)}{P(S_i)}\right)$$

$$= \sum_{j=1}^{k} \sum_{i=1}^{k} P(S_i, R_j) \log_2\left(\frac{P(S_i, R_j)}{P(S_i)P(R_j)}\right),$$

where $P(S_i/R_j)$ is the joint probability of stimulus S_i and response R_j, and $P(R_j)$ is the probability of R_j.

Given data from an AI experiment, a first-order $k \times k$ stimulus–response confusion matrix can be constructed, in which the entry in row i and column j specifies the number of times stimulus S_i led to response R_j. The maximum likelihood estimate of IT, IT_{est}, can be computed by approximating underlying probabilities with frequencies of occurrence:

$$IT_{est} = \sum_{j=1}^{k} \sum_{i=1}^{k} \left(\frac{n_{ij}}{n}\right) \log_2\left(\frac{n_{ij} \cdot n}{n_i \cdot n_j}\right),$$

where n is the total number of trials in the experiment, n_{ij} is the number of times the joint event (S_i, R_j) occurs, and $n_i = \sum_{j=1}^{k} n_{ij}$ and $n_j = \sum_{i=1}^{k} n_{ij}$ are the row and column sums. These quantities can all be derived directly from the confusion matrix obtained in the AI experiment. Unfortunately, however, IT_{est} is not only subject to statistical fluctuations but is also a biased estimate: It tends, for a limited number of trials, to overestimate IT. Furthermore, the magnitude of the bias tends to exceed the magnitude of the fluctuations (Rabinowitz, Houtsma, Durlach, & Delhorne, 1987; Rogers & Green, 1954).

Although the bias in IT_{est} can be effectively reduced to insignificance by collecting a sufficient number of trials, it is not always feasible or practical to do so when an experiment involves a large number of stimulus alternatives. Two ways of dealing with this issue have been proposed in the past. According to Miller (1954), a useful first-order correction for the bias, provided $n > 5k^2$, is to subtract $\Delta IT = (k - 1)^2/(2n\ln2)$ from IT_{est}. However, as Miller (1954) also pointed out, ΔIT often results in too large a correction when $n < 5k^2$, and many of the n_{ij} $(i \neq j)$ values are near zero. Houtsma (1983) used computer simulations to estimate the asymptotic value of IT_{est} from limited experimental data, but the method does not work well when there are large differences among the amounts of information each stimulus attribute contributes to the overall IT in a stimulus set involving many attributes (Tan, 1988). In our experiments, the number, k, of alternatives in the stimulus set was large (i.e., $57 \leq k \leq 120$), and collecting an appropriately large number, n, of trials would have been extremely time consuming (i.e., $16,245 \leq n = 5k^2 \leq 72,000$). Therefore, the strategy described below was used.

Empirical observations indicate that, for relatively large k (e.g., $57 \leq k \leq 120$) and low error rate e (i.e., $e \leq 5\%$), $IT/IS \geq (1 - 2e)$.[3] (The quantity IS, the information in the stimulus set, is equal to $\log_2 k$ when all stimuli are equally likely.) In other words, the ratio of IT over IS deviates from unity by less than twice the error rate in almost all cases (and frequently deviates by less than the error rate). Therefore, $(1 - 2e) \times IS$ can be viewed as a lower-bound estimate of IT.

In Experiments 1 and 2, where $57 \leq k \leq 120$, efforts were made to ensure high performance levels, so that $e \leq 5\%$. A conservative (i.e., a lower-bound) estimate of IT from percent-correct scores $[IT_{pc} = (1 - 2e) \times \log_2 k]$, was then computed to characterize the experimental outcome.

The maximum-likelihood estimate IT_{est} was used as an upper-bound estimate of IT. To the extent that IT_{pc} and IT_{est} are numerically close (as in the case where performance level is near perfect), either quantity can be taken as a good estimate of IT. A related quantity, 2^{IT}, is interpreted as the number of stimulus categories that can be correctly identified. It is an abstraction, since 2^{IT} is not necessarily an integer. The values of IT and 2^{IT} are used interchangeably to characterize the outcome of an AI experiment.

EXPERIMENT 1
Static Information Transfer

Experiment 1 was concerned with the development of three stimulus sets and the measurement of IT for these sets. These three sets are referred to as the 500-, 250-, and 125-msec stimulus sets, reflecting the differences in signal duration. Emphasis was placed on the 500-msec set in terms of its construction and the training of subjects.

Method

Stimuli

A set of preliminary experiments was conducted to determine which stimulus attributes could be used to construct a relatively large stimulus set with easily identifiable stimuli. (Details of the preliminary experiments and results can be found in Tan, 1996.) Three stimulus sets, with signal durations of 500, 250, and 125 msec, were developed. The attributes employed were frequency, amplitude, finger location, and (for the 125-msec stimulus set only) onset direction of motion. Each stimulus set was developed in two stages. First, a set of movement waveforms was designed by employing the attributes of frequency, amplitude, and (for the 125-msec stimulus set) onset direction of motion. These waveforms were then combined with the attribute of finger location, to create the full stimulus set. Because many of the design choices were made with the longest duration stimuli, the construction of the 500-msec set will be described in detail for each of the attributes.

The 500-msec stimulus set. As to *frequency*, it was found that subjects could naturally categorize motions over the frequency range of dc to 300 Hz into three perceptually distinct categories: *slow motion* (up to about 6 Hz), *fluttering motion* (about 10–70 Hz), and *smooth vibration* (above about 150 Hz). When components from different frequency regions were combined, the sensations associated with the single-frequency components were still discernible. Therefore, multicomponent stimuli were formed by using waveforms containing sinusoids (varying in both frequency and amplitude) from the three frequency regions (denoted by F_L, F_M, and F_H for low, middle, and high frequencies, respectively). Preliminary experiments on frequency identification indicated that subjects could reliably identify two frequencies within each of the three frequency regions (see Tan, 1996, for details). In the final design, the value of F_L was 2 or 4 Hz, the value of F_M was 10 or 30 Hz, and the value of F_H was 150 or 300 Hz.

As to *amplitude*, preliminary experiments on amplitude identification indicated that the number of amplitudes that could be reliably identified was dependent on the frequency of the waveform (see Tan, 1996, for details). Although subjects could reliably identify two amplitudes for the F_L component, they were unable to identify different amplitudes for F_M and F_H components. In the final design, the amplitude for each F_L component was 35 or 44 dB SL. The amplitude for each F_M or F_H component was fixed, because the perceptual
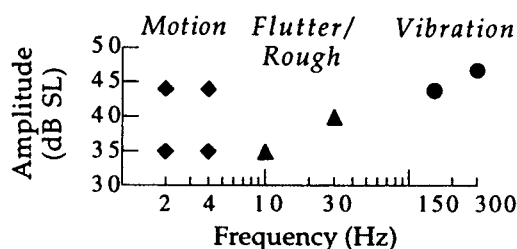
Figure 2. The eight 500-msec single-frequency waveforms. Each symbol denotes a sinusoidal signal at the specified frequency and amplitude (in dB SL). The diamonds, triangles, and circles represent signals in the F_L, F_M, and F_H frequency regions, respectively.

qualities of the F_M and F_H components were not sufficiently independent of amplitude.

On the basis of these preliminary results, eight 500-msec single-frequency sinusoidal waveforms were chosen (see Figure 2). These signals were then used to form 16 double-frequency waveforms and 6 triple-frequency waveforms by combining single-frequency signals from different frequency regions. Note that not all combinations of single-frequency components were used. For example, a 4-Hz signal was never combined with a 10-Hz signal, because the former was found to interfere with the perception of the latter. Whenever F_M and F_H components were combined, only the 300-Hz signal was used, because the F_M components were found to interfere with the identification of F_H. Finally, some amplitudes were adjusted, in order to balance the relative strengths of different signal components and to minimize fatigue from excessively strong signals. Table 1 provides a complete list of the 30 waveforms designed for the 500-msec stimulus set.

The 30 waveforms have distinctive perceptual qualities. The 2- and 4-Hz signals are perceived as slow motions, with one or two cycles at small or large amplitudes. The 30-Hz signal is very rough and seems to be beating on the fingertip. The 10-Hz signal gives a mild fluttering sensation. The 150-Hz vibration is relatively diffuse and of low pitch, whereas the 300-Hz vibration is more focused and of higher pitch. When two or three frequencies are combined, the sensations associated with the single-frequency components can still be discerned.

As to *location*, the subjects rarely made errors in identifying the stimulated digit when only one of the three digits was stimulated.

When multiple digits were simultaneously stimulated with the same waveform, the subjects were unable to distinguish two-digit signals from three-digit signals. Finally, when multiple digits were stimulated with different waveforms of the same duration, the subjects could not reliably associate a waveform with a digit correctly. In the final design, four stimulation sites were selected: Either one of the three digits (thumb, index finger, or middle finger) was stimulated, or all three of them were stimulated with the same waveform. The same set of 30 waveforms was used to stimulate any of the fingers or all the fingers. The combination of four finger configurations (i.e., only one of the digits plus all three digits) and 30 waveforms resulted in a total of 120 alternatives in the 500-msec stimulus set.

The 250-msec stimulus set. The 250- and 125-msec stimulus sets were designed using similar principles. The 250-msec stimulus set contained 30 waveforms that were very similar to those in the 500-msec set, except that the frequency of the 4-Hz components was raised to 6 Hz and that of the 10-Hz components was raised to 15 Hz (to ensure that the two F_L values could be easily discriminated and that the higher F_L value could be discriminated from the lower F_M value). A complete listing of the 30 waveforms for the 250-msec set is shown in Table 2. The same four finger configurations were used: all three digits, thumb alone, index finger alone, and middle finger alone. Therefore, there was again a total of 120 alternatives in the 250-msec stimulus set.

The 125-msec stimulus set. With a signal duration of 125 msec, the subjects could no longer reach the performance criterion of ≥95% (i.e., $e \leq 5\%$) with a set of 30 waveforms similar to those used in the 500- or 250-msec waveform sets. In order to keep performance at a high level, only one frequency value was used in each of the three frequency ranges. Onset direction of motion was introduced as an additional signal attribute. This attribute was effective for F_L components, but not for F_M or F_H components.[4] In addition, the direction attribute was only effective when movement was restricted to either side of each digit's resting position (i.e., half a cycle of a sinusoidal motion) for F_L components. The resultant 125-msec waveform set contained 19 waveforms, as is shown in Table 3. A negative sign indicates that movements started in a direction that corresponds to finger flexion. The default was to start movements in the finger-extension direction. Again, the same four finger configurations were used: all three digits, thumb alone, index finger alone, and middle finger alone. Therefore, there was a total of 76 alternatives in the 125-msec stimulus set.

The preliminary work that led to the construction of the three stimulus sets was exploratory, rather than exhaustive, in nature. A more rigorous approach toward selecting stimuli would have involved measuring the one-attribute IT associated with each stimulus

Table 1
The 30 Waveforms for the 500-msec Stimulus Set

Single frequency:	F_L	(2,35)	(2,44)
		(4,35)	(4,44)
	F_M	(10,35)	(30,40)
	F_H	(150,44)	(300,47)
Double frequency:	$F_L + F_M$	(2,35) + (10,35)	(2,35) + (30,40)
		(2,44) + (10,40)	(2,44) + (30,40)
		(4,35) + (30,40)	(4,44) + (30,40)
	$F_L + F_H$	(2,35) + (150,44)	(2,35) + (300,44)
		(2,44) + (150,44)	(2,44) + (300,44)
		(4,35) + (150,44)	(4,35) + (300,44)
		(4,44) + (150,44)	(4,44) + (300,47)
	$F_M + F_H$	(10,35) + (300,44)	(30,40) + (300,44)
Triple frequency:	$F_L + F_M + F_H$	(2,35) + (10,35) + (300,44)	(2,35) + (30,40) + (300,47)
		(2,44) + (10,40) + (300,44)	(2,44) + (30,40) + (300,47)
		(4,35) + (30,40) + (300,47)	(4,44) + (30,40) + (300,47)

Note—F_L, F_M, F_H refer to low, medium, and high frequency ranges. The units given within each pair of parentheses are the frequency in hertz and the amplitude in decibels (SL), respectively.

Table 2
The 30 Waveforms For The 250-msec Stimulus Set

Single frequency:	F_L	(2,35)	(2,44)
		(6,35)	(6,44)
	F_M	(15,35)	(30,40)
	F_H	(150,44)	(300,47)
Double frequency:	$F_L + F_M$	(2,35) + (15,35)	(2,35) + (30,40)
		(2,44) + (15,40)	(2,44) + (30,40)
		(6,35) + (30,40)	(6,44) + (30,44)
	$F_L + F_H$	(2,35) + (150,44)	(2,35) + (300,44)
		(2,44) + (150,44)	(2,44) + (300,44)
		(6,35) + (150,44)	(6,35) + (300,44)
		(6,44) + (150,44)	(6,44) + (300,47)
	$F_M + F_H$	(15,35) + (300,44)	(30,40) + (300,44)
Triple frequency:	$F_L + F_M + F_H$	(2,35) + (15,35) + (300,44)	(2,35) + (30,40) + (300,47)
		(2,44) + (15,40) + (300,44)	(2,44) + (30,40) + (300,47)
		(6,35) + (30,40) + (300,47)	(6,44) + (30,40) + (300,47)

Note—F_L, F_M, F_H refer to low, medium, and high frequency ranges. The units given within each pair of parentheses are the frequency in hertz and the amplitude in decibels (SL), respectively. Signals different from those in the 500-msec set are given in boldface.

attribute and characterizing the interattribute interactions among attributes (in pairs, triplets, etc.) in terms of the difference between multiattribute IT and the sum of one-attribute ITs for all the signal attributes involved. Although this approach is more systematic, it appeared to be extremely time consuming. An alternate approach was employed that involved selection of a few representative experimental conditions that aided in the selection of stimulus attributes and parameters. Throughout the selection process, another important criterion was that the stimuli be easy to learn and allow for high performance level (>95% correct), so as to facilitate efficient data processing (see the Data Analysis section in General Method section). Therefore, IT obtained with the three stimulus sets described above should be viewed as providing a lower bound for the static IT that can be ultimately achieved with the Tactuator.

Responses

Intuitively, it seemed that response codes should reflect the underlying structure of the stimuli: For example, each response should consist of two parts, one corresponding to stimulation site and one to stimulating waveform. A graphical response code appeared to be more manageable than text or numerical labels. Accordingly, graphic icons corresponding to the 30 waveforms were laid out as circular buttons on a digitizing tablet along with four icons, M, I, T, and ALL, corresponding to the middle finger, index finger, thumb, and all the

digits, respectively (see Figure 3 for the response codes for the 500-msec stimulus set). The icons depict the actual displacement versus time traces of the waveforms, except for the 150- and 300-Hz waveforms, which were coded by color (blue and red dots, respectively, in the actual response tablet) because they did not reproduce well at the given scale. In general, the component with the lowest frequency was the same across a row of waveform icons, and the component with the highest frequency was the same over a column of waveform icons. Some exceptions were made in order to contain the waveform icons within a relatively small area (for ease of visual search). The subjects used a stylus to press the appropriate response icon. A DEL icon was available for deleting a response, and an ENTER icon was used to terminate a trial.

The response codes for the 250- and 125-msec stimulus sets were similar to those for the 500-msec set, except for changes necessary to reflect the different signal components.

Procedure

All the subjects were trained and tested with the 500-msec stimulus set, followed by the 250-msec set, and finally with the 125-msec set. At the beginning of each experimental session, the subjects ran a practice program in which they could select any response icons and then feel the corresponding waveforms. The practice program could be terminated by the subjects when they were ready. For both

Table 3
The 19 Waveforms for the 125-msec Stimulus Set

Single frequency:	F_L	(4,35)	(4,44)
		–(4,35)	–(4,44)
	F_M	(30,40)	
	F_H	(300,47)	
Double frequency:	$F_L + F_M$	(4,35) + (30,40)	(4,44) + (30,40)
		–(4,35) + (30,40)	–(4,44) + (30,40)
	$F_L + F_H$	(4,35) + (300,47)	(4,44) + (300,47)
		–(4,35) + (300,47)	–(4,44) + (300,47)
	$F_M + F_H$	(30,40) + (300,47)	
Triple frequency:	$F_L + F_M + F_H$	(4,35) + (30,40) + (300,47)	
		(4,44) + (30,40) + (300,47)	
		–(4,35) + (30,40) + (300,47)	
		–(4,44) + (30,40) + (300,47)	

Note—F_L, F_M, F_H refer to low, medium, and high frequency ranges. The units given within each pair of parentheses are the frequency in hertz and the amplitude in decibels (SL), respectively. A negative sign indicates a reversal in movement direction. See text for further details.
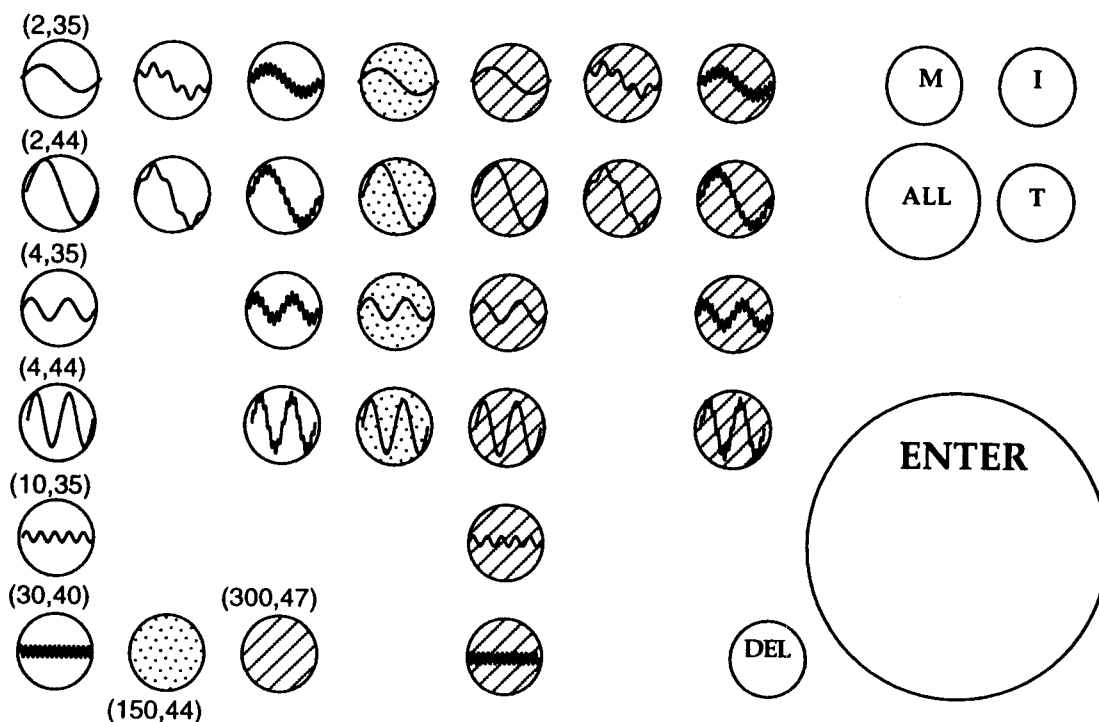
(2,35)
(2,44)
(4,35)
(4,44)
(10,35)
(30,40)   (300,47)
(150,44)

M   I
ALL   T
ENTER
DEL

Figure 3. Layout of responses for experiments using the 500-msec stimulus set. The eight single-frequency waveforms, as listed in Table 1, are labeled. The dot and line patterns represent the blue and red colors used in actual icons for the 150- and 300-Hz waveforms, respectively.

training and testing, the standard AI paradigm with trial-by-trial correct-answer feedback was employed. The number of trials per experimental run was held at twice the number of the stimuli in a stimulus set and was, thus, different for the 500- and 250-msec sets than for the 125-msec set.

There were three differences between the paradigms used for training and those used for testing. During training, each stimulus alternative was presented an equal number of times per run (i.e., randomization without replacement), to ensure equal exposure to all the signals in the stimulus set. Thus, stimulus uncertainty decreased as a function of number of trials. During testing, stimuli were presented with equal a priori probabilities on each trial (i.e., randomization with replacement). Thus, stimulus uncertainty remained constant throughout an experimental run. The second difference was that, during training, the subjects were given the option of skipping trials by not entering a valid response. Feedback was not provided on these trials, and the stimuli were presented again during the same experimental run. During testing, the subjects were required to provide a response to all the stimuli on their initial presentation. The third difference was concerned with performance criteria. During training, the subjects were required to repeat experimental runs with the same condition until either (1) a perfect run of 100% correct was achieved or (2) three runs with percent-correct scores of 95% or higher (not necessarily consecutively) were obtained.[5] During testing, three runs per stimulus set were collected for each subject.

The training procedure with the 500-msec stimulus set was different from that used with the other two sets. Initially, waveforms were delivered to the index finger alone. The 30 waveforms in Table 1 were divided into 12 groups (e.g., the first group contained the four F_L components, the second contained the two F_M components, and so forth). The subjects first practiced with and identified the 4 waveforms in Group 1, then the 6 waveforms in Groups 1 and 2, and so on until the stimulus set contained all 30 waveforms. The subjects were then trained with the same 30 waveforms presented randomly to one of the three digits (i.e., thumb, index finger, or middle finger). In this exercise, they were instructed to identify the waveforms independent of stimulation site. Finally, the subjects were trained to identify both the stimulation site and the waveforms—that is, all 120 alternatives in the 500-msec stimulus set. With the 250- and 125-msec stimulus sets, the subjects were trained to identify all the stimulus alternatives in a given set from the start. Only the training results obtained with all the alternatives in a stimulus set are presented here.

As was discussed previously, the 500-, 250-, and 125-msec stimulus sets were selected to contain easily identifiable signals for which all the subjects could reach a performance criterion of ≥95%. A high performance level, in turn, made it possible for us to calculate IT_{pc}, based on percent-correct scores, as a lower-bound estimate for IT.

Results

The learning curves for each subject and each stimulus set are shown in Figure 4. The amount of time required to reach the performance criterion varied across subjects. S_1, who was involved in designing the stimulus sets and, thus, was very familiar with the signals by the time data collection began, achieved criterion performance within one to three runs across stimulus sets. The total amount of time S_1 spent on training could not be accurately estimated, because of her involvement in many informal tests during
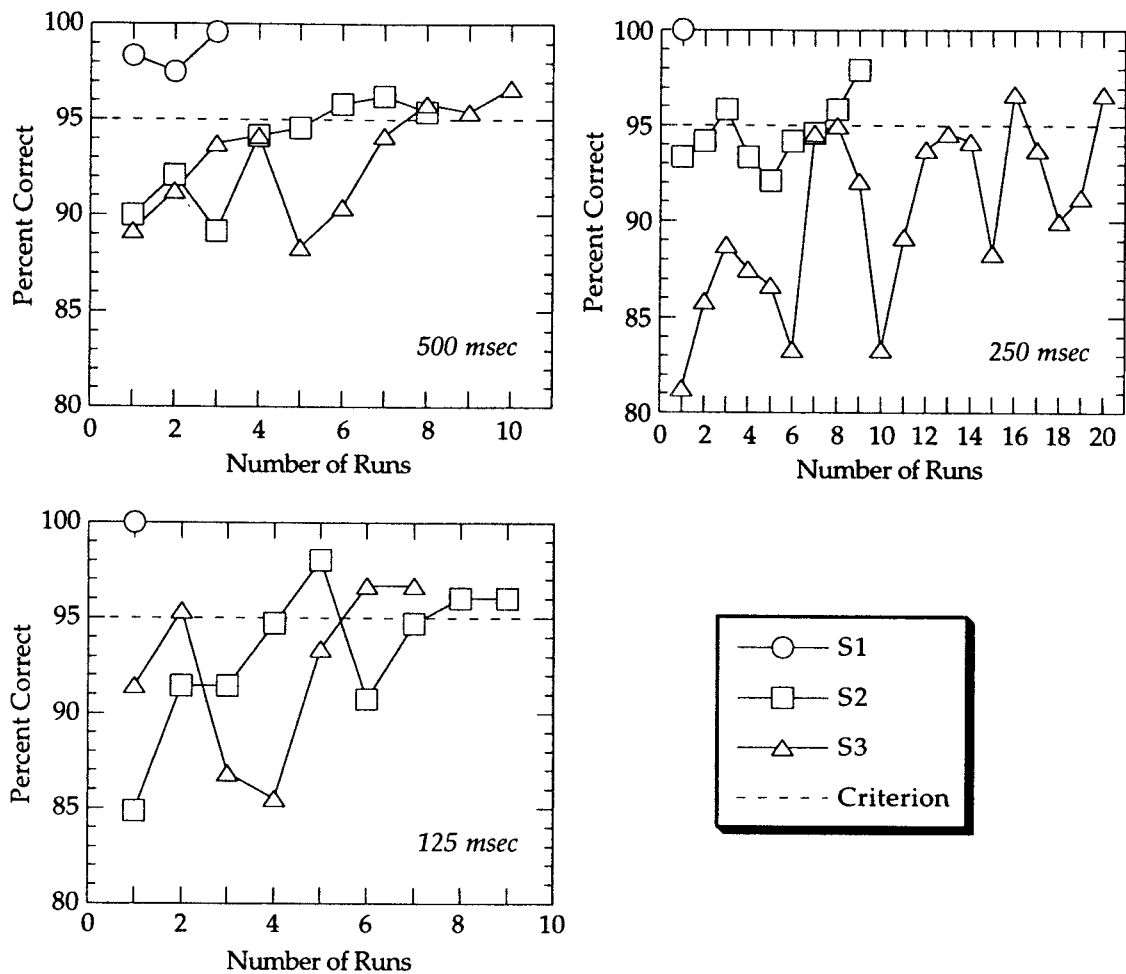
Figure 4. Learning curves for each subject with all the alternatives in the 500-, 250-, and 125-msec stimulus sets, respectively.

the stimulus development phase. The cumulative time for S_2 and S_3 to reach the performance criterion with all three stimulus sets was 19.5 and 27 h, respectively.

Results averaged over the three test runs for these subjects are shown in Figure 5. Subject S_1 maintained 99% accuracy at all three stimulus durations. S_2 achieved an accuracy between 95% and 96% with all three stimulus sets. For both of these subjects, performance level remained fairly constant across stimulus durations. Subject S_3's scores, however, showed a clear downward trend (i.e., from 97% to 91%) as stimulus duration decreased from 500 to 125 msec. Later informal testing indicated that this subject's performance level might have been improved with additional practice at the two shorter stimulus durations.

On the basis of the test runs, the average upper-bound IT estimates, IT_{est}, were 6.7, 6.7, and 6.0 bits for the 500-, 250-, and 125-msec stimulus sets, respectively. The average lower-bound IT estimates, IT_{pc}, were 6.5, 6.4, and 5.6 bits for the 500-, 250-, and 125-msec stimulus sets, re-

spectively. The corresponding $2^{IT_{pc}}$ values were 90, 84, and 49 items for the 500-, 250-, and 125-msec stimulus sets, respectively.

EXPERIMENT 2
Information Transfer Rate

The goal of experiments on IT rate was to assess the dynamic information transmission capabilities with the Tactuator. The IT rate (in bits/sec) is defined as the product of IT per presentation (in bits/item) and presentation rate (in items/sec).

Method

Identification Paradigm With Masking

The identification paradigm used in Experiment 2 incorporates both forward and backward masking as each would occur in a continuous presentation stream (see Figure 6). On each trial, the subject was asked to identify the target signal, X, sandwiched between two
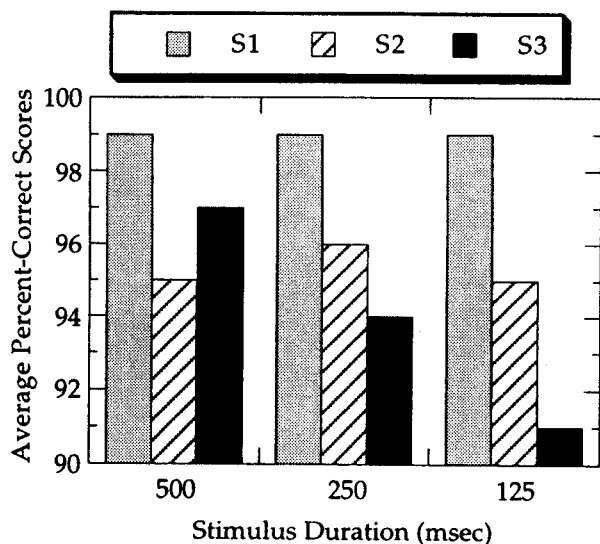
Figure 5. Percent-correct scores averaged over the three test runs for each of the 3 subjects and each of the three stimulus sets.
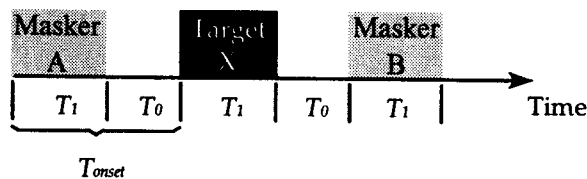

Figure 6. Diagram for the identification paradigm with both forward and backward masking.

interfering maskers, A and B. Within each trial, the duration of the target and the maskers was kept the same (T_1), and so was that of the two gaps (T_0). The time between signal onsets, T_{onset}, was simply $(T_0 + T_1)$, and the presentation rate, λ, was $1/T_{onset}$.

Stimuli and Responses

A summary of the three stimulus sets used in Experiment 2 is provided in Table 4. The number of presentation sites was limited to three: the thumb alone, the index finger alone, or the middle finger alone. (Waveforms were not applied to all three digits, as in Experiment 1, because the F_M and F_H components tended to spread in time and space.[6]) T_1 was 500, 250, or 125 msec; T_0 was 500, 400, 300, 200, 100, or 20 msec;[7] and each combination of T_0 and T_1 was tested (resulting in 18 conditions). Thus, the presentation rate, λ, ranged from 1 item/sec (for $T_0 = 500$ msec and $T_1 = 500$ msec) to 6.9 items/sec (for $T_0 = 20$ msec and $T_1 = 125$ msec).

The response codes used in Experiment 2 were the same as those used in Experiment 1.

Procedure

The experimental conditions were conducted in the same order for each of the 3 subjects. The six values of T_0 were tested in descending order for each value of T_1 (tested in order of 500, 250, and 125 msec). This testing sequence was established in order to test easier conditions prior to the more difficult conditions. With each stimulus set, the subject was first required to repeat the paradigm used in Experiment 1 (i.e., without the maskers A and B), in order to become reacquainted with the signals. All the subjects reached the performance criterion level of 95% within one or two runs. The subject was then tested using the identification paradigm with masking. On each trial, A, X, and B were each randomly selected (with replacement) from the signals that made up a given stimulus set (see Table 4). No additional timing cues were available to mark the three intervals (see note 7). The subjects were instructed to wait until all three signals had been presented before entering the response for X.[8] If the response did not have the right syntax (i.e., an icon for finger location followed by another icon for waveform), the trial was counted as an error and was not repeated. Trial-by-trial correct-answer feedback was not provided.[9] Three runs of 100 trials each were conducted at each combination of T_0 and T_1.[10] The percent-correct score on X was shown to the subject at the end of each run.

Results

For each T_0 and T_1 combination, the percent-correct scores were averaged over the three 100-trial runs. The results are presented in terms of T_{onset}, with T_1 as a parameter (Figure 7). The individual points on each curve correspond to the six T_0 values (i.e., $T_{onset} - T_1$) used with that particular T_1. For all the subjects, percent-correct scores are dependent on T_{onset}, but not on T_1 alone. In other words, there seems to be a tradeoff between T_0 and T_1. The data curves show a knee in the region $325 \leq T_{onset} \leq 450$ msec, corresponding to a presentation rate of roughly 2.2–3 items/sec. Overall, the results for S_1 and S_2 are quite similar (except for the data points at $T_1 = 125$ msec and $T_0 = 500$ msec). The data curves for S_3 reach a slightly lower plateau at a slightly larger T_{onset} value.

In order to assess the relative effects of forward and backward masking, the data shown in Figure 7 were reprocessed to determine the percentage of (incorrect) trials on which an incorrect response was identical to masker A or masker B, respectively. The data shown in Figure 8 indicate that, when percent-correct scores are not saturated (i.e., $T_{onset} < 400$ msec), subjects respond more frequently with the maskers than one would expect by chance (i.e., $1/90$ for $T_1 = 500$ and 250 msec, $1/57$ for $T_1 = 125$ msec). By comparing each subject's corresponding data plots in the left and right panels of Figure 8, it can be concluded that S_1 and S_2 tend to respond with forward masker A more often than with backward masker B. The response pattern for S_3 differed from that of S_1 and S_2 in that his use

Table 4
The Three Stimulus Sets Used in the Identification Paradigm With Forward and Backward Masking

Stimulus Set	No. Waveforms	Stimulation Site	T_1 (msec)	No. Alternatives (k)	$IS_k = \log_2 k$ (bits)
1	30	thumb, index or middle finger	500	90	6.5
2	30	thumb, index or middle finger	250	90	6.5
3	19	thumb, index or middle finger	125	57	5.8

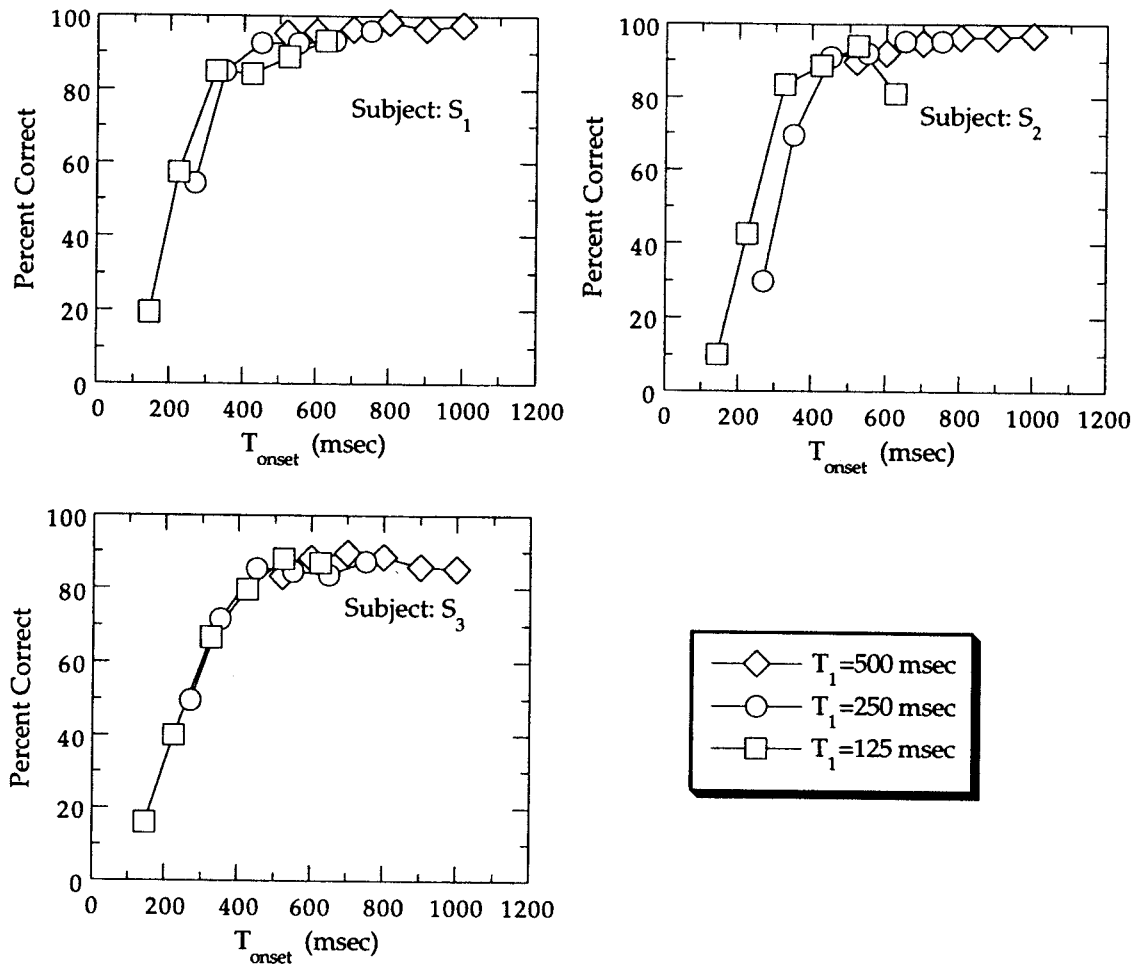Note—The waveforms can be found in Tables 1–3, respectively.

Figure 7. Average percent-correct scores for identification of X as a function of T_{onset} for each of the 3 subjects.

of forward masker A as an incorrect response was much less pronounced.

The data in Figure 7 are also used to estimate the IT rate potentially available with streams of these signals, assuming that the same percent-correct scores hold for the identification of each (consecutive) signal. The lower-bound $IS_k \times (1 - 2\bar{e})$ is used to estimate IT, where IS_k is given in Table 4 and $\bar{e}$ is the observed average error rate. The results in Figure 9 show that T_{onset} corresponding to the highest estimated IT rate is 350 msec for S_1, 325 msec for S_2, and 450 msec for S_3. In other words, the optimal presentation rate is roughly 3 items/sec for S_1 and S_2 and roughly 2.2 items/sec for S_3. Note that when the percent-correct score was below 50% and, therefore, $(1 - 2\bar{e}) < 0$, the estimated IT rate was set to 0. For S_1, a maximum IT rate of 13 bits/sec occurred at $T_1 = 250$ msec and $T_0 = 100$ msec. For S_2, a maximum IT rate of 12.1 bits/sec occurred at $T_1 = 125$ msec and $T_0 = 200$ msec. For S_3, a maximum IT rate of 10.2 bits/sec occurred at $T_1 = 250$ msec and $T_0 = 200$ msec. The maximum IT rate averaged over all the subjects was roughly 12 bits/sec.

The validity of this estimate is dependent on several assumptions. First, we assume that our identification paradigm allows us to quantitatively characterize both the forward and the backward masking that occur during continuous presentation of signal streams. Second, we assume that, with sufficient training, the subjects would eventually learn to *chunk* individual signals into meaningful *messages* that can be stored in short-term memory for later retrieval. Third, we assume that, with sufficient training, the subjects would be able to maintain the same performance level when they are required to respond to *all* the signals in a continuous input stream, as compared with when they are required to respond to only *one* signal embedded in an input stream. Although these assumptions appear to be reasonable, empirical tests are necessary to assess their validity. For example, to test the validity of the first assumption, we will train subjects to identify three successively presented signals. The results will then be compared with those obtained in this experiment. The second assumption was partly based on our previous work on Morse code reception (Tan et al., 1997). Amateur radio
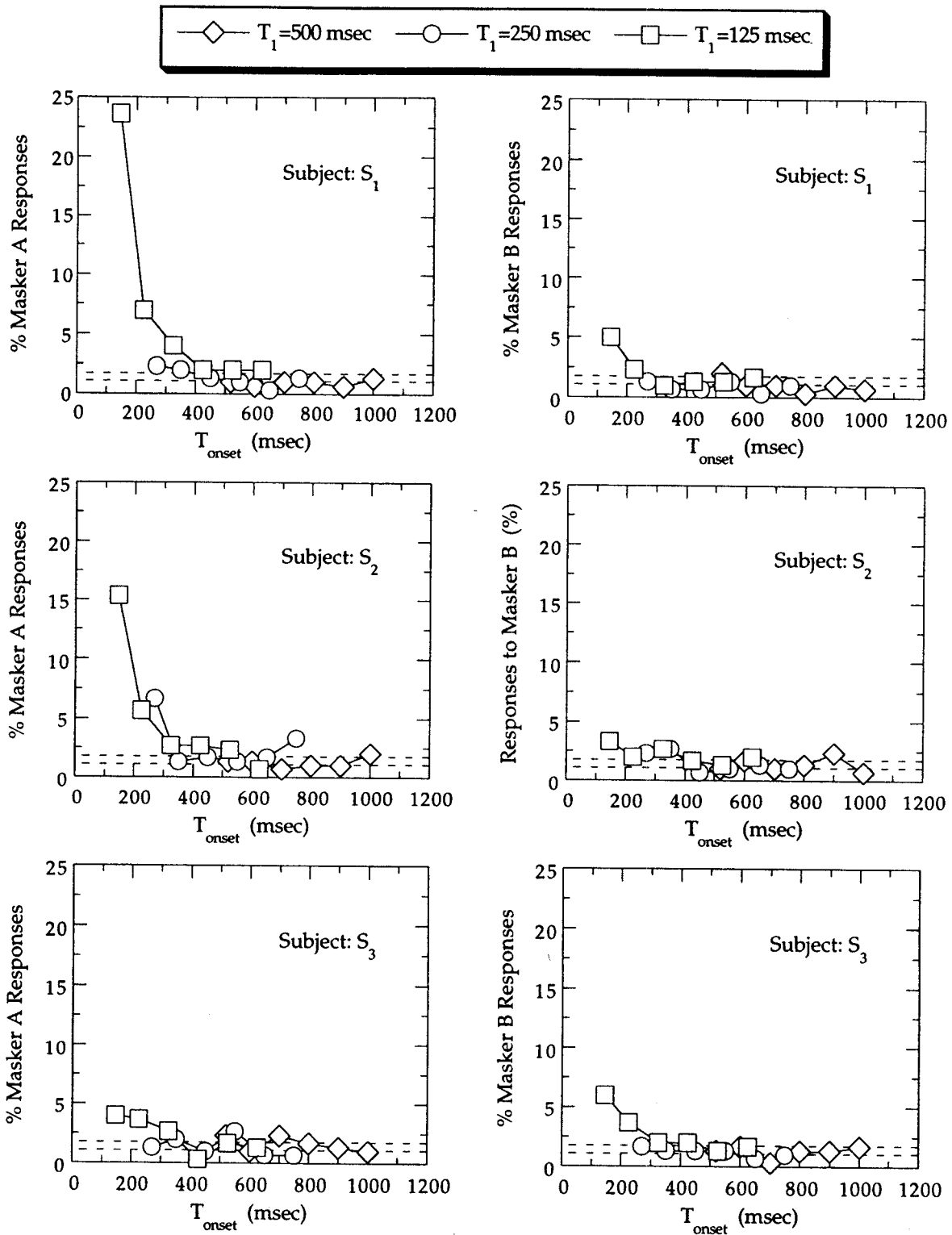
Figure 8. Results of an analysis of incorrect responses in the identification paradigm with masking for individual subjects. Left and right panels show percentage of trials on which an incorrect response was identical to maskers A and B, respectively. Dashed lines indicate chance performance of 1/90 (i.e., 1.11%) for $T_1 = 500$ and 250 msec, and of 1/57 (i.e., 1.75%) for $T_1 = 125$ msec.
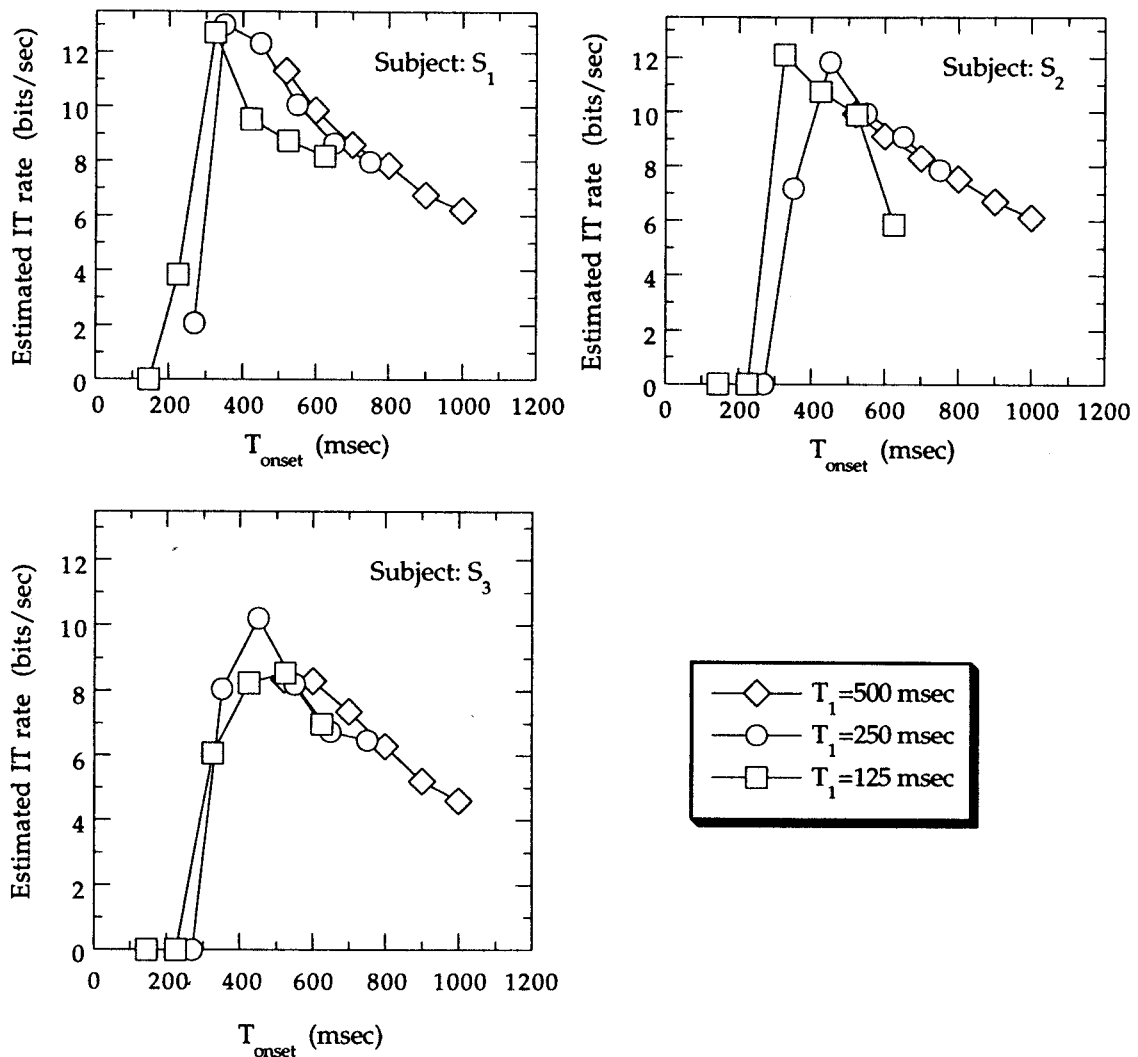
Figure 9. Estimated IT rate as a function of T_{onset} for each of the 3 subjects.

operators have long proven that it is possible to learn to chunk individual dit-dah signals into letters, words, and short messages, although the amount of training required and the eventual performance levels reached vary greatly among individuals. It would be challenging, if not impossible, to vigorously test the third assumption within the constraints of time and resources available in most studies. Even then, whether the results can be generalized to other tasks remains to be seen.

GENERAL DISCUSSION

Two series of absolute identification experiments were carried out to assess the information transmission capabilities of the Tactuator. From Experiment 1, it was found that estimated IT per stimulus presentation averaged over 3 subjects was 6.5 bits (corresponding to perfect identifi-

cation of 90 items) for the 500-msec stimulus set, 6.4 bits (corresponding to perfect identification of 84 items) for the 250-msec stimulus set, and 5.6 bits (corresponding to perfect identification of 49 items) for the 125-msec stimulus set. From Experiment 2, it was found that the optimal stimulus presentation rates were between 2 and 3 items/sec, independent of stimulus duration, and the maximum IT rate averaged over 3 subjects was about 12 bits/sec.

The data reported here, based on in-depth study of a relatively small number of subjects, may be regarded as reflecting the ultimate potential for IT and IT rate through the Tactuator. To assess overall IT and IT rate, each subject was required to perform identification tasks involving large and complex stimulus sets toward the end of this study. To prepare the subjects, the experiments were designed so that participation in each successive task was based on the successful completion of the previous task,

and throughout Experiment 1, the subjects were required to achieve a predetermined performance criterion before advancing to the next stage of the experiment. This strategy required a large time commitment on the part of each subject and, thus, limited the number of subjects who could be tested within the time frame of the study. Despite the differences in their experiences with the stimuli (i.e., S_1 was more experienced with the stimulus sets owing to her involvement in the development of the stimuli), the 3 subjects achieved similar results on IT and IT rate. The range of static IT for the 3 subjects was 6.2–6.8 bits for the 500-msec stimulus set, 6.1–6.8 bits for the 250-msec stimulus set, and 5.1–6.1 bits for the 125-msec stimulus set. The range of maximum IT rates for the 3 subjects was 10.2– 13 bits/sec.

The static IT results should be taken as a lower-bound estimate of IT for the following reasons. First, the process of stimulus design may be improved, in the sense that more alternatives may be included in a stimulus set without significantly impairing performance. Second, a conservative estimate of IT was employed that was based on percent-correct scores (rather than obtaining a direct measure of unbiased IT estimate). Third, there was some evidence that the subjects' performances could be improved by further training.

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 the capacity for processing information along unidimensional stimulus sets is limited by the magical number seven, plus or minus two (i.e., 2.3–3.2 bits). Pollack and Ficks (1954) were able to obtain IT values of between 5 and 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 developed for presentation through the Tactuator involved many stimulus aspects, with a mainly binary coding scheme. The IT value of 6.5 bits obtained with the 500-msec stimulus set appears impressive, considering the fact that the tactual system is often regarded as having a low channel capacity and, in any case, is not accustomed to receiving motional stimulation (especially at very low frequencies). It is also the highest IT that has been obtained with tactual artificial displays of any kind. For example, Sherrick (1985) reported an average IT of 2.7 bits for the identification of the rate of 400-Hz haversine bursts with redundantly varying intensities. The IT obtained from a tactile display involving vibratory intensity, frequency, and contactor area was 4–5 bits (Rabinowitz et al., 1987). Tan, Rabinowitz, and Durlach (1989) obtained an IT of 3.3 bits from the four movement channels of an artificial facial movement display (the synthetic Tadoma system). Assuming similar information transmission properties for the airflow and vibration channels

of the same synthetic Tadoma system, the total information transfer along the six channels may be estimated at roughly 5–6 bits.

The results from Experiment 1 indicate that IT was reduced by only 0.9 bit when signal duration was reduced by a factor of four, thus suggesting that higher IT rates might be achieved using the shortest duration signals. This assumption did not hold up in view of the results obtained in Experiment 2, which show that IT rate depends primarily on stimulus presentation rate, not on stimulus duration alone. The optimal stimulus presentation rate estimated from Experiment 2 results was approximately 2.2–3 items/sec, independent of stimulus duration (for signal duration T_1 in the region $125 \le T_1 \le 500$ msec). Garner (1962) and Klemmer and Muller (1953) also reported an optimal presentation rate of 2–3 items/sec, independent of stimulus uncertainty per item, from vision experiments in which subjects pressed keys in response to flashing lights. A constant presentation rate suggests that constant processing time was the principal limiting factor in processing streams of input signals.

An analysis of masking effects in Experiment 2 indicated that the subjects responded more frequently with the maskers than one would expect by chance when performance level was less than 90% correct. S_1 and S_2 tended to respond more with the forward maskers than with the backward maskers. This is inconsistent with the typical finding of greater backward than forward masking effects in pattern identification tasks (e.g., Craig, 1985). However, our identification paradigm with forward and backward masking differs from typical paradigms used in masking studies in several ways (see Craig, 1995, for a review of these paradigms). First, the effects of forward maskers (A) and backward maskers (B) on our subjects' ability to identify the targets (X) cannot be easily separated in our study. Second, our identification paradigm allows masker A (or B) and target X to be presented to the same or to different fingers, resulting in a mixture of temporal masking and selective attention paradigms. There is growing evidence that the interference in the identification of a target pattern in temporal masking and selective attention paradigms can be mainly attributed to response competition (i.e., representations of both the target and the nontarget are available, and the subject mistakenly responds with the nontarget pattern), rather than to masking (i.e., presence of the nontarget distorts the representation of the target pattern; Craig, 1995; Craig & Evans, 1995). Future studies with the stimuli developed in this study and with modified experimental paradigms will help to address these issues.

From the results of Experiment 2, we estimated IT rate averaged over 3 subjects at their respective optimal input presentation rates to be approximately 12 bits/sec. This rate can be compared to the IT rates obtained with other tactual communication devices. Using his air-driven finger stimulator, Bliss (1961) reported an IT rate of 4.5 bits/sec[11] for one experienced typist, who received letters and a few punctuation symbols (4.9 bits/symbol) with eight

fingers (excluding the two thumbs) 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[12] was obtained with 6 subjects who identified motions in six directions (i.e., $\pm x$, $\pm y$, and $\pm z$, with 2.58 bits/movement) at a presentation rate of 2.8 movements/sec. Stimulating eight fingers, as compared with only one finger, did not yield large gains in IT rate. However, during the single-finger experiment, the subject was presented with a sequence of three movements at a time; during the multifinger experiment, the subject received a sequence of 130 symbols at the specified rate and responded orally by naming the symbols as they were received. Another important factor was that the 30 symbols used in the multifinger 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.[13] Using the Optacon device (Linvill & Bliss, 1966) and English sentences as test material, Cholewiak, Sherrick, and Collins (1993) reported that their best subject was able to reach a word rate of 40 wpm, or equivalently, 5.3 bits/sec. Craig (1977) reported two sighted "extraordinary observers" who were able to read, via vibrotactile patterns generated by the Optacon, at rates of 70–100 wpm, or equivalently, 9.3–13.3 bits/sec. In a more recent study, Summers et al., (1997) reported an IT rate of 6.25 bits/sec, using a single vibrator on the fingertip and with a paradigm that required subjects to detect changes in a sequence of stimulus elements. IT rates achievable with tactual Morse code patterns have also been evaluated. Foulke and 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 note 13). The IT rate obtained in a study of Morse code reception through up–down finger motions, using conversational English material, was 2.7 bits/sec (Tan et al., 1997). The relatively lower IT rates obtained by Foulke and Brodbeck and by our study on Morse code reception (Tan et al., 1997) may be partly due to the inefficiency of the code itself.

Overall, the IT rates measured with man-made tactual displays are much lower than the rates demonstrated by natural tactual communication methods (except for results obtained with extraordinary subjects, such as those reported by Craig, 1977). Reed, Durlach, and Delhorne (1992) estimated IT rate to be 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 IT rate for fingerspelling appears to be limited by the speed at which handshapes can be produced, the IT 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 with English material, it will be possible to demonstrate comparable rates through a tactual device and through Tadoma. 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, such a finding would suggest that it is likely that the monitoring of the articulatory process per se is not the key component in the success of Tadoma.

REFERENCES

ALLUISI, E. A., MULLER, P. F., & FITTS, P. M. (1957). An information analysis of verbal and motor responses in a forced-paced serial task. *Journal of Experimental Psychology*, 53, 153–158.

BLAMEY, P. J., & CLARK, G. M. (1985). A wearable multiple-electrode electrotactile speech processor for the profoundly deaf. *Journal of the Acoustical Society of America*, 77, 1619–1620.

BLAMEY, P. J., & CLARK, G. M. (1987). Psychophysical studies relevant to the design of a digital electrotactile speech processor. *Journal of the Acoustical Society of America*, 82, 116–125.

BLISS, J. C. (1961). *Communication via the kinesthetic and tactile senses*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

BRYAN, W. L., & HARTER, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review*, 6, 345–375.

CHOLEWIAK, R. W., SHERRICK, C. E., & COLLINS, A. A. (1993). *Princeton cutaneous research project* (Rep. No. 62). Princeton, NJ: Princeton University, Department of Psychology, Cutaneous Communication Laboratory.

COWAN, R. S. C., BLAMEY, P. J., SARANT, J. Z., GALVAN, K. L., & CLARK, G. M. (1991). Perception of multiple electrode stimulus patterns: Implications for design of an electrotactile speech processor. *Journal of the Acoustical Society of America*, 89, 360–368.

CRAIG, J. C. (1977). Vibrotactile pattern perception: Extraordinary observers. *Science*, 196, 450–452.

CRAIG, J. C. (1985). Tactile pattern perception and its perturbations. *Journal of the Acoustical Society of America*, 77, 238–246.

CRAIG, J. C. (1995). Vibrotactile masking: The role of response competition. *Perception & Psychophysics*, 57, 1190–1200.

CRAIG, J. C., & EVANS, P. M. (1995). Tactile selective attention and temporal masking. *Perception & Psychophysics*, 57, 511–518.

EBERHARDT, S. P., BERNSTEIN, L. E., BARAC-CIKOJA, D., COULTER, D. C., & JORDAN, J. (1994). Inducing dynamic haptic perception by the hand: System description and some results. In C. J. Radcliffe (Ed.), *Proceedings of the American Society of Mechanical Engineers: Dynamic systems and control* (Vol. 1, pp. 345–351). American Society of Mechanical Engineers.

FOULKE, E., & BRODBECK, JR., A. A. (1968). Transmission of Morse code by electrocutaneous stimulation. *Psychological Record*, 18, 617–622.

GARNER, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.

GELDARD, F. A. (1957). Adventures in tactile literacy. *American Psychologist*, 12, 115–124.

GESCHEIDER, G. A. (1966). Resolving of successive clicks by the ears and skin. *Journal of Experimental Psychology*, 71, 378–381.

GESCHEIDER, G. A. (1967). Auditory and cutaneous temporal resolution of successive brief stimuli. *Journal of Experimental Psychology*, 75, 570–572.

HOUTSMA, A. J. M. (1983). Estimation of mutual information from limited experimental data. *Journal of the Acoustical Society of America,* 74, 1626–1629.

KLEMMER, E. T., & MULLER, J. P. F. (1953). The rate of handling information: Key-pressing responses to light patterns. *Human Factors Oper. Res. Lab. Memo Report,* 34.

LEOTTA, D. F., RABINOWITZ, W. M., REED, C. M., & DURLACH, N. I. (1988). Preliminary results of speech-reception tests obtained with the synthetic Tadoma system. *Journal of Rehabilitation Research & Development,* 25, 45–52.

LIBERMAN, A. M., & MATTINGLY, I. G. (1989). A specialization for speech perception. *Science,* 243, 489–494.

LINVILL, J. G., & BLISS, J. C. (1966). A direct translation reading aid for the blind. *Proceedings of the Institute of Electrical and Electronics Engineers,* 54, 40–51.

MILLER, G. A. (1954). Note on the bias of information estimates. In H. Quastler (Ed.), *Information theory in psychology* (pp. 95–100). Glencoe, IL: Free Press.

MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review,* 63, 81–97.

PICKETT, J. M., & PICKETT, P. H. (1963). Communication of speech sounds by a tactual vocoder. *Journal of Speech & Hearing Research,* 6, 207–222.

POLLACK, I., & FICKS, L. (1954). Information of elementary multidimensional auditory displays. *Journal of the Acoustical Society of America,* 26, 155–158.

PROCTOR, R. W., & REEVE, T. G. (EDS.) (1990). *Stimulus–response compatibility: An integrated perspective.* Amsterdam: Elsevier, North-Holland.

RABINOWITZ, W. M., HENDERSON, D. R., REED, C. M., DELHORNE, L. A., & DURLACH, N. I. (1990). Continuing evaluation of a synthetic Tadoma system. *Journal of the Acoustical Society of America,* 87, 88.

RABINOWITZ, W. M., HOUTSMA, A. J. M., DURLACH, N. I., & DELHORNE, L. A. (1987). Multidimensional tactile displays: Identification of vibratory intensity, frequency, and contactor area. *Journal of the Acoustical Society of America,* 82, 1243–1252.

REED, C. M., & DELHORNE, L. A. (1995). Current results of field study of adult users of tactile aids. *Seminars in Hearing,* 16, 305–315.

REED, C. M., DOHERTY, M. J., BRAIDA, L. D., & DURLACH, N. I. (1982). Analytic study of the Tadoma method: Further experiments with inexperienced observers. *Journal of Speech & Hearing Research,* 25, 216–223.

REED, C. M., & DURLACH, N. I. (1998). Note on information transfer rates in human communication. *Presence–Teleoperators & Virtual Environments,* 7, 509–518.

REED, C. M., DURLACH, N. I., & DELHORNE, L. A. (1992). The tactual reception of speech, fingerspelling, and sign language by the deaf-blind. *Digest of Technical Papers of the Society for Information Display International Symposium,* 23, 102–105.

REED, C. M., DURLACH, N. I., DELHORNE, L. A., RABINOWITZ, W. M., & GRANT, K. W. (1989). Research on tactual communication of speech: Ideas, issues, and findings. *Volta Review,* 91, 65–78.

REED, C. M., RABINOWITZ, W. M., DURLACH, N. I., BRAIDA, L. D., CONWAY-FITHIAN, S., & SCHULTZ, M. C. (1985). Research on the Tadoma method of speech communication. *Journal of the Acoustical Society of America,* 77, 247–257.

REED, C. M., RUBIN, S. I., BRAIDA, L. D., & DURLACH, N. I. (1978). Analytic study of the Tadoma method: Discrimination ability of untrained observers. *Journal of Speech & Hearing Research,* 21, 625–637.

ROGERS, M. S., & GREEN, B. F. (1954). The moments of sample information when the alternatives are equally likely. In H. Quastler (Ed.), *Information theory in psychology* (pp. 101–108). Glencoe, IL: Free Press.

SHANNON, C. E. (1951). Prediction and entropy of printed English. *Bell System Technical Journal,* 30, 50–64.

SHERRICK, C. E. (1985). A scale for rate of tactual vibration. *Journal of the Acoustical Society of America,* 78, 78–83.

SUMMERS, I. R., COOPER, P. G., WRIGHT, P., GRATTON, D. A., MILNES, P.,

& BROWN, B. H. (1997). Information from time-varying vibrotactile stimuli. *Journal of the Acoustical Society of America,* 102, 3686–3696.

TAN, H. Z. (1988). *Analysis of a synthetic Tadoma system as a multidimensional tactile display.* Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge.

TAN, H. Z. (1996). *Information transmission with a multi-finger tactual display.* Unpublished doctoral dissertation, Massachusetts Institute of Technology.

TAN, H. Z., DURLACH, N. I., RABINOWITZ, W. M., REED, C. M., & SANTOS, J. R. (1997). Reception of Morse code through motional, vibrotactile, and auditory stimulation. *Perception & Psychophysics,* 59, 1004–1017.

TAN, H. Z., & RABINOWITZ, W. M. (1996). A new multi-finger tactual display. In K. Danai (Ed.), *Proceedings of the American Society of Mechanical Engineers: Dynamic systems and control division* (Vol. 58, pp. 515–522). American Society of Mechanical Engineers.

TAN, H. Z., RABINOWITZ, W. M., & DURLACH, N. I. (1989). Analysis of a synthetic Tadoma system as a multidimensional tactile display. *Journal of the Acoustical Society of America,* 86, 981–988.

VERRILLO, R. T., & GESCHEIDER, G. A. (1992). Perception via the sense of touch. In I. R. Summers (Ed.), *Tactile aids for the hearing impaired* (pp. 1–36). London: Whurr.

NOTES

1. The word *dimension* has been extensively used in classical literature to describe the richness of a display (e.g., Miller, 1956). However, the concept of dimensionality has not yet been adequately defined. In particular, the dimensionality of a display is usually defined in terms of the number of independently varying *physical* variables that make up the stimuli, without considering the perceptual attributes of the stimuli. Thus, for example, it does not cover the case in our experiments, in which a mere change in one physical variable (i.e., frequency) results in a qualitative change in perception (i.e., slow motion vs. vibration). For this reason, we will use the word *attribute* throughout this paper to refer loosely to the concept of *perceptual dimensionality,* which has yet to be rigorously defined.

2. The advantage of using nonsense material is that it is relatively easy to control stimulus uncertainty. The use of meaningful test materials (such as English sentences) necessitates the assessment of the redundancy inherent in the test material.

3. A lower bound for IT/IS can be estimated by constructing a stimulus–response confusion matrix, with all correct trials evenly distributed on the k main diagonal cells and all error trials evenly distributed among the nondiagonal cells. For low error rates ($e \leq 5\%$) and a relatively large number of stimulus alternatives ($57 \leq k \leq 120$), the ratio of IT/IS to $(1 - 2e)$ ranges from 0.996 to 1.010. Thus, as a first approximation, $(1 - 2e)$ can be used as a lower bound for IT/IS.

4. For example, the 30- (an F_M component) and 300-Hz (an F_H component) signals are characterized by a general sense of roughness or smoothness, respectively. The actual movement amplitudes are sufficiently small that subjects could not discern the onset direction of these motions.

5. Such a performance criterion ensured that the subjects learned one task well before proceeding to another. It was possible to impose such a criterion in our experiments because one of the stimulus design requirements was that stimuli be easily identifiable.

6. For example, if masker A contained a 300-Hz signal component applied to all three fingers, the subjects tended to judge 300 Hz to be present in the target X. The *spreading* of F_M and F_H components was less of a problem during static IT measurements, because the interstimulus duration was relatively long.

7. According to Gescheider (1966, 1967), tactual gap detection threshold is on the order of 10 msec with relatively strong signals (e.g., 35 dB SL). A nonzero minimum value of 20 msec for T_0 ensured that there was enough *gap* between the three signal intervals.

8. This instruction was based on the consideration that some subjects may choose to respond immediately to X upon its presentation and to ignore the backward masker B completely. Since the goal of this experi-

ment was to measure *both* the forward *and* the backward masking effects, it was decided that the subjects should be required to wait until all three signals had been presented before they could respond. In tasks in which a continuous stream of signals are received, whether subjects respond *on the fly* or *behind* will depend on other factors, such as their ability to chunk information and to store messages in short-term memory (see, e.g., Tan et al., 1997).

9. Feedback was not provided for two reasons. First, all the subjects were well trained with the signals in each of the three stimulus sets. Second, requiring the subjects to attend to correct-answer feedback tended to break the *rhythm* of the run.

10. There was one exception. The percent-correct scores of the first three runs for S_2 at $T_1 = 125$ msec and $T_0 = 500$ msec were 18%, 77%, and 76%. Given the inconsistency of these three scores, one more run was conducted. The data from the first run were discarded.

11. Bliss (1961) estimated IT as IT = IS $\times$ (1 − e) but did not offer a supporting argument for this convention. Using our conservative estimate of IT = IS $\times$ (1 − 2e), his 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, however, 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.

12. Information per presentation was computed from the stimulus-response confusion matrix. Had Bliss (1961) used IT = IS $\times$ (1 − e) to compute IT on the basis of an error rate of 23%, the IT rate would have been 3.9 bits/sec.

13. The IT rate was estimated from the word rate on the basis of two assumptions. First, according to Shannon (Shannon, 1951, Figure 4), the uncertainty for strings of 8 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 8 letters. Second, it is assumed that the average word length is 4 letters/word. It follows that the information content in words is 2 bits/letter $\times$ 4 letters/word, or 8 bits/word. The IT rate is, therefore, 8 bits/word $\times$ 38 words/min, or equivalently, 5.1 bits/sec (see also Reed & Durlach, 1998).