Information Transmission
with a Multi-Finger Tactual Display

by

Hong Zhang Tan

S.B., Shanghai Jiao Tong University (1986)
S.M., Massachusetts Institute of Technology (1988)

SUBMITTED TO THE
DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 1996

Signature of Author____________________________________
Department of Electrical Engineering and Computer Science
February 29, 1996

Certified by____________________________________
Nathaniel I. Durlach
visor

Accepted by____________________________________
aler
Chairman, Departmental Committee on Graduate Students

Information Transmission with a Multi-Finger Tactual Display

by
Hong Zhang Tan

Submitted to the Department of Electrical Engineering and Computer Science

on February 29, 1996

in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

This work was motivated by our interest in using the sense of touch as an alternative communication channel for sensory substitution. Previous research has demonstrated that some deaf-and-blind individuals can receive conversational English at almost normal rates using the Tadoma method, in which the user places a hand on the face and neck of a talker and monitors the mechanical actions associated with speech production. To this day, however, no one has achieved a similar performance level with electromechanical devices developed for tactual speech communication. These devices typically do not engage the hand and use homogeneous vibrotactile arrays that lack distinctive perceptual qualities. In contrast, Tadoma makes use of the hand and a rich display, the talking face, that involves kinesthetic stimulation (low-frequency large-amplitude motions) as well as vibrotactile stimulation (high-frequency small-amplitude vibrations). A major goal of this research is to explore how information transmission can be improved by simultaneously stimulating both the kinesthetic and tactile components of the tactual system.

A multi-finger positional display, the TACTUATOR™, was developed. It consists of three independent single-contact-point actuators interfaced with the thumb, the index finger, and the middle finger. Each actuator utilizes a disk-drive head-positioning motor augmented with angular position feedback from a precision rotary variable differential transformer (RVDT). A floating-point DSP system provides real-time positional control using a digital PID controller. It is capable of delivering arbitrary waveforms within an amplitude range from absolute detection threshold to about 50 dB sensation level, and a frequency range from near DC to above 300 *Hz* (e.g., 25 *mm* slow motion with superimposed high-frequency vibration). Actuator frequency and step responses are well modeled as a second-order linear system. Distortion is low. System noise and inter-channel crosstalk are also small. Absolute thresholds measured with the stimulator are in general agreement with results from the literature. Overall, the TACTUATOR accurately follows its drive waveforms and is well suited for a variety of multi-finger tactual perceptual studies.

The information transmission capabilities with the TACTUATOR were assessed through a series of absolute identification experiments with human observers. In exploring the stimulus attributes that are most effective for producing a large set of clearly distinguishable stimuli with the TACTUATOR, it was found that subjects could naturally categorize motions over the entire frequency range into three perceptually distinctive groups: slow motion (up to about 6 Hz), a rough or fluttering sensation (about 10 to 70 Hz), and smooth vibration (above about 150 Hz). Multi-component stimuli were formed by simultaneously stimulating multiple fingers with waveforms containing sinusoids (varying in both frequency and amplitude) from the three frequency regions. Stimulation was applied to either one of three digits (thumb, index, or middle) or to all three digits simultaneously. For a stimulus duration of 500 *msec*, information transfer (IT) was 6.5 *bits* (corresponding to perfect identification of 90 stimuli); at 250 *msec*, IT was 6.4 *bits*; and at 125 *msec*, IT was 5.6 *bits*. Estimates of potential IT rates were obtained by sequencing three random stimuli and (a) having the subject identify only the middle stimulus and (b) extrapolating this IT to that for continuous streams. Stimulus durations of 125 to 500 *msec* and presentation rates of 1 to 7 *items/sec* were tested. Estimated IT rate was about 12 *bits/sec*, and optimal stimulus presentation rates were between 2–3 *items/sec* independent of stimulus duration. This IT rate is roughly the same as that achieved by Tadoma users in tactual speech communication.

In addition to the above work, several related issues were identified for further investigation: selection of stimulus uncertainty for maximizing information transfer, definition of stimulus-set dimensionality, and relationship between the capability to receive motional input sequences and one's ability to deliver the same motor outputs.

Thesis Supervisor: Nathaniel I. Durlach, Senior Research Scientist.

獻給我的父母

(This thesis is dedicated to my parents)

I thank

Jianmin, for his love and support;

Lei Lei, for bringing so much happiness into my life,
and for sleeping through the night before I did;

My parents, for their sacrifices that
made this thesis possible.

I am greatly indebted to my thesis advisors for their guidance and encouragement. I thank

Nat, for helping me to achieve more than what I expected of myself;

Bill, for guiding me through every step of my thesis work, especially the processes of hardware design and performance evaluation which were the most challenging parts for me;

Charlotte, for being a role model and for her advice on this work;

Larry and Srini, for their careful reading of this thesis and thoughtful comments.

Many individuals have assisted me over the years. I am especially grateful to:

Will D., Ken S., Louise J., Kiran D., Eric F., Tom W., Eva M., and Rhonda M., for their contribution to the design of the TACTUATOR™;

Jon S. and Joe F., for their indispensable assistance with computer programming;

Joe P. and Jon S., for their dedication and excellent work as the subjects.

The Sensory Communication Group has been, and continues to be, my home away from home. I can not imagine a more ideal environment for a "senior" graduate student.

Finally, many ideas in my thesis have been inspired by, and sometimes were the direct product of discussions with, my thesis advisors and other people who are genuinely interested in my work. Although mere thanks are inadequate, I sincerely thank all of them.

This research was supported by research grant 2 R01 DC 00126-16 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

Table of Contents

Chapter I.
Introduction

This work was motivated by an 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., non-device related) methods of tactual speech communication. Particularly noteworthy is the so-called Tadoma method that 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. Previous research has documented the remarkable abilities of experienced Tadoma users (Reed, Rabinowitz, Durlach, Braida, Conway-Fithian, & Schultz, 1985); these individuals can understand everyday speech at very high levels, allowing rich two-way conversation with both familiar and novel talkers. Conversely, attempts to develop artificial tactual speech communication devices have had only limited success, with none achieving performance anywhere near that demonstrated by Tadoma (e.g., Reed, Durlach, Delhorne, Rabinowitz, & Grant, 1989).

One problem with most previous tactual devices concerns the nature of the output display. These displays have generally been composed of multiple stimulators that deliver high-frequency vibration to the tactile sensory system. Such "homogeneous" displays have few distinctive perceptual qualities. Furthermore, for practical and/or technical reasons, the displays have rarely engaged the hand, the most sensitive and richly innervated receiving site. In contrast, Tadoma is received by the hand and a talking face is perceptually rich, simultaneously displaying various stimulation qualities that engage both the kinesthetic and tactile sensory systems.

Recognition of the need for richer tactual displays has long been evident. With the aim of developing a tactual communication system, Bliss's "reverse-typewriter" system (Bliss, 1961) was capable of delivering motional pulses to the eight fingers of both hands (excluding the thumbs) that are similar to the motions made by typists. The Sensory Communication Group at MIT has developed an artificial mechanical face display, built around a model plastic skull (Reed *et al.*, 1985), that has shown promise in conveying information important in Tadoma (Leotta, Rabinowitz, Reed, & Durlach, 1988; Rabinowitz, Henderson, Reed, Delhorne, & Durlach, 1990). As a more general display for studying haptic perception by the hand, the "OMAR" system was recently described by Eberhardt, Bernstein, Barac-Cikoja, Coulter, & Jordan (1994). It was designed to deliver kinesthetic as well as tactile stimulation to one or more fingers. Nevertheless, none of these artificial displays has, thus far, been shown to be capable of delivering tactual stimulation that can be received by human observers at a rate comparable to that achieved by Tadoma users.

In order to develop a tactual display that can be used successfully for sensory substitution, we first need a tactual display that is designed to match the perceptual capabilities of human observers. We then need to explore ways of maximizing the information transmission capabilities of such a display by careful design of stimulus and response sets. Finally, the information transfer per presentation and the information transfer rate achievable with such a device must be measured, or estimated, with human observers. Simply stated, that is the goal of this research, to go through the process from hardware development through psychophysical evaluation.

In a preliminary study, the potential for communication through the kinesthetic aspect of the tactual sense was examined in a series of experiments employing Morse Code signals. A manuscript, entitled "Reception of Morse Code Through Motional, Vibrotactile, and Auditory Stimulation", has been submitted to *Perception&Psychophysics*. It is enclosed as an appendix to this thesis (Appen. A). For the main research in this thesis, a new multi-finger tactual display was developed (Chap. 2). We aimed at a continuous frequency response so that the perception from low-frequency large-amplitude motions (i.e., kinesthetic stimulation) to high-frequency small-amplitude vibrations (i.e., vibrotactile stimulation) could be studied as a continuum. We then explored ways of constructing stimulus and response sets that were optimized for information transfer, and measured information transfer achievable with human observers (Chap. 3). The information transfer rate achievable with the multi-finger tactual display was estimated (Chap. 4).

During the course of this work, several unresolved issues that are related to this research were identified and documented (Chap. 5). Finally, a general discussion including directions for future research is presented (Chap. 6).

I-1 Background

The following is a general review of literature on the tactile and kinesthetic senses, focussed on studies concerning the human hand. The objective is to provide a basic understanding of taction that is relevant to this thesis work. The emphasis is on positional/motional stimulation of the tactual sensory system.

I-1.1 Tactile Sense

Vibrotactile stimulation of the skin surface usually consists of low-intensity, high-frequency components. Literature on absolute detection, intensity and frequency discrimination, temporal resolution, and the Vibratese language (Geldard, 1957) is reviewed here.

Receptor Mechanism / Absolute Detection Threshold

Previous physiological and anatomical experiments have identified four afferent fiber types (PC or RA II, RA I, SA II and SA I) in glabrous (nonhairy) skin of the human somatosensory periphery. Johnson & Hsiao (1992) reviewed neural mechanisms of tactual form and texture perception and proposed the following working hypothesis: The SAI system is the primary spatial system and is responsible for tactual form and roughness perception when the fingers contact a surface directly and for the perception of external events through the distribution of forces across the skin surface. The PC system is responsible for the perception of external events that are manifested through transmitted high-frequency vibration of the kind that are critical in the use of objects as tools. The RA system is responsible for the detection and representation of localized movement between skin and a surface as well as for surface form and texture when surface variation is too small to activate the SA I afferents effectively. Srinivasan, Whitehouse, & LaMotte (1990) studied the mechanism of tactile detection of slip using glass plates. They found that direction of skin stretch (impending but not actual slip) was coded solely by the SAs (whether it was SA I or SA II was not clear, since the Macacca Fascicularis monkeys they used did not have SA II fibers). The detection of slip was possible only when the glass plate had detectable surface features. Different neural mechanisms were responsible for slip detection depending on the geometry of the micro-features of the glass surface (i.e., RAs for single-dot plate; and PCs for fine homogeneous dot matrix that induced vibrations of the skin). When the surface features are of sizes greater than the response thresholds of all the receptors, redundant spatiotemporal and intensive information from all three afferent fiber types might be available for the detection of slip.

There have been numerous psychophysical studies on the absolute detection threshold of vibrotactile stimulation and its physiological substrates (e.g., Bolanowski Jr., Gescheider, Verrillo, & Checkosky, 1988; Brammer, Piercy, Nohara, Nakamura, & Auger, 1993; Gescheider, O'Malley, & Verrillo, 1983; Gescheider, Sklar, Van Doren, & Verrillo, 1985; Gescheider, Verrillo, & Van Doren, 1982; Labs, Gescheider, Fay, & Lyons, 1978; Verrillo, 1963). Psychophysical evidence that four channels participate in the perceptual process was presented in Bolanowski Jr., *et al.* (1988). In a series of experiments involving selective masking of the various channels and modification of the skin-surface temperature, four psychophysical channels were defined: P (Pacinian), NP I (non-Pacinian I), NP II and NP III. Table I-1 summarizes the major findings in Bolanowski Jr., *et al.* (1988) and previous work done by these researchers at the Institute for Sensory Research at Syracuse University (Gescheider *et al.*, 1983; Gescheider *et al.*, 1985; Gescheider *et al.*, 1982; Labs *et al.*, 1978; Verrillo, 1963).

Representative absolute detection thresholds from Bolanowski, Gescheider, Verrillo, & Checkosky (1988) (see Fig. I-1) show thresholds that are constant (at 26 dB relative to 1 μm peak, i.e., 40 μm peak-to-peak) up to about 3 *Hz*, decreasing at a rate of about −5 dB/octave up to 30 *Hz* and, then, −12 dB/octave up to 300 *Hz*, after which threshold increases.

It is important to realize that the absolute sensitivity at a particular frequency as measured psychophysically is normally determined by the "channel" having the lowest threshold, this being a function of stimulus conditions such as body site, stimulator size, duration, skin-surface temperature, and static indentation. Also, because of the substantial overlapping of sensitivities among the four channels, at suprathreshold levels and with non-sinusoidal stimuli having broad

TABLE I-1. Summary of the properties of the four mechanoreceptors.

Channel	Pacinian	NP I	NP II	NP III
frequency response range (*Hz*) and shape	40 – 800 U-shaped, min @ 300	3 – 100 almost flat	15 – 400 U-shaped	<0.4 – >100 similar to NPI
threshold (re: 1 μm) & slope at low freq.	< –20 dB @ 300 *Hz* –12 dB/oct	28 dB @ 3 *Hz* –5 dB/oct	10 dB @ 300 *Hz* –6 dB/oct	28 dB @ 0.4 *Hz* almost flat
frequency over which threshold is lowest	> 30 *Hz*	3 – 30 *Hz*	none	< 3 *Hz*
sensory attribute	vibration	flutter	unknown	pressure[a]
temperature dependence	yes	yes	yes	yes
temporal summation	yes	no	yes	no
spatial summation	yes	no	unknown	no
physiological substrate	PC	RA (Meissner)	SA II	SA I

---

[a] The perceptual quality associated with stimulation below 3 *Hz* is dependent upon stimulus intensity and stimulation site. When amplitudes are small and stimulation site can not be moved, as was true in Bolanowski *et al.* (1988) where absolute detection thresholds on thenar eminence were measured, a sense of pressure is perceived. When large amplitudes at very low frequencies are applied to the fingerpads, however, slow motions are perceived. (Footnote by the author.)

frequency spectra (e.g., pulse, ramp, noise), perceptual quality may be determined by the combined inputs from the four channels.

There is evidence that psychophysical detection threshold might be age-related. For example, Brammer *et al.* (1993) found that the threshold mediated by the Pacinian receptors at the fingertips decreased in sensitivity at an average rate of 2.6 dB per 10 years' increase in age, whereas there is little effect of age on SA and RA thresholds.

Intensity Discrimination

The dynamic range of the vibrotactile system is limited; it goes from detection threshold to roughly 55 dB above this threshold (i.e., 55 dB SL) beyond which vibrations become unpleasant or painful (Verrillo & Gescheider, 1992). The earliest study on intensity discrimination thresholds was probably done by Knudsen (1928). The right index fingertip was tested at 64, 128, 256, and 512 *Hz* with sinusoidal vibrations. It was found that the intensity JND (measured as 20 log [(A+ΔA)/A], where A=vibration amplitude and ΔA=ampiltude increment) was independent of

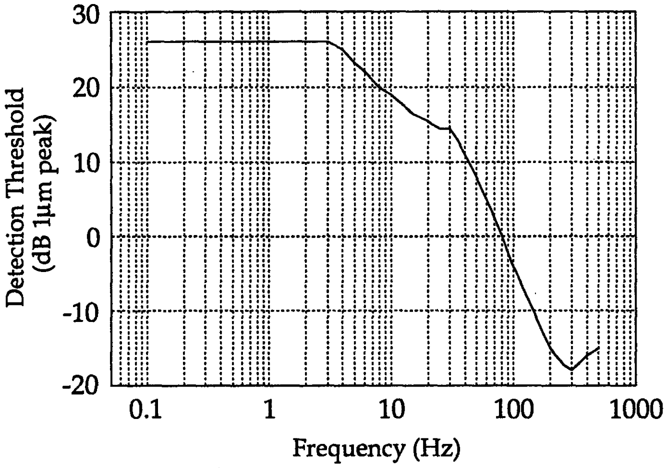

Figure I-1. Absolute detection thresholds for sinusoidal stimuli (from Bolanowski *et al.*, 1988).

frequency, decreased as intensity increased, and approached 0.4 dB when intensity was over 35-40 dB SL. In the fifties, researchers at the University of Virginia measured intensity discrimination on the ventral thorax (i.e., the chest) as part of the efforts to design the Vibratese language (Geldard, 1957). Intensity JNDs were found to decrease from 3.5 to 1.2 dB when the intensity of 60-*Hz* vibrations increased from 20 to 46 dB SL. In the seventies, Craig (Craig, 1972; Craig, 1974) studied vibrotactile difference thresholds for intensity in the absence and presence of background masking vibrations. The test stimulus was a 160 *Hz* vibration delivered to the right index finger. The masking stimulus was a 160 *Hz* vibration delivered to the right little finger. As expected, the lowest JND was obtained with no background vibration and the JND increased as the level of the masker increased from 2.0 to 20 μm peak to peak. Independent of the presence or level of the masker, the JND approached 2.0 dB when the reference intensity increased from 1 to 20 dB SL. In a more recent study (Gescheider, Bolanowski Jr., Verrillo, Arpajian, & Ryan, 1990), the thenar eminence was stimulated by a 25 or 250 *Hz* sinusoid, a narrow-band noise centered at 250 *Hz*, or a wideband noise. Of the three methods of stimulus presentation used, the continuous-pedestal method (i.e., an intensity increment imposed upon a continuous background 'pedestal' of vibration rather than on pedestals of brief duration) produced the lowest threshold. With this method, it was found that the JND decreased from 2.5 dB to 0.7 dB as stimulus intensity increased from 4 to 40 dB SL, independent of the power spectrum and frequency of the stimulus.

In summary, intensity JNDs decrease as intensity increases, are roughly independent of frequency, and range between 0.4 and 3.5 dB.

Frequency Discrimination

The frequency range over which the threshold for vibrotactile stimulation can be meaningfully measured is from DC to roughly 1 *kHz* (Verrillo & Gescheider, 1992). As Geldard pointed out (Geldard, 1957; Geldard, 1960), the task of measuring the frequency JND is not a simple one, due to the fact that perceived vibratory *pitch* depends on both the intensity and the frequency of the stimulation (more so than auditory pitch). For example, when vibratory frequency is fixed and a subject's attention is directed at pitch, a decrease (or increase) in pitch is perceived when the intensity is increased (or decreased). Thus, control for differences in subjective intensity (i.e., vibratory loudness) and for contaminating transients at onset and offset points of the stimulus envelope is crucial for *pitch* JND measurements. In this review, we use the term *frequency JND* for results obtained without equalizing vibratory loudness, and *pitch JND* for measurements made with equal-loudness vibratory stimuli.

The results on frequency discrimination JNDs are mixed. Knudsen (1928) obtained an average frequency JND (in terms of $\Delta f/f$) of 15 to 30% using 34 dB SL vibrations at 64 – 512 *Hz* delivered to the index fingertip (duration of stimuli was not documented). Mowbray & Gebhard (1957) found that the frequency JND increased from 2 to 8% when the repetition rate of intermittent mechanical pulses increased from 1 to 320 *Hz*. Pulse duration increased from 1.5 *msec* at 320 *Hz* to 7.5 *msec* at 1 *Hz*, and intensity of stimulation varied from 17 to 26 dB SL. Note that because the subjects felt vibrations through a rod between two fingers, these numbers may be lower than those from a standard one-finger experiment. The first author to use stimuli of equal subjective intensity for frequency discrimination was probably Goff (1967). Sinusoidal 1-*sec* long vibrations were delivered to the index finger. When frequency increased from 25 to 200 Hz, the pitch JND (in terms of $\Delta f/f$) increased from 18 to 36% for stimuli at 20 dB SL and from 31 to 55% for stimuli at 35 dB SL (relative to absolute detection threshold at 100 *Hz*). Rothenberg, Verrillo, Zahorian, Brachman, & Bolanowski (1977) explored the possibility of displaying the fundamental frequency of speech to the skin through frequency-varying (warbled) stimulus patterns. The volar forearm between palm and wrist was tested with four types of 1-*sec* long stimuli: constant-frequency sinusoids and pulse trains (equal subjective magnitude at 14 dB SL), and sinusoids and pulses with time-varying frequencies (at 20 dB SL). Over the frequency range of 10 to 300 Hz, the pitch JND increased from 15 to 25% for constant-frequency sinusoids, and from 10 to 35% for constant-frequency pulses. However, warble-tone stimuli improved discrimination at higher frequencies: the pitch JND decreased from 40 to 9% for warble-tones, and from 50 to 20% for warble-pulses. Another study (Franzen & Nordmark, 1975) reported a very low pitch JND of 3% at 30 dB SL. However, the methodology was non-standard and it was questionable if the authors had measured the uncertainty about the JND or the JND itself (see Rothenberg *et al.*, 1977, p.1004). Formby, Morgan, Forrest, & Raney (1992) measured vibrotactile frequency resolution on the thenar eminence using 250 *Hz*, 800 *ms* (with 10 *ms* rise/fall ramps), 100% sinusoidally amplitude modulated carriers. The overall level of each stimulus was randomized over a range of ±5 *dB* around a mean level of about 25 *dB* SL. A 2AFC adaptive procedure with 200 *ms* interstimulus interval was used. They found Weber fractions of 30 to 40% for modulation frequencies of 5 to 60 *Hz*, and much higher Weber fractions of 64 and 76% for modulation frequencies of 80 and 100 *Hz*.

In summary, the results from several studies (Knudsen, 1928; Goff, 1967; Rothenberg *et al.*, 1977; Formby *et al.*, 1992) with constant-frequency stimuli are consistent and indicate relatively poor frequency/pitch JNDs (10 to 76%) that increase with frequency over a range of 5 to 512 Hz. The study by Rothenberg (Rothenberg *et al.*, 1977) showed that when frequency-varying stimuli were used, the pitch JND decreased with frequency. These studies also indicated that the JND increased when intensity cues were eliminated by careful matching of the subjective magnitude of the frequencies being discriminated. Other studies cited above were less consistent with this picture, presumably due to differences in experimental methodology.

Temporal Resolution

Using 60-*Hz* vibrations on the ventral thorax, Geldard (1957) found that duration JNDs increased monotonically (and almost linearly) from 50 to 150 *msec* when duration increased from 0.1 to 2.0 sec. It was estimated that there were roughly 25 JNDs within the duration range tested.

Gescheider (1966) measured the time difference (Δt) between the onset of two clicks necessary for non-fused perception (defined as two temporally separated sensations, or when a rough rather than a smooth sensation was perceived). Measurements were obtained using the method of limits on either bilateral index fingertips, ipsilateral ring and index fingertips, or a single area on the index fingertip. The intensity difference between the first and the delayed stimuli ($\Delta A = A_{delayed} - A_{first}$) varied from –15 to 20 dB, with the more intense stimulus kept at 35 dB SL. Mean Δt threshold *vs.* ΔA curves for the three stimulation sites tested were U-shaped with minimal thresholds occurring at $\Delta A = 5$ dB. This minimum Δt was 12.5 *msec* for bilateral index fingertips, and 10.0 *msec* for ipsilateral ring and index fingertips, or single area on the index fingertip. (Similarly determined auditory resolution was around 1.6–1.8 *msec*.) The tactile threshold of 10.0 *msec* was very close to the response duration of the vibrator to a 1-*msec* square wave. It is unclear, therefore, whether the minimum threshold was limited by the experimental apparatus. In Gescheider (1967), two additional experiments were conducted with ipsilateral ring and index fingertips stimulation. In the first experiment, Δt decreased from 50 *msec* to 10 *msec* when $A_{delayed} = A_{first}$ varied from 10 to 35 dB SL. In the second, Δt decreased monotonically from 50 *msec* to 22 *msec* when $A_{delayed}$ increased from 10 to 35 dB SL and A_{first} was fixed at 20 dB SL. When $A_{delayed}$ was fixed at 20 dB SL and A_{first} increased from 10 to 35 dB SL, the Δt vs. A_{first} curve was U-shaped with a minimum of 30 *msec* at A_{first}=15-20 dB SL. These results were explained by hypothesizing suppressive effects of the first stimulus on the neural response produced by the delayed stimulus.

Van Doren, Gescheider, & Verrillo (1990) used a 2AFC gap-detection paradigm to measure tactile temporal resolution on the right thenar eminence as a function of age (8–75 yrs old). The stimulus contained either two 350-*msec* bursts separated by a gap, or a continuous burst with the same total duration. The bursts were either 256-*Hz* sinusoidal vibrations or bandpassed (250–500 *Hz*) noise. In any experimental run, the duration of the gap was fixed, and the threshold was measured in terms of the lowest stimulus amplitude necessary to detect the gap using a tracking paradigm. The data from noise stimuli agreed well with the data from clicks obtained by Gescheider (1967). Sinusoidal thresholds (in dB SL) were lower than noise thresholds, especially at shorter gaps. The threshold for detecting short gaps increased with age for noise stimuli, but not for sinusoidal stimuli. It was argued that the effects of age on gap detection may be due to multiple processes.

A Vibrotactile Communication System: The Vibratese Language

In the fifties at the University of Virginia, a vibratory communication system was developed and tested on three subjects (Geldard, 1957). The system consisted of five calibrated vibrators placed at the four corners and the center of a rectangle on the chest. Three intensities (*soft*, *medium*, and *loud* within 20 to 400 μm), three durations (0.1, 0.3, and 0.5 sec), and the five loci, all absolutely identifiable, formed a 45-element system. The frequency of vibrations was fixed at 60 *Hz*. Fig. I-2 illustrates the coding of the so-called *Vibratese language* which was designed to transmit single letters and digits as well as the most common English words. Note that locus, the most *distinctive* cue provided by the system, was used to encode the five vowels at the shortest duration and highest intensity. Long durations belonged to numerosity. Frequently occurring letters were assigned to the shortest duration in consideration of communication speed. The five elements corresponding to the medium intensity and longest duration were not used.

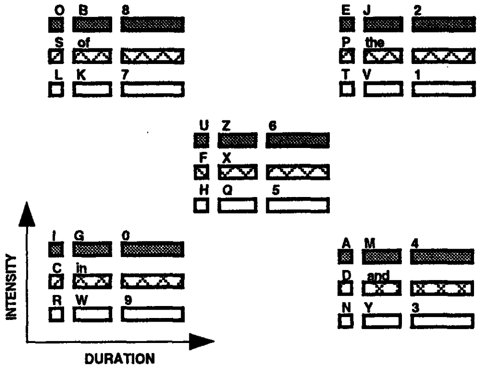

Figure I-2. Coding of the Vibratese language. Each group of nine symbols belongs to a single vibrator that varies in intensity (3 steps) and duration (3 steps). The five vibrators display all letters, all numerals, and the most frequently encountered short words. (From Geldard, 1957, Fig. 3)

The sending system was built around a typewriter and the maximum sending speed was estimated to be 67 words per minute by assuming five-letter words as the standard. Subjects learned the code for single-letter presentation in about 12 *hours*. The best subject reached a receiving plateau rate of 38 words per minute (words and sentences).

I-1.2 Kinesthetic Senses

The term kinesthesia refers to our awareness of body postures, movements, and muscle tensions. Only studies on the perception of joint positions and motions induced by external sources are reviewed here. The stimuli used in these studies are usually low-frequency signals.

The physiological mechanisms underlying kinesthetic sensing have been the topic of controversy since the turn of the century. It is now generally accepted that muscle receptors are the most likely candidates for kinesthetic detectors. The role of tactile receptors (i.e., the detection of skin stretch) is not clear and is thought to signal joint movement but perhaps not joint position. Joint receptors are found to be responsive only when the joint is near the extremes of flexion or extension. See Clark & Horch (1986) for an excellent general review of the literature on kinesthesia.

Passive Movement-Detection Threshold

Most studies on perception of passive movements involve rotating a joint smoothly at a constant velocity with minimal vibration and without visual feedback. Subjects usually have lower thresholds (i.e., better acuity) for detecting a joint movement than for identifying the direction of motion. Reported detection threshold values range from 0.20 to 6.10°, depending on the joints tested and the rate of rotation used. When detection of angular rotation is considered, proximal joints like the hip and shoulder show a greater sensitivity to movement than distal joints like those in the fingers (Clark & Horch, 1986). However, if performance is defined in terms of linear displacements of the endpoint (i.e., the fingertip for the arm), the distal joints are superior to the proximal joints (Hall & McCloskey, 1983).

The detection of passive joint movement is dependent upon the rate of rotation. Excursions that go unnoticed at one speed may become readily detectable at a faster speed. Clark, Burgess, & Chapin (1986) demonstrated that the PIP (proximal interphalangeal) and MCP (metacarpal interphalangeal) joints of the human index finger differ in their ability to detect passive joint movements. With the PIP joint, a subject's ability to detect a small change in joint position (e.g., 5°) was impaired when the rate of rotation was progressively reduced from 128 to 2 $°/min$. With the MCP joint, however, subjects could detect small (2.5°) flexion-extension displacements of the joint with no appreciable decrement in performance from almost 100% accuracy when the rate of joint rotation decreased from 128 to $1°/min$. Therefore, it was suggested that humans have a *static-position sense* with the MCP joints, but only a *movement sense* with the PIP joints. In general, movement has a more vivid character than position.

Limb Position / Movement Discrimination

In a typical limb-position matching experiment, the subject is asked to match the position of the reference limb with the other limb. The reference limb is either self- or passively-positioned and either self- or passively-maintained. Results are characterized by accuracy (i.e., offset of the mean of the errors), and precision or variability (i.e., standard deviation around the mean of the errors). Studies show that self-positioned limbs are matched more accurately than passively-positioned ones regardless of whether the reference limb is self- or passively-maintained. There is a tendency for precision to be better at the most flexed and extended joint positions than at the middle of the joint position range. There is also a tendency for accuracy to be best at the middle of the test range and biased towards the middle at other joint positions (see Clark & Horch, 1986). Jones & Hunter (1992a) measured differential thresholds for limb movement by asking subjects to compare the standard deviation (SD) of two 6-*sec* 15-*Hz* bandlimited Gaussian displacement perturbations delivered simultaneously to the two arms anchored at the elbows. They found a Weber fraction of 8%, which meant that subjects could resolve a difference as small as 5 μm between two perturbations when the reference SD was set at 50 μm. Jones & Hunter (1992b) also report a similar Weber fraction of 8% for position discrimination.

We have found that the differential threshold for the PIP and MCP joint angles is around 2.5° using active finger motions (unpublished results). The threshold is 2.0° for the wrist and elbow joints, and decreases to 0.8° for the shoulder joint (Tan, Srinivasan, Eberman & Cheng, 1994). As in the case of joint movement detection, proximal joints are more sensitive to position changes than distal ones when performance is defined in terms of joint-angle resolution. Proximal joints are less sensitive when displacements of the endpoint are considered.

I-1.3 Devices that Stimulate the Kinesthetic Senses

As mentioned earlier, most tactile communication devices employ vibrotactile stimulation characterized by high-frequency low-amplitude signals. The only two devices of which we are aware that stimulate the kinesthetic as well as the tactile components of the tactual sensory system are reviewed here. One of them was developed by Bliss (1961) at the Massachusetts Institute of Technology. Bliss's Ph.D. thesis was concerned with the development of a tactual communication system via an "inverse typewriter". It touched upon many important issues that are encountered in this thesis work. A detailed review of his thesis is provided. The other device, called OMAR, was developed recently by researchers at Gallaudet University (Eberhardt *et al.*, 1994). A brief review of the limited information currently published on OMAR will also be provided.

Review of Bliss (1961)

Four basic psychophysical experiments were conducted to guide the development of Bliss's stimulator. They were amplitude discrimination, direction identification, finger location identification, and pattern recognition with visual and tactual senses. The first three are reviewed here.

<u>*Discrimination of Passive Finger Movement Amplitudes*</u>

The stimuli were generated by sending a position-pulse to a servo motor with potentiometer feedback. The length of the linkage interfacing the servo motor and the finger was unspecified. A two-interval discrimination paradigm was used. The difference limen (DL) was defined as the movement amplitude difference between a standard stimulus (ST) and a comparison stimulus that was noticed 50% of the time. Weber fraction was computed as DL/ST. For STs of 0.70 and 0.57 *mm*[1] (with two subjects per ST, unspecified number of trials, pulse durations of 68.8 and 110 *msec*, respectively, rise time of 15 *msec*, and up-down finger motions with the unspecified finger extended), the average Weber fraction was found to be around 8% (Exps.1–4, p. 40–41, Bliss, 1961). For an ST of 0.70 *mm*, one subject was tested with pulse durations of 50, 90, and 200 msec, and the resultant Weber fractions were 5.9, 5.3, and 5.9%, respectively (Exps. 5–7, p.40–41, Bliss,

---

[1] All units were converted to metric units for consistency.

1961). This one subject was doing consistently better in the second set of experiments than under comparable conditions in the first set of experiments. In a third set of experiments (Exps. 8–10, p.40–41, Bliss, 1961), the same single subject was tested at an ST of 0.74 *mm* with three different finger postures (with pulse duration of 95 *msec*) and the Weber fraction was found to be 4.2 and 4.6% with up-down motions with the knuckle (I assume that Bliss meant the PIP joint) bent and the unspecified finger extended, respectively, and 8.1% for sidewise motions. Once again, this subject's performance improved by another 2% under conditions comparable to those tested in the first two sets of experiments.

To summarize, the Weber fraction was 4 to 8% for ST of 0.57 to 0.74 *mm*, pulse duration of 50 to 200 *msec*, and up-down as well as sidewise finger motions. More subjects need to be tested with systematically varied parameters in order to establish these results.

With a very sketchy description, the author reported a Weber fraction of 18% for pulse duration with a reference duration of 160 *msec*. The author also concluded from another briefly described experiment that the subject seemed to be able to *"accurately detect a change in the area or energy of the pulse, but he can not discriminate between a change in pulse height and a change in pulse duration."* This is an interesting result that warrants further investigation.

IT and IT rate for Passive Finger Movement Directions

A device similar to a typewriter key powered by three mutually perpendicular solenoids was used to generate motions of approximately 4.7 *mm* in six directions (see Fig. I-3). Two sets of experiments were conducted with the right index finger. The first set of experiments employed a one-interval absolute identification paradigm (unclear if feedback was given) with single 70 *msec* movement pulses. Information transfer (IT) was found to be 1.58 *bits* (46 trials, no errors) when directions 1, 2 and 3 were used, 1.57 *bits* (96 trials) when directions 1, 2, 5 and 6 were used, 1.43 *bits* (284 trials) when directions 1, 2, 3, 4 and 5 were used, and 1.54 *bits* (279 trials) when all six directions were used (number of subjects unknown in these experiments). Therefore, a maximum IT of 1.58 *bits* could be achieved when the movement directions were orthogonal to each other in a 3D space.

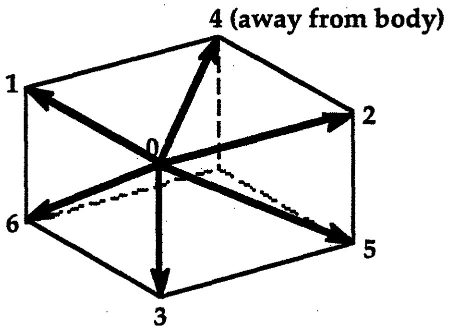

Figure I-3. The six directions of finger movement studied by Bliss. "0" is the rest position.

In the second set of experiments, two position pulses were delivered to the right index finger in rapid succession. Using all six directions, the IT rate was 3.4 *bits/sec* at a presentation rate of 2.2 *movements/sec,* and 3.52 *bits/sec* at a presentation rate of 3.7 *movements/sec.* Three higher presentation rates were tested, but data in terms of IT rate were not given.

In another experiment, a device capable of moving in the ±x, ±y and ±z directions was used to deliver a sequence of three movements at a rate of 2.8 *movements/sec.* The resultant IT rate was 4.7 *bits/sec* (6 untrained subjects and a total of 180 trials) suggesting that a higher IT rate could be achieved by moving the fingers in mutually orthogonal directions.

Identification of Finger Locations

In this experiment, pairs of fingers (the index, middle, and fourth fingers of each hand) were moved upwards for approximately 3.2 *mm* with a duration of 10 *msec.* Two subjects were asked to identify the pair of fingers being moved and each received 120 presentations. The data were presented in a 6x6 confusion matrix. It was not clear how Bliss converted the subject's response in terms of finger pairs to entries in the confusion matrix in terms of single finger locations. Nevertheless, Bliss concluded that *"(1) Most errors in finger localization result from confusion between adjacent fingers on the same hand; and (2) There were more errors involving the middle finger of each hand than any of the other four fingers used in the experiment."* (Bliss, 1961, p. 52)

The Air-Driven Finger Stimulator and Results of the Typewriter Presentation

Based on the above results, an air-driven "reverse-typewriter" was developed. This display was extremely clever; a picture of it can be found on p. 73 of Bliss (1961). Basically, the stimulator consisted of eight finger rests arranged in two groups. The user could place the two hands on these finger rests in a manner similar to typing. Each finger rest was capable of moving in ±x, ±y and ±z directions. Since each axis could be in the +, –, or neutral position, there were a total of 27 (3x3x3) states per finger rest. It was not clear how much displacement was achievable along any given axis, although it was mentioned that the amplitude of motion began to decrease rapidly at speeds greater than 15 *words/min.*

Several types of presentations were tried with the pneumatic finger stimulator. The type most relevant to this thesis is the *typewriter presentation.* In this presentation, the fingers were moved in a way similar to the active motions of a typist. For example, finger movements toward the body indicated the characters corresponding to the bottom row characters on a typewriter. Some modifications were used in order to incorporate all alphanumeric codes. For instance, lower and upper cases were indicated by the simultaneous movement of three fingers, etc.

In one set of experiments, 42 random triplets composed of the six letters *e, t, n, a, o* and *i* were presented to eight subjects. The average IT was 1.75 *bits/letter* out of a maximum possible 2.58 *bits/letter.* In another experiment, 30 symbols (the alphabet, comma, period, space, and upper case) were presented in random order with equal probability to one subject (with less than 15 hours of practice). Six sequences of 130 symbols each were delivered at a rate of 0.5 to 1.5 *letters/sec.* The subject responded verbally by naming the symbols as they were received. The IT rate, computed as the product of percent-correct scores, presentation rate (*letters/sec*) and information per symbol (4.91 *bits/letter*), reached a maximum of 4.5 *bits/sec* at a presentation rate of 1.32 *letters/sec.* Bliss argued that a higher IT rate could be achieved if the subject received more training, if more alternatives per symbol were used, and if the codes for different symbols were more evenly distributed among the fingers tested. The bandwidth of the device itself may have also been a limiting factor. Bliss commented that the typewriter presentation was easy to learn, especially if the subject had a previous knowledge of typing; but no comparison of training data from typists and non-typists was provided.

Review of Eberhardt et al. (1994)

OMAR was motivated by the desire to use kinesthetic as well as vibrotactile stimulation to present speech information as a supplement to lipreading. Its actuator was based on the head-positioning motor of a Micropolis hard-disk drive. A linear potentiometer was used as a position sensor, and an adjustable dashpot was used to assure system stability as the loading condition changed abruptly (e.g., the finger left the device briefly). Although no force sensor was used, the error gain, hence the stiffness of the system, was programmable. Sufficiently bandlimited low-frequency high-amplitude movements were closed-loop controlled. According to the authors (Eberhardt *et al.*, 1994), the system operated open-loop at vibration frequencies "due to limited dynamic range of the position potentiometer and drive circuit." Frequency responses for movements (i.e., 0.5 to 20 *Hz*) and for vibration (i.e., at 100, 200, 400, and 800 *Hz*) at several signals levels were presented. From the magnitude-gain plot for movements, the system appeared to be nonlinear. The −3 dB bandwidth at low frequencies ranged from about 19 to 10 *Hz* for nominal amplitudes from 10 to 40 *mm*. From the amplitude plot for vibrations (measured with an accelerometer), the magnitude gains (assuming 0 dB gain at 0.5 *Hz*) were approximately −36 dB at 100 *Hz*, −50 dB at 200 *Hz*, −56 dB at 400 *Hz*, and −58 dB at 800 *Hz*. As an example of how OMAR was utilized in psychophysical experiments, the psychometric function for vibration onset asynchrony (VOA) was measured. Subjects were asked to judge the asynchrony of a vibration and a movement (maximum displacement: 28.7 *mm*, rise time: unclear). The VOA time for judging that vibration started before movement was between −38 to −75 ms (a negative sign means vibration leads movement). Additional psychophysical studies are being conducted with the OMAR system. For example, Craig and Rinker (informal presentation at the Tactile Research Group meeting in Los Angeles, Nov. 9, 1995) described frequency and amplitude discrimination experiments with one or two fingers (of the same or different hands).

This page is intentionally left blank.

Chapter II.
Development of a Multi-Finger Positional Display

A multi-finger positional display was required to present various positional and movement information to the passive human fingers. The challenge was to find a single actuator that could operate over a large amplitude-frequency range so that the continual perceptual change from kinesthetic stimulation to vibrotactile stimulation could be studied. As a kinesthetic display, the device should be capable of delivering relatively large motions at very low frequencies without introducing perceptible vibrations. As a vibrotactile display, it should be capable of delivering small-amplitude vibrations at vibratory frequencies. The device should also be capable of delivering motions at intermediate frequencies with intermediate amplitudes.

II-1 Design Specifications

Table II-1 lists the design specifications for the multi-finger positional display, The TACTUATOR™. Additional comments are provided below for some of the specifications.

- *Spec. #1*

The TACTUATOR was to be interfaced with three of the five digits of the hand (i.e., the thumb, the index and the middle fingers). The thumb was chosen because it moves somewhat independently of the other fingers. The index and middle fingers were chosen because they are used with the thumb to perform most hand functions. The ring and little fingers were not considered at this time for two reasons. First, they play less important roles in daily tasks. Second, including them would significantly increase the complexity of hardware design and the final cost. Through modular design, however, it is possible to include these two fingers in the future.

TABLE II-1. Design specifications for the TACTUATOR.

No.	Category	Specs
1	No. of fingers to be moved	3 (thumb, index, and middle)
2	movement trajectory	along a line
3	controlled/sensed variable	position
4	excitable bandwidth and overall dynamic range	300 *Hz* 96 *dB*
5	maximum range of motion	25 *mm*
6	movement pattern	arbitrary within specified range of motions and frequencies
7	extraneous vibration	no perceptible high frequency noise when moving at low frequencies
8	backlash	none or minimize
9	contact site	fingerpad
10	attachment	none (i.e., the finger is not strapped to the device)
11	geometry and orientation	•with the forearm rested horizontally and the wrist kept at its neutral position, finger movements should simulate the opening/closing of a fist. •structure should fit the LEFT hand; desirable if it fits the right hand as well.
12	load characteristics	relaxed human fingers
13	audible noise	not crucial (masking noise can be used if necessary)
14	safety	use mechanical stops to limit range of motion
15	size	•moving parts should fit into the palm •no limit on other parts that are out of the way
16	flexibility	should accommodate hands of different sizes

- *Spec. #3*

What makes the TACTUATOR unique is that the *position* of the fingerpad is sensed and controlled. Many haptic interfaces use position information to control force.

- *Spec. #4*

The specification on bandwidth and dynamic range was determined using the detection threshold data from the Institute for Sensory Research at Syracuse University. The threshold *vs.* frequency plot is reproduced in Fig. II-1 using data estimated from Fig. 1 in Bolanowski, Gescheider, Verrillo, & Checkosky (1988). The threshold curve is fairly constant up to about 3 *Hz*. It then decreases at a rate of -5 dB per octave up to 30 *Hz*, and at an increased rate of -12 dB per octave up to 300 *Hz*. Beyond that, detection threshold rises again. Since the frequency response of most electromechanical systems falls off at higher frequencies, it seemed reasonable to work towards an excitable bandwidth of 300 *Hz*. The dynamic range of 96 dB was determined by noting that (1) the difference between the detection thresholds at very low frequencies and at 300 *Hz* is about 44 dB, (2) the intensity range of the vibrotactile system is limited to about 55 dB above the detection threshold, beyond which the vibrations become very unpleasant or painful (Verrillo & Gescheider, 1992), and (3) the dynamic range of a 16-*bit* A/D or D/A converter cannot exceed 96 dB.

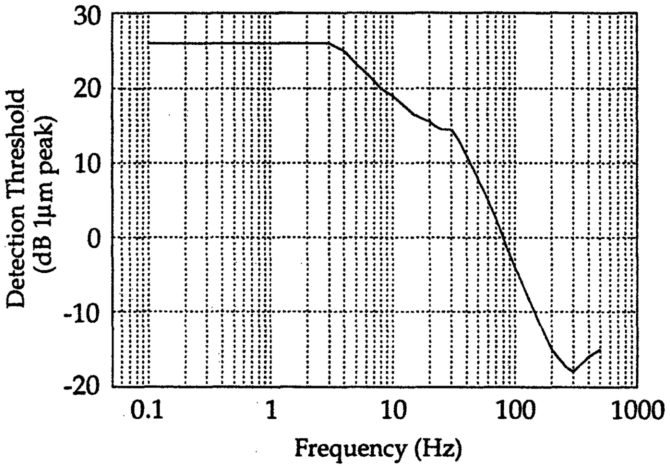

Figure II-1. Detection threshold from Bolanowski *et al.* (1988).

- *Spec. #5*

Through preliminary experimentation, we found that the index fingertip can be comfortably moved over roughly 25 *mm* (peak-to-peak displacement) at 2 *Hz* or slower. The comfortable range-of-motion decreased quickly as frequency increased. Therefore, the maximum range of motion for the TACTUATOR was set to 25 *mm*.

- *Spec. #11*

The TACTUATOR was designed mainly to fit the left hand for the following reasons. First, our experiments on tactual reception of Morse code indicate that the skill of receiving motion with one hand is readily transferrable to the other hand with minimal additional training; thus, using the non-dominant hand should not compromise one's performance. Second, assuming that most subjects are right-handed, interfacing the non-dominant hand with the device leaves the dominant hand free for entering responses. The speed at which one can enter a response is crucial to the experiments on information transmission rate.

- *Spec.#16*

Ideally, the TACTUATOR should be modifiable to fit either the left or the right hand of various sizes.

II-2 Hardware/Controller Design and Configuration

II-2.1 An Overview

The overall system is shown in Fig. II-2. There are three independent motor assemblies that are interfaced with the thumb (channel #1), the index finger (channel #2) and the middle finger (channel #3), respectively. One angular position sensor is attached to the moving parts of each of the three motor assemblies. The position sensor transforms the angular position of each actuator to a DC voltage, which is then sampled by a corresponding analog-to-digital converter. Each converter outputs a 16-bit integer at a $4\ kHz$ sampling rate. Within a TMS320C31-based DSP environment, each sampled sensor voltage is compared to a reference voltage. A 16-*bit* digital command signal is then computed from this error signal using a proportional-integral-differential controller. This command signal is converted by a corresponding 16-bit digital-to-analog converter, amplified by a power amplifier, and sent to the actuator. This process completes one cycle of the closed-loop control. The important system components are discussed below.

II-2.2 The Motor Assembly

The head-positioning motor from a Maxtor hard-disk drive was selected as the actuator because of its high bandwidth and its smooth operation at very low frequencies. Fully-assembled hard-disk

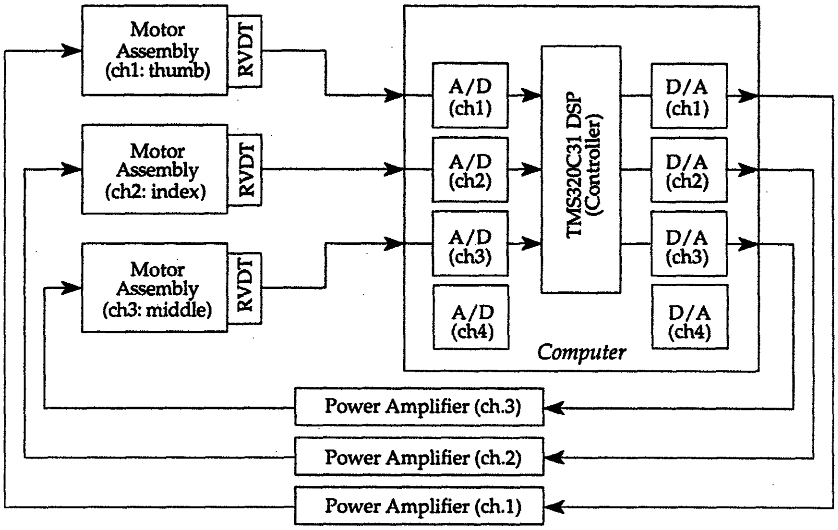

Figure II-2. The sampled-data system.

drives were stripped of electronic components. The original casing was cut so that only the head-positioning motor, its bearing and supporting structures remained. Additional hardware was designed around this remaining structure to position it in the desired orientation, to provide an interface site for the fingerpad (see "Interface with the Fingerpads" below), and to align the angular position sensor with its bearing (see "The Sensor" below).

The actuator has two built-in mechanical stops which limit its range of motion to slightly less than 30°. With an armature of length 50 *mm*, the achievable range of motion is 26 *mm*.

Electromechanical Model

Fig. II-3 is a schematic diagram of the armature-controlled DC motor. The torque delivered by the motor, T, is proportional to the input current i_a, and the ratio of T over i_a, K, is called the *motor torque constant*. The armature-winding has a small resistance, R_a, and a small inductance, L_a. The back emf voltage, e_b, is proportional to the velocity of the motor, $\frac{d\theta}{dt}$, and the ratio of e_b over $\frac{d\theta}{dt}$, K_b, is called the *back emf constant*. Finally, J and B denote the equivalent moment of inertia and the equivalent viscous-friction coefficient of the motor and load referred to the motor shaft, respectively.

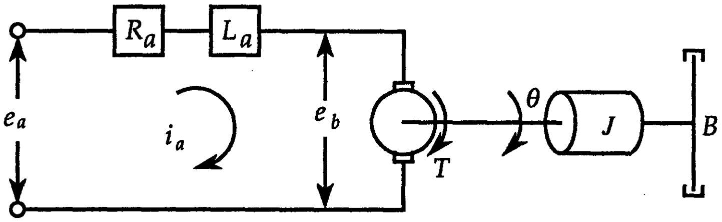

Figure II-3. Schematic diagram of the armature-controlled dc motor.

Measured with an LCR meter, $R_a = 4\ \Omega$, and $L_a = 0.3\ mH$. The torque constant, K, was estimated to be 0.2 $N \cdot m/A$ by applying known weights to the end of the arm and measuring the input current needed to hold that weight. The back emf constant, K_b, is equal to K when metric units are used. The remaining two parameters J and B could not be measured easily.

II-2.3 Interface with the Fingerpads

We considered strapping the fingerpads to the armature of the motors, but felt that such a mechanism might introduce backlash. A thimble design would probably work fine with large-amplitude slow motions, but not with small-amplitude vibrations. The final design places the fingerpads of the thumb, index finger and middle finger on aluminum pins (diameter: 4.75 *mm*) that are press-fit into the armature of the motors (see Fig. II-4). The trajectory for the thumb and that for the index or middle finger are perpendicular to each other. This configuration keeps the wrist at its neutral position and maintains a natural hand posture. This setup has worked very well for the large ranges of amplitudes and frequencies used in this study.

II-2.4 The Sensor

The sensor is a rotary variable differential transformer (Schaevitz, R30A). This sensor was chosen on the basis of its compact size (27 *mm* diameter and 22 *mm* height), high response bandwidth (1 *kHz* nominal), excellent linearity (0.09%, 0.12%, and 0.23% of full scale displacement for the three factory-calibrated R30As we have), and virtually infinite resolution (due to electromagnetic coupling of mechanical input to electrical output). The R30A works with the ATA-101 (Schaevitz) which is a power-line-operated instrument that provides excitation, amplification, and demodulation for the R30A. For our application, the excitation frequency was set to 10 *kHz* to

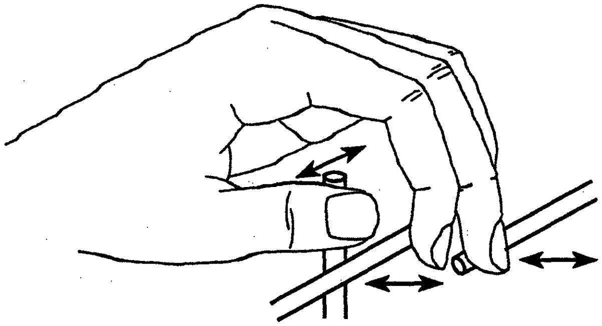

Figure II-4. Diagram illustrating fingerpad placement on the moving parts of the motor assemblies.

achieve the 1 *kHz* nominal -3 dB bandwidth. The output voltage range was adjusted to be ±3 *volt* for full range of motion of the actuator (i.e., slightly less than 30°) to match the input voltage range of the analog-to-digital converters. In addition, the three ATA-101s for the three channels were configured in a master/slave arrangement to synchronize the frequencies of the individual oscillators, thereby minimizing heterodyning interference (i.e., crosstalk) between the three channels.

The mounting of the sensor required particular care. Whereas the center of the bearing of the actuator is stationary and its armature is free to rotate, the center core of the R30A is the moving part. Therefore, the peripheral rotation of the actuator has to be converted to the motion of the sensor shaft. Four custom parts were designed to assist sensor mounting: a dummy sensor, a sensor support, a motion link, and two sensor fixtures. The dummy sensor is used to position the sensor support. The dummy sensor has the same dimensions as the R30A except that it has a longer and threaded shaft that can be screwed into the actuator bearing. This insures that the dummy sensor is always co-concentric with the actuator. The sensor support is then fixed relative to the dummy sensor. The sensor support defines the position of the body of the sensor, but allows it to be rotated so that the null position can be adjusted. Once the sensor support is fixed relative to the actuator, the dummy sensor is replaced by the R30A which is then assured to be co-concentric with the actuator bearing. The motion link serves to connect the armature to the shaft of the R30A. Finally, two sensor fixtures are used to stabilize the sensor body once its null-position has been carefully adjusted.

II-2.5 The DSP Board and I/O Modules

The TMS320C31 board (Spectrum Signal Processing) is a 2/3 length PC/AT format real-time applications platform based around the TMS320C31 32-bit floating-point digital signal processor from Texas Instruments. Two Burr-Brown daughter modules are fit into the AMELIA (Application ModulE Link Interface Adapter) sites on the C31 board, providing a total of four input and four output channels. Three input and three output channels are used for normal operation. Each daughter module has 2 input and 2 output channels using 16-bit successive approximation converters. Its sampling rate is programmable up to 200 *kHz* on inputs and 500 *kHz* on outputs. For normal operation, a sampling rate of 4 *kHz* is used and all input and output channels are synchronized. The input voltage range is ±3 *volt* maximum. The input channels include sample-and-hold amplifiers, 4th-order active Butterworth anti-alias filters, and low noise buffering. The output voltage range is also ±3 *volt* maximum. The output channels include 4th-order Butterworth reconstruction filters and low noise output buffering. The cutoff frequencies of all the 4th-order Butterworth filters are determined by interchangeable resistor packs. For normal operation, the value of all the resister packs is 39 *kOhm* in order to achieve a cutoff frequency of 1.55 *kHz* for all the filters. With this cutoff frequency, the group delay of the 4th-order Butterworth filters is fairly constant up to 300 *Hz* and averages 279 *μsec* within this frequency range.

II-2.6 The Power Amplifier

The Crown D-150A power amplifier (Crown International, Inc.) is a voltage-to-voltage power amplifier with a flat frequency response and near zero phase shift within the frequency range of interest (DC to 300 *Hz*). Although originally designed for driving loudspeakers, it is well suited to drive the hard-disk head-positioning motors with a typical resistance of 4 *ohms* and a negligible inductance (0.3 *mH*). Unlike pulse-modulated power amplifiers, the Crown D-150A introduces little additional noise or harmonic distortion. For its normal operation, the gain of the power amplifiers for all three channels is set to 2. Because the output of the digital-to-analog converters is limited to ±3 *volt*, this ensures that the input voltage to the motor is limited to ±6 *volt*.

II-2.7 Other Supporting Structures

The three motor assemblies are placed on a stool of height 50 *cm*. Foam padding is used between the motor assemblies and the surface of the stool to absorb vibration. The relative positions of the three motor assemblies can be easily adjusted. The motor assemblies are enclosed by a wooden box with an arm support. The wooden box has an opening on the top so that the thumb, index and middle fingers can rest on the moving parts of the actuators. Foam padding covers the surface of the wooden box and the arm support for subject's comfort. Finally, felt materials are used between the feet of the stool and the floor to further isolate the whole structure.

II-2.8 The PID Controller

A digital PID (positional-integral-differential) controller is used. Fig. II-5 is the signal flow chart for a single channel. $C(z)$ is the Z-transform of the digital controller, $G(s)$ is the Laplace-transform of the motor assembly, $S(s)$ is the Laplace-transform of the sensor unit, K (=2) is the gain of the power amplifier. The reference signal, r_k, is either generated by the computer or stored in memory in digital form. The input to the digital controller is the error signal, $e_k = r_k - y_k$, where y_k is the digitized sensor signal corresponding to the angular position of the actuator, $y(t)$. The output of the digital controller, u_k, is converted to an analog signal $u(t)$, amplified by K, and applied to drive the motor. Random disturbances to the motor assembly and the sensor are denoted by $w(t)$ and $v(t)$, respectively.

The controller parameters (K_p, K_i, and K_d for the proportional, integral, and differential terms, respectively) were determined by the Ziegler-Nichols PID stability-limit tuning method (Franklin, Powell, & Workman, 1990). Initially, K_i and K_d were set to 0. Then K_p was gradually increased until continuous oscillation occurred. The value of the gain (K_u) and the period of the oscillation (P_u) were recorded. The PID controller parameters were then set to $K_p = 0.6K_u$, $K_i = 2K_p/P_u$, and $K_d = K_p \cdot P_u/8$. The relevant parameters for the three motor assemblies are listed in Table II-2. The three motor assemblies have almost identical controllers.

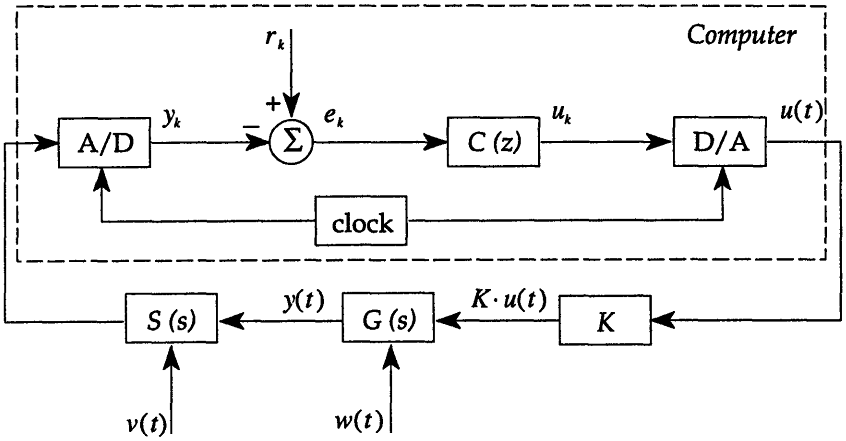

Figure II-5. Signal flow chart for a single channel.

TABLE II-2. Controller parameters for the three motor assemblies.

Motor Assembly	K_u	P_u	K_p	K_i	K_d
#1	3.30	32.0 *msec*	1.98	123.75	0.007920
#2	3.15	33.6 *msec*	1.89	112.50	0.007938
#3	3.30	33.6 *msec*	1.98	117.86	0.008316

Digital implementation of the PID controller

Fig. II-5 is a diagram of the digital PID controller. The proportional term of the digital controller is simply $K_p \cdot e_k$. The integral term of the digital controller is $K_i \cdot T_s \cdot \sum e_k$, where T_s is the sampling period and $\sum e_k$ the running sum of error signals. The integral term is reset to 0 whenever the magnitude of the running sum of errors exceeds 0.3 *volt*. The differential term of the digital controller is $v_k \cdot K_d / T_s$, where v_k is the lowpass-filtered version of the velocity estimate v'_k, $v'_k = e_k - e_{k-1}$ (the T_s term is incorporated into the K_d/T_s term). Because v'_k was noisy and caused "buzzing", it was filtered with a digital 2nd-order Butterworth filter, $B(z)$, with a cutoff frequency of 300 *Hz* (a lower cutoff frequency made the overall system hard to stabilize). The difference equation for the Butterworth filter is:

$$v_k = 1.3602 \cdot v_{k-1} - 0.5201 \cdot v_{k-2} + 0.04 \cdot v'_k + 0.08 \cdot v'_{k-1} + 0.04 \cdot v'_{k-2}.$$

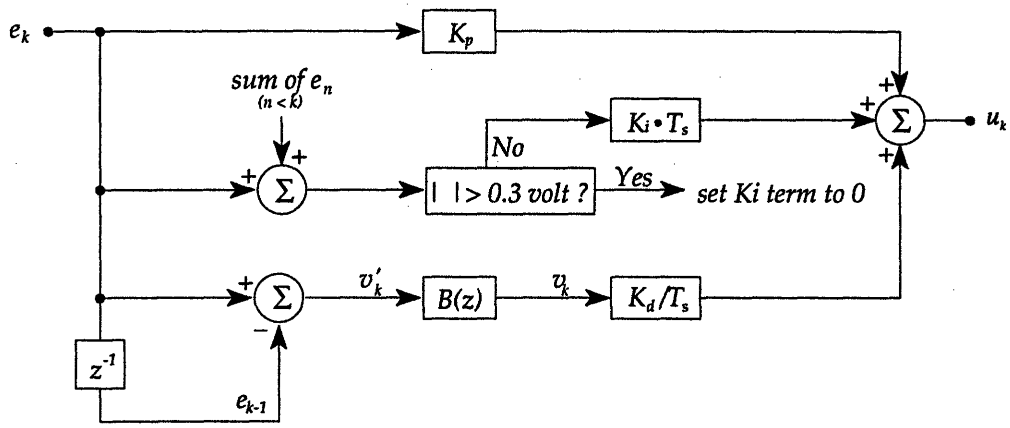

Figure II-6. The digital PID controller.

It turned out that the effect of the integral term was negligible in the sense that the overall system frequency response and step response were hardly affected by the integral term given the parameters summarized above. Therefore, we effectively have a digital PD controller.

II-3 Performance Measurements

In general, the original design specifications outlined earlier are satisfied or exceeded. The following measurements characterize the system performance further. Unless otherwise specified, the results shown in the following subsections were obtained with a Hewlett-Packard spectrum analyzer (HP 35660A Dynamic Signal Analyzer) using a continuous signal. The default input to the spectrum analyzer is the reference signal, and the default output is the sensor output signal.

Conversion between sensor voltage and fingerpad displacement

The spectrum analyzer measures signals in terms of *dB re 1 V rms (dB V rms).* In order to relate our measurements of sensor voltage to detection thresholds, we need to establish the conversion between these units and *dB re 1 μm peak.* For a sinusoidal signal $A\sin(2\pi Ft)$, 0 *dB μm peak* is equivalent to $A = 1$ *μm*, or $A\ rms = 1/\sqrt{2}$ *μm*. Note that a full sensor output range of ±3 *volt* corresponds to the full range of motion of 25.4 *mm*. Thus 0 *dB μm peak* is equivalent to

$$\left(\frac{1}{\sqrt{2}} \times 10^{-3} \times \frac{6}{25.4}\right) Vrms = 1.67 \times 10^{-4} Vrms\text{, or, - }75.54\ dB\ V\ rms.$$

Equivalently, 0 *dB V rms* is equivalent to ≈ 76 *dB μm peak.*

Sensation levels (denoted dB SL) are defined as the signal level relative to the detection threshold, computed (equivalently) either in *dB V rms* or *dB μm peak* units.

II-3.1 Frequency Response

The random source signal generated by the spectrum analyzer was sampled with the spare A/D and used as the reference signal r_k. Its level was set to 100 *mvolt.* A frequency response was measured as the ratio of the spectrum of the sampled sensor reading, y_k, and the spectrum of r_k. The command signal, $u(t)$, monitored on an oscilloscope, was mostly within ±1 *volt* and never exceeded the ±3 *volt* limit (i.e., no "clipping" occurred). The three channels exhibit very similar frequency responses. Fig. II-7 shows the frequency response of channel 3 in terms of magnitude response and group delay, measured from 0.5 *Hz* to 400.5 *Hz*. Overall, the closed-loop system behaves similar to a 2nd-order system with a -3 *dB* bandwidth of 50 *Hz* and a roughly 12 dB/octave roll-off rate at higher frequencies. The resonance frequencies of the three channels are between 28.5 and 30.5 *Hz* with resonance peaks of 4.1 to 4.5 *dB*. The largest group delay occurs at 32.5 *Hz* and is 14 *msec* for all three channels. Insofar as the systems are linear, a desired output magnitude at any frequency can be achieved by compensating for the magnitude response as shown in Fig. II-7. These systems are, however, not of minimum phase because the group delays are non-zero (i.e., ≈ 2.5 *msec*) as frequency approaches ∞.

II-3.2 Closed-Loop System Linearity

The linearity of the closed-loop system was checked in two ways. First, levels of the sensor signals were measured at a wide range of sensation levels. The extent to which a plot of sensor signal level *vs.* input signal level (both in dB units) follows a straight line of unit slope determines the linearity of the system. Measurements were taken from channel 2 under both loaded and unloaded conditions. For the loaded condition, the index finger rested lightly over the moving bar of channel 2. The reference signals for channels 1 and 3 were the 100 *mvolt* random signals

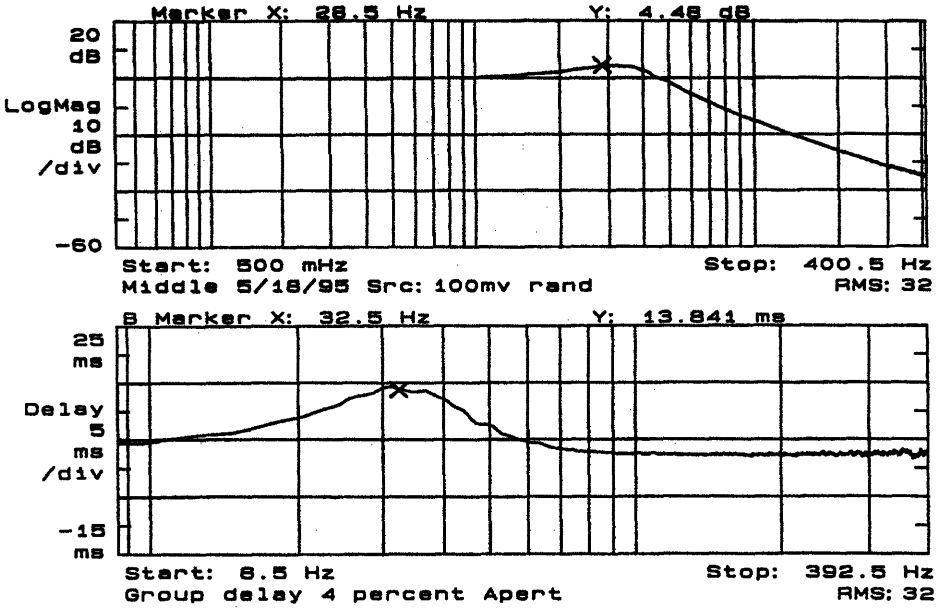

Figure II-7. Typical frequency response of the closed-loop systems measured with a spectrum analyzer (80 dB range). Above: Magnitude response. Below: Group delay.

generated by the spectrum analyzer. There was virtually no crosstalk due to noise excitation of these two channels since measurements on channel 2 were hardly affected by the presence of the noise. Results for data taken at 2, 20 and 200 *Hz* for motion levels ranging from 2 to 56 dB SL under both loaded and unloaded conditions are shown in Fig. II-8. Also shown are the best-fitting unit-slope straight lines (in the least-square-error sense). Results for measurements taken at 2 *Hz* were offset by 20 dB in Fig. II-8 for clarity. All measurements are highly linear as shown by the high correlation-coefficients (0.996 - 0.999). The effect of loading can be characterized by the differences in the intercepts of the best-fitting unit-slope lines which were 1.5 dB, 2.7 dB, and 0.1 dB for data at 2 *Hz*, 20 *Hz*, and 200 *Hz*, respectively. Finally, output levels at 200 *Hz* were saturated at the highest drive level (i.e., ≈56 dB SL) for both loaded and unloaded conditions.

The second method of checking the linearity of the closed-loop system involved measuring the system step-response and performing simulations in MATLAB. The step response was measured by recording the sensor signal from channel 2 with its reference signal set to a ± 0.2 *volt* 4 *Hz* square wave. This amplitude value was sufficiently small that no saturation of the command

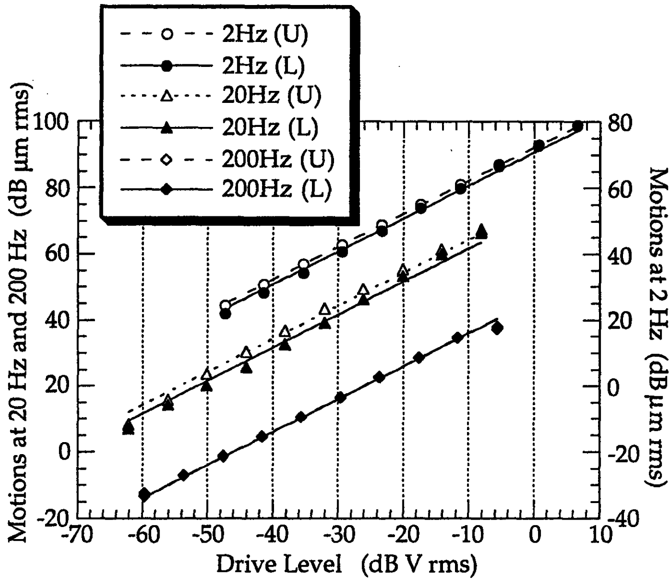

Figure II-8. Input-output relationship for channel 2 at three frequency values with best-fitting unit-slope lines. "U" and "L" denote unloaded and loaded conditions, respectively.

signal occurred. The "measured" step response in Fig. II-9 shows one half cycle of the normalized recorded sensor signal. The "measured" magnitude gain in Fig. II-9 is replotted from the upper panel of Fig. II-7. Simulations were performed by computing the magnitude gain and the step response of the closed-loop system with a 2nd order system with no zeros. The "simulated" curves in Fig. II-9 show the results of simulation using a 2nd order system with a pair of poles at $-65 \pm 200i$. Most of the features in magnitude gain and step response of the closed-loop system are captured by a 2nd order system model. This provides further evidence for the overall linearity of the closed-loop system. Note that the lower panel of Fig. II-9 shows a delay of 10 sampling periods, i.e., 2.5 *msec*, which is consistent with the group delay measurements shown earlier. Therefore, the closed-loop systems can be characterized approximately as a minimum-phase 2nd-order systems plus an excess delay of 2.5 *msec*.

II-3.3 Noise Characteristics

The noise floor of the closed-loop system was measured. Measurements were taken at the sensor output with the reference signals of all three channels set to zero. The sensor output included

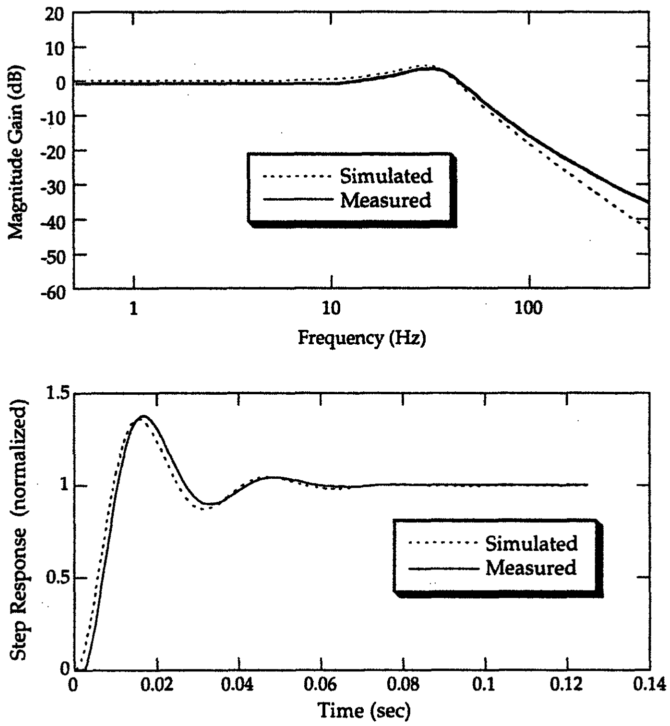

Figure II-9. Simulating magnitude gain and step response with a 2nd order system.

mechanical noise of the motor (associated with the closed loop system) as well as electrical noise with the sensor. Fig. II-10 shows the measurement from channel 1 which has the highest level of 60 *Hz* power-line noise among the three channels. It can be seen that the most prominent components of the noise spectrum are associated with the line frequency of 60 *Hz* and its harmonics at 180 and 300 *Hz*. The level of the 60 *Hz* component is - 72, - 73 and - 79 dB Vrms for channels 1, 2 and 3, respectively. The detection thresholds measured by Rabinowitz, Houtsma, Durlach, & Delhorne (1987) and Bolanowski *et al.* (1988) are plotted on top of the noise spectrum for comparison. It is clear that the noise spectrum levels are mostly below the absolute detection thresholds except for power-line components. In the worst case (i.e., around 60 *Hz*), the noise level is about 8 dB SL above the detection threshold measured by Rabinowitz *et al.* (1987).

To separate the mechanical noise from electrical noise, the above measurements were repeated with input to the motor fixed at 0 *volt* and the moving parts of the three channels fixed. The

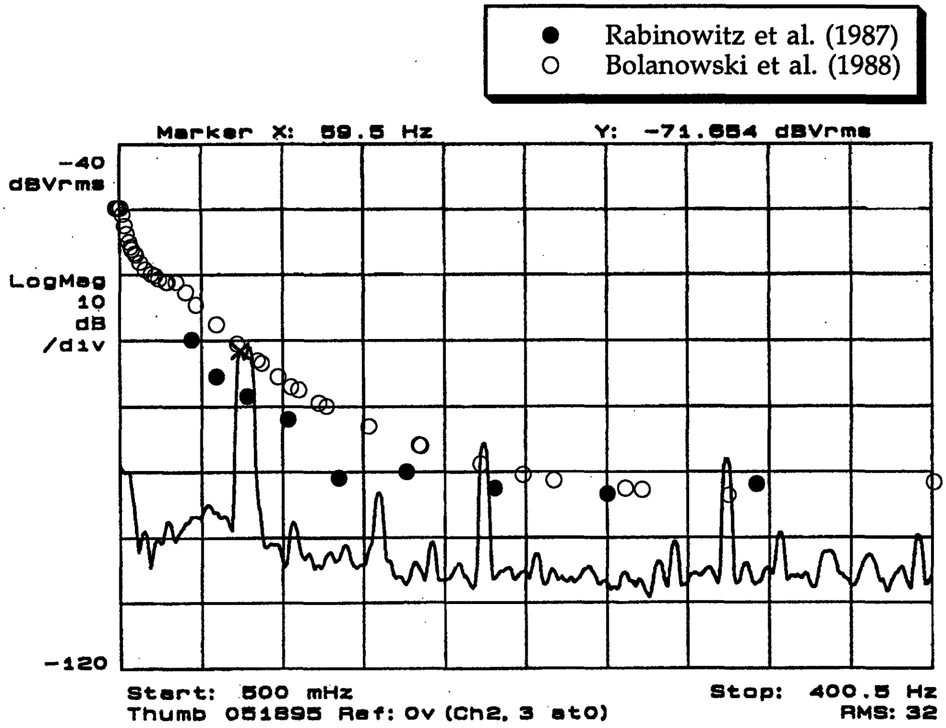

Figure II-10. Noise spectrum compared to detection thresholds.

spectrum for channel 1 was essentially the same as that in Fig. II-10 except for a 7 dB drop in the level of the 60 *Hz* component. The same was true with channel 2 and 3 with an average drop of 7.5 dB in the level of the 60 *Hz* component. Therefore, most of the background noise is electrical.

II-3.4 Harmonic Distortion

Overall system distortion produced by the motor and sensor was assessed using single tone inputs. The reference signal for the target channel was Asin(2πFt) with *F* ranging from 1 *Hz* to 300 *Hz*. The amplitude was adjusted for each frequency so that the output level (i.e., R30A reading) was roughly 56 dB SL. The levels at the fundamental frequency and at the 2nd up to 6th harmonics were recorded along with the levels at 60 and 180 Hz. All measurements were conducted with unloaded and loaded conditions. For the loaded condition, the thumb rested lightly over the moving bar of channel 1. (Because channel 1 showed the worst noise characteristics in previous measurements, the detailed harmonic measurements were conducted on this channel to reveal the worst case.) The reference signals for the other two channels were derived from a 100 *mVrms* random noise generated by the spectrum analyzer and sampled with the spare A/D. The results are presented in Fig. II-11 (shown in two panels for clarity). The upper panel shows the results for the 2nd, 3rd and 4th harmonics along with the sensor output level and absolute detection threshold – all in *dB V rms* units. The bottom panel shows the results for the 5th and 6th harmonics. Note that the harmonics are plotted at their actual frequencies. For instance, the 2nd, 3rd, 4th, 5th and 6th harmonics of a 100 *Hz* signal are plotted at 200, 300, 400, 500 and 600 *Hz*, respectively. The absolute detection thresholds are plotted at the measurement frequencies. Therefore, the harmonic levels can be directly compared with the detection thresholds plotted at the same frequency. The levels of the power-line components at 60 *Hz* and 180 *Hz* are essentially independent of measurement frequencies. When they are neither the fundamental nor one of the harmonics frequencies, the levels of the 60 *Hz* and 180 *Hz* components average -72 and -85 dB V rms, respectively. They are close to, or below, the detection thresholds at 60 *Hz* and 180 *Hz*, respectively, under both loaded and unloaded conditions.

In Fig. II-11, the data points for the absolute detection thresholds are taken from Bolanowski *et al.* (1988) for frequencies up to 500 *Hz*, and from Lamore (1984) for frequencies of 1 *kHz* and 2 *kHz*. As expected, the fundamental output curve is above the detection threshold curve by roughly 56 dB. In the upper panel, the levels of harmonics 2-4 are at least 40 dB below the output signal level for the unloaded condition (open symbols). For the loaded condition (filled symbols), however, greater distortion occurs. The maximum distortion occurs with the 2nd harmonics near 60 *Hz*. This arises because fundamental frequencies of 30 Hz nearly coincide with the system's resonant frequency. The closed-loop gain diminishes near resonance, and finger loading results in asymmetric compression of the sinusoidal stimulus, thereby increasing the 2nd harmonic distortion. However, even in this case the distortion is more than 30 dB below the fundamental output level (and tactual masking may further reduce any effect of this distortion).

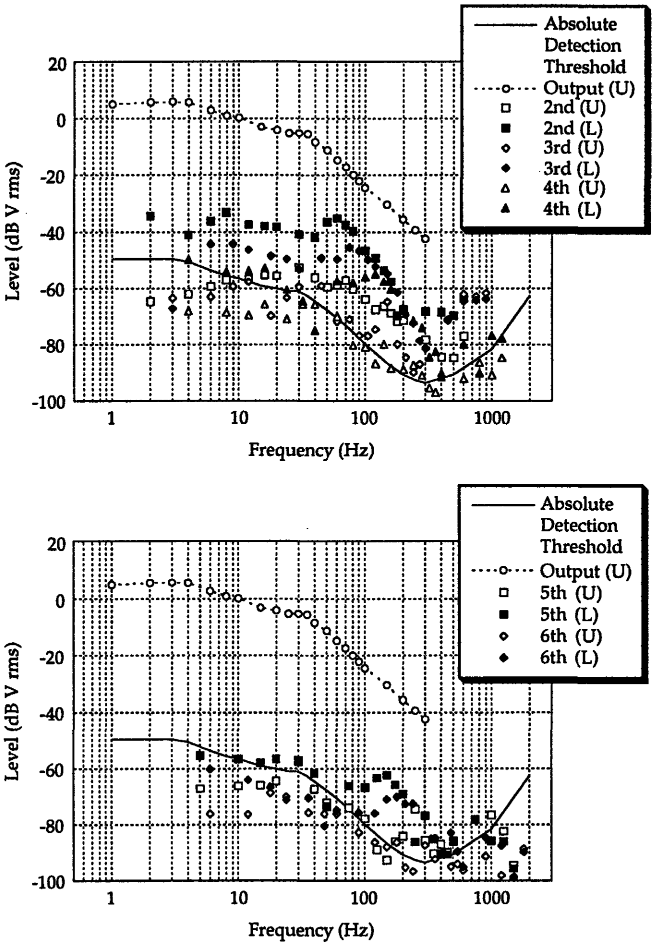

Figure II-11. Levels of sensor output signals and harmonics compared with detection thresholds. "U" and "L" denote unloaded and loaded conditions, respectively.

The lower panel shows that the 5th and 6th harmonics are at least 60 dB below the fundamental output levels. They are close to, or below, the absolute detection thresholds below 70 *Hz*, and never exceed - 60 dB *V rms* (or ≈ 15 dB μm peak).

The single frequency measurements were also used to obtain estimates of the system frequency response (circles in Fig. II-12). At an output level of 56 dB SL, there is close agreement in the responses under the unloaded and loaded conditions, except for a slightly smaller resonance peak under the loaded condition. The magnitude response derived from the above single-frequency measurements matches that obtained from the spectrum analyzer (shown as the solid line in Fig. II-12, reproduced from the upper panel in Fig. II-7) except for frequencies above 150 *Hz*. This is due to output signal saturation at these frequencies (the maximum signal level achievable is ≈ 55 dB SL at 200 *Hz*, ≈ 53 dB SL at 250 *Hz* and ≈ 51 dB SL at 300 *Hz*). In terms of subjective comfort, an output level of 56 dB feels too strong at mid to high frequencies. Therefore, the above measurements reveal the worst possible case.

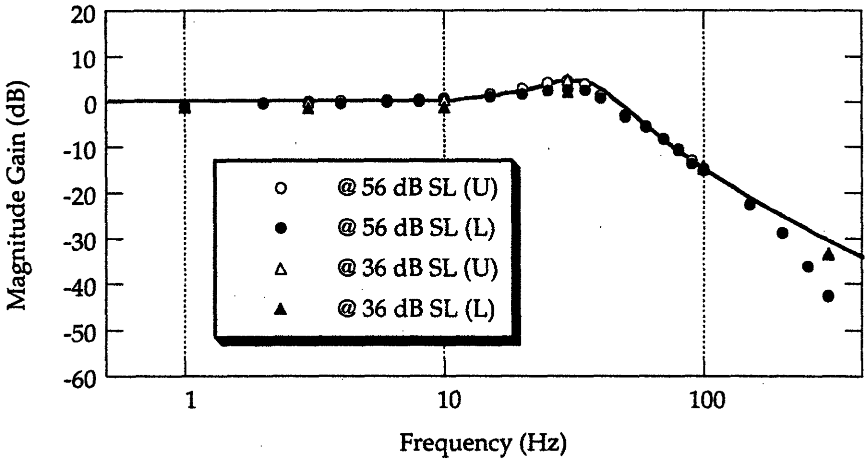

Figure II-12. Magnitude response derived from single-frequency measurements (at 56 and 36 dB SL output levels) compared with that from spectrum analyzer measurements (solid line).

To avoid output limitations, measurements were obtained on channel 1 with sensor output level kept at roughly 36 dB SL for selected frequencies (i.e., 1, 3, 10, 30, 100 and 300 *Hz*). Estimates of the system frequency response derived from these "small-signal" data (triangles in Fig. II-12) are in close agreement with that obtained with the random noise inputs over the entire frequency range. The harmonic distortion results in Fig. II-13 are mostly below the corresponding detection thresholds.

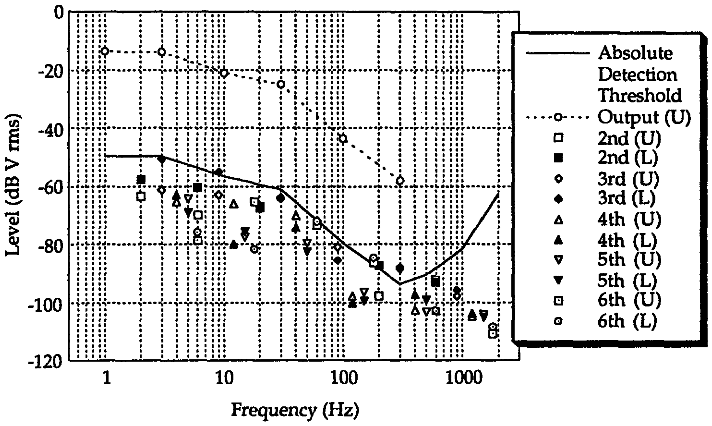

Figure II-13. Harmonics measurements at an output level of 36 dB SL.

II-3.5 Crosstalk

A sine of 2, 20, or 200 Hz was used as the reference input for channel 1. Moderate (35 dB SL) and high (55 dB SL) level tones were applied to that channel. The sensor outputs from channels 2 and 3 were measured while their reference inputs were kept at zero. The spectral component at the frequency corresponding to that of the reference signal for channel 1 was recorded and expressed in dB re motion on channel 1 (see Table II-3). At 2 *Hz*, movements on one channel cause very little crosstalk in the other two channels even at an output level of 55 dB SL (i.e., ± 11 *mm* on channel 1 and < ± 0.05 μm on channels 2 and 3). At 20 Hz, crosstalk is about −80 to −70 dB and at 200 Hz, it increases to about −40 dB. Clearly higher frequencies generate more crosstalk in other channels.

TABLE II-3. Crosstalk measurements. An asterisk indicates that the level is at the noise floor.

Frequency of the test signal (ch. 1)	Level of the test signal on ch. 1	Relative level of the spectral component at the test frequency (ch. 2)	Relative level of the spectral component at the test frequency (ch. 3)
2 Hz	55 dB SL	−107 dB*	−112 dB*
20 Hz	35 dB SL	−73 dB*	−68 dB*
20 Hz	55 dB SL	−83 dB	−77 dB
200 Hz	37 dB SL	−38 dB*	−43 dB*
200 Hz	55 dB SL	−37 dB	−46 dB

II-3.6 Spectrum of the Sum of Sinusoidal Inputs

The motion that results when a channel was driven with a sum of two or three sinusoidal inputs was assessed by its spectrum. All measurements were done on channel 2 with reference inputs to channels 1 and 3 kept at zero. Fig. II-14 shows the motion (i.e., sensor output) when 20 *Hz* and 200 *Hz* tones, each at 36 dB SL, were applied. The primary spectral peaks are at 20 *Hz* and 200 *Hz* (the signal frequencies). The component at 40 *Hz* (the 2nd harmonic of the 20 *Hz* signal) is approximately 45 dB below the 20 *Hz* signal. Components at 60 *Hz* (3×20 *Hz*) and 180 *Hz* (200 *Hz* – 20 *Hz*) are also evident, but they are at levels of residual power-line noise (described above).

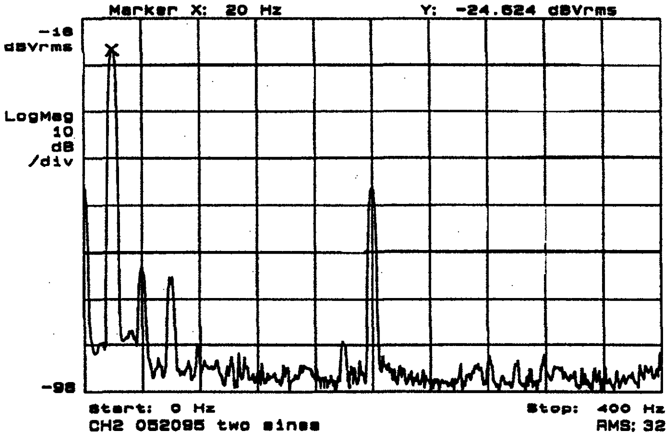

Figure II-14. Spectrum of the sum of two sinusoids at 20 and 200 *Hz*.

Fig. II-15 shows the sensor output spectrum when 200 *Hz* and 225 *Hz* tones, each at 47 dB SL, were applied. The primary spectral peaks are at 200 *Hz* and 225 *Hz* (the signal frequencies). The component at 25 *Hz* (225 *Hz* – 200 *Hz*) is approximately 20 dB below the primary signal level. The component at 50 *Hz* (the 2nd harmonic of the 25 *Hz* component) is over 30 dB below the signal level. Components at 60 *Hz* and 180 *Hz* are at levels of residual power-line noise.

Fig. II-16 shows the sensor output spectrum when 2 *Hz*, 30 *Hz* and 300 *Hz* tones, at 53, 49 and 47 dB SL respectively, were applied. This is one of the signals used in subsequent psychophysical experiments. Because of the spectrum analyzer's limited resolution, the sensor output was measured with a frequency span of 50 *Hz* (top panel) to show the spectral details of the 2 *Hz* and

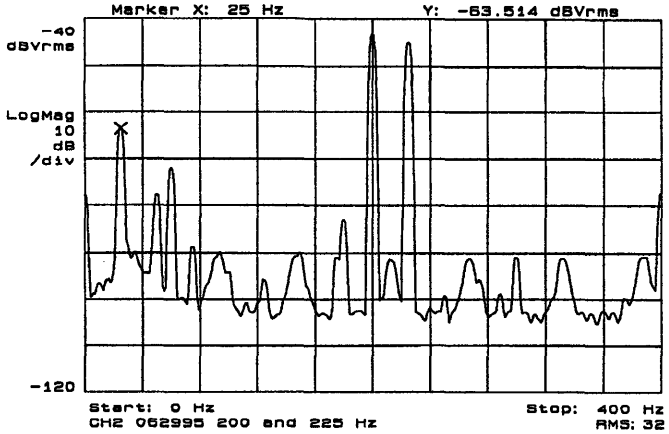

Figure II-15. Spectrum of the sum of two sinusoids at 200 and 225 *Hz*.

30 *Hz* components, and with a span of 400 *Hz* to show the 300 *Hz* component. The upper panel shows that the dominant peaks are at 2 *Hz* and 30 *Hz* (the signal frequencies) at the desired output levels. There are also components at 26 *Hz* (30 *Hz* – 2x2 *Hz*), 28 *Hz* (30 *Hz* – 2 *Hz*), 32 *Hz* (30 *Hz* + 2 *Hz*) and 34 *Hz* (30 *Hz* + 2x2 *Hz*) that are ≈ 40 dB below the 30 *Hz* component. The lower panel shows, in addition to the peaks at 30 *Hz* and 300 *Hz* (the signal frequencies), peaks at 60 *Hz* (2x30 *Hz*) and 330 *Hz* (30 *Hz* + 300 *Hz*). Overall, the peaks at the signal frequencies dominate the spectrum.

II-3.7 Absolute Detection Thresholds

Finally, as a behavioral performance verification, the absolute detection thresholds with the TACTUATOR were measured with a one-interval forced-choice paradigm. On each trial, the amplitude of the signal was either zero (i.e., no signal) or A, chosen randomly with equal *a priori* probabilities. The subject was instructed to report whether the signal was present. For each frequency tested, the values of A were chosen to be around the expected threshold. The absolute detection threshold was estimated to be the amplitude that corresponded to ≈ 70% correct performance. The results measured on the index fingers of two subjects (S_1 and S_4) were quite consistent (Fig. II-17). They were interpolated to form the absolute detection threshold curve for the TACTUATOR from 2 *Hz* to 300 *Hz* (solid line in Fig. II-17). These thresholds were 9 dB above those reported by Bolanowski *et al.* (1988) for frequencies below 30 *Hz*, and the same as those

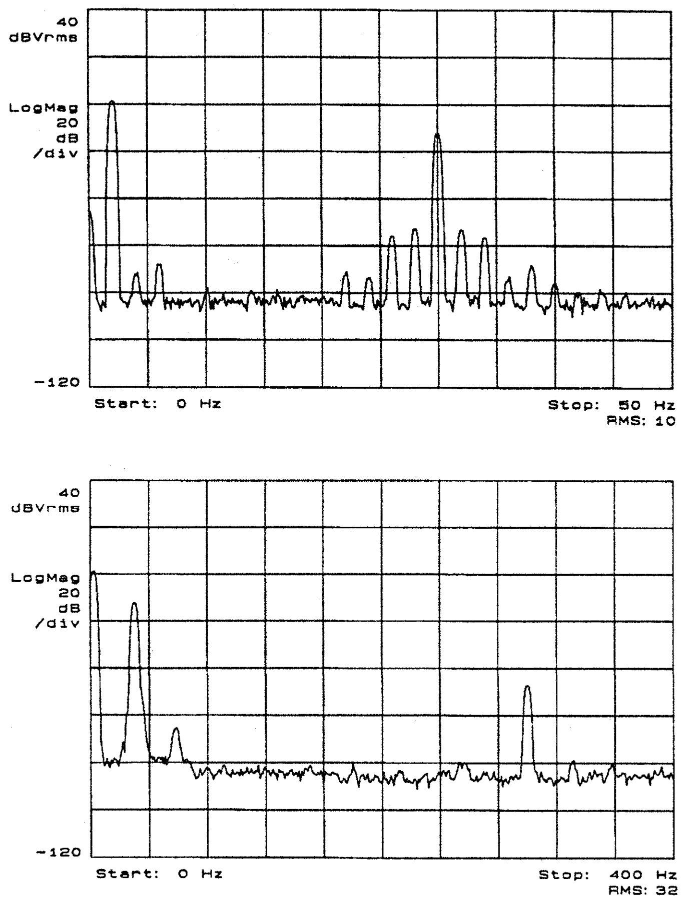

Figure II-16. Spectrum of the sum of three sinusoids at 2, 30 and 300 *Hz* measured with the spectrum analyzer with a frequency span of 50 *Hz* (upper panel) and 400 *Hz* (lower panel).

reported by Bolanowski *et al.* (1988) for frequencies above 60 *Hz*. Thresholds for the thumb and the middle finger for S_1 were measured at selected frequencies. In general, the absolute detection thresholds were quite similar for the three digits.

So far, we have based our sensation level calculations on the absolute detection thresholds reported by Bolanowski *et al.* (1988) (see Fig. II-1). In the rest of this thesis, the interpolated new thresholds shown in Fig. II-17 are used to define sensation levels in terms of the differences between signal levels and the absolute detection thresholds at the corresponding frequencies. A more important measure is tactual loudness based on subjective assessments of signal levels.

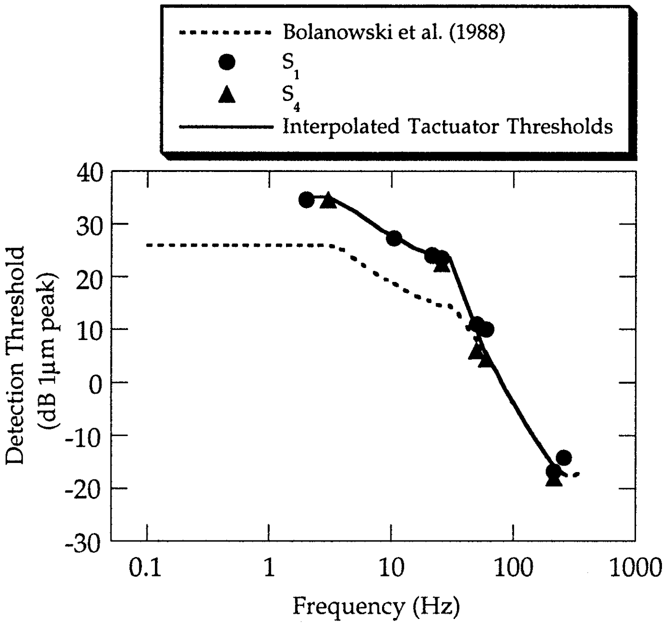

Figure II-17. Absolute detection thresholds for the index finger with the TACTUATOR.

Verrillo, Fraioli & Smith (1969) measured the equal loudness contours at 10 stimulus intensities and 10 frequencies (Fig. II-18). Each equal loudness curve defines the combinations of frequency and intensity that result in judgments of equal tactual loudness. The curves are nearly parallel, particularly for sensation levels (at 250 *Hz*) above 15 dB. The maximum discrepancy between sensation levels (at 250 *Hz*) and loudness contours for frequencies below 300 *Hz* is about 3 dB at 40 *Hz* (i.e., loudness appears to grow more rapidly at low sensation levels for low frequencies relative to the 250 *Hz* signal). It appears that sensation level is a good approximation to tactual loudness. Therefore, no efforts were made to equalize the tactual loudness of our equal-sensation-level test signals.

II-3.8 Summary

The above measurements indicate that the TACTUATOR serves as a linear positional display throughout its generating range. The useful overall dynamic range of the system exceeds 96 dB. This follows from noting that stimuli of +82 dB *μm* peak (+6.5 dB V rms) can be delivered at low frequencies and threshold stimuli near −14 dB *μm* peak (−90 dB V rms) can be delivered near

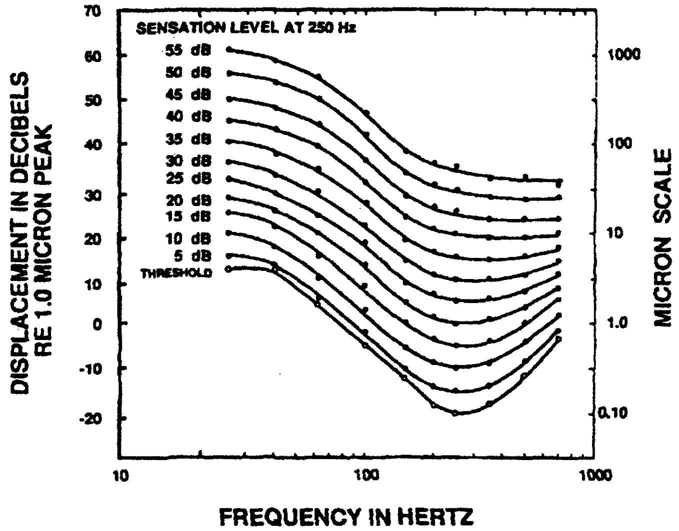

Figure II-18. Curves of equal loudness contours reproduced from Verrillo *et al.* (1969).

250 *Hz*. Distortion is generally low. Background noise, including electrical and mechanical components, as well as crosstalk between different channels, is also small. Absolute thresholds measured with the TACTUATOR are in general agreement with those reported in literature. Therefore, the TACTUATOR is well suited for a variety of multi-finger tactual perceptual studies.

This page is intentionally left blank.

Chapter III.
Static Information Transmission

This chapter is concerned with the development of three stimulus sets and the measurement of information transfer per presentation with these stimuli. These stimulus sets are referred to as the 500-*msec*, 250-*msec*, and 125-*msec* stimulus sets reflecting the differences in their signal durations. Emphasis was placed on the 500-*msec* stimulus set in terms of its construction and subject training. In Sec. III-1, we give the background for this work. In Sec. III-2, we describe the 500-*msec* stimulus set and its corresponding response set. The probe experiments used for the construction of this stimulus set are summarized in Appendix B. In Sec. III-3 – III-5, the information transfer measurements with the three stimulus sets are presented and compared.

III-1 Background

This section provides an overview of the absolute identification (AI) paradigm, the computation of information transfer (IT), issues concerning IT estimation from experimental data, and principles for maximizing IT with human observers.

III-1.1 The Absolute Identification (AI) Paradigm

The AI paradigm of interest to us in this set of experiments involves a set of *k* stimuli S_i, $1 \le i \le k$, a set of *k* responses, R_j, $1 \le j \le k$, and a one-to-one mapping between the stimuli and responses. The stimuli are presented one at a time in random order with equal *a priori* probabilities and the subject is instructed to respond to each stimulus presentation with the response defined by the one-to-one mapping, i.e., to identify which of the *k* stimuli was presented. Without loss of generality, we can assume that the stimuli and responses are labeled such that the response corresponding to the stimulus S_i under the one-to-one mapping is R_i. In other words, R_i is the correct response to S_i.

In some experiments, the subject is provided with trial-by-trial correct-answer feedback. That is, on each trial the subject is informed of the correct response after making his or her own response. 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").

The stimulus set is said to be one-dimensional if only one attribute of the stimulus (e.g., intensity) is varied. In this case, it is only the value of this variable that needs to be identified by the subject. The stimulus set is said to be multi-dimensional if more than one attribute of the stimulus (e.g., intensity and frequency) is varied. In some experiments with multi-dimensional stimulus sets, the subject is required to identify the values of all the stimulus attributes that are varied in the stimulus set (e.g., both the intensity and frequency of the stimulus). In others, the subject is required to identify the values of only a subset of the attributes that are varied and to ignore the other attributes. In this case, the stimulus attributes to which the subject must attend and respond are referred to as "target" attributes; the attributes that are varied but are ignored in specifying the response set are said to be "roved." (Stimulus attributes that are not target attributes are referred to as "background" attributes, and background attributes that are not roved are referred to as "fixed," even though their values may be changed in proceeding from one experiment to the next.) Thus, for example, if both intensity and frequency are varied within the stimulus set, but only intensity is to be identified, the experiment would be referred to as an intensity identification experiment with roving frequency. (Note that in AI experiments with roving parameters, the entities S_i referred to above are not actually individual stimuli but rather classes of stimuli.) The extent to which identification performance for a given attribute is degraded by roving another attribute provides a measure of the perceptual interaction between the two attributes.

Independent of the particular type of AI experiment under consideration, we will summarize the results in terms of the first-order stimulus-response matrix, i.e., the $k \times k$ matrix in which the entry in row *i* and column *j* specifies the number of times stimulus S_i led to response R_j. In other words, we will assume that the trials are statistically independent and ignore all possible sequential effects.

Further discussion of such notions as stimulus-set dimensionality, perceptual interaction, and stimulus-response compatibility, is included in later sections when we discuss the issue of maximizing information transfer.

III-1.2 Information Concepts and Computation

Information is something we get when we learn something we didn't know before. Any communication act provides information only insofar as it reduces a condition of ignorance or uncertainty about the state of things under consideration. The amount of information in a stimulus set (IS) is defined by the weighted sum of $\log_2 P(S_i)$:

$$IS = -\sum_{i=1}^{k} P(S_i)\log_2 P(S_i)\ , \qquad \text{(Eqn. 1)}$$

where $P(S_i)$ is the *a priori* probability of stimulus S_i, and k is the number of alternatives in the stimulus set. Information transfer measures the increase in information about the signal transmitted resulting 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, as above, $P(S_i)$ is the *a priori* probability of S_i. The average information transfer IT is thus given by

$$IT = \sum_{j=1}^{k}\sum_{i=1}^{k} P(S_i, R_j)\log_2\left(\frac{P(S_i/R_j)}{P(S_i)}\right) = \sum_{j=1}^{k}\sum_{i=1}^{k} P(S_i, R_j)\log_2\left(\frac{P(S_i, R_j)}{P(S_i)\,P(R_j)}\right), \qquad \text{(Eqn. 2)}$$

where $P(S_i, R_j)$ is the joint probability of stimulus S_i and R_j, and $P(R_j)$ is the probability of R_j. Note that the direction of communication is not important in computing IT because of the symmetry of S_i and R_j in the above equation. 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.

III-1.3 Issues in IT Estimation

The first issue in obtaining a reliable estimate of IT for a given stimulus set concerns the total number of trials to be collected. The maximum likelihood estimate of IT from a confusion matrix is computed by approximating underlying probabilities with frequencies of occurrence:

$$IT_{est} = \sum_{j=1}^{k}\sum_{i=1}^{k}\left(\frac{n_{ij}}{n}\right)\log_2\left(\frac{n_{ij}\cdot n}{n_i\cdot n_j}\right), \qquad \text{(Eqn. 3)}$$

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, IT_{est} is not only subject to statistical fluctuations, but it is also a biased estimate: it tends, for a limited number of trials, to overestimate IT. Further, the magnitude of the bias tends to greatly exceed the magnitude of the fluctuations (Rogers & Green, 1954; Rabinowitz, Houtsma, Durlach, & Delhorne, 1987).

According to Miller (1954), a useful first-order correction for the bias provided $n>5k^2$ is to subtract $\Delta IT = \frac{1}{2n\cdot\ln 2}(k-1)^2$ from IT_{est}. Miller (1954) also pointed out, however, that when $n < 5k^2$ and many of the n_{ij} $(i \neq j)$ values are near zero (i.e., transmission is good), ΔIT often results in too large a correction. 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 amount of information each stimulus attribute contributes to the overall IT in a stimulus set involving many attributes (Tan, 1988). Thus, the best way of obtaining a reliable estimate of IT is still, if at all possible, to collect sufficient data to satisfy the constraint $n > 5k^2$.

In our experiments, when the number of alternatives in the stimulus set was small (i.e., $k \leq 10$), at least $n=5k^2$ trials were conducted for each subject. When the number of alternatives in the stimulus set was large (i.e., $30 \leq k \leq 120$) and collecting $5k^2$ trials appeared too time-consuming (i.e., $4{,}500 \leq n = 5k^2 \leq 72{,}000$), a different strategy was used. According to Rabinowitz (1995, personal communication and unpublished data), IT_{est} tends to reach an asymptote faster if the performance level is high. In the extreme case, when identification performance is perfect, very few trials are needed to determine that IT = IS = $\log_2 k$ (assuming that the k alternatives are equally likely). As the percent-correct score decreases, IT decreases but the exact relationship depends upon the distribution of errors. For relatively large k (>10) and low error rate e (i.e., $0 \le e \le 0.05$, $95\% \le$ percent-correct score $\le 100\%$), the ratio of IT over IS deviates from 100% by less than twice the error rate in almost all cases (and frequently deviates by less than the error rate). Therefore, a conservative estimate of IT from percent-correct scores, denoted IT_{pc}, is given by

$$IT_{pc} = IS \times (1 - 2e) .$$

III-1.4 Principles for Increasing Information Transfer with Human Observers

Given a one-dimensional stimulus set, the information transfer for human observers is limited to roughly 2.3 to 3.2 *bits* corresponding to roughly 7±2 perfectly identifiable stimuli (Miller, 1956). This limit can be overcome by employing multi-dimensional stimulus sets. Therefore, the most important thing to do in increasing information transfer is to use stimuli with as many dimensions as possible. In this section, we discuss the principles for increasing IT for each dimension, the principles for recruiting additional dimensions, and the related issues of redundant coding, stimulus-uncertainty selection, and stimulus-response (SR) compatibility.

The principles for increasing IT for any dimension are (1) to use the entire variable range, and (2) to space stimuli in equal perceptual units. According to Braida & Durlach (1972, Fig. 4d and Fig. 8), IT increases monotonically with range of intensity for auditory intensity perception. Durlach, Delhorne, Wong, Ko, Rabinowitz, & Hollerbach (1989) also showed that IT increases with range of length for manual length identification by the finger-span method. Given the maximum stimulus range, the values of stimuli should ideally be equally spaced in terms of JNDs. If Weber's law holds for the stimulus variable under consideration (e.g., tonal intensity, force), equal perceptual distances can be accomplished by spacing stimuli logarithmically across the entire range.

For a multi-dimensional stimulus set, it is generally true that the greater the number of stimulus attributes, and the smaller the perceptual interaction among these attributes, the higher the value of IT. First, the stimulus set should include as many dimensions as possible. Although the concept of dimensionality is still not well defined (see Chap. V for discussion of this issue), it is generally true that higher dimensionality is associated with a larger number of stimulus attributes. In this sense, human faces constitute a good example of a rich display; i.e., a display with a large number of dimensions. One can easily recognize hundreds of faces because there are many facial features that contribute to the overall "look" of a face. Pollack & Ficks (1954) obtained an IT of 7 *bits* (>> 3.2 *bits*) by employing an eight-dimensional auditory display. Thus it is generally fruitful to vary as many attributes as possible in a stimulus set. In addition, the perceptual interaction between stimulus attributes should be minimized. The effectiveness of an additional stimulus attribute can be judged by the additional IT it brings compared with the additional IS it contributes to the stimulus set. Each of the eight auditory dimensions employed by Pollack & Ficks (1954) contributed 1 *bit* to the overall IS, and roughly 0.77–0.97 *bits* to the overall IT. It has also been well established that a higher IT can be obtained by using a few steps along many stimulus attributes than using many steps along a few dimensions. Pollack & Ficks (1954) showed, with six-dimensional auditory displays, that overall IT increased by 25% when the coding went from binary to trinary, but only by an additional 5% when it went from trinary to quinary. In general, extreme subdivision of a stimulus dimension does not appear warranted.

Redundant coding can be used to increase IT without necessarily increasing IS at the same time. Eriksen & Hake (1955) showed that whereas IT for unidimensional identification of size, hue or brightness of visual stimuli is between 2.3 to 3.1 *bits*, IT for identifying visual stimuli with these three attributes varying in concert is 4.1 *bits*. (In all cases, the number of stimuli was either 17 or 20.) One is thus led to the following question: given a set of stimulus attributes, is it better to vary the attributes independently or in a totally correlated way in order to increase IT? Lockhead (1966) provided some data on this issue for visual identification of line length and position. When stimulus duration was 200 *msec* and the display was well lit, IT was 1.1 *bits* for length identification, 1.0 *bit* for position identification, 1.2 *bits* when the two were perfectly correlated, and ≥ 1.7 *bits* (a lowerbound estimate) when the two were varied independently. It thus seems that higher IT can be achieved by varying stimulus attributes independently. In this thesis, emphasis was placed on discovering effective stimulus attributes and determining the number of steps along each attribute. We did not explore the option of varying stimulus attributes redundantly.

Selection of stimulus uncertainty also affects the IT that can be achieved. Because the IT *vs.* IS curve is usually thought of as being monotonic and having an asymptotic value, in addition to the fact that it is always below the straight line defined by IT = IS, the general rule of thumb is to select an IS value that is higher than the expected IT. In our first set of probe experiments on amplitude identification with fixed or roving frequencies, an IS value that was at least 1 *bit* higher than the expected IT was selected. However, we then discovered (based on a limited amount of data) that the relationship between IT and IS may not be monotonic. In particular, it appeared that IT decreased as IS increased above our estimate of the maximum IT achievable. It may therefore be more efficient to select several IS values around the expected IT in order to reveal the maximum IT. This issue is discussed further in Sec. B-2 and Chapter V.

Finally, stimulus-response compatibility also affects IT. The term "stimulus-response (S-R) compatibility" was popularized by the research of Fitts and his colleagues, in which assignments of stimuli to responses were manipulated. In one of their studies, Deininger & Fitts (1955) studied three stimulus sets and three corresponding response sets in a perceptual-motor task where the subject was instructed to move a stylus along a certain path when a particular stimulus appeared. They found that performance is best (in terms of reaction time and errors) when the response set is spatially congruent with the stimulus set *and* the matching of the points in the stimulus space to those in the response space is spatially consistent. The authors demonstrated that when the mapping between stimuli and responses was more compatible, subjects performed faster with less errors. The phenomena of S-R compatibility, however, are not restricted to situations involving physical correspondence between the stimulus and response sets, as is evident in a recent book of reviews on this topic (Proctor & Reeve, 1990). It is generally accepted that compatibility effects reflect basic cognitive processes (i.e., mental representations and translations between them) that influence human performance in a wide variety of situations. Although the relative compatibility between two groups of stimulus-response sets can be determined by subject's performance in terms of speed and error rate, there are no universal rules for the design of the optimal response set for a given stimulus set. In our experiments, many attributes (e.g., frequency, amplitude, and site of stimulation, etc.) were associated with a stimulus. A response set that reflects the salient features of the stimulus set would hopefully lead to higher performance faster. Although it is not obvious that a higher S-R compatibility necessarily leads to a higher IT plateau after extensive training, it does seem obvious that it reduces the training time required to approach such a plateau.

Further discussion of the characteristics of stimulus sets that are likely to lead to high IT is presented in Sec. V-2.

III-2 The Stimulus and Response Sets for a Stimulus Duration of 500-*msec*

A series of probe experiments were conducted to determine the effective stimulus attributes that can be used in constructing a relatively large stimulus set with easily identifiable stimuli. It was found that subjects could naturally categorize motions over a frequency range of near DC to 300 *Hz* into three perceptually distinctive categories: slow motion (up to about 6 *Hz*), fluttering motion (about 10 *Hz* to 70 *Hz*), and smooth vibration (above about 150 *Hz*). Therefore, multi-component stimuli were formed by simultaneously stimulating multiple fingers with waveforms containing sinusoids (varying in both frequency and amplitude) from the three frequency regions. The number of values to be used with each stimulus attribute was determined by employing the absolute identification paradigm with fixed and roving backgrounds. It was found that subjects could reliably identify two frequencies within each of the three frequency regions, two amplitudes with the low-frequency component, and one amplitude (i.e., fixed amplitude) with the mid- and high-frequency components, provided that masking is minimized by carefully balancing the signal strengths of components from different frequency regions. Based on the results obtained from the probe experiments and the intuitions gained from running these experiments, the 500-*msec* stimulus set was constructed. A corresponding response code was designed that reflected the underlying structure of this stimulus set. This section provides an overview of the stimulus and response sets with a duration of 500 *msec*. A more detailed description of the probe experiments can be found in Appendix B.

III-2.1 The 120 Stimuli in the 500-*msec* Stimulus Set

The structure of the 500-*msec* stimulus set is illustrated in Fig. III-1. Each stimulus (S) is defined by which finger(s) (L_i, i=1,2,3 for thumb, index finger, and middle finger, respectively) are stimulated

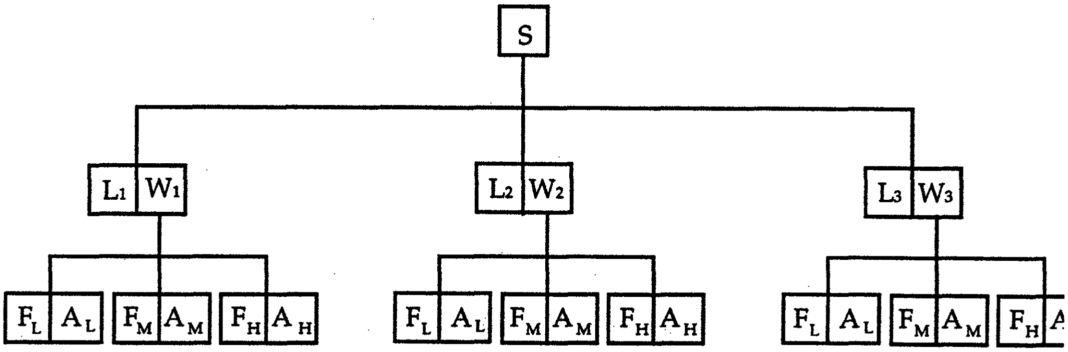

Figure III-1. The structure of a stimulus.

with which waveform (W_i, i=1,2,3). The value of L_i was 1 if the corresponding finger was stimulated and 0 otherwise. Four (4) stimulation sites were employed: either one of the three fingers was stimulated, or all three of them were stimulated with the same waveform. In other words, the possible combinations of (L_1, L_2, L_3) were (1,0,0), (0,1,0), (0,0,1) or (1,1,1), and in the case of (1,1,1), W_1=W_2=W_3. Each waveform, of which there were 30, was a broadband signal containing three sinusoidal components with frequencies and amplitudes denoted by (F_L, A_L), (F_M, A_M), and (F_H, A_H). The same set of waveforms was used to stimulate any of the fingers or all fingers. The combinations of 4 finger locations and 30 waveforms resulted in a total of 120 alternatives in the 500-*msec* stimulus set.

Among the 30 waveforms, each of which had a 10 *msec* rise-fall time, eight (8) used a single frequency (i.e., the amplitudes for two of the three components were zero), sixteen (16) used two frequencies (i.e., the amplitude for one of the three components was zero), and six (6) used three frequencies. Among the 8 single-frequency waveforms (Fig. III-2), 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.* The amplitude for each F_L was 35 dB SL or 44 dB SL. The amplitude for each F_M or F_H was fixed because the perceptual qualities of the middle- and high-frequency components were not independent of amplitude. Double- and triple-frequency waveforms were constructed by combining single-frequency elements in different frequency regions (see Table III-1 for a complete listing). 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 middle and high-frequency components were combined, only the 300-*Hz*

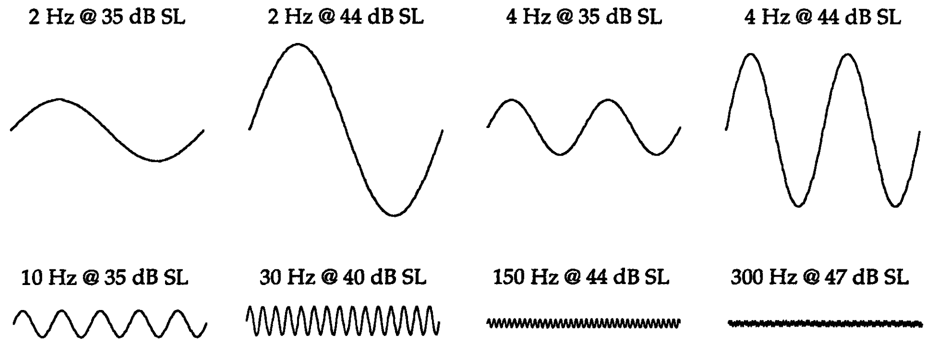

Figure III-2. Single-frequency waveforms.

TABLE III-1. The 30 waveforms for the 500-*msec* stimulus set. Units are (*Hz*, dB SL).

single-frequency waveforms	(2, 35) (2, 44) (4, 35) (4, 44)	Group1
	(10, 35) (30, 40)	Group2
	(150, 44) (300, 47)	Group3
double-frequency waveforms	(2, 35)+(10, 35) (2, 35)+(30, 40) (2, 35)+(150, 44) (2, 35)+(300, 44)	Group 4
	(2, 44)+(10, 40) (2, 44)+(30, 40) (2, 44)+(150, 44) (2, 44)+(300, 44)	Group 5
	(4, 35)+(30, 40) (4, 35)+(150, 44) (4, 35)+(300, 44)	Group 6
	(4, 44)+(30, 44) (4, 44)+(150, 44) (4, 44)+(300, 47)	Group 7
	(10, 35)+(300,44) (30, 40)+(300, 44)	Group 8
triple-frequency waveforms	(2, 35)+(10, 35)+(300, 44) (2, 35)+(30, 40)+(300, 47)	Group 9
	(2, 44)+(10, 40)+(300, 44) (2, 44)+(30, 40)+(300, 47)	Group 10
	(4, 35)+(30, 40)+(300, 47)	Group 11
	(4, 44)+(30, 40)+(300, 47)	Group 12

signal was used because the middle-frequency 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 due to excessively strong signals.

The 30 waveforms have distinctive perceptual qualities. The 2-*Hz* and 4-*Hz* signals are perceived as slow motions with 1 or 2 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 is relatively mild, and gives rise to a wobbling sensation when combined with a 2-*Hz* signal. The 150-*Hz* vibration is relatively diffused and of lower pitch. The 300-*Hz* vibration is more focused and of higher pitch. When two or three frequencies are combined, the sensations associated with single-frequency components can still be discerned.

III-2.2 The Response Code

It was a challenge to design a response set and a stimulus-response mapping that was compatible with the 120 stimuli. Intuitively, it seemed that the response set should reflect the underlying structure of the stimulus set; e.g., each response should consist of two parts, one corresponding to stimulation site, and one to stimulating waveform. After preliminary experimentation, it seemed that a graphical response code might work better 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 fingers, respectively (Fig. III-3). A "DEL" icon was available for deleting responses if the subject felt that the wrong icon was accidentally pressed. An "ENTER" icon was used to terminate a trial. 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 across a column of waveform icons. Some exceptions were made in order to contain the waveform icons to a relatively small area (for ease of visual search). Subjects used a stylus to pick the appropriate response icons by pressing on them. Since the 150 *Hz* and 300 *Hz* waveforms did not reproduce very well at that scale, they were represented by blue and red dots in the actual response tablet.

III-3 Information Transfer Measurements with the 500-*msec* Stimulus Set

III-3.1 General Methods

Three subjects (S_1, S_2 and S_3) were trained and tested. S_1 (the author) is a 30 year old female graduate student at MIT; S_2 is a 42 year old male who also participated in our earlier study on

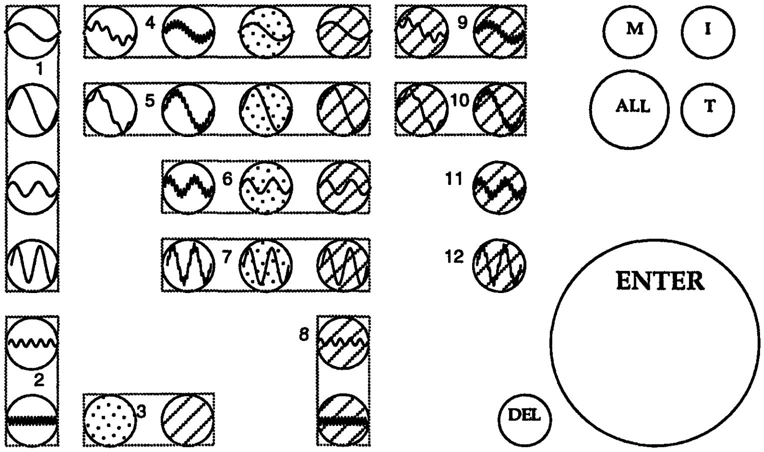

Figure III-3. Layout of responses for experiments using the 500-*msec* stimulus set. The numbers correspond to the grouping listed in the rightmost column of Table III-1. (The dot and line patterns represent the blue and red colors used in actual icons for the 150 *Hz* and 300 *Hz* waveforms, respectively.)

tactual reception of Morse code; and S_3 is a 20 year old undergraduate student at MIT. All subjects are right-handed with no known tactual impairments of their hands.

During all experiments, the TACTUATOR was visually blocked from the subjects. Subjects wore earplugs and earphones with pink noise to eliminate auditory cues. (The TACTUATOR produces no audible noise except at 300 *Hz*.)

For both training and testing, the standard AI paradigm with trial-by-trial correct-answer feedback was employed. There were two differences, however, between the paradigms used for training and testing. During training, each stimulus alternative was presented an equal number of times per run (i.e., randomization without replacement). This ensured that subjects had an equal opportunity to learn all the signals in the stimulus set. The side effect was that stimulus uncertainty decreased as a function of number of trials. During testing, stimuli were presented with equal *a priori* probabilities on each trial (randomization with replacement). Thus, stimulus uncertainty remained the same throughout an experimental run. The other difference was that during training, subjects were allowed to skip trials. Skipped trials were repeated later on. During testing, subjects were required to respond to all trials.

Training was conducted for all stimulus durations before testing was done. There was a two months gap between the end of training and the beginning of testing for subjects S_1 and S_2. Subject S_3 was tested immediately after he completed training. Results for training are presented in terms of percent-correct scores and total number of hours; results for testing are summarized in terms of IT. The discussion of the AI paradigm and IT computation presented earlier is applicable to the testing procedures. Because of the relatively large number of alternatives in the stimulus set (i.e., up to 120), IT was computed as $IT_{pc} = IS_k \times (1 - 2 \times \bar{e})$, where $IS_k = \log_2 k$, k is the number of equally-likely alternatives in the stimulus set),[1] and $\bar{e}$ is the average error rate. IT_{est} was also computed. However, because the amount of experimental data is relatively limited, IT_{est} should be treated as an upperbound.

III-3.2 Training Results

Subjects learned to identify the 120 alternatives in the 500-*msec* stimulus set in a number of steps. They were first trained to identify the 30 waveforms on the index finger, then to identify the same 30 waveforms when applied randomly to any one of the three fingers, and finally to identify both the finger locations and the waveforms of all 120 alternatives in the 500-*msec* stimulus set. For each stimulus set, training was terminated when a subject reached the performance criterion of either one run of 100% correct or three runs with percent-correct scores of 95% or higher (not necessarily consecutively).

Waveform Identification on the Index Finger

The waveforms in the 500-*msec* stimulus set were divided into 12 groups (see Table III-1 and Fig. III-3): the waveforms in the first group contained only F_L components; those in the second group contained only F_M components; and those in the third group contained only F_H components. Groups 4 to 8 contained double-frequency waveforms. Groups 9 to 12 contained triple-frequency waveforms. Subjects first practiced with and identified the 4 waveforms in group 1, then the 6 waveforms in groups 1 and 2, then the 8 waveforms in groups 1 to 3, and so on until the stimulus set contained all 30 waveforms.

---

[1] $IS_k = \log_2 k$ is derived by substituting $P(S_i) = 1/k$ in Eqn. 1.

During practice, subjects could choose to feel any waveform on the index finger by selecting the corresponding waveform icon on the response tablet. Practice was self-terminated when they a subject felt ready to run the identification experiments. Each stimulus alternative was applied exactly 5 times to the index finger for each experimental run. In other words, the number of valid trials per run, upon which percent-correct scores were computed, was 20 for group 1, 30 for groups 1 and 2, etc. Subjects had to reach the performance criterion before new waveforms were added to the stimulus set. Results were summarized either as 100% or by averaging the last three percent-correct scores (Table III-2).[1] Also, the total number of hours of training was recorded for S_2 and S_3.[2] The three subjects were able to reach the performance criterion with the 30 waveforms in the stimulus set with an average accuracy of 100% (S_1), 100% (S_2, 9 *hours*), and 96% (S_3, 15 *hours*).

TABLE III-2. Average percent-correct scores from waveform identification on the index finger.

No. of Stimuli	S_1	S_2	S_3
4 (Group 1)	-	100%	100%
6 (Groups 1-2)	-	100%	100%
8 (Groups 1-3)	-	100%	96%
12 (Groups 1-4)	-	96%	100%
16 (Groups 1-5)	-	98%	97%
19 (Groups 1-6)	-	100%	98%
22 (Groups 1-7)	-	96%	97%
24 (Groups 1-8)	-	96%	96%
26 (Groups 1-9)	-	97%	96%
28 (Groups 1-10)	-	95%	95%
29 (Groups 1-11)	-	97%	93%
30 (Groups 1-12)	100%	100%	96%

Waveform Identification with Roving Fingers

To quickly check whether the 30 waveforms were readily identifiable when applied to the thumb or the middle finger, the subjects practiced with and identified these waveforms again when they

---

1. Because S_1 was highly experienced with these stimuli, she was only tested with all 30 waveforms in the stimulus set.
2. Total number of hours of training could not be accurately estimated for S_1 because she was involved in the development of all stimulus sets and, therefore, was over-exposed to the stimuli.

were applied randomly to any one of the three fingers. There were 90 alternatives (30 waveforms × 3 finger locations) in the stimulus set. The response set still consisted of the 30 waveform icons. During practice, the subject could select any combination of stimulation site and waveform by picking the "M", "I" or "T" icon followed by a waveform icon. During a training run, however, only the waveform response was required. Each of the 30 waveforms was applied exactly twice to each of the three fingers during each run, resulting in a a total of 180 non-skipped trials per run.

It was noticed that because of the difference in range of motion of the three fingers, adjustment in signal amplitude was needed to equalize loudness perception of the low-frequency signals. Informal testing was done in which S_1 was presented with one of the 4 waveforms in group 1 on the index finger, and a signal of the same frequency on the thumb or the middle finger. S_1 could adjust the amplitude of the signal on the thumb or the middle finger until it felt equally "loud" to the one on the index finger. It was found that equal tactual loudness could be achieved by increasing the amplitude of the 2 *Hz* and 4 *Hz* signal components by 2 dB for the middle finger and decreasing it by 2 dB for the thumb. With this modification, all subjects were able to reach the performance criterion with an average accuracy of 99% (S_1), 98% (S_2, 1 *hour*), and 94% (S_3, 1 *hour*). These results indicated that the thirty waveforms could be well identified using any one of the three fingers.

Identification of All 120 Alternatives

Subjects were now ready to be trained with all 120 alternatives in the 500-*msec* stimulus set. Each stimulus alternative was applied twice during each training run, resulting in a total of 240 non-skipped trials per run. Subjects were instructed to first select the "M", "I", "T" or "ALL" icon for site of stimulation, then the waveform icon corresponding to the stimulus as a response. All three subjects were able to reach the performance criterion with an average accuracy (over the last three runs) of 98% (S_1), 96% (S_2, 3.5 *hour*) and 96% (S_3, 3.5 *hour*).

Learning curves for the three subjects are presented in Fig. III-4. All subjects were able to reach the performance criterion within 10 training runs.

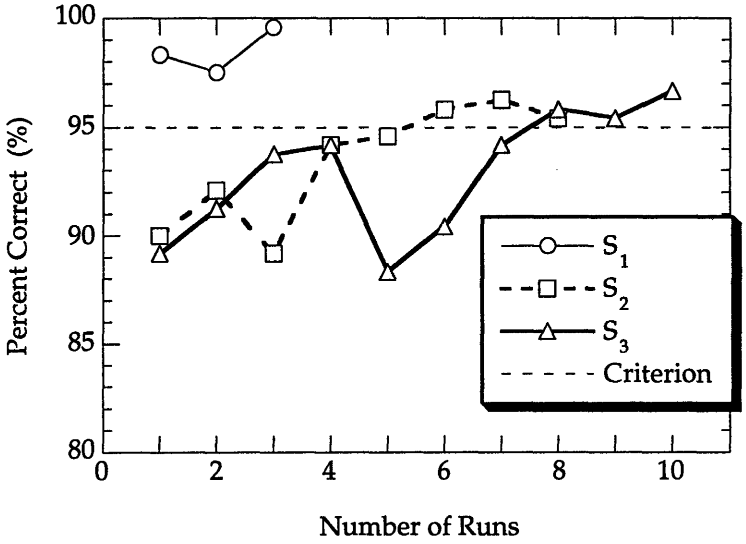

Figure III-4. Learning curves for each subject with all 120 stimuli in the 500-*msec* stimulus set.

The number of skipped trials for the last three runs were 12/26/22 for S_1, 3/0/3 for S_2, and 48/51/37 for S_3. S_2 rarely skipped a trial; when he did, either too few or too many response icons were in the response, indicating that efforts were made to respond to those trials. S_1 and S_3, however, skipped many trials intentionally (i.e., their responses were blank).

III-3.3 Test Results

During final testing, three runs (each containing exactly 240 trials) were collected for each subject. Results are summarized in Table III-3. The IT_{pc} averaged over all three subjects is 6.5 *bits*, corresponding to 90 perfectly identified items. The values of IT_{est} are not much higher than IT_{pc}, mainly because they cannot exceed the IS value of 6.9 *bits*.

TABLE III-3. Information Transfer with the 500-*msec* stimulus set.

Subject	*Percent-Correct Scores*	IT_{est}	*Average*	IT_{pc}	*Average* IT_{pc}
S_1	99%, 99%, 98%	6.8 *bits*	99%	6.8 *bits*	**6.5 *bits***
S_2	94%, 95%, 96%	6.6 *bits*	95%	6.2 *bits*	
S_3	96%, 97%, 97%	6.6 *bits*	97%	6.5 *bits*	

III-4 Information Transfer Measurements with the 250-*msec* Stimulus Set

III-4.1 Stimuli

The 250-*msec* waveform set contained 30 waveforms that were very similar to those in the 500-*msec* waveform 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*. With these changes, the two F_L values of 2 and 6 *Hz* could be easily discriminated, as were the higher F_L value of 6 *Hz* and the lower F_M value of 15 *Hz*. A complete listing of the 30 waveforms is shown in Table III-4. The same four finger locations were used: all fingers, thumb alone, index finger alone, and middle finger alone. Therefore, there were again a total of 120 alternatives in the stimulus set. The response code shown in Fig. III-3 was modified to take account of the waveform changes.

TABLE III-4. The 30 waveforms for the 250-*msec* stimulus set. Units are (*Hz*, dB SL). Signals that are different from those in the 500-*msec* stimulus set are underlined.

single frequency	(2, 35) (2, 44) (6, 35) (6, 44) (15, 35) (30, 40) (150, 44) (300, 47)
double frequency	(2, 35)+(15, 35) (2, 35)+(30, 40) (2, 35)+(150, 44) (2, 35)+(300, 44) (2, 44)+(15, 40) (2, 44)+(30, 40) (2, 44)+(150, 44) (2, 44)+(300, 44) (6, 35)+(30, 40) (6, 35)+(150, 44) (6, 35)+(300, 44) (6, 44)+(30, 44) (6, 44)+(150, 44) (6, 44)+(300, 47) (15, 35)+(300,44) (30, 40)+(300, 44)
triple frequency	(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)

III-4.2 Training Results

The training procedure was the same as that used with the 500-*msec* stimulus set. Subjects practiced with and identified all 120 stimuli in the 250-*msec* stimulus set. Each stimulus was presented twice during each training run. All subjects were able to reach the performance criterion with an average accuracy (over the last three runs) of 100% (S_1), 96% (S_2, 3.5 *hours*), and 93% (S_3, 6 *hours*).

Learning curves for the three subjects are presented in Fig. III-5. All subjects were able to reach the performance criterion within 20 training runs.

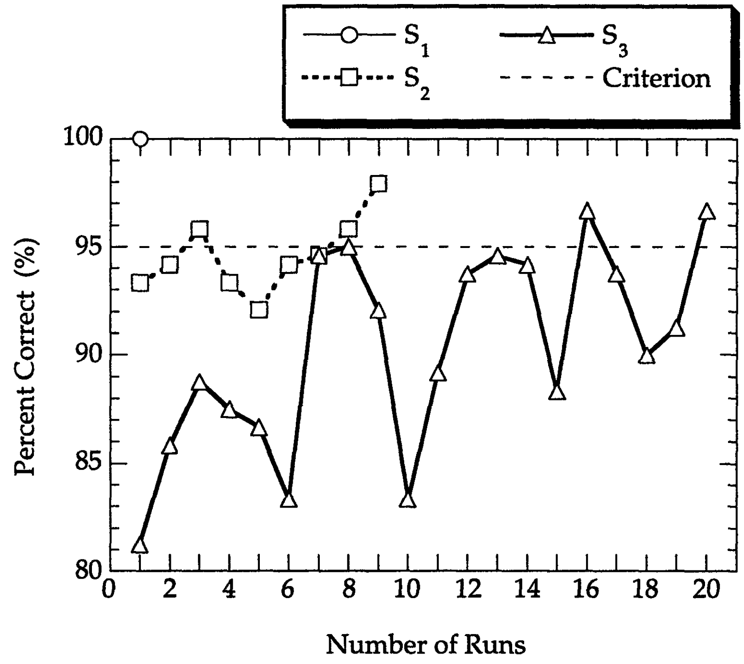

Figure III-5. Learning curves for each subject with all 120 stimuli in the 250-*msec* stimulus set.

The number of skipped trials for the last three runs were 22 for S_1 (one run of 100%), 4/0/1 for S_2, and 38/25/63 for S_3. Again, S_2 rarely skipped a trial; but S_1 and S_3 skipped many trials.

III-4.3 Test Results

During final testing, there were exactly 240 trials per experimental run. After the initial three runs, it appeared that S_2 and S_3 had not yet reached a performance plateau (Fig. III-6). Therefore, a total of 10 runs were collected with these two subjects. Scores for the last three runs are averaged and summarized in Table III-5. The IT_{pc} averaged over all three subjects is 6.4 *bits*, corresponding to 84 perfectly identified items. Thus, there is very little loss (i.e., 0.1 *bit*) in IT by shortening the stimulus duration by a factor of 2. The values of IT_{est} are, again, not much higher than IT_{pc}, mainly because they cannot exceed the IS value of 6.9 *bits*.

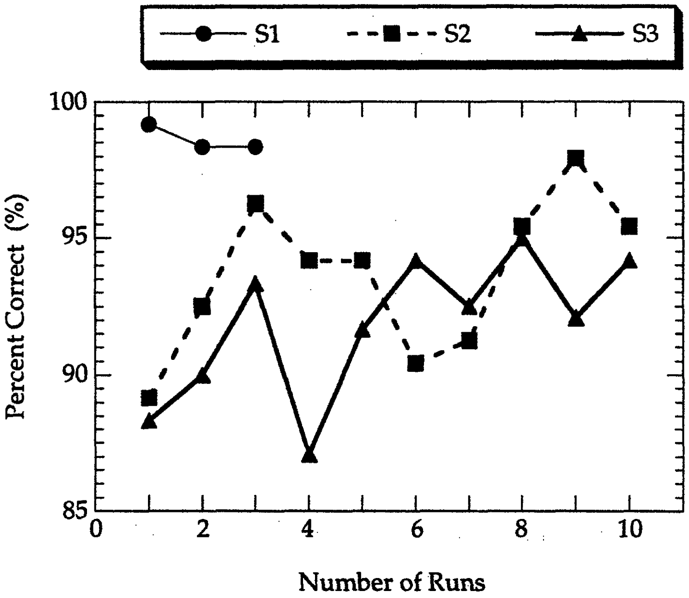

Figure III-6. Test results in terms of percent-correct scores for each subject with the 250-*msec* stimulus set.

TABLE III-5. Information Transfer estimated from the last three test runs conducted with the 250-*msec* stimulus set.

Subject	*Percent-Correct Scores*	IT_{est}	*Average*	IT_{pc}	*Average* IT_{pc}
S_1	99%, 98%, 98%	6.8 *bits*	99%	6.8 *bits*	**6.4 *bits***
S_2	95%, 98%, 95%	6.7 *bits*	96%	6.3 *bits*	
S_3	92%, 95%, 94%	6.6 *bits*	94%	6.1 *bits*	

III-5 Information Transfer Measurements with the 125-*msec* Stimulus Set

III-5.1 Stimuli

With a signal duration of 125 *msec*, subjects could no longer reach the performance criterion of $\geq$ 95% with 30 waveforms that were similar to the ones in the 500-*msec* 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. Direction of motion was also introduced as an additional signal attribute.[1] The resultant 125-*msec* waveform set contained 19 waveforms, as shown in Table III-6. A negative sign indicates that movements started in a direction that corresponds to finger flexion. The default used in this and all previous experiments was to start movements in the finger-extension direction. The direction attribute was only effective with the F_L components.

TABLE III-6. The 19 waveforms for the 125-*msec* stimulus set. Units are (*Hz*, dB SL).

single frequency	(4, 35) −(4, 35) (4, 44) −(4, 44) (30, 40) (300, 47)
double frequency	(4, 35)+(30, 40) −(4, 35)+(30, 40) (4, 44)+(30, 40) −(4, 44)+(30, 40) (4, 35)+(300, 47) −(4, 35)+(300, 47) (4, 44)+(300, 47) −(4, 44)+(300, 47) (30, 40)+(300, 47)
triple frequency	(4, 35)+(30, 40)+(300, 47) −(4, 35)+(30, 40)+(300, 47) (4, 44)+(30, 40)+(300, 47) −(4, 44)+(30, 40)+(300, 47)

Again, the same four finger locations were used: all fingers, thumb alone, index finger alone, and middle finger alone. Therefore, there were a total of 76 alternatives in the stimulus set. The response code shown in Fig. III-3 was modified accordingly.

III-5.2 Training Results

The training procedure was the same as that used before with the 500-*msec* and 250-*msec* stimulus sets. Each stimulus was presented twice during each training run. All subjects were able to reach the performance criterion with an average accuracy (over the last three runs) of 100% (S_1), 96% (S_2, 2.5 *hours*), and 96% (S_3, 1.5 *hours*).

---

[1] Direction of movement was not an effective signal attribute for longer signal durations.

Learning curves for the three subjects are presented in Fig. III-7. All subjects were able to reach the performance criterion within 9 training runs.

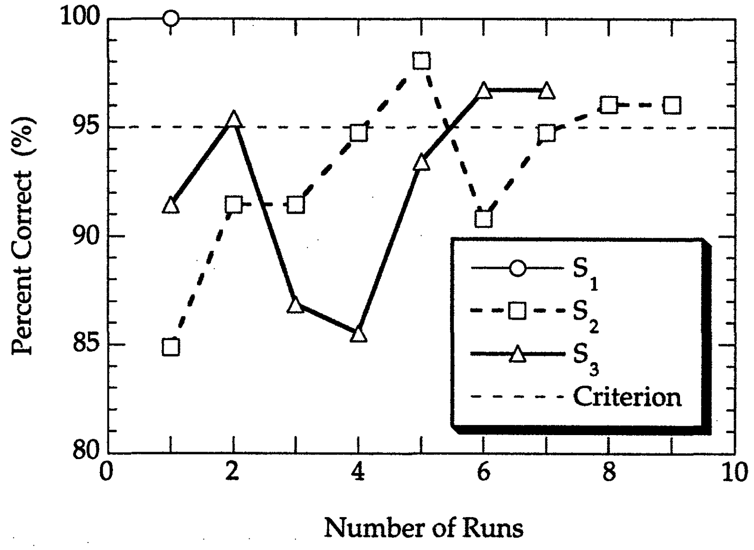

Figure III-7. Learning curves for each subject with all 76 stimuli in the 125-*msec* stimulus set.

The total number of trials skipped were 43 for S_1 (one run of 100%), 1/0/0 for S_2, and 59, 47, 41 for S_3. Although S_2 continued to skip very few trials, the percentage of skipped trials increased dramatically for S_1 and S_3.

III-5.3 Test Results

During final testing, there were exactly 152 trials per experimental run. After the initial three runs, it was not clear if S_2 and S_3 had reached a performance plateau (Fig. III-8). Therefore, one more run was collected with these two subjects. Scores for the last three runs are averaged and summarized in Table III-7. The IT_{pc} averaged over all three subjects is 5.6 *bits*, corresponding to approximately 50 perfectly identified items. Overall, there is an approximately 1 *bit* decrease in IT when the signal duration was reduced from 500 (or 250 *msec*) to 125 *msec*. The values of IT_{est} are, again, not much higher than those of IT_{pc} and limited by the IS value of 6.2 *bits*.

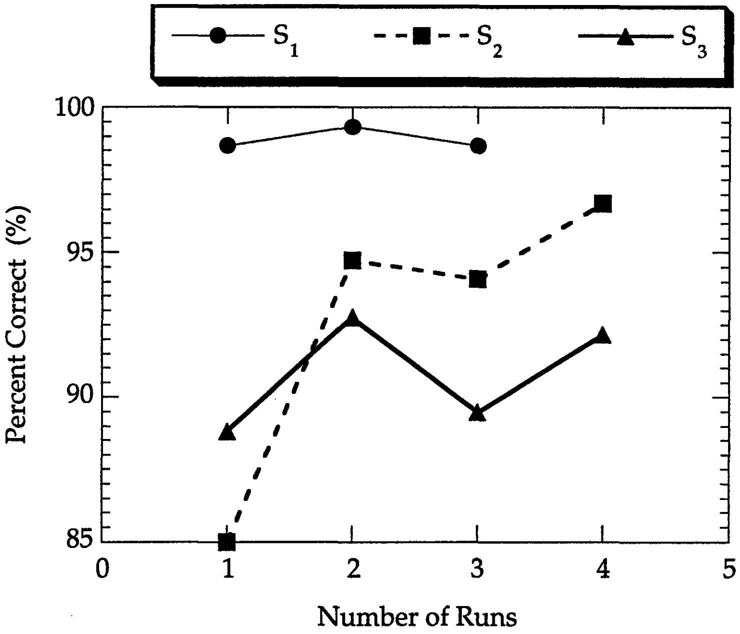

Figure III-8. Test results in terms of percent-correct scores for each subject with the 125-*msec* stimulus set.

TABLE III-7. Information Transfer with the 125-*msec* stimulus set.

Subject	*Percent-Correct Scores*	IT_{est}	*Average*	IT_{pc}	*Average* IT_{pc}
S_1	99%, 99%, 99%	6.1 *bits*	99%	6.1 *bits*	**5.6 *bits***
S_2	95%, 94%, 97%	6.0 *bits*	95%	5.6 *bits*	
S_3	93%, 89%, 92%	5.9 *bits*	91%	5.1 *bits*	

Chapter IV.
Information Transfer Rate

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

IV-1 Background

It takes a very long time to train a subject to receive continuous streams of encoded information. The first step in training is to learn to recognize the individual signals that make up the continuous presentation stream. This process, as we have shown earlier, 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 time is minimized. This prepares the subject for the next stage of organizing basic signals into meaningful "chunks" that can be stored in short-term memory and retrieved later on. 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 absolute limit on the information each chunk can contain. Therefore, the goal is to maximize the chunk sizes in *bits/chunk*. This process can take many years. Also, there is evidence that reaching a temporary plateau in performance does not necessarily imply completion of the training process. For instance, 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 reception curves were interpreted as evidence that a student of telegraphy first learned to receive individual letters, then developed the skills to receive common words as the basic units, and eventually, after many years of full-time practice on the job, learned to receive short phrases. Our subjects were not trained extensively on chunking for two reasons. First, our signals are nonsense materials and not particularly well suited for chunking. The advantage of having nonsense materials is that it is relatively easy to control stimulus uncertainty. Once meaningful materials are used for testing, the redundancy inherent in the codes needs to be assessed. Second, even if we had used meaningful materials, it is not clear that sufficient training could be given in a period commensurate with this thesis.

In view of these problems, our strategy in assessing information transfer rate was to measure identification performance in the context of other signals using an identification paradigm with both forward and backward masking.

A general understanding of the conditions under which optimal IT rate occurs is useful for selecting the values of stimulus uncertainty and presentation rate. According to Garner (1962), given sets of stimuli with a range of stimulus uncertainties, maximum IT rate occurs when stimulus uncertainty is at its maximum. Therefore, the three stimulus sets developed earlier should be used in their entirety. According to Klemmer & Muller (1953), the optimal presentation rate is two to three *items/sec* independent of the stimulus uncertainty in the items. To find the optimal presentation rate for our setup, the presentation rate was varied over a large range. Finally, since our subjects were already trained on the 500-*msec*, 250-*msec* and 125-*msec* stimulus sets, we were also able to examine the dependence of performance level on signal duration.

IV-2 Experimental Paradigm: Identification with Masking

The identification paradigm used in this portion of our research incorporates both forward and backward masking as it would occur in a continuous presentation stream (Fig. IV-1). On each trial, the subject was asked to identify the target X, sandwiched between two interfering maskers A and B. The duration of the target and maskers was kept the same (T_1). The duration of the two gaps was also kept the same (T_0). The time between signal onsets, T_{onset}, was simply (T_0+T_1) and the presentation rate, λ, was $1/T_{onset}$. Note, that, our notion of possible masking effects is meant to be general, including opportunities for peripheral and central masking effects.

Table IV-1 shows the three stimulus sets used in these experiments. Three finger locations were used: the thumb alone, the index finger alone, or the middle finger alone. We decided not to

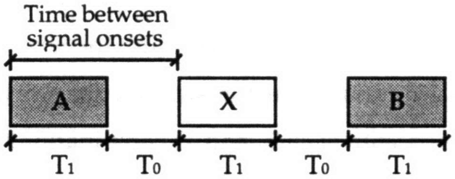

Figure IV-1. Diagram for identification paradigm with both forward and backward masking.

TABLE IV-1. The three stimulus sets.

Stimulus Set	No. of Waveforms	T_1 (*msec*)	No. of Alternatives (*k*)	IS_k (*bits*)
#1	30 (Table III-1)	500	30 × 3	6.5
#2	30 (Table III-1)	250	30 × 3	6.5
#3	19 (Table III-1)	125	19 × 3	5.9

apply waveforms to all three digits because the F_M and F_H components tended to spread in time and space.[1] T_1 was either 500 *msec*, 250 *msec*, or 125 *msec*. T_0 was 500, 400, 300, 200, 100, or 20 *msec*.[2] Each combination of T_0 and T_1 was tested. The presentation rate, λ, ranged 1-1.9 *items/sec* for T_1 = 500 *msec*, 1.3-3.7 *items/sec* for T_1 = 250 *msec*, and 1.6-6.9 *items/sec* for T_1 = 125 *msec*.

Data were collected with the same three subjects (S_1, S_2, and S_3) used in our static tests for all 18 conditions (3 T_1 × 6 T_0). Each subject was first tested with T_1 = 500 *msec* and descending values of T_0, then with T_1 = 250 *msec*, and then T_1 = 125 *msec*. This ensured that the subjects had maximum training on easier tasks before they were tested with the more difficult ones. With each stimulus set, the subject was required to repeat the one-interval AI paradigm (without the maskers A and B) conducted earlier for static IT measurements in order to get familiar with the signals again. The subject was then tested with the identification paradigm with masking. On each trial, A, X, and B were each randomly selected from the same stimulus set. No additional timing cues were available to mark the three intervals. Subjects had to wait until all three signals were presented

---

[1] For example, if masker A contained a 300 *Hz* signal component applied to all three digits, subjects tended to judge the 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 inter-stimulus duration was much longer than T_0.

[2] According to Gescheider (1966, 1967), 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 intervals.

before entering the response for X. If the response did not have the right syntax (i.e., an icon for the finger location followed by an icon for the waveform), the trial was counted as an error and was not repeated later on. Trial-by-trial correct-answer feedback was not provided.[1] Three runs of 100 trials each were performed with each T_1 and T_0 combination.[2] The percent-correct score on X was shown to the subject at the end of each run.

IV-3 Results

For each T_1 and T_0 combination, the three percent-correct scores were averaged. Results are presented in terms of T_{onset}, with T_1 as a parameter (Fig. IV-2). 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 subjects, percent-correct scores are dependent on T_{onset}, but not on T_1 alone. In other words, there seems to be a trade-off between T_1 and T_0. The data curves show a knee in the region $325 \le T_{onset} \le 450$ *msec*, corresponding to a presentation rate of roughly 2.2 to 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.

To estimate the IT rate potentially available with streams of these signals, we will assume 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 IV-1 and $\bar{e}$ is the observed average error rate. The results, shown in Fig. IV-3, indicate an optimal presentation rate of 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 rates averaged over all

---

[1] Feedback was not provided for two reasons. First, all subjects were well trained with all signals in all three stimulus sets by now. Second, requiring the subjects to attend to correct-answer feedback tended to break the "rhythm" of the run whenever an error was indicated.

[2] 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.

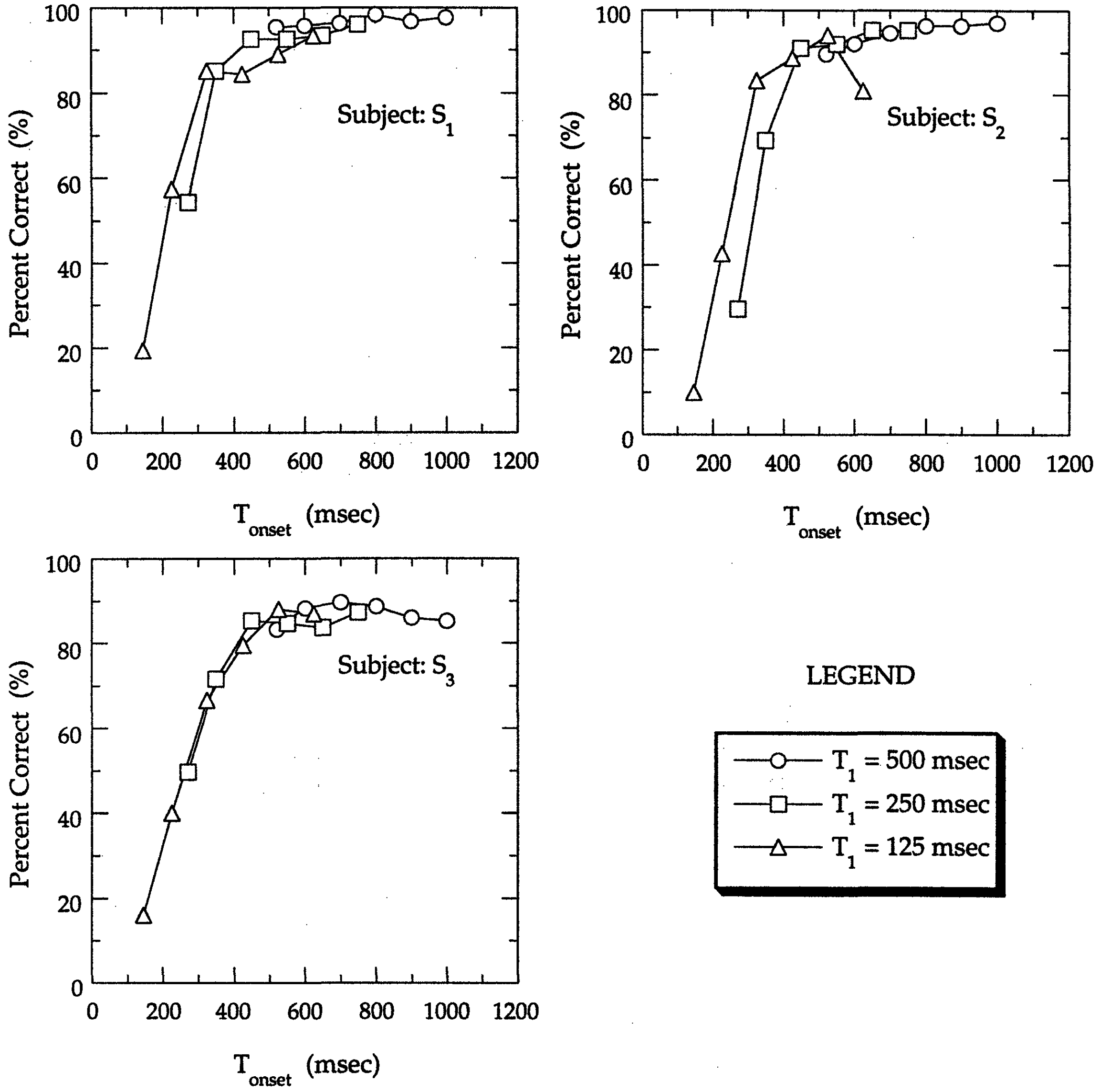

Figure IV-2. Percent-correct scores for identification of X as a function of T_{onset} for all subjects.

subjects was about 12 ***bits/sec.*** The validity of this estimate is largely dependent upon the assumption that with sufficient training, these subjects would eventually learn to "chunk" individual presentations into longer "messages" so that continuous streams of information can be received at a similar rate. Although we believe this assumption to be a reasonable one, it

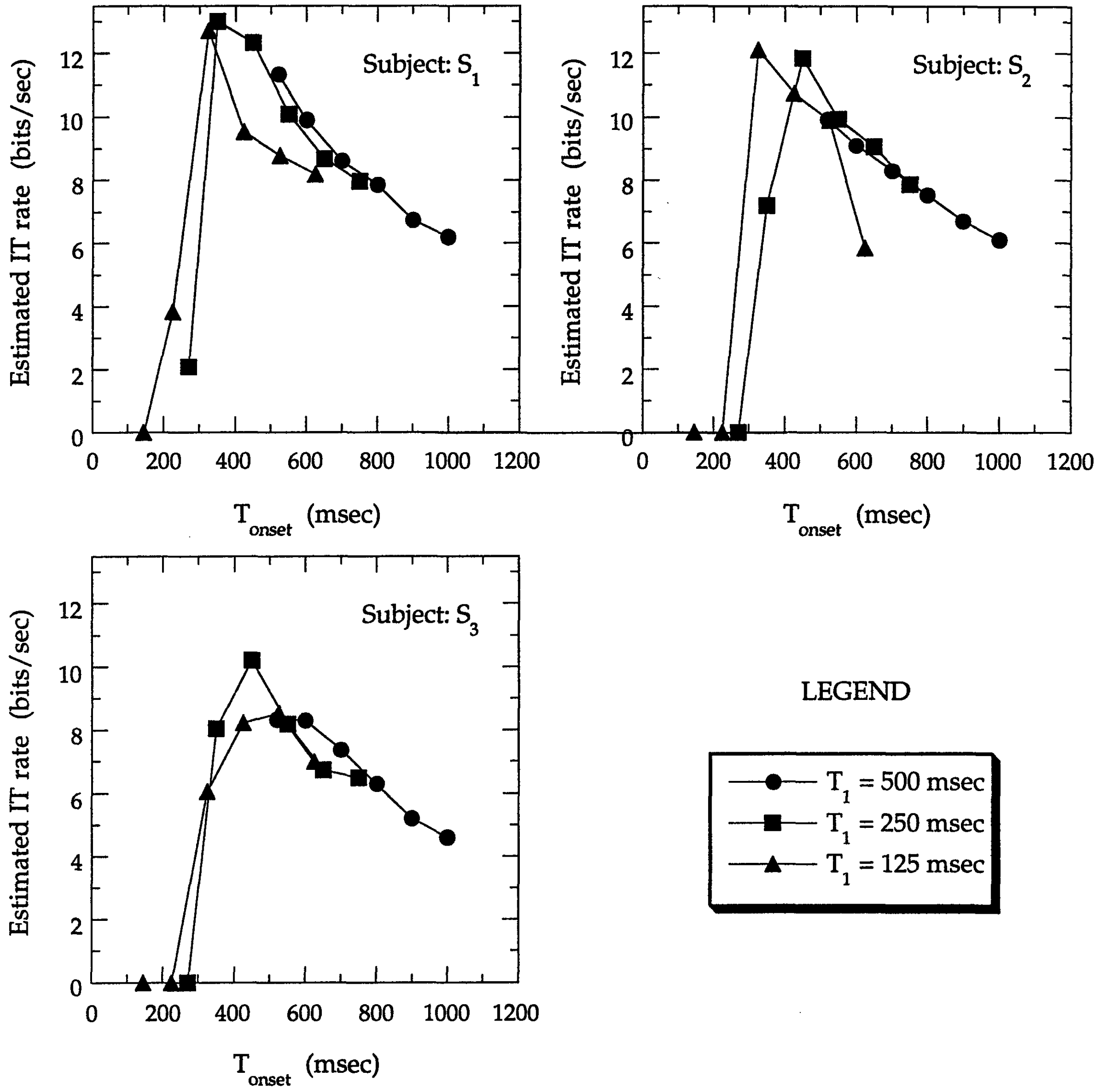

Figure IV-3. Estimated IT rates as a function of T_{onset} for all subjects.

obviously needs to be tested empirically. Because such a test would consume a very long training period, we were not able to include it in this thesis work.

Chapter V. Unresolved Issues

Three issues are discussed in this chapter: (1) selection of stimulus uncertainty for estimation of maximum information transfer, (2) definition of stimulus-set dimensionality in the context of increasing information transfer, and (3) relationship between motor output and reception of motional input sequences.

V-1 Selection of Stimulus Uncertainty

Consider first the case of a univariate stimulus set. It is well established that for such a set, IT is limited by the magic number 7±2 (Miller, 1956). Moreover, as illustrated in Fig. V-1 (left panel), it is generally accepted that IT increases monotonically with IS when IS is small and then plateaus when IS is large (see, for example, Garner, 1962, p. 75). With this picture in mind, attempts to determine the maximum IT (i.e., the plateau level) usually involve selecting an IS that is large relative to the expected value of IS at the knee of the IT *vs.* IS curve. However, two results suggest that this strategy may not be optimal.

First, we have observed in our probe experiments (see Appen. B) that IT tends to decrease slightly as IS is increased beyond the point at which the subject begins to make a significant number of identification errors (Fig. V-1, right panel). This implies that IT is not a monotonically increasing function of IS, and that to determine the maximum value of IT, it is necessary to judiciously select IS values around the expected value of maximum IT and to iteratively refine these choices. These IS values are likely to be much lower than those chosen by the usual approach of selecting IS>>IT.

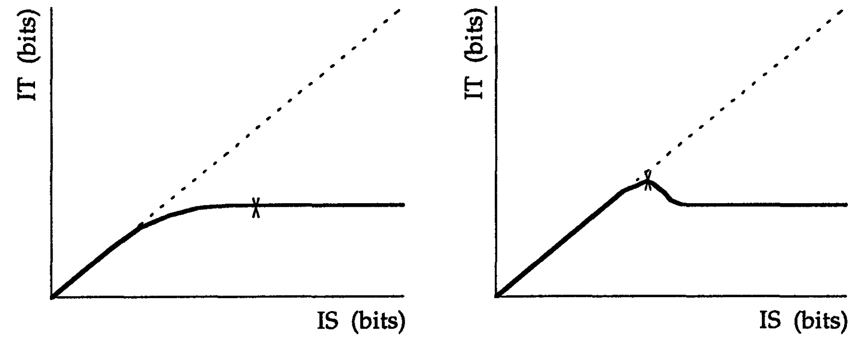

Figure V-1. Schematic illustrations of IT *vs.* IS relationships. The dashed lines indicate where IT=IS.

Second, using a smaller IS can greatly reduce the number of trials required to estimate IT, thereby increasing test efficiency. The saving in the total number of trials needed to obtain a relatively unbiased estimate of IT can come from two sources. The first source is the smaller IS. As discussed in Sec. III-1.3, Miller (1954) argued that at least $5k^2$ ($k = 2^{IS}$) trials are needed in order to effectively eliminate the bias in the estimate of IT. Based on this criterion, an IS of 5 *bits* would require 5120 trials (!), a number that is difficult to achieve in a practical experiment (Fig. V-2). Assuming, for example, that the maximum IT for a certain variable is around 3 *bits*, measuring IT with IS values of 1, 2, 3 and 4 *bits* (which would reveal the peak in IT) requires a total of 1,700 trials, which is much less than the number of trials needed if we selected an IS of 5 *bits*.

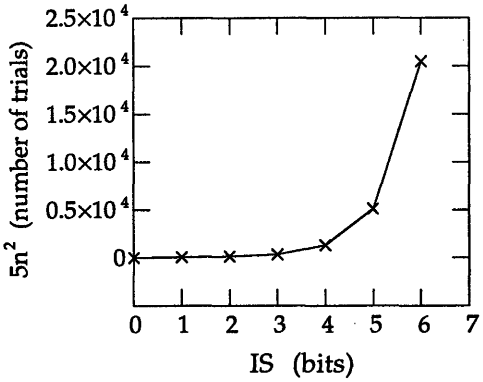

Figure V-2. Total number of trials needed for an unbiased estimate of information transfer.

Another result of using IS near, and not much greater than, IT is that overall identification accuracy is very high. This produces a further saving in the testing time that is required. According to Rabinowitz (1995, personal communication and unpublished data), estimated IT converges faster as a function of total number of trials when the overall performance level is high. In the extreme case, if performance were perfect, then the stimulus-response confusion matrix is purely diagonal and it converges after few trials.[1] When information transfer is estimated with a small IS (see the "cross" in the right panel of Fig. V-1), performance is close to being perfect (i.e., the "cross" is close to the dashed line). This means additional savings in total number of trials in order to obtain a good estimate of information transfer, a saving which can be very large. For example, our estimate of IT ≈ 6.5 *bits* with IS = 6.9 *bits* (with 500-*msec* stimuli) were obtained with 720 trials; Miller's criterion would have required at least 72,000 trials.

The existence of a maximum IT near where the IT *vs* IS curve deviates from a straight line of unit slope can be found in some earlier studies (e.g., Garner, 1953; Braida, 1969), but not in others (e.g., Pollack, 1952). We suspect that when this peak does exist, it is relatively small (i.e., probably less than 0.5 *bits* above the asymptotic value). Thus, in terms of estimating maximum IT, it does not lead to serious error if one used the asymptotic value. However, as noted above, the selection of stimulus uncertainty has a substantial impact on the number of trials needed to obtain an unbiased estimate of maximum IT. The effect of stimulus uncertainty on number of trials is even more evident when one considers the case of multivariate stimuli and larger maximum IT values.

To investigate this issue further, we need to first verify that a maximum IT exists with an intermediate value of IS for different sensory modalities under a variety of experimental conditions. We can then explore the possible sources of this local maximum.

V-2 Dimensionality of a Stimulus Set

As mentioned previously in Sec. III-1.4, it is generally well accepted that the magnitude of information transfer is related to the "richness" of a display. The word "dimension" has been used

---

[1] Conversely, when IT/IS is small, the confusion matrix has many more entries (approaching k^2), that must be accurately determined.

extensively in classical literature to describe the richness of a display. For example, Miller (1956) showed with empirical results that information transfer for unidimensional stimulus sets is limited to 2–3 *bits.* Pollack & Ficks (1954) showed that a much higher IT could be obtained with an eight-dimensional auditory display. A closer examination of the way the word "dimension" is used in these two classical papers, however, reveals that dimension, and unidimensionality, are not adequately defined. According to Pollack & Ficks (1954), "when a stimulus aspect can be manipulated independently, it is often called a dimension of an elementary auditory display" (p. 155). This notion that dimensionality is defined by the number of independently manipulated variables in the display appears to be accepted implicitly by many investigators in the field. However, as will be seen below, in order for the notion of dimensionality to be truly useful in the domain of psychophysics, it is essential that it be defined in terms of perceptual properties of the stimulus set, not the physical properties.

In the following paragraphs, we consider three factors related to the "richness" of stimulus sets and to the amount of information transfer that is likely to be obtained with these sets. All of these factors need to be considered in the search for a definition of stimulus-set dimensionality that is both rigorous and relevant to the amount of information transfer that can be achieved when human observers are required to identify stimuli in the stimulus set.

One important factor in considering the "richness" of a display concerns the perceptual segmentation of a stimulus parameter. Consider, for example, identifying the site of a single-point stimulus (e.g., a pin prick) applied to some position along a line running up the back of one's body from ankle to neck. Although the subject might confuse points within a single body segment (e.g., along the calf of the legs), it is reasonable to think that very few mistakes would occur across the boundaries of a segment (e.g., between a point on the calf and one on the thigh). In such a case, it seems likely that Miller's 7±2 constraint would be applicable only to the locations within a given segment and that the total number of distinguishable locations would be roughly 7 times the number of segments. Moreover, in this case, the boundaries appear to be defined by our ability to label the body surface verbally. Boundaries can also be defined by salient changes in the sensation that is elicited as the physical parameter is varied. Consider the case in which one is listening to a broadband click presented binaurally with an interaural time delay τ, and the subject's task is to identify τ. In the region $0 \leq \tau \leq 1$ *msec*, the subject will hear a single fused image within the head whose location will move from the center of the head to the leading ear as τ varies from 0 *msec* to 1 *msec*, and the identification response will be based on the lateralization of this image. In the region 1 *msec* $\leq \tau \leq$ 10 *msec*, the image will remain at the leading ear, but its character will change: as τ increases from 1 *msec* to 10 *msec*, the image will become increasingly rough and complex, eventually breaking up into two clicks that are temporally as well as spatially separated (i.e., a click to the leading ear followed by a click to the lagging ear). Before the click breaks up into two distinct clicks, the identification response will be based on the roughness or complexity of the image. Finally in the region $\tau \geq 10$ *msec*, where two distinct temporally separated clicks are heard, and the identification response will be based on subjective estimates of the time duration between the clicks. In this example, most errors will occur within one of the three τ segments, but not across their boundaries, because of the distinctive perceptual qualities that are associated with these ranges of τ. In general, in considering the potential IT that can be obtained from these two examples, it is the number of perceptually distinct segments, not the number of physical variables, that is important. If the number of perfectly identifiable items that can be achieved per segment is l and there are m segments, then the total number of perfectly identifiable items will be roughly $m \cdot l$. In other words, segmentation leads to an additive increase in the number of perfectly identifiable items.

A second important factor in considering the "richness" of a display is the number of physical parameters that can be varied within a single-component stimulus. Traditionally, researchers have tried to increase IT by independently varying the physical parameters of a single-component stimulus. For example, greater IT can be obtained by varying both the frequency and the amplitude of a vibration than by transmitting either variable alone. In this case, the overall IT will be determined not only by the IT per parameter and the number of parameters, but also by the perceptual interaction between the parameters. For instance, "tactual loudness" is dependent on both the frequency and the amplitude of a vibration (Verrillo *et al.*, 1969). A more extreme case occurs in binaural lateralization. Not only is lateralization influenced by interaural amplitude difference as well as interaural time delay, but it is exceedingly difficult to distinguish between the two: these two physical variables appear to give rise to one perception, i.e., lateralization. In general, however, when the perceptual interaction among parameters is small, as in the case of the multidimensional auditory displays employed by Pollack & Ficks (1954), the potential IT will be roughly equal to the product of the IT per parameter and the number of parameters. In other words, independently varying multiple parameters of a single-component stimulus can potentially lead to a multiplicative increase in the number of perfectly identifiable items.

The third important factor in considering the "richness" of a display concerns the use of multi-component stimuli. For example, one could apply two simultaneous single-point stimulations to different locations on the surface of the body. Or, as we have done in our main experiments, one can make use of both multiple stimulation sites (the different fingers) and multiple frequency components (the different stimulating waveforms).

In this thesis research, we have attempted to make use of all three means for optimizing information transfer. We identified three perceptually distinctive frequency segments (denoted F_L, F_M, and F_H, respectively) within the single physical variable of stimulation frequency. In the 500-*msec* stimulus set, single-frequency waveforms were used with multiple number of frequencies in each of the three segments. Within the F_L segment, frequency and amplitude of the waveform were varied independently. Double- and triple-frequency waveforms were used by combining single-frequency elements from different frequency segments. Finally, these waveforms were applied simultaneously to the thumb, index, or middle fingers, or to all three fingers. However, we do not yet know how to characterize our stimulus sets in terms of "dimensionality".

V-3 Relationship between Motor Output and Motional Inputs

One issue we had looked into during the study of tactual reception of Morse code was the extent to which one's ability to receive motional input sequences is dependent upon one's experience in outputting similar motions. Ideally, we want to look at a variety of daily tasks that require skilled motor output (e.g., typing), construct an apparatus that can deliver similar motions to the hand (e.g., a reverse typewriter), recruit subjects who are skilled with the output task as well as those who are inexperienced, and study how well they can learn to receive the motional inputs. The difference in performance between the experienced and inexperienced subjects (as revealed, for example, in the learning curves) may be an indication of the influence motor output skills have on reception of motional stimulation.

In the study described in Appendix A, we studied how well subjects could receive motional stimulation delivered to the fingertip as up-down displacements. The movement patterns were designed according to the way a "straight keyer" is used to send Morse code. The signal was on whenever the fingertip was down. The timing and duration of the signal was determined by the Morse code for the letter being transmitted. Experienced Morse code operators and inexperienced subjects were trained to receive single letters, random letter sequences, words, and sentences in Morse code through such up-down finger motions. We found that experienced subjects generally performed better than inexperienced subjects. The biggest difference between the two subject groups was their ability to process words and sentences. The experienced subjects were very good at "chunking", a skill they developed via auditory reception of Morse code over many years. The inexperienced subjects, however, could barely receive words at rates above 10 *wpm* and could not receive sentences at all. It seems clear that the main differences between the experienced and inexperienced subjects were related to central processing abilities. Note that most Morse code operators are highly trained in both sending (manually) and receiving (auditorily) the code. Their ability to process continuous *input* streams of Morse code with audio tones obviously contributed to their ability to do the same with up-down finger motions. Thus, in this case, it was not possible to attribute the performance difference in the two subject groups directly to the experienced subjects' ability to *output* Morse code manually.

In future studies of this issue, we will need to account for a number of factors when considering the difference between subjects who are highly trained with certain motor output tasks and those who are not. The first factor is the nature of the code involved in such tasks. For example, radiotelegraphy and stenography require knowledge of special codes with which the average person is not familiar. Typing, however, uses a code with which many people are familiar. We should avoid using tasks that require special knowledge of codes, because it will be difficult to determine if a difference in performance between experienced and naive subjects is due to the difference in their ability to output motions, their knowledge of the code, or both.

The second factor is the readiness of an experienced subject's ability to transfer skills acquired through one modality to the other. For example, Morse code operators are highly skilled at receiving the code auditorily. To the extent that they can learn to "hear" Morse code through the tactual channel, they may process tactual Morse code very differently than a naive subject would.

Finally, the issue of training must be considered. A person who is good at a motor output task has probably been trained extensively (for months and years). A laboratory subject cannot be expected to receive the same kind of training with an experimental device. Thus, we need a strategy to be able to compare performances among different subjects despite the different amount of training they have received.

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 × 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 $125 \le T_1 \le 500$ 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 $IT = IS \times (1-e)$, but did not explain why. Using our conservative estimate of $IT = IS \times (1-2e)$, 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 $IT = IS \times (1-e)$ 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.

Chapter VII. Bibliography

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review,* 93 (2), 154–179.

Bliss, J. C. (1961). Communication via the kinesthetic and tactile senses. Ph.D. Dissertation, Dept. of Electrical Engineering, M.I.T.

Bolanowski Jr., S. J., Gescheider, G. A., Verrillo, R. T., & Checkosky, C. M. (1988). Four channels mediate the mechanical aspects of touch. *The Journal of the Acoustical Society of America,* 84 (5), 1680–1694.

Braida, L. D. (1969). Intensity Perception in Audition. Ph.D. Dissertation, Massachusetts Institute of Technology.

Braida, L. D., & Durlach, N. I. (1972). Intensity perception II. Resolution in one-interval paradigms. *The Journal of the Acoustical Society of America,* 51(2), 483–502.

Brammer, A. J., Piercy, J. E., Nohara, S., Nakamura, H., & Auger, P. L. (1993). Age-related changes in mechanoreceptor-specific vibrotactile thresholds for normal hands. *The Journal of the Acoustical Society of America,* 93, p. 2361.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language. The acquisition of a hierarchy of habits. *The Psychological Review,* 6(4), 345–375.

Campbell, S. L. (1993). Uni- and multidimensional identification of rise time, spectral slope, and bandwidth. Ph.D. Dissertation, University of Washington.

Cholewiak, R. W., Sherrick, C. E., & Collins, A. A. (1993). Princeton Cutaneous Research Project (No. 62). Princeton University.

Clark, F. J., Burgess, R. C., & Chapin, J. W. (1986). Proprioception with the proximal interphalangeal joint of the index finger: Evidence for a movement sense without a static-position sense. *Brain,* 109, 1195–1208.

Clark, F. J., & Horch, K. W. (1986). Kinesthesia. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance: Sensory processes and perception (p. 13/1 – 13/62). New York: Wiley.

Cornett, R. O. (1967). Cued speech. *American Annals of the Deaf*, 112, 3–13.

Craig, J. C. (1972). Difference threshold for intensity of tactile stimuli. *Perception & Psychophysics*, 11, 150–152.

Craig, J. C. (1974). Vibrotactile difference thresholds for intensity and the effect of a masking stimulus. *Perception & Psychophysics*, 15, 123–127.

Deininger, R. L., & Fitts, P. M. (1955). Stimulus-response compatibility, information theory, and perceptual-motor performance. In H. Quastler (Ed.), Information Theory in Psychology: Problems and Methods (p. 317–341). Glencoe, Illinois: The Free Press.

Delhorne, L. A., Besing, J. M., Reed, C. M., & Durlach, N. I. (1990). Tactual cued speech as a supplement to speechreading. *American Speech and Hearing Association*, 30, 73.

Durlach, N. I., Delhorne, L. A., Wong, A., Ko, W. Y., Rabinowitz, W. M., & Hollerbach, J. (1989). Manual discrimination and identification of length by the finger-span method. *Perception & Psychophysics*, 46(1), 29–38.

Durlach, N. I., Tan, H. Z., MacMillan, N. A., Rabinowitz, W. M., & Braida, L. D. (1989). Resolution in one dimension with random variations in background dimensions. *Perception & Psychophysics*, 46(3), 293–296.

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. *Proceedings of the American Society of Mechanical Engineers*, 55(1), 345–351.

Eriksen, C. W., & Hake, H. W. (1955). Multidimensional stimulus differences and accuracy of discrimination. *Journal of Experimental Psychology*, 50 (3), 153–160.

Formby, C., Morgan, L. N., Forrest, T. G., & Raney, J. J. (1992). The role of frequency selectivity in measures of auditory and vibrotactile temporal resolution. *The Journal of the Acoustical Society of America*, 91(1), 293–305.

Foulke, E., & Brodbeck Jr., A. A. (1968). Transmission of Morse Code by Electrocutaneous Stimulation. *The Psychological Record*, 18, 617–622.

Franklin, G. F., Powell, J. D., & Workman, M. L. (1990). Digital Control of Dynamic Systems (2nd ed.). Reading, MA: Addison-Wesley Publishing Company.

Franzen, O., & Nordmark, J. (1975). Vibrotactile Frequency Discrimination. *Perception & Psychophysics*, 17, 480–484.

Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, 46, 373–380.

Garner, W. R. (1962). Uncertainty and Structure as Psychological Concepts. New York: John Wiley and Sons, Inc.

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

Geldard, F. A. (1960). Some neglected possibilities of communication. *Science*, 131 (3413), 1583–1588.

Gescheider, G. A. (1966). Resolving of successive clicks by the ears and skin. *Journal of Experimental Psychology*, 71(3), 378–381.

Gescheider, G. A. (1967). Auditory and cutaneous temporal resolution of successive brief stimuli. *Journal of Experimental Psychology*, 75(4), 570–572.

Gescheider, G. A., Bolanowski Jr., S. J., Verrillo, R. T., Arpajian, D. J., & Ryan, T. F. (1990). Vibrotactile intensity discrimination measured by three methods. *The Journal of the Acoustical Society of America*, 87 (1), 330–338.

Gescheider, G. A., O'Malley, M. J., & Verrillo, R. T. (1983). Vibrotactile forward masking: Evidence for channel independence. *The Journal of the Acoustical Society of America*, 74 (2), 474–485.

Gescheider, G. A., Sklar, B. F., Van Doren, C. L., & Verrillo, R. T. (1985). Vibrotactile forward masking: Psychophysiology evidence for a triplex theory of cutaneous mechanoreception. *The Journal of the Acoustical Society of America*, 78 (2), 534–543.

Gescheider, G. A., Verrillo, R. T., & Van Doren, C. L. (1982). Prediction of vibrotactile masking functions. *The Journal of the Acoustical Society of America*, 72 (5), 1421–1426.

Goff, G. D. (1967). Differential discrimination of frequency of cutaneous mechanical vibration. *Journal of Experimental Psychology*, 74 (2), 294–299.

Hall, L. A., & McCloskey, D. I. (1983). Detections of movements imposed on finger, elbow and shoulder joints. *Journal of Physiology*, 335, 519–533.

Henderson, D. R. (1989). Tactile Speech Reception: Development and Evaluation of an Improved Synthetic Tadoma System. M.S. Thesis, Massachusetts Institute of Technology.

Houtsma, A. J. M. (1983). Estimation of Mutual Information from Limited Experimental Data. *The Journal of the Acoustical Society of America*, 74 (5), 1626–1629.

Johnson, K. O., & Hsiao, S. S. (1992). Neural mechanisms of tactual form and texture perception. *Annual Review of Neuroscience*, 15, 227–250.

Jones, L. A., & Hunter, I. W. (1992a). Differential thresholds for limb movement measured using adaptive techniques. *Perception & Psychophysics*, 52 (5), 529–535.

Jones, L. A., & Hunter, I. W. (1992b). Human operator perception of mechanical variables and their effects on tracking performance. *Proceedings of the Winter Annual Meeting of the American Society of Mechanical Engineers*, DSC-Vol.42, 49–53.

Jorden, E. H. (1962). Beginning Japanese (Part 1). New Haven: Yale University Press.

Klemmer, E. T., & Muller, Jr. P. F. (1953). The rate of handling information — key-pressing responses to light patterns. Human Factors Oper. Res. Lab. Memo Report 34.

Knudsen, V. O. (1928). Hearing with the sense of touch. *Journal of General Psychology*, 1, 320–352.

Labs, S. M., Gescheider, G. A., Fay, R. R., & Lyons, C. H. (1978). Psychophysical tuning curves in vibrotaction. *Sensory Processes*, 2, 231–247.

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 and Development*, 25(4), 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.

Lockhead, G. R. (1966). Effects of dimensional redundancy on visual discrimination. *Journal of Experimental Psychology*, 72 (1), 95–104.

Loomis, J. M., & Lederman, S. J. (1986). Tactual Perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance: Cognitive processes and performance (p. 31/1 – 31/41). New York: Wiley.

Miller, G. A. (1954). Note on the bias of information estimates. In H. Quastler (Ed.), *Information Theory in Psychology*, 95–100.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *The Psychological Review*, 63 (2), 81–97.

Mowbray, G. H., & Gebhard, J. W. (1957). Sensitivity of the skin to changes in the rate of intermittent mechanical stimuli. *Science*, 125, 1297–1298.

Pollack, I. (1952). The information of elementary auditory displays. *The Journal of the Acoustical Society of America,* 24, 745–749.

Pollack, I., & Ficks, L. (1954). Information of elementary multidimensional auditory displays. *The Journal of the Acoustical Society of America,* 26(2), 155–158.

Proctor, R. W., & Reeve, T. G. (Eds.) (1990). Stimulus-Response Compatibility: An Integrated Perspective. *Advances in Psychology,* 65. North-Holland: Elsevier Science Publishers B.V.

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, S88.

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. *The Journal of the Acoustical Society of America,* 82 (4), 1243–1252.

Reed, C. M., Durlach, N. I., Delhorne, L. A., Rabinowitz, W. M., & Grant, K. W. (1989). Special monograph on sensory aids for hearing-impaired persons. *The Volta Review,* 91, 65–78.

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,* XXIII, 102–105.

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. *The Journal of the Acoustical Society of America,* 77(1), 247–257.

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,* 101–108.

Rothenberg, M., Verrillo, R. T., Zahorian, S. A., Brachman, M. L., & Bolanowski, S. J. J. (1977). Vibrotactile frequency for encoding a speech parameter. *The Journal of the Acoustical Society of America,* 62, 1003–1012.

Sherrick, C. E., & Cholewiak, R. W. (1986). Cutaneous Sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance: Sensory processes and perception (p. 12/1 – 12/58). New York: Wiley.

Srinivasan, M. A., Whitehouse, J. M., & LaMotte, R. H. (1990). Tactile detection of slip: surface microgeometry and peripheral neural codes. *Journal of Neurophysiology,* 63 (6), 1323–1332.

Tan, H. Z. (1988). Analysis of synthetic Tadoma system as a multidimensional tactile display. M.S. Thesis, Department of Electrical Engineering and Computer Science, MIT.

Tan, H. Z., Rabinowitz, W. M., & Durlach, N. I. (1989). Analysis of a synthetic Tadoma system as a multidimensional tactile display. *The Journal of the Acoustical Society of America,* 86(3), 981–988.

Tan, H. Z., Srinivasan, M. A., Eberman, B. S., & Cheng, B. (1994). Human factors for the design of force-reflecting haptic interfaces. *Proceedings of the Winter Annual Meeting of the American Society of Mechanical Engineers,* Chicago, 1994.

Van Doren, C. L., Gescheider, G. A., & Verrillo, R. T. (1990). Vibrotactile temporal gap detection as a function of age. *The Journal of the Acoustical Society of America,* 87(5), 2201–2206.

Verrillo, R. T. (1963). Effect of contactor area on the vibrotactile threshold. *Proceedings of the Winter Annual Meeting of the American Society of Mechanical Engineers,* 35, 1962–1966.

Verrillo, R. T., Fraioli, A. J., & Smith, R. L. (1969). Sensation magnitude of vibrotactile stimuli. *Perception & Psychophysics,* 6, 366–372.

Verrillo, R. T., & Gescheider, G. A. (1992). Perception via the sense of touch. In I. R. Summers (Ed.), Tactile Aids for the Hearing Impaired (p. 1–36). London: Whurr Publishers.

Appendix A. Tactual and Auditory Morse Code Reception[1]

---

[1] This manuscript has been submitted to *Perception&Psychophysics.*

Reception of Morse Code Through Motional, Vibrotactile, and Auditory Stimulation

Hong Z. Tan, Nathaniel I. Durlach, William M. Rabinowitz, Charlotte M. Reed,
& Jonathan R. Santos[1]

ABSTRACT

The potential for communication through the kinesthetic aspect of the tactual sense was examined in a series of 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 letter, 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. Results are compared to those obtained in other research with tactual communication devices.

INTRODUCTION

In this paper, we focus on the ability to receive information through motional stimulation (i.e., the kinesthetic sense). Our long term goals are: (1) to study the kinesthetic sense as a communication channel, (2) to compare performance through the kinesthetic sense with that through other senses, and (3) to compare the ability to receive motional stimulation with the ability to produce the same movement patterns.

Most studies of tactual communication have focused on the cutaneous / tactile sensory system (see Geldard, 1973; Kaczmarek, Webster, Bach-y-Rita, & Tompkins, 1991). In contrast, research on the kinesthetic sensory system is extremely limited (see Clark & Horch, 1986, for a review). Bliss (1961) investigated the use of the kinesthetic sense as a communication channel in experiments employing an air-driven finger stimulator that was constructed as a reverse typewriter. The stimulator consisted of eight finger rests arranged in two groups on which the user could place the fingers of both hands in a manner similar to typing on the home row. Each stimulator was capable of simulating motions corresponding to the active movements of a typist's fingers in reaching the upper and lower rows on a keyboard. In one set of experiments, 42 random triplets composed of the letters *e, t, n, a, o* and *i* were presented to eight subjects. The average information transfer was

---

[1] This research was supported by research grant number 2 R01 DC 00126-16 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We are greatly indebted to Phil Temples, K9HI and Joe Parskey, NJ1P for their expertise and dedication in working as the experienced subjects.

1.75 *bits/letter* out of a maximum possible 2.58 *bits/letter.* In another experiment, 30 symbols (the alphabet, comma, period, space, and upper case) were presented in random order with equal probability to one subject (with less than 15 hours of practice). Six sequences of 130 symbols each were delivered at a rate of 0.5 to 1.5 *letters/sec.* The subject responded verbally by naming the symbols as they were received. The information rate, computed as the multiplication of percent correct, presentation rate (*letter/sec*) and information per symbol (4.91 *bits/letter*), reached a maximum of 4.5 *bits/sec* at a presentation rate of 1.32 *letters/sec.*

In a more recent study (Eberhardt et al., 1994), a two-degree-of-freedom (up-down and front-back) finger stimulator named OMAR was developed to provide slow motion as well as vibration to a finger through a single actuator. Early experiments demonstrated that some subjects were able to judge onset asynchronies of vibration and movement with such a system.

In the present study, we investigated the ability to receive information through up-down finger motions. In order to assess communication rate, a code was needed to convert the up-down finger motions into meaningful messages. The International Morse code was chosen because it is a well-established code and its learning patterns have been well studied. Bryan & Harter (1899) followed students of telegraphy for over half a year and tested their ability to send and receive Morse code weekly. They found that while the students' ability to send the code improved monotonically, their ability to receive the code reached several plateaus and eventually exceeded that of sending. The plateaus in the reception curves were interpreted as evidence that a student of telegraphy first learned to receive individual letters, then developed the skills to receive common words as the basic units, and eventually learned to receive short phrases after many years of practice. By employing highly skilled Morse code operators as subjects in the current study, it was possible to take advantage of their previous experience in chunking coded messages. Inexperienced subjects were also trained and tested for comparison. The fact that Morse code is used to both send and receive information enables us to investigate the relationship between the ability to receive motional stimulation and the ability to produce such motions. Finally, Morse code can be adapted to many sensory modalities. Although Morse code is traditionally received through the auditory channel, hearing-impaired ham operators have put their hands on speakers to receive Morse code through the tactual channel. We compared subjects' ability to receive the Morse code through motional, vibrotactile, and auditory stimulation using common tasks.

METHODS

Morse Code is a temporal sequence of patterns in which each letter of the alphabet has its own unique pattern. Patterns consist of elements (dot = one unit = U; dash = three units = $3U$) and pauses. Morse-code reception was studied for motional, vibrotactile, and auditory stimulation as a function of presentation rate R in words per minute (*wpm*), which is related to the duration of U (in *msec*) by $R = 1200/U$. A more complete description of Morse Code is provided in the Appendix attached at the end.

Subjects

Two experienced Morse Code operators from the Boston Amateur Radio Club (subjects E1 and E2) and two inexperienced MIT students (subjects N1 and N2) participated in the experiments. E1 and E2 were both males, aged 38 and 40, and were licensed as extra-class ham radio operators. N1 was a 28-year old female and N2 an 18-year old male. Three of the subjects (E2, N1, and N2) were right-handed and one (E1) was left-handed. Except for N1, who was also the experimenter, all subjects were paid on an hourly basis.

Tasks

The reception of Morse Code through motional, vibrotactile, and auditory stimulation was studied using four tasks in the following order: single-letter identification, three-letter random-sequence identification, common-word identification, and sentence reception. Table A-1 lists the testing conditions in chronological order. All four subjects participated in each experiment except that (1) only the experienced subjects were tested with sentences (because the inexperienced subjects were unable to perform this task), and (2) the experienced subjects were not trained auditorily with the single-letter and three-letter sequences (because they were already experienced with the reception of Morse Code through auditory stimulation).

TABLE A-1. EXPERIMENTAL CONDITIONS AND ORDERING (M: MOTIONAL, V: VIBROTACTILE, A: AUDITORY)

MODE	*TASK*	*SUBJECTS*
M	1-letter	E_1, E_2, N_1, N_2
M	3-letter	E_1, E_2, N_1, N_2
M	words	E_1, E_2, N_1, N_2
M	sentences	E_1, E_2
V	1-letter	E_1, E_2
V&A interleaved	1-letter	N_1, N_2
V	3-letter	E_1, E_2
V&A interleaved	3-letter	N_1, N_2
V	words	E_1, E_2
V&A interleaved	words	N_1, N_2
V	sentences	E_1, E_2
A	sentences	E_1, E_2
A	words	E_1, E_2

Single-letter identification. On each trial, the subject was presented (through motional, vibrotactile, or auditory stimulation, as described below) with the Morse code for one of the 26 letters of the alphabet. The subject was instructed to respond with one of the 26 letters on a computer keyboard, and then trial-by-trial correct-answer feedback was provided by displaying the correct response on a computer screen. Each run consisted of 130 presentations of single letters in random order with each of the 26 letters presented exactly 5 times. The duration of each run varied from 5 to 20 min depending on the response time of the subject. Each subject started from the lowest rate of stimulus presentation and was allowed to proceed to the next higher rate only after the completion of (1) one run with a perfect score of 100%, (2) at least three runs with scores over 95% (not necessarily consecutively), or (3) roughly ten or more consecutive runs with similar scores (i.e., a clear plateau). Four rates were tested: 12, 16, 20, and 24 *wpm*, except for motional stimulation where the rate of 20 *wpm* was not used.

Three-letter identification. On each trial, the subject was presented (through motional, vibrotactile, or auditory stimulation) with the Morse code of a three-letter nonsense word (with each letter chosen randomly with equal *a priori* probabilities from the 26 letters), instructed to respond with a three-letter sequence, and then shown the correct response. The letters were separated by a pause of duration *3U*. The subject could either "copy on the fly" (i.e., the response to the first letter was entered while the second letter was being presented) or "copy behind" (i.e., the subject waited until all three letters were presented before entering the response). Each run consisted of 52 presentations of three-letter sequences in random order, such that each letter of the alphabet was presented exactly 6 times. A response was considered correct only if all three letters were identified correctly in the correct order. Each subject started from the lowest rate of stimulus presentation and was allowed to proceed to the next higher rate only after the completion of (1) one run with a perfect score of 100%, (2) three runs with scores over 90% (not necessarily consecutively), or (3) roughly ten or more consecutive runs with similar scores (i.e., a clear plateau). Four rates were tested: 12, 16, 20 and 24 *wpm*. All subjects chose the copy-behind method of responding.

Common-word identification. The material consisted of 600 words obtained from the corpus of The American Heritage Word Frequency Book (Carroll, Davies, & Richman, 1971). The selection of words was based on rate of occurrence and minimum length. All the stimuli occupy ranks between 1000 and 5300 per million and contain at least 7 letters. Two randomizations of the 600 words into twelve 50-item word lists were constructed and employed in the testing such that all lists from the first randomization were presented prior to lists from the second randomization. The subjects were told before the experiment that the test material consisted of common English words. On each trial, the subject was presented (through motional, vibrotactile, or auditory stimulation) with the Morse code of one word from a chosen list, instructed to respond by typing out a response word (either by "copying on the fly" or "copying behind"), and then shown whether the response was "right" or "wrong".[2] The letters within a word were separated by a pause of duration 3U. Each run consisted of one list (i.e., 50 words). Different rates were selected for experienced and inexperienced subjects with each of the three types of stimulation in order to obtain a wide range of percent-correct scores as a function of stimulus presentation rate. Each subject performed at least three runs per stimulus presentation rate (unless the performance was 0% or above 90%, in which case only one run was conducted), and proceeded from the lowest to the highest rate.

Sentence reception. The test material consisted of CUNY sentence lists commonly used for speech and hearing research (Boothroyd, Hanin, & Hnath, 1985). Each of the 60 lists contains 12 sentences

---

[2] Because each word was presented again later, the subjects were not shown the correct word when a mistake was made.

arranged by topic (e.g., food, animals, weather, etc.). Each sentence in a list consists of 3 to 14 common English words and each list contains exactly 102 words. The same list was never used twice with the same subject. The difficulty levels of these sentences were estimated to be equivalent to fifth to sixth grade reading levels. Prior to the experiment, the subjects were told that the test material consisted of conversational sentences but were not informed of the topics. On each trial, the subject was presented (through motional, vibrotactile, or auditory stimulation) with the Morse code of one sentence from a chosen list, instructed to repeat the sentence verbally, and only given informal feedback (e.g., the experimenter revealed specific words in the sentence if the subject asked). Letters within a word were separated by a pause of duration 3U, and words within a sentence were separated by a pause of duration 7U. The subject could either respond "on the fly" or after the entire sentence had been presented. Each run consisted of one list (i.e., 12 sentences). At the end of a run, the experimenter counted the number of words the subject was able to repeat regardless of the ordering and ignored extra words in the response. The overall word score was computed as the number of correctly-repeated words divided by 102, the total number of words in each CUNY-sentence list. Different rates were selected for the three types of stimulation in order to obtain a wide range of percent-correct scores as a function of stimulus presentation rate. Each subject was tested with at least three lists at each rate and proceeded from the lowest to the highest rate.

During all experiments, the subject was informed of the overall percent-correct score at the end of each run. Each experimental session lasted 1 to 2 hours. Subjects were free to take breaks between runs at their own pace. The experienced subjects generally completed two sessions per week. The inexperienced subjects completed three or more sessions per week.

Instrumentation and Procedure

Motional stimulation. A device designed to move the fingertip up and down was constructed around a permanent magnet servo motor with feedback from a tachometer and an optical encoder (Fig. A-1). A Plexiglas lever was attached to the motor shaft. The subject rested the index fingertip lightly over a roller which was snug-fit into a hole on the lever. The distance from the center of the motor shaft to that of the roller was 40 *mm*. The roller served to control the point of contact and to accommodate any relative motions between the finger and the lever. The system parameters were adjusted so that the position-step response was critically damped, with a rise-time of approximately 20 *msec*.

The waveforms used to drive the motor were two-level square waves. Fig. A-2 shows the waveform for the letter "P". Each waveform started with an inter-letter pause of *3U* followed by the appropriate dot-dash pattern for that letter. For the typical arrangement of the stimulator system, a downward motion at the fingertip indicated the onset of a dot or a dash. The actual vertical displacement of the fingertip was adjusted to be ≈10 *mm*. This was found to be the largest amplitude that felt comfortable at the highest rate tested (i.e., 24 *wpm*) through preliminary experimentation. With the finger pressing lightly on the roller, the overall position of the roller (and lever) shifted downwards by 1-2 *mm*, but the relative up-down motion was otherwise unchanged.

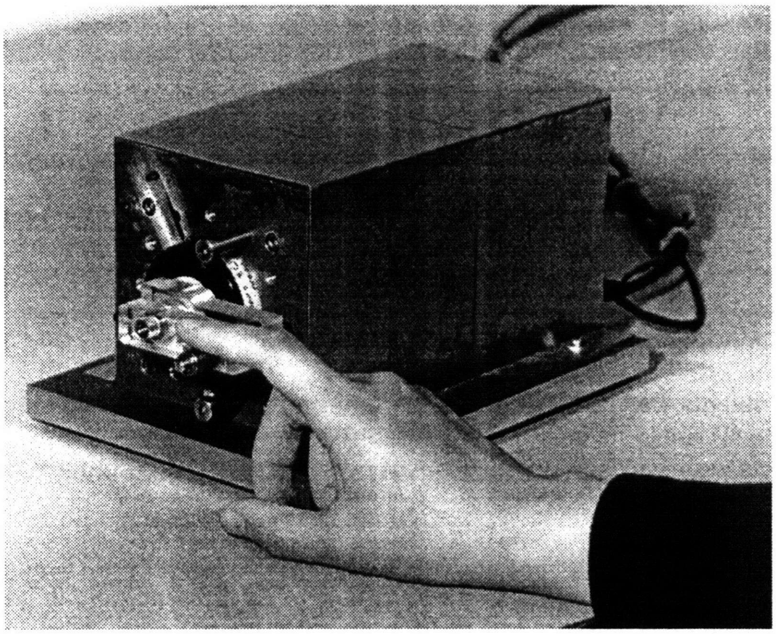

Figure A-1. The experimental apparatus. The finger is rested on a roller placed 40 *mm* from the center of the rotor. The two shoulder screws above and below the Plexiglas bars serve as the mechanical stops.

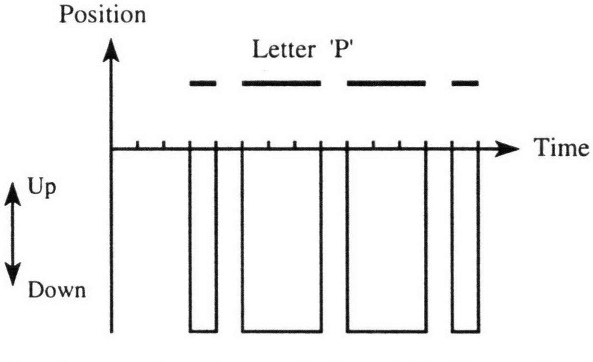

Figure A-2. Waveform used to deliver the letter "P" for motional stimulation.

The apparatus was always hidden from view. Subjects wore earphones with acoustic noise to mask any auditory cues from the apparatus. Stimuli were presented to the index finger of the dominant hand of each subject. The standard posture was to rest the fingertip lightly on top of the roller and follow the up-down motions of the roller. In general, subjects were encouraged to use a consistent posture throughout all experiments, although alternative postures were employed by some subjects under some conditions.[3] Presentation rates ranged from 4 - 24 *wpm* across the various tasks corresponding to a range in *U* from 300 - 50 *msec.* Before the experiments began, the inexperienced subjects were provided with a brief training period (averaging 3.6 *hours*) to associate letters with the movement patterns.

Vibrotactile stimulation. Stimulation was applied through an electrodynamic minishaker (Alpha-M AV-6). A 200-*Hz* sinusoidal signal gated by the square wave shown in Fig. A-2 was applied to the minishaker. The presence and the duration of the vibration indicated the presence and the duration of a dot or a dash. The subject placed the index finger of the dominant hand on the top of a flat contactor (9 *mm* in diameter) that was fit to the minishaker. The 200-*Hz* vibration was presented at a nominal level of ≈50 dB SL. Presentation rates ranged from 8-40 *wpm* across the various tasks, corresponding to a range in *U* from 150 to 30 *msec.*

During the experiments, the minishaker was placed inside a wooden box lined with sound-absorbing foam to (1) shield it visually from the subject and (2) attenuate the sound caused by the vibration. Subjects wore earphones with acoustic noise to mask any residual auditory cues from the minishaker.

Auditory stimulation. Morse-code sequences were presented diotically via headphones using the same 200-*Hz* signals that were applied to the minishaker. The presence and the duration of an auditory tone indicated the presence and the duration of a dot or a dash. For stimulus presentation rates above 56 *wpm*, a 5-*ms* Hanning window was applied to the rising and falling portion of the signals to reduce "clicks". The subject could adjust the overall gain so that the earphone signal "felt comfortably loud". Presentation rates ranged from 12 - 73.85 *wpm* across the various tasks, corresponding to a range in *U* of 100 - 16.25 *msec.*

Data Analysis and Reduction

For each subject, task, type of stimulation, and presentation rate, a learning curve was constructed in which the percent-correct score was plotted as a function of run number. Based on the learning curve, decisions on when to terminate were made on the basis of the criteria described for each task in the methodology section. The learning-curve data were reduced by averaging percent-correct scores (a) across the final three runs at each presentation rate, and then (b) across experienced subjects E1 and E2 and across inexperienced subjects N1 and N2.

---

3. Subjects were discouraged, but not prohibited, from experimenting with non-standard settings. They were asked to document all deviations from the standard setup in a log book and to discuss them with the experimenter at the end of the session. For the single-letter identification experiment, E_1 used the downward motion at the first rate of 12 *wpm*, but switched to the upward motion after starting the 16 *wpm* condition. His performance in terms of percent correct scores was measured to be about 30% higher for upward motions than for downward motions. He was thus permitted to use a set of waveforms with a polarity opposite to that shown in Fig. A-2 for all subsequent experiments. In addition, E_1 switched to a smaller range of motion (i.e., fingertip displacement was decreased to 5 *mm*) to reduce fatigue in the three-letter random-sequence identification experiment. Subject E_2 used the standard posture but preferred a larger range of motion after beginning the 24 wpm condition. After demonstrating an improvement in performance, he was permitted to decrease signal attenuation by 4 dB (i.e., fingertip displacement was increased to 15 *mm*) for all subsequent experiments.

RESULTS

Single-Letter Identification

Typical learning curves for motional stimulation are shown in Fig. A-3. The upward arrow in E2's graph indicates the time at which E2 increased the range of motion from 10 to 15 *mm* (see note 2). Apparently, the increase in movement amplitude had little effect on the overall characteristics of the learning curve at 24 *wpm*. With this simplest task and with motional stimulation, experienced and inexperienced subjects exhibit similar learning curves. As expected, both types of subjects started with lower percent-correct scores and took longer to reach performance criterion as presentation rate increased.

The percent-correct scores averaged over the last three runs for motional (M), vibrotactile (V), and auditory (A) stimulation are shown individually and in summary form in Fig. A-4. Whereas the performance of the two experienced subjects E1 and E2 was quite similar for all tests conducted, the performance of N1 was sometimes much better than that of N2. Nevertheless, averaging the data for the two inexperienced subjects does not affect our general conclusions. Therefore, in the remainder of this paper, only the summary graphs will be presented. From the summary graph in Fig. A-4, it is observed that the experienced subjects achieved the performance criterion of 95% correct at all rates tested with the motional and vibrotactile stimulation.[4] The inexperienced subjects were not able to achieve the performance criterion at rates above 16 *wpm* with motional or vibrotactile stimulation. Their performance with auditory stimulation, however, was nearly perfect at all rates tested. In general, it is clear that (1) the experienced subjects performed better than the inexperienced subjects, and (2) performance of the inexperienced subjects with auditory stimulation was better than that with motional or vibrotactile stimulation.

Three-Letter Random-Sequence Identification

The percent-correct scores averaged over the last three runs for motional (M), vibrotactile (V), and auditory (A) stimulation are shown in Fig. A-5. The experienced subjects achieved the performance criterion of 90% correct only at the lower rates of 12 and 16 *wpm* for motional and vibrotactile stimulation.[3] The inexperienced subjects were not able to achieve the performance criterion at any rate with motional stimulation and only at the slowest rate of 12 *wpm* with vibrotactile stimulation. However, their performance with auditory stimulation reached performance criterion at all rates tested. Thus, it is clear that (1) this task is more difficult than the single-letter identification task for both subject groups, (2) the experienced subjects performed better than inexperienced subjects, (3) performance of the inexperienced subjects with auditory stimulation was better than that with vibrotactile stimulation, and (4) performance with vibrotactile stimulation was better than that with motional stimulation.

---

4. Had the experienced subjects performed this task with auditory stimulation, they would have achieved nearly perfect scores at all rates tested.

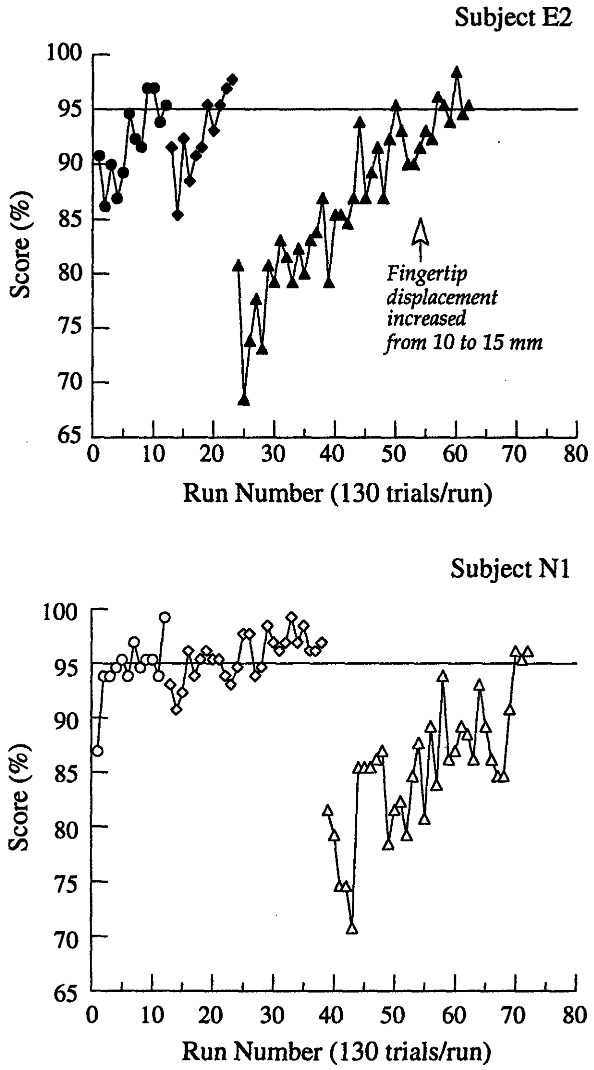

Figure A-3. Learning curves for motional stimulation from the single-letter identification test for one experienced subject (E2, filled symbols, above) and one inexperienced subject (N1, open symbols, below) at 12 *wpm* (circles), 16 *wpm* (diamonds), and 24 *wpm* (triangles). Horizontal lines indicate the performance criterion of 95%.

Common-Word Identification

This is the only task where both subject groups were tested with all three modes of stimulation. The percent-correct scores averaged over the last three runs for motional (M), vibrotactile (V), and auditory (A) stimulation are shown in Fig. A-6. Percent-correct word scores decreased with stimulus presentation rate at average rates of 5%/*wpm* (M), 5%/*wpm* (V), and 3%/*wpm* (A) for experi-

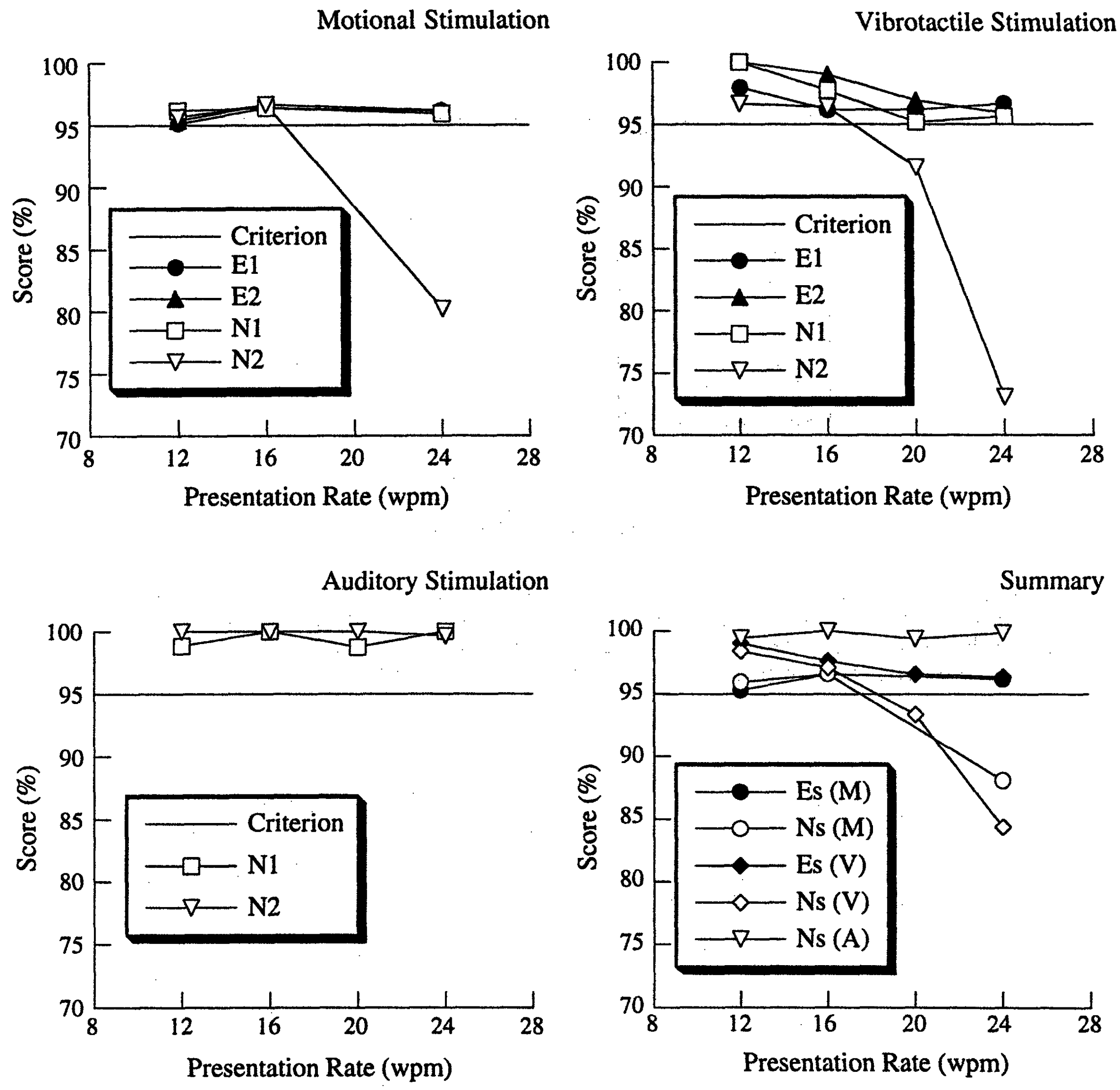

Figure A-4. Percent-correct scores averaged over final three runs from the single-letter identification test as a function of presentation rate. Individual subject results are shown in separate panels for motional, vibrotactile, and auditory stimulation. Results from each type of stimulation are summarized in the final panel by averaging across scores from the experienced subjects (Es) and from the inexperienced subjects (Ns). Horizontal lines indicate the performance criterion of 95%.

enced subjects, and 6%/*wpm* (M), 7%/*wpm* (V), and 5%/*wpm* (A) for inexperienced subjects. The presentation rates corresponding to 50% correct scores were 22 *wpm* (M), 31 *wpm* (V), and 51 *wpm* (A) for experienced subjects, and 11 *wpm* (M), 16 *wpm* (V), and 25 *wpm* (A) for inexperienced subjects.

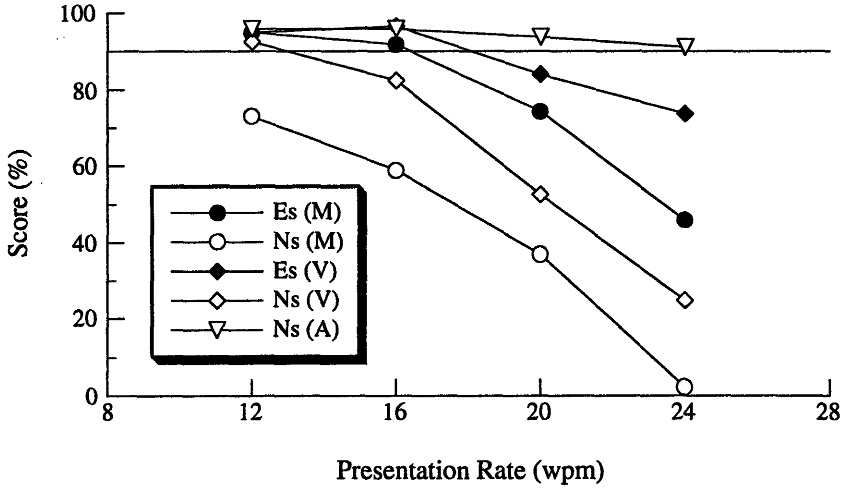

Figure A-5. Percent-correct scores averaged over final three runs from the three-letter random-sequence identification test as a function of presentation rate. Data are shown for experienced (filled symbols) and inexperienced (open symbols) subjects with motional (circles), vibrotactile (diamonds) and auditory (triangles) stimulation. Horizontal line indicates the performance criterion of 90%.

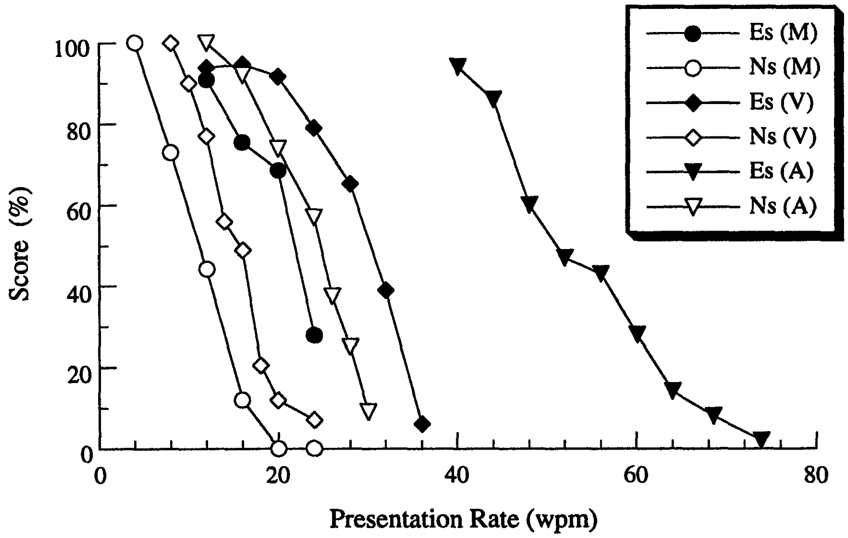

Figure A-6. Percent-correct scores averaged over final three runs from the common-word identification test as a function of presentation rate. Data are shown for experienced (filled symbols) and inexperienced (open symbols) subjects with motional (circles), vibrotactile (diamonds) and auditory (triangles) stimulation.

The results indicate that (1) in general, subjects' performance with auditory stimulation was *much* better than that with vibrotactile stimulation, which, in turn, was better than that with motional stimulation, and (2) experienced subjects performed better than inexperienced subjects with all three types of stimulation.

As another metric of performance, the equivalent word rate γ was calculated as the product of percent-correct score and stimulus presentation rate. (Cholewiak, Sherrick, & Collins, 1993, refer to this measure as the *correct words per minute.*) A maximum γ was associated with each test (see Fig. A-7). As stimulus presentation rate increased, γ increased initially, but was limited by the highest achievable γ (i.e., the presentation rate). After γ reaches the maximum, there is a trade-off between presentation rate and percent-correct scores in that γ remained at the maximum level with increasing presentation rate. After that, γ decreased as presentation rate increased. The maximum γ scores averaged across experienced subjects were 14, 19, and 38 *wpm* with motional, vibrotactile and auditory stimulation, respectively. The corresponding scores averaged across inexperienced subjects were 6, 9, and 15 *wpm*, respectively.

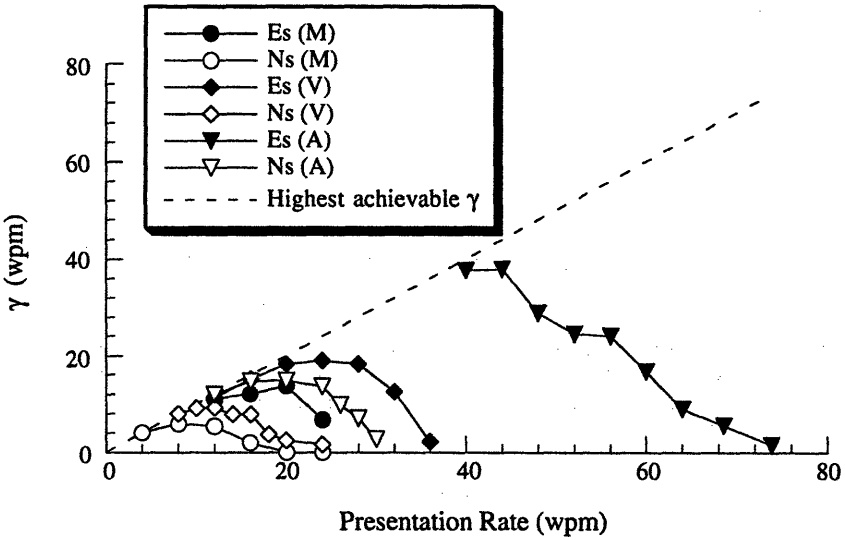

Figure A-7. Equivalent word rates γ (*wpm*) from the common-word identification test for experienced (filled symbols) and inexperienced (open symbols) subjects with motional (circles), vibrotactile (diamonds), and auditory (triangles) stimulation. Dashed line indicates the highest achievable equivalent word rate (i.e., the presentation rate).

Sentence Reception

The inexperienced subjects were unable to perform this test with any of the stimulation types; hence, only the experienced subjects were tested. The percent-correct scores averaged over the last three runs for motional (M), vibrotactile (V) and auditory (A) stimulation are shown in Fig. A-8. As stimulus presentation rate increased, performance decreased. Performance with auditory stimulation was *much* better than that with vibrotactile stimulation, which, in turn, was better than that with motional stimulation. Percent-correct word scores decreased with stimulus presentation rate at average rates of 4%/*wpm* (M), 3%/*wpm* (V), and 2%/*wpm* (A). The presentation rates corresponding to 50% correct scores were 25 *wpm* (M, extrapolated), 32 *wpm* (V), and 59 *wpm* (A). The average maximum γ scores were 18, 21, and 43 *wpm* with motional, vibrotactile, and auditory stimulation, respectively. The slightly higher γ achieved with this test compared with that achieved with the common-word identification test is probably due to the increased redundancy in the test material.

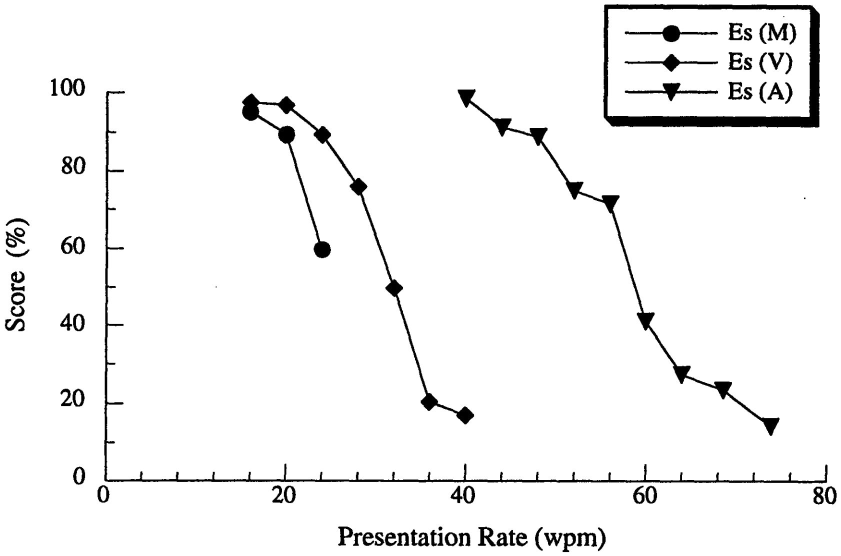

Figure A-8. Percent-correct scores averaged over final three runs from the sentence reception test as a function of presentation rate. Data averaged across the experienced subjects are shown for motional (circles), vibrotactile (diamonds), and auditory (triangles) stimulation.

DISCUSSION

We have tested experienced and inexperienced Morse code operators on their ability to receive Morse code through motional, vibrotactile, and auditory stimulation using single-letter, three-letter, common-word, and conversational-English test materials. In order to compare subjects' performance across modalities and tasks, the equivalent word rates (γ) were computed for all cases. These results are shown in Fig. A-9. The asterisks on top of the columns for the single-letter identification tests indicate that these γ values might have been higher if stimulus presentation rates over 24 *wpm* had been used. On the average, excluding data from the single-letter identification tests, the ratio of the equivalent word rates for vibrotactile stimulation to that for motional stimulation (γ_v:γ_m) was 1.2 for the experienced subjects and 1.5 for the inexperienced subjects. The ratio of the equivalent word rates for auditory stimulation to that for motional stimulation (γ_a:γ_m) was 2.6 for the experienced subjects and 2.5 for the inexperienced subjects. The ratio of the equivalent word rates for auditory stimulation to that for vibrotactile stimulation (γ_a:γ_v) was 2.2 for the experienced subjects and 1.7 for the inexperienced subjects. Overall, auditory reception of Morse code is about twice as fast as tactual (i.e., motional or vibrotactile) reception of the code for both subject groups.

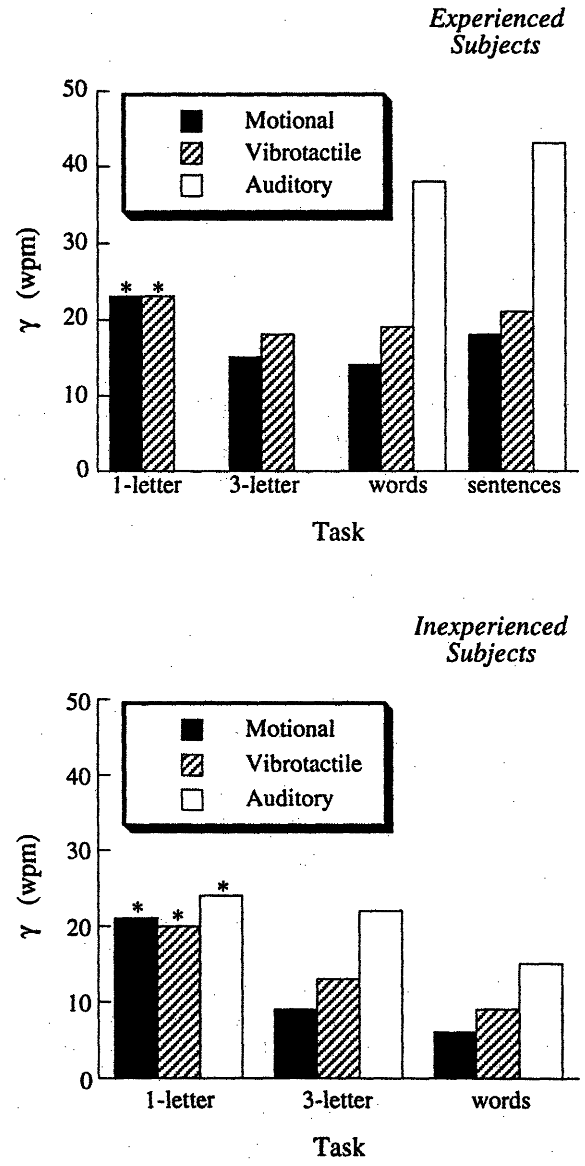

Figure A-9. Equivalent word rates γ (*wpm*) for each of the tasks and modes of stimulation. Upper panel presents results averaged across the two experienced subjects, and lower panel presents results averaged across the two inexperienced subjects.

The difference in the auditory and tactual rate of Morse code reception may be explained in terms of the unit signal length and the temporal properties of taction and audition. In general, the auditory system responds faster and more accurately to dynamic stimulation than the tactual system. For instance, Gescheider (1966) reported that the time difference necessary for resolving 2 successive events was 1.8 *msec* for equally loud binaural clicks and 10 *msec* for pulses applied to the fingertip. Our results can be compared quantitatively with those obtained by Lechelt (1957) on auditory and tactile numerousness perception using binaural clicks and 2-*msec* square-wave mechanical taps to the left middle finger with trains of 2 to 9 signals presented at rates of 3 to 8 *items/sec*. He found that whereas auditory counts were nearly perfect for all conditions tested, cutaneous counts tended to underestimate the actual number of signals. Cutaneous counts were about 90% of the actual number of signals at a presentation rate of 8 *items/sec*. Simplifying the Morse code as a series of dots (e.g., the code for "H" is *dit-dit-dit-dit*), a constant rate of 8 *items/sec* corresponds to a *U* value of 63 *msec*, or equivalently, 19 *wpm*. Despite the difference in signal duty cycles between Lechelt's study and ours, this is consistent with the equivalent word rate of 18 and 21 *wpm* with motional and vibrotactile stimulation, respectively, achieved by the experienced subjects.

The difference in performance between the two subject groups is evident in that whenever both subject groups performed the same tasks, experienced subjects attained higher values of γ than the inexperienced subjects. The inexperienced subjects were simply unable to perform some of the tasks, despite the fact that each subject received a total of 70-80 hours of training. The fact that the experienced subjects had more than 20 years of experience with the Morse code gave them several advantages over the inexperienced subjects. First, the experienced subjects were able to process finger motions at letter and word levels. The subjects reported that they could "hear" the code while feeling the motions on their fingers. This transfer of learning from the tactual sense to the auditory sense, a modality these subjects were highly trained on, allowed them to have more time to concentrate on the content of the message rather than focusing on the identification of single letters. Differences in the response strategy of the two subject groups for the common-word test material illustrate this point. The strategy of inexperienced subjects was to type out the responses letter by letter and then edit the string of letters into meaningful words. The experienced subjects, however, would either type out a whole word or skip a trial if they failed at word recognition. These subjects occasionally made spelling errors indicating again that they were focusing on words rather than letters. Second, the experienced subjects were well trained with "chunking" of letters into meaningful words or messages. They reported that during the reception of a word, they were constantly predicting the next letter based on letters already presented. This ability to hold letters in short term memory until they are incorporated into a meaningful unit is the result of years of practice. Finally, both of the experienced subjects used the straight key to send Morse code element by element before the more efficient iambic keyer became available. Their ability to send Morse code manually might have contributed to their ability to receive the code tactually.

To follow up the last point, a supplementary test was performed to determine the speed at which the experienced subjects could send Morse Code. They were tested with the straight key since its element-by-element mode corresponds directly to the mode used in our reception tests. The resulting speed for manually sending the Morse code of CUNY sentences was 23 *wpm* for each experienced subject.[5] This is consistent with the equivalent word rates obtained from sentence-reception tests with motional and vibrotactile stimulation (18 and 21 *wpm*, respectively; see the top panel of Fig. 9) for these experienced subjects.

The information transfer rates of several tactual communication methods can be compared. For natural methods of tactual communication, Reed, Durlach, & Delhorne (1992) estimated information transfer rates to range from 7.5 bits/sec for fingerspelling to 12-14 bits/sec for Tadoma and tactual sign language. Based on results obtained in the present study, the information transfer rate for receiving Morse code using conversational English material through motional and vibrotactile stimulation is roughly 2.7 bits/sec.[6] 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 roughly 1.3 *bits/sec* (according to note 5). These relatively low rates of tactual reception of Morse code are most likely limited not only by subjects' reception rate, but by major inefficiencies in the code; i.e., the bit-wise coding of information, the 3:1 dash-dot ratio, and the wasteful silences between dots and dashes. With the standard timing pattern for Morse code, the average duration across the 26 letters is roughly 8*U*. At a presentation rate of 20 *wpm* (i.e., $U = 60$ *msec*), the average duration for a letter is 480 *msec*.

Using his pneumatic reverse typewriter, Bliss (1961) reported that one experienced typist was able to receive letters and a few punctuation symbols at a rate of 4.5 *bits/sec* with a stimulus presentation rate of 1.32 symbols/sec and a stimulus uncertainty of 4.9 bits/presentation. Using the Optacon device (Linvill & Bliss, 1966) and English sentences as test material, Cholewiak et al. (1993) reported that their best subject was able to reach a word rate of 40 *wpm*, or 5.4 *bits/sec* (according to note 5). Using the display for the Vibratese language, Geldard (1957) reported that one subject was able to handle 38 wpm, or 5.1 *bits/sec* (according to note 5). These information-transfer rates are higher than those obtained here for Morse code. In making such a comparison, however, it should be noted that whereas our apparatus conveys Morse code through a 1-*bit* display, Bliss's device encodes letters and punctuation with each finger movement, the Optacon employs 108 stimulating pins (6×18, according to Fig. 47-1 in Cholewiak et al., 1993), and the Vibratese was coded using five vibrators with letters, numerals, and some short words as the basic elements.

---

5. In these tests, the subjects used a straight key oscillator (MFJ-557 from Tucker Electronics & Computers), the output of which was connected to a cassette recorder. Each subject was asked to send manually the Morse code of five CUNY sentence lists. They were instructed to (1) send as fast as they could assuming an excellent receiver, (2) not correct for any mistakes, and (3) take breaks only between sentences. The recording was then timed and scored by another ham radio operator. The sending speed for each sentence was computed as the number of words in the sentence divided by total time. The results were then averaged and multiplied by the overall percent-correct scores.
6. The information transfer rate was estimated as follows. The CUNY sentences contain 102 words per 12 sentences, thus averaging 8.5 word/sentence. According to Shannon (1951, Fig. 4), strings of that length have between 1.2 and 2.1 bits/letter. Using 2 bits/letter as the upper bound and 4 letters/word (from CUNY sentence statistics) as the average word length in the corpus, we estimated the information content to be 2 bits/letter × 4 letter/word, or 8 bits/word. Assuming that the experienced subjects can receive Morse codes of CUNY sentences reliably at 20 wpm (see top panel of Fig. A-9) through motional and vibrotactile stimulation, we conclude that the information transfer rate is 8 bits/word × 20 word/min, or equivalently, 2.7 bits/sec.

We are currently investigating the feasibility of communication through combined tactile and kinesthetic stimulation on multiple fingers using a novel multi-finger positional display. It is expected that by improving the encoding scheme as well as the display, we can achieve information rates comparable to those demonstrated by natural methods of tactual communication.

REFERENCES

ARRL (1993). *The ARRL Handbook for Radio Amateurs (7th Ed.).* The American Radio Relay League, Newington, CT 06111.

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

Boothroyd, A., Hanin, L., & Hnath, T. (1985). A sentence test of speech perception: Reliability, set equivalence, and short term learning. *Speech and Hearing Science Report,* **RC110** (City University New York).

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage Word Frequency Book.* New York: American Heritage Publishing Co., Inc.

Cholewiak, R. W., Sherrick, C. E., & Collins, A. A. (1993). *Princeton Cutaneous Research Project,* **62.** Princeton University.

Clark, F. J., & Horch, K. W. (1986). Kinesthesia. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Sensory processes and perception* (Vol. 1, pp. 13/1 - 13/62). New York: Wiley.

Eberhardt, S. P., Coulter, D. C., Bernstein, L. E., Barac-Cikoja, D., & Jordan, J. (1994). Inducing dynamic haptic perception by the hand: system description and some results. In *Proceedings of Winter Annual Meeting of the American Society of Mechanical Engineers: Dynamic Systems and Control,* **55**, 345-351.

Foulke, E., & Brodbeck Jr., A. A. (1968). Transmission of Morse Code by Electrocutaneous Stimulation. *The Psychological Record,* **18**, 617-622.

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

Geldard, F. (Ed.). (1973). *Conference on Cutaneous Communication Systems and Devices.* The Psychonomic Society, Inc.

Kaczmarek, K. A., Webster, J. G., Bach-y-Rita, P., & Tompkins, W. J. (1991). Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering,* **38**, 1-15.

Lechelt, E. C. (1957). Some stimulus parameters of tactile numerousness perception. In F. Geldard (Ed.), *Cutaneous Communication Systems and Devices,* 1-5. The Psychonomic Society, Inc.

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.

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.

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

APPENDIX

The International Morse Code

The International Morse Code is the original modulation method used in Amateur Radio. The two basic elements of Morse code are dot (sounded *dit*) and dash (sounded *dah*). It is usually received auditorily with fixed-frequency tones (usually between 500 *Hz* and 1500 *Hz*) indicating the presence and timing of *dits* and *dahs.* Unique combinations of *dits* and *dahs* specify the letters of the alphabet, numerals, punctuation marks, and procedure signals. For this study, we used letters only. A complete list of Morse Code for letters appears in Fig. A-10 with short and long bars indicating *dits* and *dahs,* respectively.

The length of a *dit, U,* is the basic unit of time in Morse Code. The duration of a *dah* is *3U.* Within a letter, the pause between adjacent elements is *U.* The space between letters is *3U.* The space between words or groups is *7U.* These relationships are illustrated in Fig. A-11.

The rate of Morse Code is expressed in terms of words per minute (*wpm*). The length of a "standard" word is defined as *50U.* The word "PARIS" is of this length and is used to accurately set transmission speed. The relationship between the length of a *dit, U,* and the rate of transmission, *R,* is:

$$U(second) = 60/[R(wpm) \times 50] \qquad \text{(Eq. 4)}$$

or, equivalently,

$$U(millisecond) = 1200/[R(wpm)]\,. \qquad \text{(Eq. 5)}$$

For instance, at 12 *wpm,* the duration of a *dit* is 100 msec and that of a *dah* is 300 *msec.*

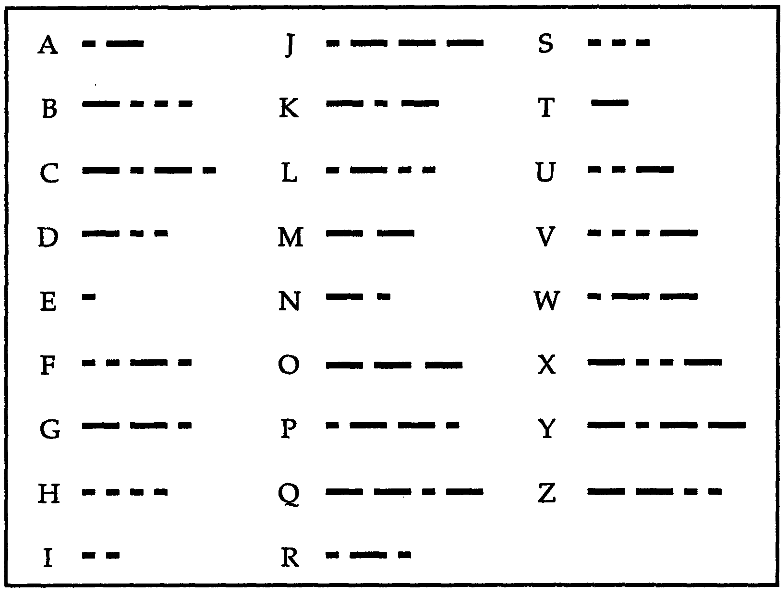

Figure A-10. Morse Code for Letters of the Alphabet

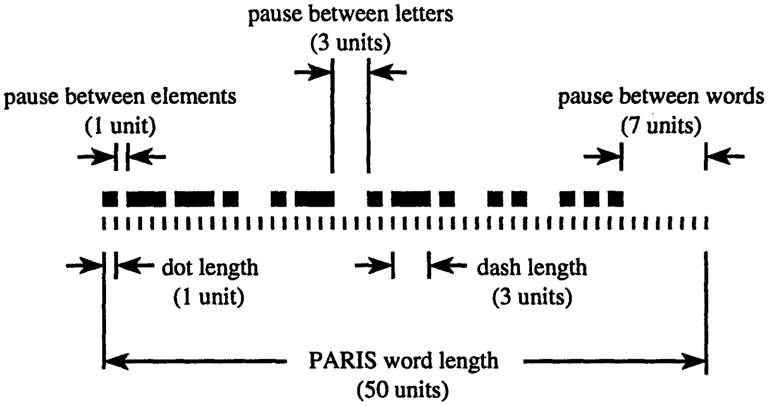

Figure A-11. Diagram of timing in International morse Code. (Adopted from *The ARRL Handbook for Radio Amateurs*.)

Appendix B.

Ideas and Experiments Used for the Construction of the 500-msec Stimulus Set

This appendix summarizes the ideas and preliminary experiments that led to the selection of the stimulus attributes and their values in the 500-msec stimulus set. Although these experiments, all of which were conducted on the index finger, were carefully designed and carried out with a substantial number of trials, they in no sense constituted an exhaustive exploration. The resulting 500-*msec* stimulus set was based partly on the results of these experiments and partly on the intuitions gained from the experience of running these experiments.

B-1 Stimulus Attributes

Because any waveform can be represented by a sum of sinusoidal components in a Fourier analysis and because the TACTUATOR was very well suited to generate sinusoids, sinusoidal waveforms were used as the basic signal elements. The relevant variables for stimulus generation were thus frequency and amplitude of the waveform, and stimulation site. The range of frequency was from DC to 300 *Hz*. The range of amplitude was from absolute detection threshold to about 50 dB SL across the whole frequency range. Stimulation site could be any combination of the thumb, the index finger, and the middle finger of the subject's left hand.

The TACTUATOR is capable of stimulating the kinesthetic and vibrotactile aspects of the tactual sense, as well as the sensation associated with the intermediate frequencies and amplitudes. To the extent that the associated sensations are distinctive and discernible when waveforms of different frequencies/amplitudes are combined, these frequency/amplitude ranges can be regarded as separate stimulus attributes. The following observations were made. First, it is clear that a 2 *Hz* slow motion is qualitatively different from a 300 *Hz* smooth vibration. Also, many researchers have commented that stimulation in the frequency range of 30 to 50 *Hz* feels "rough", "fluttering" and "unsteady", indicating a possible middle frequency region that gives rise to a qualitatively distinctive perception. Second, when a movement defined by the sum of sinusoids at 2 *Hz*, 30 *Hz* and 300 *Hz* of the same duration is presented to a finger, a slow motion due to the 2 *Hz* component, a rough motion due to the 30 *Hz* component, and a vibration due to the 300 *Hz* component, can all be perceived. Thus, components at different frequencies can be perceptually distinctive when they are properly combined. This led to the conclusion that the frequency range of DC to 300 *Hz* should be divided into low-, middle- and high-frequency regions, thereby serving as three stimulus attributes. Amplitudes were considered as attributes attached to the three frequency attributes.

In order to define the three frequency ranges and to examine the perceptual qualities of single-frequency motions across the frequency range of DC to 300 *Hz*, the following experiment was conducted. Single-frequency motions at 24 dB SL were presented in random order to one subject (S_1) at 22 frequencies ranging from 2 *Hz* to 258 *Hz* (Fig. B-1), with each frequency being presented exactly 10 times. The subject was instructed to characterize each motion as belonging to one of the following five perceptual categories: 1 - slow motion (i.e., kinesthetic sense), 3 - rough motion/ fluttering, 5 - vibration (i.e., vibrotactile sense), 2 - between 1 and 3, and 4 - between 3 and 5. No feedback was provided. A motion was defined to belong to categories 1, 3 or 5 if it was judged so at least 9 times out of the 10 presentations (the corresponding frequencies are indicated by filled diamonds in Fig. B-1). Otherwise a motion was categorized as belonging to the in-between category of 2 or 4 (open diamonds in Fig. B-1). There was a clear F_M range from 10 to 70 *Hz*. The limits for F_L and F_H, however, appeared to be relative. For example, if the 2 *Hz* motions were removed from the stimulus set, 3 *Hz* motions were judged as slow ones. If the amplitude of the 104 *Hz* motions was reduced, such motions tended to be judged as smooth vibrations. In general, if there was an adequate space between the highest F_L and the lowest F_M values, then the motions in each range were found to be perceptually distinctive. The same can be said about the F_M and F_H ranges.

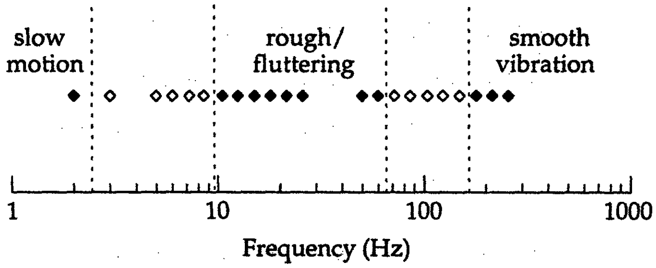

Figure B-1. Results of frequency range categorization.

For the variable of stimulus site, the thumb, the index finger, and the middle finger were regarded as three stimulus attributes. Therefore, there were a total of nine (9) stimulus attributes: F_L (low frequency), A_L (amplitude for F_L), F_M (middle frequency), A_M (amplitude for F_M), F_H (high frequency), A_H (amplitude for F_H), L_1 (thumb), L_2 (index finger), and L_3 (middle finger). The following experiments were aimed at determining the number of values along each stimulus attribute.

B-2 Selection of Amplitude Values

These experiments examined the IT for signal amplitude under fixed- and roving-frequency conditions using the AI paradigm. In the first experiment, amplitude identification was studied with roving frequency (Fig. B-2). There were a total of 22 frequency values ranging from 2 to 258 *Hz*. Each frequency had 10 associated amplitude levels ranging from 16 to 52 dB SL, resulting in a total of 220 single-frequency signals. The duration of each stimulus was fixed at 500 *msec* with a 10 *msec* rise-fall time. Each dot in Fig. B-2 indicates one signal at a specified frequency and amplitude. All signals with the same sensation level shared the same response code. On each trial, one of the 220 signals was chosen randomly. The subject's task was to respond to the level of the signal independent of its frequency. One subject (S_1) was tested. Trial-by-trial correct-answer feedback was provided. Each experimental run consisted of 220 trials (i.e., each of the 220 signals was presented exactly once). Three runs were conducted. The results formed a 10 by 10 stimulus-response confusion matrix (i.e., there were 10 alternatives in the stimulus set as far as the target attribute was concerned). Based on a total of 660 trials, IT_{est} was found to be 1.4 *bits*, or 2–3 items.

In the second experiment, amplitude identification was studied with frequency fixed at 2 *Hz* (i.e., the 10 "dots" in the first column in Fig. B-2 were used as stimuli). Two subjects (S_1 and S_3) were tested, each completing a total of 1000 trials. In this experiment, IT_{est} was 1.7 *bits* for S_1 and 1.8 *bits* for S_3, corresponding to 3 items. Thus amplitude identification was slightly better when the frequency of the stimulus was held fixed.

In the third experiment, the second experiment was repeated with four signals at 2 *Hz* corresponding to the four amplitude values at 16, 28, 40 and 52 dB SL (i.e., the range of amplitude levels in the second and third experiments was the same). Both subjects achieved 100% correct with 200 trials, i.e., IT_{est} was 2.0 *bits*. A comparison of the results of the second and third experiments indicates that subjects achieved a *higher* IT with a *smaller* number of alternatives in

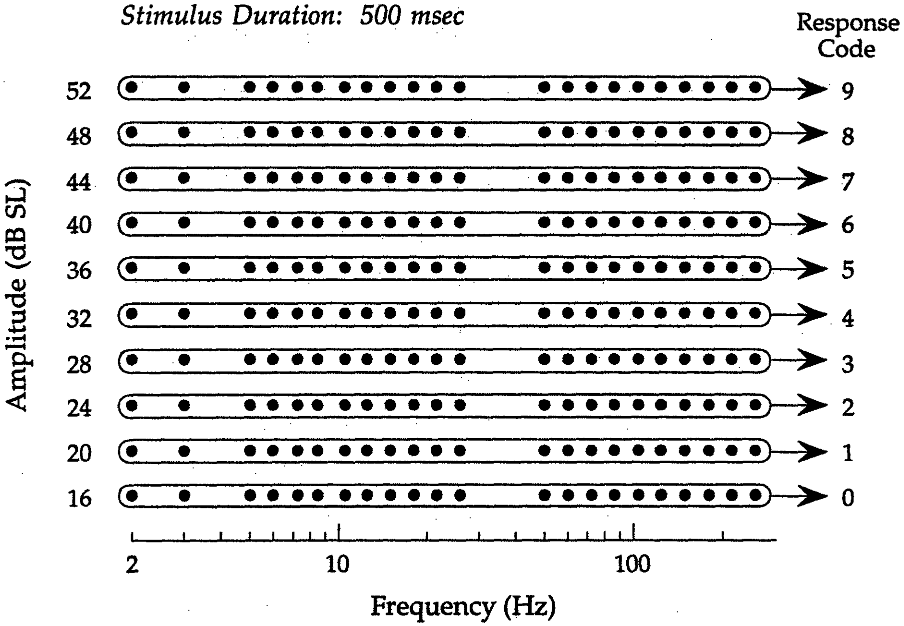

Figure B-2. Diagram for amplitude identification experiment with roving frequencies.

the stimulus set. This finding was inconsistent with the generally accepted view that the relationship between IT and IS is monotonic and reaches a plateau for large IS values.

In order to examine whether the same non-monotonic relationship between IS and IT existed under roving background conditions, the first experiment was repeated with four sensation levels of 16, 28, 40 and 52 dB SL (i.e., only the "dots" in the four rows corresponding to these sensation levels were used as stimuli). One subject (S_1) was tested. A total of 176 trials was collected to form a 4 by 4 stimulus-response confusion matrix. IT_{est} was 1.7 *bits*, or 3 items. This was slightly higher than the IT_{est} obtained from the first experiment.

From these four experiments, it was concluded that the number of amplitude values that can be identified per frequency was probably between 2 to 3 items.

B-3 Selection of F_L Values

In these experiments, the target attribute was the low-frequency component (filled circles in Fig. B-3). The background attributes were middle frequency (open triangles in Fig. B-3), high frequency (open square in Fig. B-3), and the amplitudes associated with all three frequency ranges. The values of F_L used were 2, 3, 4, and 5 *Hz*; the values of F_M used were 11, 15, 22, and 30 *Hz*; the values of F_H used were 110, 150, 220, and 300 *Hz*; and the amplitude values were 17, 26, 35, and 44 dB SL. The highest signal level was kept at 44 dB SL so that the 50 dB system dynamic range was not exceeded when multi-frequency signals were used. One subject (S_1) was tested. In the most general case, low frequency identification was studied with all background attributes roving. Each stimulus was a triple-frequency signal of the form $A_L\sin(2\pi F_L t) + A_M\sin(2\pi F_M t) + A_H\sin(2\pi F_H t)$, where (F_L, A_L), (F_M, A_M) and (F_H, A_H) were randomly selected from the 16 low-, middle- and high-frequency signals shown in Fig. B-3, respectively. The duration of each stimulus was fixed at 500 *msec* with a 10 *msec* rise-fall time. The subject's task was to respond to the frequency of the low-frequency component according to the response code specified in Fig. B-3. Trial-by-trial correct-answer feedback was provided. A total of 128 trials was collected and used to form a 4 by 4 stimulus-response confusion matrix. IT_{est} was 0.3 *bits*, or 1 item, indicating that the subject was not able to distinguish among the four low frequencies when the other elements of the stimulus were roved as described.

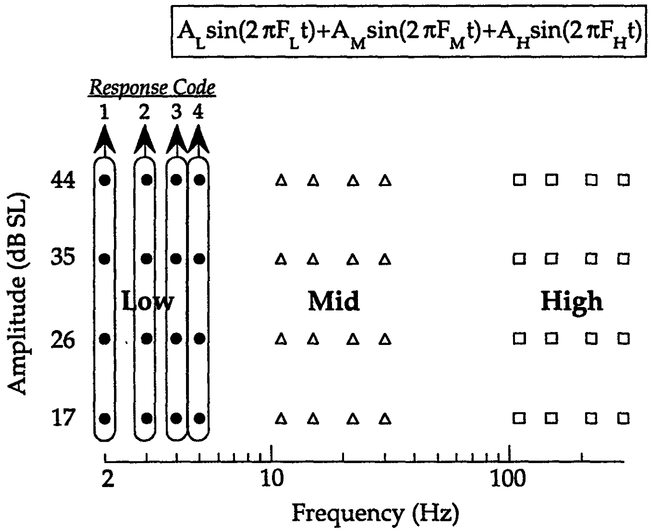

Figure B-3. Diagram for low frequency identification experiments with roving backgrounds.

Since the subject was clearly able to do better than 0.3 *bits* at identifying the low frequency F_L when the stimulus was of the form $A_L\sin(2\pi F_L t)$ and A_L was fixed, attention was turned towards finding the combinations of target and background attributes that made low frequency identification possible. With this goal in mind, the above experiment was repeated with double-frequency signals to examine the effect middle- or high-frequency components had on the perception of F_L. In the first of such two experiments, each stimulus was of the form $A_L\sin(2\pi F_L t) + A_M\sin(2\pi F_M t)$, where (F_L, A_L) and (F_M, A_M) were randomly selected from the 16 low- and middle-frequency signals shown in Fig. B-3, respectively. The estimated IT from a total of 240 trials was 0.6 *bits*, still relatively poor. Examination of the confusion matrix revealed that 87% of the errors were made with the combination $A_L = 17$ dB SL and $A_M = 44$ dB SL. By excluding trials involving this combination, we were able to increase IT to 1.2 *bits*. In other words, it appeared that the middle-frequency component tends to make low-frequency identification impossible unless A_M is kept at a moderate level. In the second experiment, each stimulus was of the form $A_L\sin(2\pi F_L t) + A_H\sin(2\pi F_H t)$, where (F_L, A_L) and (F_H, A_H) were randomly selected from the 16 low- and high-frequency signals shown in Fig. B-3, respectively. In this case, the estimated IT from a total of 498 trials was 1.2 *bits*. Furthermore, excluding trials with $A_L = 17$ dB SL did not improve IT.

Finally, the original low-frequency identification experiment with triple-frequency stimuli was repeated with A_L kept at 44 dB SL. The middle and high frequency components were still randomly selected from their respective corresponding 16 alternative shown in Fig. B-3. In this case, the estimated IT from a total of 100 trials was 1.2 *bits*.

These experiments indicated that essentially two categories of low-frequency stimulation were identifiable in the presence of roving amplitudes and randomly-selected middle- and high-frequency components. However, the amplitude of the low-frequency stimulation must be kept high relative to the level of the middle-frequency components. Interference from high-frequency components appeared to be small.

B-4 Selection of F_M and F_H Values

The motions generated by (F_L, A_L) were slow enough that the number of cycles could be easily counted. The motions generated by (F_M, A_M), however, were too fast to be counted. Therefore, judgment of F_M was based mainly upon the perceptual qualities of such motions. For example, a 30 *Hz* motion was so rough that it seemed to be beating on the finger. A 10 *Hz* motion felt unsteady and gave a wobbling sensation when superimposed on a slow motion. Such perceptions were not invariant, however, when A_M was randomized. Therefore, A_M was fixed at 35 dB SL in the middle frequency identification experiments. The subject (S_1) could not identify more than one F_M value when the stimuli were of the form $A_M\sin(2\pi F_M t) + A_L\sin(2\pi F_L t)$, and (F_L, A_L) was randomly picked from the corresponding 16 alternatives shown in Fig. B-3. The subject could reliably identify 2 values of F_M when the stimuli were of the form $A_M\sin(2\pi F_M t) + A_H\sin(2\pi F_H t)$, and (F_H, A_H) was randomly picked from the corresponding 16 alternatives shown in Fig. B-3.

Identification of F_H was also found to be impossible (i.e., IT = 0.33 *bits*) when A_H was randomized using single-frequency stimuli of the form (F_H, A_H). When A_H was fixed at 35 dB SL, IT was 0.98 *bits*. The remaining experiments were performed with A_H fixed at 44 dB SL. The reason for setting A_H at a higher level than A_M was that (F_M, A_M) tended to mask the high-frequency component when it was too strong, but (F_H, A_H) did not mask the middle-frequency component as much. For the one subject tested (S_1), high frequency identification was about 1 *bit* when stimuli were of the form $A_H\sin(2\pi F_H t) + A_L\sin(2\pi F_L t)$, and (F_L, A_L) was randomly selected from the 16 alternatives shown in Fig. B-3. However, it was impossible to identify more than one high-frequency component when the stimuli were of the form $A_H\sin(2\pi F_H t) + A_M\sin(2\pi F_M t)$, and (F_M, A_M) was randomly selected from the 16 alternatives shown in Fig. B-3.

B-5 Selection of L Values

According to the diagram in Fig. III-1, different waveforms could be used to stimulate different fingers. In order to simplify the response code, we decided to use only one waveform to stimulate any finger(s) in one stimulus. From informal experimentation, it was found that subjects rarely made errors in identifying the finger location when single finger stimulation was used. When more than one finger was used, because some signals (e.g., the 30-*Hz* and the F_H components) tended to "spread" in time and space, it was sometimes difficult to determine which fingers were being stimulated. When all fingers were stimulated with the same waveform, however, it was very easy to differentiate it from single-finger stimulation. It was thus decided that either a single finger would be stimulated, or all three fingers would be stimulated with the same waveform.

B-6 Construction of the 500-msec Stimulus Set

The above described probe experiments led us to the following conclusions. Two levels of signal amplitude could be reliably identified when the frequency of the signal was randomized. Two F_L values could be reliably identified when middle- and high-frequency components were superimposed on the low-frequency component *and* A_M was modest. Two F_M values could be reliably identified when a high-frequency component was superimposed on it and A_M was held fixed. Low-frequency components interfered with the identification of F_M. Two F_H values could be reliably identified when the high-frequency component was superimposed on a low-frequency component and A_H was held fixed. Middle-frequency components interfered with the identification of F_H.

Based on these results, 30 waveforms were constructed. Each waveform was a broadband signal of the form $W=(F_L,A_L)+(F_M,A_M)+(F_H,A_H)$. They were combined with the 4 stimulation sites to form the 120 alternatives in the 500-*msec* stimulus set, in the form of $S=(L_1,W_1)+(L_2,W_2)+(L_3,W_3)$. These signals were selected with the intent of all being perfectly identifiable. Given our finding that selecting a large IS value may not be the best way to estimate the maximum IT achievable, our strategy was to start with a stimulus set with as many easily identifiable alternatives as possible. If subjects failed to perform well within a short period of training, then the stimulus set would be pruned until performance level was high again.