

ANALYSIS OF SYNTHETIC TADOMA SYSTEM AS
A MULTIDIMENSIONAL TACTILE DISPLAY

by

HONG ZHANG TAN

B.S., Shanghai Jiao Tong University
(1986)

Submitted to the Department of
Electrical Engineering and Computer Science
in Partial Fulfillment of the
Requirements for the
Degree of
MASTER OF SCIENCE IN ELECTRICAL ENGINEERING
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 1988

©Massachusetts Institute of Technology, 1988, All rights reserved.

Signature of Author: _____
Department of Electrical Engineering and Computer Science, May 1988

Certified by: _____
Nathaniel I. Durlach, Thesis Supervisor

William M. Rabinowitz, Thesis Supervisor

Accepted by: _____
Arthur C. Smith, Chairman
Departmental Committee on Graduate Students



ANALYSIS OF SYNTHETIC TADOMA SYSTEM AS A MULTIDIMENSIONAL TACTILE DISPLAY

by
HONG ZHANG TAN

Submitted to the Department of
Electrical Engineering and Computer Science
on May 6, 1988 in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

ABSTRACT

The Tadoma method is a means of speechreading based on tactile monitoring of the articulatory process. A "synthetic" Tadoma system, involving an artificial face with six facial actions, has been developed as a first order approximation to the natural Tadoma system. Experiments were conducted to explore the information transmission characteristics of the synthetic Tadoma system in terms of the four facial movements it incorporates: upper lip in-out, lower lip in-out, lower lip up-down, and jaw up-down movements.

Discrimination experiments showed that the just-noticeable-differences associated with each movement is about 9% of the reference displacement. One-dimensional absolute identification experiments produced, on the average, 1.7 bits of information transfer. Four-dimensional identification experiments produced information transfers in the range of 3-4 bits. Of the four dimensions considered, performance on the lower lip up-down movement was most affected, and performance on the jaw up-down movement was least affected, by simultaneous roving movements on the other dimensions.

Temporal order discrimination experiments indicated that subjects were able to discriminate different temporal orders of two movements when the onset time was separated by 3-6 msec. However, the performance was both movement and roving-range dependent.

Thesis Supervisors: Nathaniel I. Durlach and William M. Rabinowitz.

ACKNOWLEDGMENTS

I am greatly indebted to my advisors Nat and Bill for their constant encouragement, support and understanding, and for their guidance throughout the study.

I am grateful to my husband Jianmin for his work as an extraordinarily consistent subject.

I would like to thank Rosalie, Diane, and Matthew for their countless help.

Being foreign, I feel at home at CBG with many wonderful people — Charlotte, Larry, Lou, Yoshiko, Debbie, Nora, Mike, Harry, Pat Z., Pat P., Barbara, Janet, Lorraine, Karen, Carrin, Ken, Steve and Paul.

Horace Smith and Marilyn Pierce at the EECS graduate office deserve special thanks for their administrative support.

I owe my genuine gratitude to Dr. Ge Yao Chu for his generosity in establishing the Chu Fellowship which provided unrestricted financial aid to myself and many other Chinese students. Without this generous fellowship I would not have been able to study in the United States.

Although mere thanks are inadequate, I thank all of them.

Contents

1. Introduction	1
2. The Synthetic Tadoma System	4
3. Experiments	7
3.1 Discrimination Experiments	8
3.1.1 Preliminary Discrimination Experiments	8
3.1.2 Main Discrimination Experiments	11
3.2 Identification Experiments	15
3.2.1 1-D Identification with Fixed Background	15
3.2.2 1-D Identification with Roving Background	16
3.2.3 4-D Identification	17
3.3 Discrimination between Different Temporal Orders of Movements . .	19
3.3.1 Preliminary Experiments	20
3.3.2 Temporal Order Discrimination with Fixed Displacements . .	22
3.3.3 Temporal Order Discrimination with Roving Displacements .	23
4. Discussion	25
A. On Decision Model and Information Theory	29
A.1 d' and JND	29
A.1.1 A Model for 2I-2AFC Experiments	29
A.1.2 Definition of d' and JND	30
A.1.3 Computation of d' and JND	31
A.2 IT	32
A.2.1 Information and Uncertainty	32
A.2.2 Information Transfer in an Identification Experiment	33

B. Estimation of Information Transfer from Limited Experimental Data	35
B.1 Computer Simulation Algorithm	36
B.2 Results	38
B.3 Discussion	39

List of Figures

2.1	Displacement-counts relations for the four movements.	5
3.1	An example of how JND was computed.	9
3.2	Three stimulus designs for preliminary experiments.	10
3.3	Four stimulus designs for temporal-order experiments.	21
3.4	Preliminary results for temporal-order experiments.	22
A.1	A schematic illustration of the model.	31
B.1	An example of computer simulation.	37
B.2	Data processing for 1D identification tests on lio.	40
B.3	Data processing for 1D identification tests on lud.	41
B.4	Data processing for 1D identification tests on jud.	42
B.5	Data processing for 1D roving identification tests (HZT).	43
B.6	Data processing for 1D roving identification tests (JT).	44
B.7	Results of 4D identification experiments.	45
B.8	Relation of ITs from 1D and 4D identification experiments.	47

List of Tables

2.1	Straight-line fits to the measurements of D-C relation.	6
3.1	Preliminary results for discrimination experiments.	11
3.2	Results of discrimination with fixed background.	12
3.3	Results of discrimination with roving background.	14
3.4	Results of 1-D identification with fixed background.	16
3.5	Results of 1-D identification with roving background.	17
3.6	Results of temporal order discrimination with fixed displacements. .	23
3.7	Results of temporal order discrimination with roving displacements.	24

Chapter 1

Introduction

The study of tactile communication of speech serves several purposes. First, for those who are both blind and deaf, such study will help make possible the use of another sense modality for speech communication. Second, the study will lead to more insight into how the tactile sense works. Finally, the study will improve our basic understanding of speech communication.

Tactile speech communication was once thought impossible. Recent work, however, has demonstrated that well-trained deaf-blind people can perceive speech through tactile monitoring of the articulatory process at nearly normal rates (see [10,15,16]). In this method, referred to as Tadoma, the person receiving speech places his or her hand on the face and neck of the talker and, in the absence of auditory and visual input, monitors actions associated with the speech-production process to understand speech. In the typical hand position, the thumb rests across the middle of the lips and the fingers fan out across the face and neck.

The Tadoma method is superior to any other tactile display with respect to the understanding of running speech. Some characteristics of Tadoma that might be responsible for good performance are:

- a face is a “rich” multidimensional tactile display
- the display directly represents the articulatory process
- the hand is an exceptionally sensitive body part for tactile perception.

The synthetic Tadoma system[16], a first order approximation to the natural Tadoma system, was built to facilitate the study of cues used by Tadoma readers . The display portion of the synthetic Tadoma system (the artificial face) is a plastic anatomical-model with six facial actions that can be computer controlled. The specific actions included are:

- upper lip in-out movement (uio)
- lower lip in-out movement (lio)
- lower lip up-down movement (lud)
- jaw up-down movement (jud)
- laryngeal vibration
- oral airflow.

Each of these actions can be controlled independently. Leotta [8] evaluated the synthetic Tadoma system by comparing speech-segment discrimination performance on the synthetic system to results obtained from natural Tadoma. Results with the synthetic system and natural Tadoma were generally similar, although some deficiencies in the synthetic system were revealed.

In this thesis, efforts will be made to characterize the artificial face as a multidimensional tactile display using elementary, non-speech stimuli. All of the movement dimensions will be studied separately and jointly with respect to information transmission. The laryngeal vibration and air flow systems were being modified during the research and, as a result, were not examined.

To fully understand the success of natural Tadoma for the perception of running speech, rates of information transfer through the tactile sense should be examined. This requires study of dynamic as well as static stimuli on the synthetic Tadoma system. The latter, however, will be the major focus of this thesis.

Assessment of information transmission was carried out in two steps. First, discrimination experiments were conducted to investigate the basic resolution of the human tactile sense with respect to the four movements uio, lio, lud and jud. Identification experiments were then performed to determine how much information could be received through the movement systems. According to studies by Pollack, Garner, and others (see [5,9,11,12,13]), information transfer for unidimensional stimuli is limited to roughly 2 bits. However, by employing several dimensions (as in Tadoma), this limit can be greatly increased. In general, the amount of

information transfer depends not only on the number of dimensions, but on how the different dimensions interact.

The goal of temporal order discrimination experiments was to examine the limits of temporal resolution for the Tahoma movements. The task required the judgment of temporal orders. Hirsh and Sherrick ([6]), in their study concerning the minimum time that can separate two events so that their temporal order could be resolved, found that approximately 20 msec was needed to insure judgments that were 75% correct.

Chapter 2

The Synthetic Tadoma System

The synthetic Tadoma system [16] consists mainly of an artificial face which is built around an anatomical-model plastic skull. The four facial movements (uio, lio, lud, and jud) are controlled by servomotors which are operated in a closed-loop, position-control system with shaft velocity and position feedback provided. Two solenoid valves control the airflow from a tank of compressed air through different nozzles mounted behind the upper-front TEETH¹ of the skull. A bone-conduction vibrator is mounted in the area where the fingers would normally touch the neck to simulate the laryngeal vibration.

The LIPS are rubber tubes. The two ends of the upper (lower) LIP are attached to the upper (lower) back TEETH of the skull. Flexible metal cables in semi-rigid plastic casings, which are attached to the middle of the two LIPS at one end and to the motors at the other end, are used to produce LIP protrusions and the lower LIP up-down movement. For movement of the JAW relative to the skull, a rotational joint was constructed in the vicinity of the condylar process and a drive linkage was connected from the interior of the JAW to the JAW servomotor. The four facial movements are the focus of the study in this thesis.

The FACE is normally set in a reference or neutral position. This position is defined such that the JAW and the lower LIP are fully up (MOUTH closed), and the upper and lower LIPS are fully in (no protrusions). All movements begin and end in this state.

The displacements of the four movements are under computer control. For the current system, the displacements are most conveniently specified in terms of *counts*. The transform from counts to actual displacements for each of the four

¹I shall use SMALL CAPS font to refer to the corresponding parts on the synthetic Tadoma system throughout the thesis.

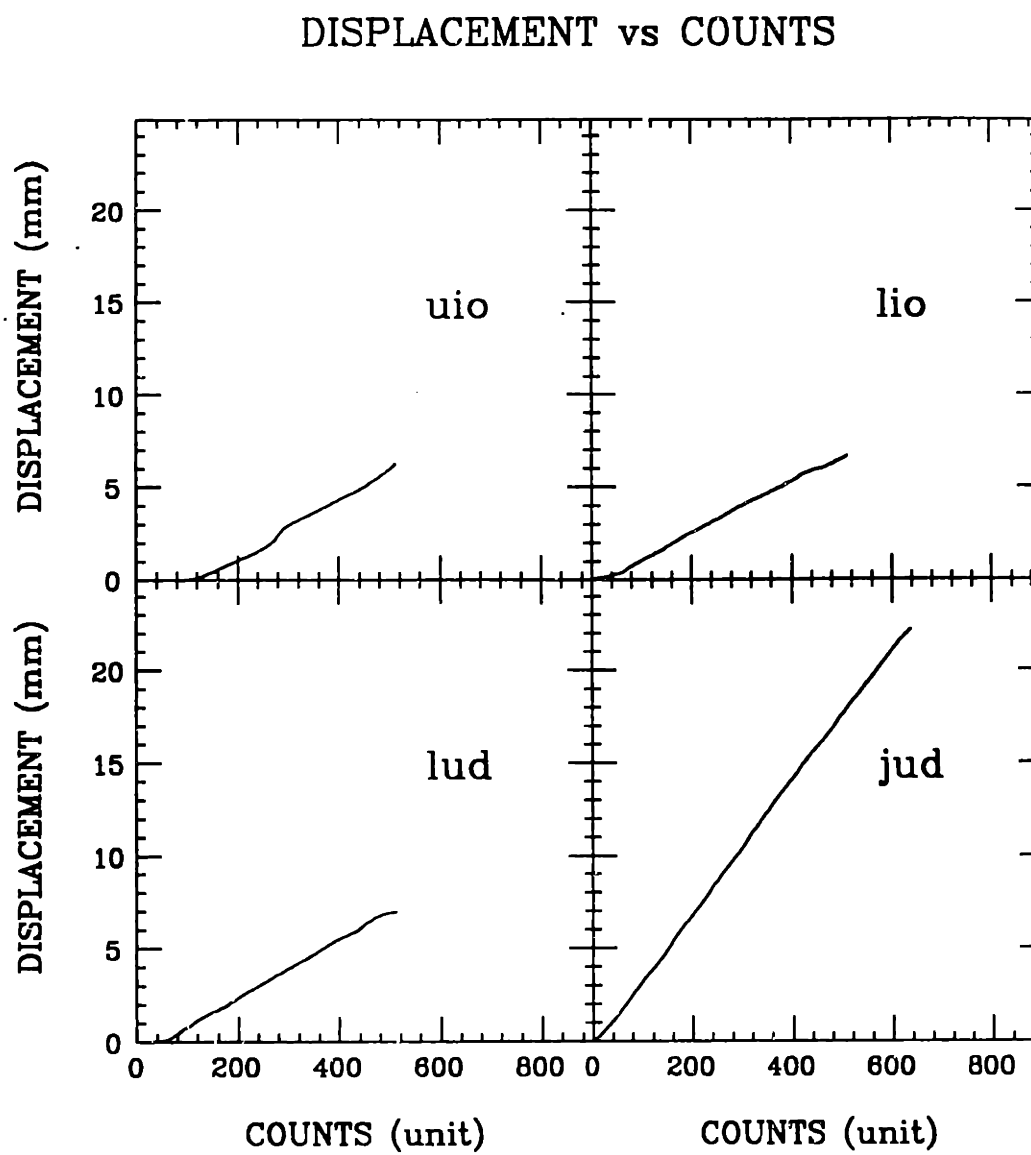


Figure 2.1: Displacement-counts relations for the four movements.

Movements	Slope k	Intercept C_0	σ
uio	0.016	130	0.122
lio	0.014	20	0.083
lud	0.015	46	0.072
jud	0.036	8	0.113

Table 2.1: Straight-line fits to the measurements of D-C relation.

movements is plotted in Figure 2.1. These measurements were made as follows. A microscope was positioned to view the entire displacement range of the movement system. A grid paper was also placed in the visual field and was used as a ruler. The measurements were then made by making a particular component move from its neutral position to a steady-state position (specified by counts) and reading off the corresponding displacement from the grid paper. We estimate the the precision of the measurements to be $\frac{2}{10}$ mm. The maximum displacements are 6-7 mm for the uio, lio and lud movements, and 24 mm for the jud movement.

The values of slope (k) and intercept on the counts-axis (C_0) for the straight lines that best fit these measurements, as well as the corresponding values of the rms deviations σ (along the displacement-axis), are presented in Table 2.1. The displacement (D) as a function of counts (C) can be expressed by $D = k(C - C_0)$, with $C \geq C_0$, for each of the four movements. From Figure 2.1, we see that the JAW moves linearly from the beginning. The other three channels all have zero segments at the beginning of the curves, which is caused by backlash in cable drives. The uio channel has the largest such segment. Efforts were made to correct this problem and we believe that the uio channel moves more linearly now.

For the reader's convenience, the rest of the thesis shall always use millimeters as a measure of displacement instead of counts. The conversion from counts to millimeters is based on Table 2.1.

Chapter 3

Experiments

The three main experiments completed were Discrimination, Identification and Temporal-Order experiments.

Two subjects, HZT (female, the author) and JT (male), were used in the study. Both are hearing and sighted MIT students, and neither had previous exposure to the Tadoma method. In all experiments, the subject sat with his/her elbow resting on the table on which the FACE was placed. The subject put his/her hand on the FACE with the thumb resting across the middle of the LIPS and the fingers fanning out across the CHEEK (the typical hand position of Tadoma users). The hand was kept on the FACE during the interstimulus intervals (except in the preliminary discrimination tests). The subject faced away from the FACE so that the movements could not be sensed visually. Also, the subject wore earmuffs (and ear plugs when necessary) to eliminate possible auditory cues from the FACE.

The waveforms used to drive the FACE were computer generated displacement steps composed of a 300 ms constant portion and 150 ms (Hanning) transition portions at the beginning and the end of the waveforms. The 600 ms waveform was then scaled by desired amplitudes to produce desired displacements. All waveforms differed only in their scaled amplitudes.

All experimental runs started with an informal training period. The subjects could start the experimental run when they felt ready. Each run consisted of 100 trials. Trial-by-trial feedback was always provided.

The four FACIAL movements will be referred to as “displacements”, “movements”, “parameters”, “channels”, or “dimensions” interchangeably. The terms “target” and “background” shall be used to clarify the different roles the four movements play in an experiment.

3.1 Discrimination Experiments

All discrimination experiments used the symmetric 2I-2AFC (two-interval, two-alternative-forced-choice) paradigm. Each subject completed two runs for each reference (D) and increment (ΔD) pair. For each D , three ΔD s were used. Thus, each subject completed six runs (600 trials) for each reference. The order in which these six runs were performed was randomized and not repeated for different subjects. Performance was measured, for each of the three ($D, \Delta D$) pairs, by forming the appropriate 2×2 matrix (pooled from two runs) and estimating the sensitivity index d' and response bias β . The response bias β was found to be sufficiently small to be ignored (within 10% of the corresponding d'). As found in many studies on sensory discrimination (e.g., [3]) and in our discrimination experiments, the dependence of d' on the increment ΔD for a fixed reference displacement D can be described by a straight line through the origin. Thus the performance for a fixed D can be summarized by the slope $\delta' = d'/\Delta D$ of the straight line (i.e., the sensitivity per millimeter). The δ' was estimated by averaging the three $d'/\Delta D$ ratios obtained above. The just-noticeable difference JND, defined by the performance criterion $d'=1$, is given by $JND=1/\delta'$. All the results for the discrimination experiments are presented in the form of JND. Figure 3.1 shows an example of how JND was computed. As can be seen, the $d'(\Delta D)$ relation can be well described by a straight line through origin.

3.1.1 Preliminary Discrimination Experiments

It was not very clear at first whether a subject should keep his/her hand on the FACE throughout the whole experimental run (simulating what the Tadoma users do) or only touch the FACE briefly when the facial feature was in the desired position. The purpose of the preliminary study was to compare these two methods. Only subject HZT participated in this study.

As mentioned in Chapter 1, the study of static stimuli was the focus of this thesis. However, when a subject kept his/her hand on the FACE during a complete movement (from neutral position to desired position, then back to neutral position),

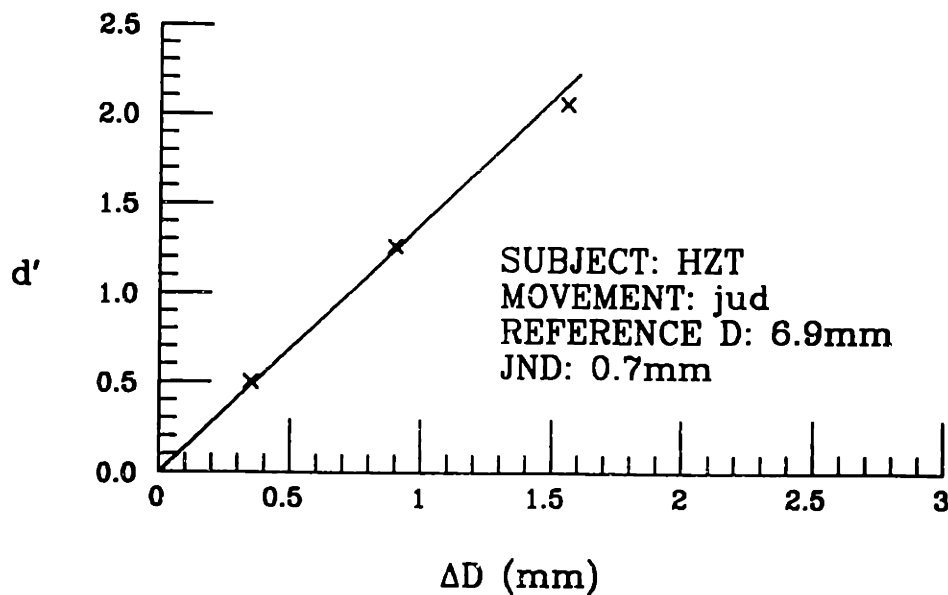


Figure 3.1: An example of how JND was computed.

he/she not only picked up the static amplitude cue from the steady portion of the stimulus but also a dynamic velocity cue from the transition. (i.e., since the transition duration was constant, the movement velocity to and from the steady-state displacement varied in proportion with the steady-state displacement.) This potential velocity cue could be eliminated by not letting subjects touch the FACE during transition movements.

Three kinds of stimuli were designed to reveal the effect of possible velocity cues. In scheme 1 (which was ultimately selected for the main experiments), the transition time was held at 150 ms and the total duration of the waveform was held at 600 ms (see Figure 3.2). The subject kept her right hand on the FACE during the whole stimulus (using the same posture Tadoma users employ). Both velocity and amplitude cues were available to the subject in this scheme.

In scheme 2, the transition time remained at 150 ms, but the total duration of the waveform was increased to 1100 ms (see Figure 3.2). Following the end of the rising portion of the waveform, the word "ON" appeared on the terminal screen, followed by the word "OFF" 300 ms later. The subject was instructed to put her hand on the FACE when she saw the word "ON" and take it away from the FACE

schemes	scheme 1	scheme 2	scheme 3
d'	1.41	1.41	1.44

Table 3.1: Preliminary results for discrimination experiments.

when she saw “OFF”. The additional 500 ms was sufficient to allow the subject to remove her hand from the FACE without feeling the transition portion at the end of the waveform. (Also, the transition portion at the beginning of the waveform could never be felt.) In this scheme, the subject’s hand touched the FACE for roughly 300 ms and only the steady-state amplitude cue was available to her.

In scheme 3, the steady-state portion of the displacement was kept at 300 ms (as in scheme 1), but the transition time for the two stimuli was adjusted such that the slopes of the rising and falling portions were the same for both displacements. The subject kept her hand on the FACE during the whole stimulus presentation. Here, a possible duration cue (total duration of the waveform) as well as an amplitude cue was available to the subject.

All three of these schemes were tested for jud6.9¹ and jud7.7 (i.e., for a reference $D = 6.9 \text{ mm}$ and $\Delta D = 0.8 \text{ mm}$). Results of these tests, presented in Table 3.1, show that all three schemes lead to essentially equal values of d' . The equivalence of schemes 1 and 3 indicates either that velocity is not a cue or, if velocity is a cue, that the amplitude cue is partially masked by the transitions and that the two effects cancel.

Since the three schemes yielded similar results, we elected to use scheme 1 (the scheme closest to the method Tahoma users employ) for subsequent experiments.

3.1.2 Main Discrimination Experiments

Our goal was to investigate the resolution of the human tactile sense for the four parameters uio, lio, lud and jud.

¹This kind of code shall be used a lot in this thesis. The first three letters refer to the movement channel, followed by a real number in millimeters indicating the displacement of the movement.

Target Channel	Reference D (mm)	JND/D (%)		Ave across Subjects	Ave across References
		HZT	JT		
uio	1.7	9.5	9.6	9.6	10.2
	5.1	9.9	11.6	10.8	
lio	0.3	8.2	6.0	7.1	9.0
	3.2	9.4	10.0	9.7	
	6.0	9.4	11.2	10.3	
lud	0.1	7.4	7.2	7.3	9.3
	3.1	11.5	10.6	11.1	
	6.1	9.3	9.6	9.5	
jud	0.4	9.0	9.0	9.0	8.8
	1.2	11.3	9.8	10.6	
	1.5	10.2	9.4	9.8	
	3.3	8.4	8.7	8.6	
	6.9	11.0	6.2	8.6	
	14.1	8.0	8.2	8.1	
	21.3	7.3	6.8	7.1	

Table 3.2: Results of discrimination with fixed background.

Two types of discrimination experiments were performed. In the first, we applied movements to only one of the four channels (referred to as the “target” channel) and kept the other three channels (referred to as the “background” channels) at their neutral positions. In the second, random amplitude movements were applied to all of the three background channels. The purpose of doing this pair of experiments was to examine how resolution in one channel is influenced by random movements in other channels.

Discrimination with Fixed Background

The waveform used to drive the FACE was the same as in scheme 1 of Figure 3.2. The results for 1-D discrimination experiments are summarized in Table 3.2. Instead of listing the JNDs associated with each reference displacement D , the corresponding Weber fractions JND/ D in percentage are given in the table.

The intersubject differences are generally very small (the only substantial difference occurring for jud6.9). For the uio and lio movements, the Weber fraction tends to increase slightly as D increases. For the jud movement, however, the Weber fraction tends to decrease slightly as D increases (except for jud0.4). Overall, the Weber fractions are roughly constant (9%) and independent of the reference displacements D . These results are very important to the design of subsequent experiments and shall be referred to later in this thesis.

Discrimination with Roving Background

This experiment differed from the discrimination experiment with fixed background in that the movement on the target channel was accompanied by simultaneous random movements on the three background channels. The displacement on each of the background channels was chosen independently from four possible levels. Before each experiment, the subject was informed of the target channel and instructed not to change hand position² in order to avoid interference produced by the roving movements in the background channels. They were also told to ignore the three roving channels as much as possible and pay attention only to the target channel.

In order to eliminate the effect of training, we repeated the discrimination experiment with fixed background at the same time we performed the discrimination experiments with roving background. For each reference-increment pair, each subject first finished the six runs with roving background and then (sometimes after a short break) the six runs with fixed background. All twelve runs were always completed on the same day.

Only one reference displacement was chosen for each of the four movement channels (see Table 3.3). The JNDs associated with this reference for both roving and nonroving discrimination experiments were computed and compared.

The results, presented in Table 3.3, show that the $\frac{JND_{roving}}{JND_{fixed}}$ varies between 1.4 and 6.2 and is always substantial. Notice that this ratio is ∞ in the "lud" row. This

²The uio, lio, and lud movement could only be felt by the thumb. The jud movement could be sensed by both the thumb and palm. When uio was not the target channel, the subject could lower his/her thumb such that it only felt the lower LIP of the FACE. This was not permitted in the experiments and was carefully monitored.

Target Channel	Reference D(mm)	Subject HZT			Subject JT			average	
		$\frac{JND_r}{D}\%$	$\frac{JND_l}{D}\%$	$\frac{JND_f}{JND_l}$	$\frac{JND_r}{D}\%$	$\frac{JND_l}{D}\%$	$\frac{JND_r}{JND_l}$	$\frac{JND_l}{D}\%$	$\frac{JND_r}{JND_l}$
uio	1.1	25.7	6.4	4.0	20.3	6.5	3.1	6.5	3.6
lio	2.5	33.8	6.5	5.2	47.6	7.7	6.2	7.1	5.7
lud	2.3	∞	7.0	∞	∞	6.9	∞	7.0	∞
jud	6.9	13.1	5.9	2.2	15.1	10.5	1.4	8.2	1.8

Note: JND_r is short for JND_{roving} , and JND_f is short for JND_{fixed} .

Table 3.3: Results of discrimination with roving background.

denotes the fact that neither subject could resolve the difference of the two stimuli even when the increment was as big as the reference. The data in column $\frac{JND_{fixed}}{D}$ for both subjects, compared with similar ones in Table 3.2, show some training effect (except for subject JT's performance for jud6.9). Finally, while the averaged $\frac{JND}{D}$ percentage in Table 3.2 is largest for uio and smallest for jud, the averaged $\frac{JND_{fixed}}{D}$ percentage in Table 3.3 is smallest for uio and largest for jud. Thus, the improvement in performance is greatest in uio and least in jud.

It's interesting to note that although the $\frac{JND_{roving}}{JND_{fixed}}$ ratios are not exactly the same for the two subjects, the dependence of this ratio on target channel is the same for both subjects. The JND for lud was most affected by roving movements, followed by lio, then uio, and finally jud. This ordering appears reasonable in the light of the following observations. The lud movement can only be felt by sensing the position of the lower LIP on the thumb. Both lio and jud movements also change the lower LIP position. When all these movements are present in a single presentation, subjects can't separate out the lud movement. The lio movement is felt by sensing the lower LIP protrusion. When the lud and jud movements are present, the lower LIP touches the subject's thumb at different locations, causing the protrusion to be felt differently. The uio movement is relatively well separated from the other three movements. It can be felt by sensing the upper LIP protrusion against the thumb without direct interference from any other movement. However, when the lower LIP protrusion exceeds that of the upper LIP, the thumb may be moved away from the

upper LIP, causing degradation in performance. Finally, the jud movement is least affected because it can be felt by the subject's palm as well as thumb.

3.2 Identification Experiments

Unidimensional as well as four-dimensional information transfer characteristics of the FACE in terms of the four movement channels were examined. All identification experiments employed a one-interval absolute identification procedure with trial-by-trial feedback. For each identification test, subjects were informed of the number of alternatives in the stimulus set and received brief training in associating the alternatives with a set of integers in a natural order. Upon receiving a stimulus, the subject was required to respond with the corresponding integer. A stimulus-response confusion matrix was obtained, and information transfer (IT) was computed. All ITs are bias-corrected (see Appendix B).

3.2.1 1-D Identification with Fixed Background

In this experiment, we examined the limit of information transfer for unidimensional stimuli. Three channels were examined: lio, lud and jud. Preliminary results showed that the uio channel yielded results similar to those of the lio channel. To save time, data on uio were not completed.

Sixteen stimuli were used for the jud channel and 14 stimuli for both the lio and lud channels. The displacements of the stimuli were equally spaced on a logarithmic scale (consistent with Weber's law). The 0 mm displacement was included as one of the stimuli. The displacements of the stimuli spanned 24 mm for jud and 7 mm for the lio and lud channels.

The ITs, summarized in Table 3.4, vary between 1.68 and 1.99 bits. The entries in column "trials" are the number of trials completed by each subject for each target channel. It should be noted that although the IT entries in Table 3.4 are the results obtained from different total number of trials, they are all bias-corrected. Thus direct comparison is valid. The intersubject differences are 0.01, 0.04, and 0.19 bits

Target Channel	Number of stimuli	Subject HZT		Subject JT		average IT(bits)
		IT(bits)	trials	IT(bits)	trials	
lio	14	1.77	1200	1.76	1300	1.77
uio†	14	1.77	—	1.76	—	1.77
lud	14	1.68	1200	1.72	1300	1.70
jud	16	1.99	1200	1.80	1600	1.90

†We assumed that IT for lio and uio were the same.

Note: All "IT" entries are bias-corrected.

Table 3.4: Results of 1-D identification with fixed background.

for lio, lud, and jud respectively. For both subjects, the ITs are the highest for the jud channel, followed by the lio (and uio) channel, and the lud channel. This order is reasonable because the jud channel has a much bigger range of displacement than all other three channels, and includes more JNDs (about 50) in its stimulus set. According to the study by Braida and Durlach [1] in the domain of auditory intensity perception, IT is limited by sensory as well as memory noise when the range of stimuli includes less than 50 JNDs. It seems reasonable to say that IT for the jud channel is less limited by sensory noise than the other channels. This issue will be discussed further in Chapter 4.

3.2.2 1-D Identification with Roving Background

The purpose of this experiment was to examine the interaction of movement channels. The channel being identified will be referred to as the target channel and the other three channels as the roving background channels. This study parallels the 1-D discrimination with roving background.

The stimulus set used for the target channel was the same as that used in the 1-D identification experiments with fixed background. Simultaneous roving movements were applied to the three background channels. The displacement for each of the background channels was chosen independently from four possible levels. Before each experimental session, each subject was informed of the target channel. They

Target Channel	Number of stimuli	Subject HZT			Subject JT			average $\frac{IT_{fixed}}{IT_{roving}}$
		IT_{roving}	IT_{fixed}	$\frac{IT_{fixed}}{IT_{roving}}$	IT_{roving}	IT_{fixed}	$\frac{IT_{fixed}}{IT_{roving}}$	
uio	14	1.37	1.77	1.3	1.06	1.76	1.7	1.5
lio	14	0.85	1.77	2.1	0.89	1.76	2.0	2.1
lud	14	0.45	1.68	3.7	0.56	1.72	3.1	3.4
jud	16	1.44	1.99	1.4	1.38	1.80	1.3	1.4

Table 3.5: Results of 1-D identification with roving background.

were instructed not to change their usual hand position on the FACE and to ignore the roving movements as much as possible. Each subject completed 1000 trials for each of the four target channels. The results of this experiment (IT_{roving}) are summarized in Table 3.5. For convenience of comparison, IT_{fixed} are reproduced from the “IT(bits)” columns in Table 3.4.

The intersubject differences in IT_{roving} are 0.31, 0.04, 0.11, and 0.06 bits for the uio, lio, lud and jud channels respectively. The IT_{roving} for lud channel is roughly 0.5 bits for both subjects and this indicates that neither subject was able to identify even two stimulus categories in this channel. Referring back to Table 3.3, one notes that subject’s performance on the discrimination task was also substantially degraded by simultaneous roving movements in other channels. By comparing the order of the ratios $\frac{IT_{fixed}}{IT_{roving}}$ in Table 3.5 with that of the ratios $\frac{JND_r}{JND_f}$ in Table 3.3, the results are seen to be consistent in that the lud channel is most affected by simultaneous movements in other channels, followed by lio, uio, and then jud.

3.2.3 4-D Identification

In this experiment, we examined the limit of information transfer for 4-D stimuli. The number of alternatives used for each channel was four. Thus the total number of alternatives was 256. The four displacement values for each channel included 0 mm, were equally space on a logarithmic scale, and spanned 6 mm for uio, 7 mm for lio and lud, and 23 mm for jud channel. As found in the 1-D identification experiments

with roving background, the amount of information obtainable from a single channel never exceeds 1.5 bits. Thus, four alternatives on each channel was thought to be sufficient. The displacement for each channel was chosen, independently, out of its four possible alternatives with equal probability. Then a 4-D stimulus was formed and presented to the subject. Upon receiving the stimulus, the subject was instructed to respond with an integer between 1 and 4 for each of the four movement channels. As usual, trial to trial feedback was given. Each subject received about one hour of training in associating the integers 1 to 4 with each of the four movements on each of the four channels. A total of 5000 trials was collected for each subject. No obvious improvement from session to session (due to training effect) was observed.

Performance was measured by pooling the 5000 trials for each subject to form a 256 by 256 confusion matrix from which the information transfer was computed. Efforts were made to obtain an unbiased estimate of IT in this case. A detailed description of the data processing is given in Appendix B. The resulting information transfer was 3.8 bits for subject HZT and 3.1 bits for subject JT. Referring back to Table 3.5, one notes that the sum of IT_{fixed} over the four channels, roughly 7 bits for each subject, is twice as large as the 4-D IT. The fact that multi-dimensional identification yields much lower IT than the sum of 1-D results is found by many researchers (e.g., see [14]). The degradation of the 4-D results can be due to interactions among channels as well as to increased memory noise. In this regard, note from Table 3.5 that the sum of IT_{roving} over the four channels is 4.1 bits for subject HZT and 3.9 bits for subject JT. These numbers are very close to the ITs obtained from the 4-D identification experiments.

One further observation could be made here. During the experiments, both subjects felt that extra short term memory was needed to transform sensations to responses. Recall that the stimulus presentation consisted of four movements aligned in time. The subjects were to respond to uio, lio, lud and jud movements sequentially. Thus they had to "memorize" the "feelings" caused by the movements (in whatever way) as they responded to them one by one. Although this extra load on short term memory should degrade the information transfer for the 4-D stimuli,

the data demonstrate that this degradation was very limited: the information transfer measured in the 4-D identification experiments is very close to the sum of the information transfer obtained from the 1-D roving identification experiments. In other words, the main cause of the degradation of information transfer in the 4-D experiments (below the above-measured 7 bits) is the interference among stimulus channels, the same interference as that measured by comparing fixed and roving discriminations.

3.3 Discrimination between Different Temporal Orders of Movements

The following experiments explored the ability to discriminate between different temporal orders of a pair of movements. These experiments were only an exploratory probe.

We first performed preliminary experiments to help design experimental stimuli. We then did discrimination experiments with both fixed and roving displacements. The experiments were performed using a single-interval, two-alternative-forced-choice paradigm with trial-by-trial feedback. A pair of movements, separated in their onset time by Δt , was presented to the subject with either movement delayed with equal probability. For each combination of movement pair and Δt , each subject completed two runs (200 trials). Three different Δt were chosen for each movement pair. The order in which the six runs were performed was randomized and not repeated for the two subjects. Performance was measured by forming the appropriate 2×2 matrix for each Δt and estimating the sensitivity index d' and response bias β . As already mentioned in the discrimination experiments, the response bias β was small enough to be ignored. The function $d'(\Delta t)$ was again well approximated by a straight line through origin (see Figure 3.4); thus the performance from the six runs could be summarized by the just-noticeable difference JND.

3.3.1 Preliminary Experiments

In designing the experiments, many parameters needed to be determined. In terms of timing, we could make onset time of the movements asynchronized, offset time asynchronized, or both. In terms of displacements of the movements, we could use fixed or roving displacements. We designed four schemes to test various combinations of the parameters.

For the four movement channels, there were six pairs of movements. The pair *lio&jud* was used for the preliminary study. Only subject HZT participated in this study.

The four stimulus designs are depicted in Figure 3.3. Schemes 1, 2, and 3 used fixed displacements and scheme 4 used roving displacements. The waveforms A and B used in scheme 1 were onset-time asynchronized and offset-time synchronized; the waveforms used in scheme 2 were both onset-time and offset-time asynchronized; and the waveforms used in scheme 3 were onset-time synchronized and offset-time asynchronized. Whereas schemes 1 and 3 provided a duration cue in addition to an onset or offset cue, scheme 2 provided both an onset and offset cue but no duration cue. The waveforms used in scheme 4 were similar to those in scheme 2 in that they were both onset-time and offset-time asynchronized; however, they both had roving displacements. The roving range was 0.4 to 5.3 mm for the *lio* channel and 1.5 to 14.1 mm for the *jud* channel. The roving amplitude for a particular presentation was chosen from four possible levels with equal probability. The roving displacements for the two waveforms were chosen independently. For the four schemes, the transition time at either the beginning or the end of the waveforms was fixed at 150 ms for all the waveforms. Duration of the waveforms are specified in Figure 3.3.

The data are plotted in Figure 3.4. The data could be fit by two straight lines. The line which fits the data of scheme 1, 2, and 3 corresponds to a JND of 9 ms. The line which fits the data for scheme 4 corresponds to a JND of 35 ms. The results indicate that it didn't make any difference whether we made the onset-time of movements asynchronized, or offset-time asynchronized, or both. The results also show that roving amplitudes for the movements degraded the subject's ability to

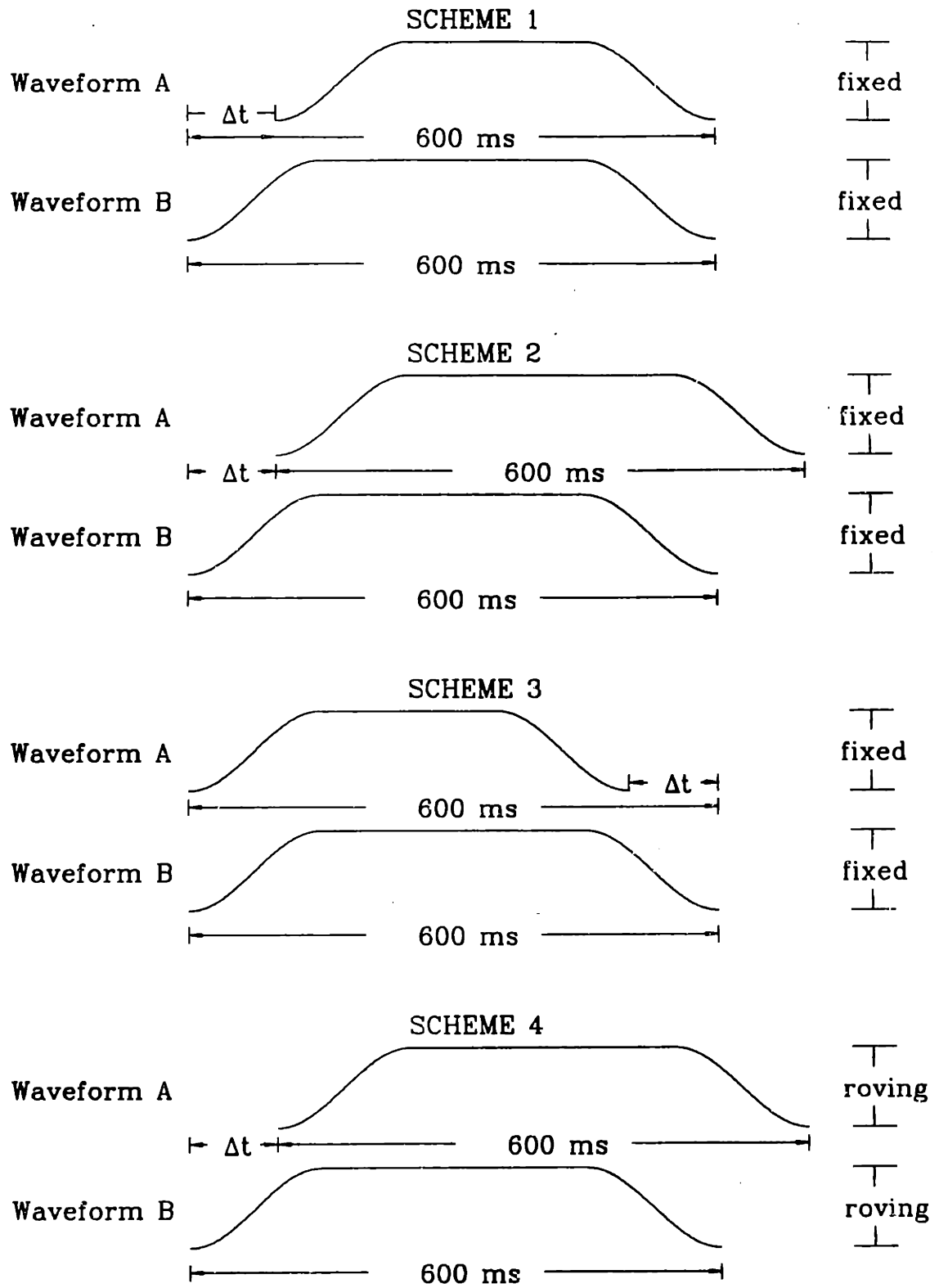


Figure 3.3: Four stimulus designs for temporal-order experiments.

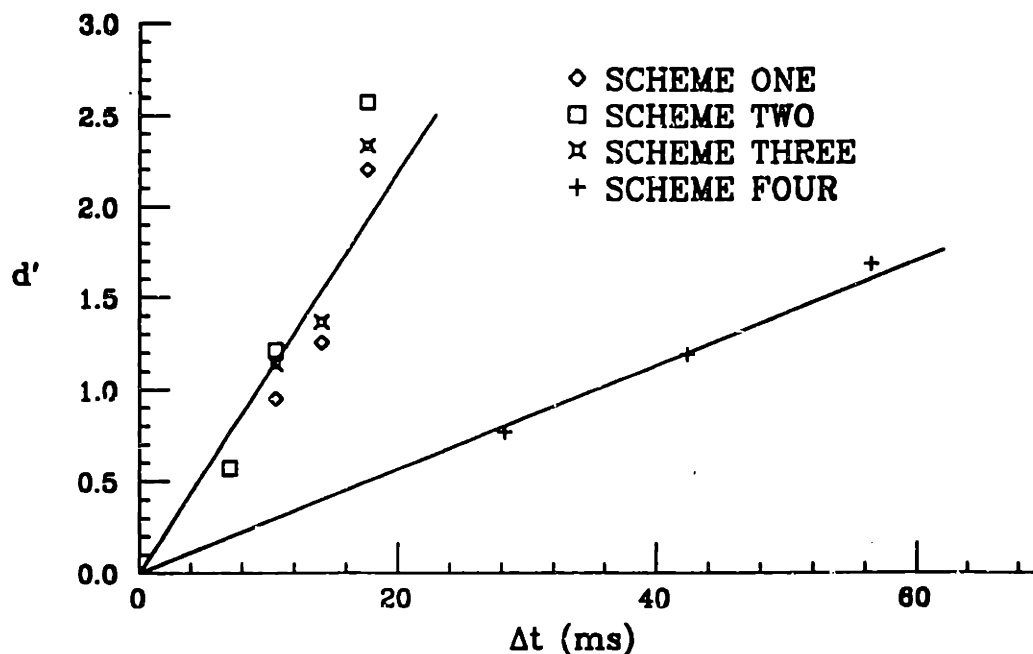


Figure 3.4: Preliminary results for temporal-order experiments.

discriminate different temporal orders of movements.

The above conclusions led to the following two experiments. One experiment employed scheme 1 to study temporal order resolution for all six pairs of movements (see Section 3.3.2). The other experiment examined the effect of the range of roving amplitudes on temporal order resolution (see Section 3.3.3).

3.3.2 Temporal Order Discrimination with Fixed Displacements

The temporal relation of the stimulus waveforms was the same as in scheme 1 described in Section 3.3.1 (also see Figure 3.3). In this experiment, we used fixed displacements (3.2 mm) for each of the four movements. The results for this experiment are summarized in Table 3.6.

The six columns in Table 3.6 are arranged such that the averaged JNDs increase from left to right. The largest intersubject differences occur when one of the paired movements is jud (the last three pairs in the table). The average JNDs for these three pairs are the largest (6.2 ms). The first pair, uio&lio, has the smallest JND

Movement Pair	uio& lio	lio& lud	uio& lud	uio& jud	lio& jud	lud& jud
HZT: JND(ms)	3.2	4.2	4.7	5.3	4.9	5.4
JT: JND(ms)	3.2	4.0	4.2	6.8	7.5	7.4
average JND(ms)	3.2	4.1	4.5	6.1	6.2	6.4

Table 3.6: Results of temporal order discrimination with fixed displacements.

(3.2 ms). This is probably because the two movements are most similar. The other two pairs (the second and the third) have an average JND of 4.3 ms. More experiments need to be done to characterize the dependence of JND on movement pairs. Overall, the subject's ability to discriminate between different temporal orders of movements was exceptionally good (3–6 ms) compared to other results from similar studies involving other types of parameters (e.g., see [6]).

3.3.3 Temporal Order Discrimination with Roving Displacements

We wanted to explore further how roving displacements affect a subject's ability to discriminate between different temporal orders of movements. The movement pair we chose to examine was lio&jud. Only subject HZT was involved in this experiment.

Some relevant data were already available to us. In Section 3.3.2, we found that with fixed displacements, JND for the movement pair lio&jud was 4.9 ms for subject HZT. In Section 3.3.1, we found that when roving range was 0.4–5.3 mm for lio and 1.5–14.1 mm for jud (scheme 4), the JND for the pair lio&jud was 35 ms for HZT. In order to eliminate the effect of training in comparing data, we redid these two experiments.

A smaller roving range was used in this experiment. The roving range was 2.5–4.6 mm for the lio channel and 6.9–12.3 mm for the jud channel.

The data, summarized in Table 3.7, suggest that a subject's ability to discrimi-

Displacements of JND	fixed displacements	small roving range	large roving range
JND(ms)	4.9	8.4	21.8

Table 3.7: Results of temporal order discrimination with roving displacements.

nate between different temporal orders of movements is systematically affected by roving displacements of the paired movements. The larger the roving range, the greater the influence. More subjects are needed to generalize the above conclusion. As for training effects, subject HZT did not improve in discriminating temporal order with fixed displacement (always 4.9 ms), but she improved a lot in discriminating temporal order with the bigger roving range (JND decreased from 35 to 21.8 ms).

Chapter 4

Discussion

Our data from the discrimination experiments roughly obey Weber's law. However, the recent study on length discrimination by Durlach et al. (see [3]) reported that Weber's law was violated. Our experiment mainly tested the sensitivity along a thumb, while theirs used the finger-span method. Thus, the difference between the two tests with regard to Weber's law might be due to differences in the physiological nature of the tasks. The Weber fraction we found in discrimination with fixed background is roughly 9% (refer to Table 3.2). However, due to training effect, this fraction was only 7% when we redid part of the tests (refer to Table 3.3). As reported by Durlach et al., the JND in length measured in discrimination experiments by the finger-span method is roughly 1 mm for reference lengths of 10 to 20 mm, which is roughly comparable to the results we obtained.

Our absolute identification experiments are most comparable to the study by Rabinowitz *et al.* (see [14]). Both studies used multidimensional tactile displays as experimental apparatus and measured identification performance by information transfer (IT).

Our one-dimensional identification experiments yielded 1.6–2.0 bits of IT, which is slightly higher (by 0.1–0.3 bits) than the results obtained by Rabinowitz *et al.* The corresponding number of categories that are identifiable on one dimension (3–4) is, however, still somewhat below that implied by “the magic number 7 ± 2 ” (see [9]).

According to the study by Braida and Durlach [1] in the domain of auditory intensity perception, IT is limited by both sensory and memory noise when the range of stimuli is modest. The IT is dominated by memory noise and saturates at 2.2 bits when the range of stimuli includes more than 50 JNDs. From our one-dimensional discrimination experiments, it is estimated that the range of stimuli

for our one-dimensional identification experiments included about 30–50 JNDs.¹ As mentioned in the paper by Rabinowitz *et al.*, the number of JNDs in their tactile intensity range (27 dB) was only 15–25. In general, when the number of JNDs is modest, IT increases as the number of JNDs increases. This could explain why our one-dimensional IT is slightly higher than that of Rabinowitz *et al.*, but slightly lower than “the magic number 7 ± 2 ”.

Our four-dimensional identification experiments yielded, on the average, 3.5 bits of IT, while the three-dimensional experiment of Rabinowitz *et al.* produced 4 bits. As indicated by our one-dimensional roving d' discrimination and identification experiments, the lud movement on the synthetic Tadoma system was most affected by simultaneous roving movements on other movement channels. The IT loss caused by simultaneous roving movements was 70% for this variable. The fact that the IT_{roving} for lud was only 0.5 bits for both subjects (refer to Table 3.5) indicates that little information was available from the lud movement in a multi-dimensional experiment. For the remaining three dimensions, uio, lio and jud, simultaneous roving movements caused an IT loss of 32%, 51%, and 26% respectively.² According to Figure 3 in the paper by Rabinowitz *et al.*, the variable “contactor area” was almost “orthogonal” to the other variables. The loss of IT for the other two variables was about 50% for intensity and 25% for frequency.³ Overall, the three variables employed by Rabinowitz *et al.* were more “independent” than the four movements we used. This might explain why the IT they obtained from their three-dimensional tests was relatively high compared to the IT we obtained from our four-dimensional tests.

The 3.5 bits of IT obtained from the movement systems on the synthetic Tadoma system might be substantially lower than the IT the synthetic Tadoma system is capable of transmitting, since oral airflow and laryngeal vibration are likely to be

¹For instance, the JNDs for jud movement was on the average 9% of references. The range of jud movement was 0–24 mm. If we choose 0.1 mm as the minimum reference displacement, then the jud movement covered about 50 JNDs.

²These percentages are computed from Table 3.5 by $(IT_{fixed} - IT_{roving})/IT_{fixed}$.

³The IT_{roving} data are not available from their study. These percentages are estimated from the corresponding pooled projections in their paper.

substantially independent of the movement systems. It is possible that these two channels could each add 1–2 bits of IT to the movement systems.

Another interesting observation from the identification experiments concerns the relations that exist between the 1D and 4D ITs. In particular, the sum of the 1D ITs obtained with a fixed background (about 7 bits) is higher than the 4D IT (about 3.5 bits). This is reasonable because we know that the four facial parameters are dependent; thus some degradation in 4D IT should be expected. The sum of the 1D ITs obtained with a roving background (about 4 bits) is, however, very close to the 4D IT. One might argue that 4D identification requires extra short term memory for the complex response coding and therefore should show an additional degradation in IT beyond that shown in the 1D roving identification experiments. Our data suggest that this effect is very limited. If the relation $IT(\text{multi}D) = \sum IT_{\text{roving}}(1D)$ can be shown generally true, one could have a more efficient way of estimating multi-dimensional IT (i.e., by performing only the 1D roving identification experiments).

The results of our temporal-order discrimination experiments showed that the subjects were able to discriminate different temporal orders of movements when the onset-time of a pair of movements was separated by 3–6 ms (with fixed displacements). This JND is very small compared to that obtained by Hirsh and Sherrick [6], which was roughly 20 msec. The big difference between our data and theirs can be accounted for by the psychophysical method employed in the two studies. In our temporal order discrimination test, trial-by-trial feedback was given. It was reported by both subjects who participated in our temporal order discrimination experiment that once they were familiar with the stimuli, they were able to respond without feeling exactly which of the movements was delayed. Instead, they made their judgments according to the “overall rhythm” and/or “overall trajectory” of the movements. In the study by Hirsh and Sherrick, no feedback was given; thus, it is likely that subjects based their judgments on feeling exactly which action was delayed. In other words, when feedback was provided, a subject can easily find a trick to base the judgment on the difference between the integrated movements, which might be easier than feeling exactly which movement was delayed. This explains why our JND is much smaller than that obtained in the classical study

of Hirsh and Sherrick. The idea that two tactile patterns that are presented in close temporal and spatial proximity are integrated into a composite representation upon which subjects base their decisions is suggested by many researchers (see [4]). The fact that a subject's ability to discriminate between different temporal orders of two movements was movement-dependent (our JNDs span a range of 3–6 msec) suggested that the distinguishability of the two integrated representations of the two temporal orders was also movement-dependent.

Further temporal resolution experiments could be performed to explore the role feedback plays in such tests. As a natural extension, sequential discrimination and identification experiments can be performed using longer sequence of actions. With the oral airflow and laryngeal vibration systems installed on the synthetic Tadoma system, it will also be interesting to perform discrimination and identification tests on these two parameters, as well as 6-D identification experiments to explore the ultimate limit of IT the synthetic Tadoma system is capable of transmitting. Further work concerns the generality of the hypothesis $IT(\text{multi}D) = \sum IT_{\text{roving}}(1D)$. More experiments need to be done to examine the above relation in the tactile domain as well as other sensory modalities like audition. We expect to develop a model which is capable of predicting multi-dimensional identification results given data from discrimination as well as 1D identification (with a fixed or roving background) experiments.

Appendix A

On Decision Model and Information Theory

A.1 d' and JND

Our purpose here is to introduce the “sensitivity index” d' and the just-noticeable-difference (JND) in the context of 2I-2AFC (two intervals, two-alternative-forced-choice) experiments. The following notes are based on [2].

A.1.1 A Model for 2I-2AFC Experiments

A 2I-2AFC experiment is one in which:

- There are two admissible signal sources S_1 and S_2 and two admissible presentations, $U_1 = (S_2, S_1)$ and $U_2 = (S_1, S_2)$, each of which is a temporally ordered pair;
- There are two admissible responses R_1 and R_2 ;
- On each trial, the experimenter presents U_1 or U_2 randomly with a priori probabilities $P(U_1) = P(U_2) = 0.5$
- The subject is instructed to respond R_1 when U_1 is presented and R_2 when U_2 is presented.

The model for a 2I-2AFC experiment is:

- There is a probabilistic transition mechanism between stimuli and their internal presentation — a real variable X with known Gaussian densities $p(X/S_i)$, $i=1$ or 2 . The mean of $p(X/S_i)$ is M_i , and the variance is σ^2 for both densities;

- Each stimulus presentation (of the form U_1 or U_2) determines an ordered pair of values (X_1, X_2) , where X_1 is the value of X determined from the first member of the pair (S_2 in U_1 and S_1 in U_2), X_2 is the value of X determined from the second member of the pair (S_1 in U_1 and S_2 in U_2). X_1 and X_2 are statistically independent;
- The subject observes the pair (X_1, X_2) and forms the decision variable $Y = X_2 - X_1$;
- There exists a fixed cut-off value C (the “criterion”) on the Y axis. The subject responds R_1 if and only if $Y < C$ and R_2 if and only if $Y \geq C$;
- The statistics of X are independent of all aspects of the experiment except S_1 and S_2 .

Figure A.1.1 provides a schematic illustration of the model for $M_2 > M_1$.

The two conditional probabilities $p(Y/U_i)$ ($i=1$ or 2) are also Gaussian densities with mean \overline{M}_i and variance σ being:

$$\begin{aligned}\overline{M}_1 &= M_1 - M_2 \\ \overline{M}_2 &= M_2 - M_1 \\ \sigma &= 2\sigma^2\end{aligned}\tag{A.1}$$

From Equation A.1, $\overline{M}_2 = -\overline{M}_1$. The conditional probabilities $P(R_2/U_2)$ and $P(R_2/U_1)$ depend on C , σ , \overline{M}_1 and \overline{M}_2 . If we assume $\sigma = 1$ we could uniquely recover C and \overline{M}_i , given the conditional probabilities $P(R_2/U_i)$. An estimate of $P(R_2/U_i)$ can be obtained from experimental data $f(R_2/U_i)$, the conditional frequencies that subject responded R_2 when presented U_i .

A.1.2 Definition of d' and JND

The so-called “sensitivity index” d' is defined as $d' = (M_2 - M_1)/\sigma$, which is independent of the criterion C and measures the separation of the conditional

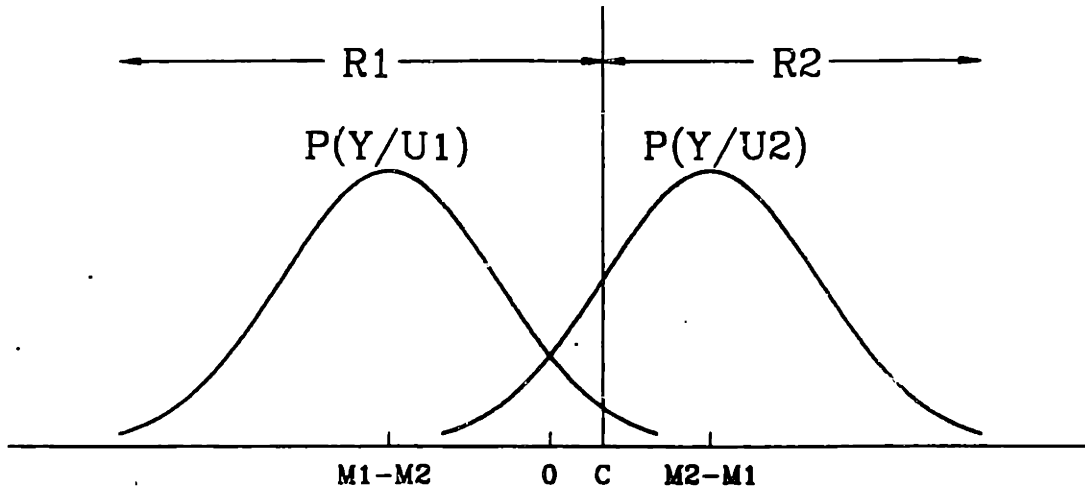


Figure A.1: A schematic illustration of the model.

probability densities $p(X/S_1)$ and $p(X/S_2)$. The d' depends on the nature of the stimuli. It's the distinguishability of S_1 and S_2 .

Note that directly from the experimental data, we can only compute $\bar{d}' = (\bar{M}_2 - \bar{M}_1)/\bar{\sigma}$ which is associated with the probability densities $p(Y/U_1)$ and $p(Y/U_2)$. However, from Equation A.1 we see that a simple relation exists between d' and \bar{d}' .

$$\bar{d}' = \sqrt{2}d' \quad (\text{A.2})$$

The JND is defined as a parameter difference¹ for stimuli S_1 and S_2 such that $d'=1$. When $\sigma = 1$ and $C=0$, JND corresponds to the point where the conditional probability $P(R_2/U_2)$ (also called the detection probability) is 70%, which has been used as a criterion point by many people.

A.1.3 Computation of d' and JND

In our experiments, we had a program which computed d' ($d' = \bar{d}'/\sqrt{2}$) given the experimental data $f(R_2/U_2)$ and $f(R_2/U_1)$. We tried to adjust the stimuli

¹JND could be a frequency difference when S_1 and S_2 are two tones at different frequencies.

parameters such that d' was within the range of 0.5–2.6 which corresponds to the case when the two probability densities $p(Y/U_1)$ and $p(Y/U_2)$ were neither too close nor too apart. The experiment was thus neither too hard nor too easy for the subject.

According to many studies (e.g., [3]), $d' = k \cdot \Delta$, where k is a constant and Δ is the parameter difference for stimuli. From the definition for JND, we have $JND=1/k$. In our experiments, we usually collected three d' data points under each condition, computed the best-fitting k (slope of a straight line that best fits the data) and then JND.

A.2 IT

The word IT stands for information transfer (also called mutual information between stimuli and responses).

A.2.1 Information and Uncertainty

Information is something we get when some person or machine tells us 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. Thus information occurs only if there exists some a priori uncertainty, and the *amount* of information is determined by the *amount* by which the uncertainty has been reduced.

The measure for uncertainty is defined as $U = \log_2 k$, where k is the number of categories or possible outcomes. This measure satisfies the two conditions that (a) it is monotonically related to the number of possible outcomes and, (b) each successive event adds the same amount of uncertainty and thus makes available the same amount of information. This elementary definition of information assumes that all of the k outcomes are equally likely.

In the case of a univariate distribution, the *average* uncertainty associated with

a discrete probability distribution $P(x)$ is a weighted sum given by

$$U(x) = - \sum P(x) \log_2 P(x) \quad (\text{A.3})$$

This is also called a Shannon measure of average information.

A.2.2 Information Transfer in an Identification Experiment

In a typical identification experiment, a series of stimuli is presented, one stimuli at a time in random order. The subject makes an absolute (categorical) response of some predesignated type to each stimulus. A confusion matrix (a stimulus-response matrix) is then formed with stimuli S_i as the rows and responses R_j as columns, with the cell entry indicating the number or proportion of times that each stimulus category is called each response category.

Suppose that the number of possible stimuli and responses is both n . It follows from Equation A.3 that the average uncertainty in stimuli (IS) is:

$$IS = - \sum_{i=1}^n P(S_i) \log_2 P(S_i) \quad (\text{A.4})$$

The average uncertainty in responses (IR) is:

$$IR = - \sum_{j=1}^n P(R_j) \log_2 P(R_j) \quad (\text{A.5})$$

For the discrete bivariate distribution $P(S_i, R_j)$, the *joint uncertainty* is defined as:

$$U(S_i, R_j) = - \sum_{j=1}^n \sum_{i=1}^n P(S_i, R_j) \log_2 P(S_i, R_j) \quad (\text{A.6})$$

And the *maximum joint uncertainty* is limited by the sum of IS and IR:

$$U_{max}(S_i, R_j) = IS + IR \quad (\text{A.7})$$

It is now clear that the information transfer IT, which is the amount of information received by the subject, is the amount by which the joint uncertainty for the discrete bivariate distribution $P(S_i, R_j)$ is reduced. This means that IT equals the

difference of $U_{max}(S_i, R_j)$ and $U(S_i, R_i)$. From Equations A.4, A.5, A.6 and A.7, we have

$$\begin{aligned} IT &= U_{max}(S_i, R_j) - U(S_i, R_j) \\ &= \sum_{j=1}^n \sum_{i=1}^n P(S_i, R_j) \log_2 \left(\frac{P(S_i, R_j)}{P(S_i)P(R_j)} \right) \end{aligned} \quad (A.8)$$

A maximum likelihood estimate for IT from a confusion matrix can be obtained using the following equation.

$$IT_{est} = \sum_{j=1}^n \sum_{i=1}^n \left(\frac{n_{ij}}{n} \right) \log_2 \left(\frac{n_{ij} \cdot n}{n_i \cdot n_j} \right) \quad (A.9)$$

where n_{ij} is the number of the joint event (S_i, R_j) in a sample of n trials, and $n_i = \sum_{j=1}^n n_{ij}$ and $n_j = \sum_{i=1}^n n_{ij}$. These quantities can all be derived from the confusion matrix obtained from an absolute identification experiment.

Appendix B

Estimation of Information Transfer from Limited Experimental Data

The title of this appendix is borrowed from Houtsma's paper (see [7]). We begin our discussion by briefly summarizing the key points in his paper.

The problem raised in Houtsma's paper was "how to obtain a reliable estimate of information transfer (IT) from identification data for a large set of alternative stimuli, while keeping the number of required experimental trials within the realm of reality." Recall that we used IT_{est} (see Equation A.9 in Appendix A) to estimate IT from a confusion matrix. This is a biased (but asymptotically unbiased) estimate of true IT. In order to get a good estimate of IT, a sufficiently large number of trials have to be collected. Miller (1954) demonstrated a method for computing the bias in IT_{est} for data samples in which the number of trials is at least five times the number of cells in the confusion matrix. This means that when the number of alternatives for stimuli is 100, at least 50,000 trials are needed, which is beyond reality. Houtsma proposed a practical approach to address this problem. An absolute identification experiment was simulated with varying numbers of trials and varying amounts of response noise. The result of the simulation is a group of curves of estimated IT as a function of number of trials. The results of an actual identification experiment (with limited number of trials) were then shown to closely follow a particular simulation curve (i.e., that associated with a particular response noise). Houtsma suggested that a good estimate of asymptotic IT for the experiment could be obtained as the asymptotic IT from this nearest simulation curve.

We applied this general idea to process our data obtained from 1D and 4D absolute identification experiments. When we tried to fit our 1D and 4D data to the computer simulated curves of estimated IT as a function of number of trials, we

found that (1) we needed finer scales for response noise (Houtsma used a uniformly distributed random integer and the width of this distribution could only be integral) in order to have a close fit; (2) our 4D results deviated considerably from the simulated curve. It should be pointed here that in dealing with the data from a 3D identification experiment with 5 alternatives on each of the three channels, Houtsma simulated this test as a 1D identification experiment with number of alternatives being 125. This model might be too simplified.

We then made some modification to Houtsma's algorithm: (1) We used a Gaussian (instead of a uniform) distribution to simulate response variance, thus the computer simulated curves could be made arbitrarily close. (2) We simulated 4D identification experiments as multidimensional identification experiments.

The details of our algorithm, our principal results, and discussion of some relevant issues appear in the following three sections.

B.1 Computer Simulation Algorithm

A computer program was developed to permit simulating 1D, 2D, 3D, or 4D identification experiments. We'll give a brief outline for the algorithm in the case of simulating a 4D identification experiment.

Before each simulation, the program is given the following parameters: (1) the number of stimulus-response alternatives for each of the four dimensions (K_i , $i=1,2,3,4$); (2) the standard deviations of Gaussian (response) noise for each of the four dimensions (σ_i , $i=1,2,3,4$); (3) maximum number of trials (L_{max}) to be generated.

On each trial, for each dimension, an integer X_i ($i=1,2,3,4$) is chosen with equal probability in the range $1 \leq X_i \leq K_i$ ($i=1,2,3,4$). For each dimension, a response $Y_i = X_i + nint(G_i)$ ¹ ($i=1,2,3,4$) is then generated, where G_i is a Gaussian random variable with standard deviation σ_i . Note that since the values of G_i are chosen independently, we are inherently assuming independence of the four dimensions. Trials in which Y_i came out larger than K_i or smaller than 1 were dealt with by

¹ $nint(G_i)$ means the nearest integer to G_i .

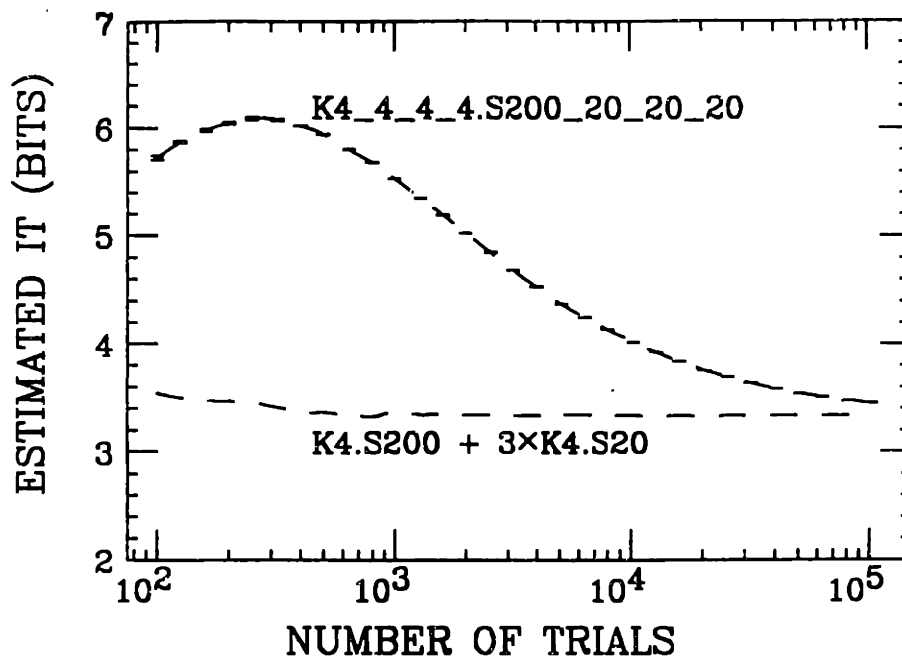


Figure B.1: An example of computer simulation.

assigning Y_i either K_i or 1. The above obtained X_i and Y_i ($i=1,2,3,4$) were encoded into a confusion matrix of size $K = K_1 \times K_2 \times K_3 \times K_4$ by the following equations:

$$S = X_1 + (X_2 - 1) \times K_1 + (X_3 - 1) \times K_2 \times K_1 + (X_4 - 1) \times K_3 \times K_2 \times K_1$$

$$R = Y_1 + (Y_2 - 1) \times K_1 + (Y_3 - 1) \times K_2 \times K_1 + (Y_4 - 1) \times K_3 \times K_2 \times K_1$$

where S and R refer to stimulus and response. Thus a confusion matrix of size K was generated for specified number of trials. The above process is repeated until L_{max} trials are generated.

A curve of estimated IT as a function of number of trials is then obtained by performing 16 simulations using same parameters, and averaging the results. Figure B.1 shows such a curve (the upper curve) with its mean and error bars (± 1 standard deviation) plotted. Note that the simulation exhibits a very small variance which diminishes quickly as the number of trials increases.

The notation used in Figure B.1 is typical of the many figures we shall show in this appendix. We shall always use dashed lines to plot computer simulation results

and solid lines to plot experimental results. For simulations, we label each curve with $K_m.S_{100 \times \sigma}$, where m (an integer for the 1D case and a 4-tuplet of integers for the 4D case) is the size of the confusion matrix and σ (a real number for the 1D case and a 4-tuplet of real numbers for the 4D case) is the standard deviation of the Gaussian distribution. Thus the label “K4.4.4.4.S200.20.20.20” used for the upper curve plotted in Figure B.1 means that this curve was generated by simulating a 4D identification experiment with 4 alternatives on each of the four dimensions (the size of the confusion matrix was thus 256) and the standard deviation of Gaussian noise was 2, 0.2, 0.2, and 0.2, respectively. The lower curve in the figure is the sum of four 1D simulations each having 4 stimulus-response alternatives and response noises of 2, 0.2, 0.2, and 0.2, respectively. Note that the asymptotic IT (at ∞ trials) for 4D simulations such as that in Figure B.1 is not obvious. It takes more than 100,000 trials for the curve to converge to an asymptotic value. Nevertheless, because the simulation algorithm, by construction, forms the 4D response from response noises that are independent for each dimension, it follows directly that $IT(4D, \infty) = \sum_{i=1}^4 IT_i(1D, \infty)$. This is demonstrated by the two curves in Figure B.1. The 1D $IT(\infty)$ s are easy to estimate since they converge at much lower number of trials (several hundreds of trials). Hence, the $IT(4D, \infty)$ for the simulation follows directly. This procedure was used to obtain asymptotic IT estimates for the 4D results.

B.2 Results

We now show how we obtained the unbiased estimates of ITs for our experimental data. Three kinds of identification experiments were performed (see Section 3.2). The first was 1D identification with a fixed background. The second was 1D identification with a roving background. The third was 4D identification.

Recall that in the 1D identification experiments with a fixed background, we measured IT for three channels: *lio*, *lud* and *jud*. In Figures B.2, B.3, and B.4, the experimental results (solid curves) and the corresponding “closest” simulation curves (dashed curves) are shown for the two subjects HZT and JT on each of the

three channels. In 1-D identification experiments with a roving background, we measured the IT for uio, lio, lud and jud channels. These data and the corresponding “closest” simulation curves are shown in Figures B.5 and B.6. The “closest” simulation curves were chosen by eye, giving greatest weight to the form of the curve of empirically estimated IT associated with large values of the number of trials (i.e., we assumed that the early portion of the data was most noisy in terms of IT). The IT of the simulation curve at 10,000 trials was then taken as the asymptotic IT estimate for the experimental results. Fits are generally excellent, especially after 300 trials.

For the 4D identification experiments, a 4D simulation was performed. These results, shown in Figure B.7, are more deviant from the experimental data. These curves are, however, the best fits we were able to find in the following three respects: First, our 1D simulations with response noise adjusted to produce the same asymptotic value of IT at 100,000 as that of the 4D simulations are steeper than the 4D simulations. Thus the 1D simulations are more deviant from the empirical data and, hence, are inferior to the 4D simulations. Second, since we know from our 1D identification experiments with a roving background that the lud channel provides little information, it seems natural to use a set of $\sigma_1 - \sigma_4$ with one value much larger than the other three. Simulations conducted with such parameters were found to be slightly steeper than those obtained from using uniform σ_i ($i=1,2,3,4$) and, therefore, were more deviant from the experimental results. Third, we tried using different sets of $\sigma_1 - \sigma_4$ with a variety of modest ratios among the noise components. The resulting curves differed very little from each other. As a result of all these tests, the simulation curves shown in Figure B.7 seem to “best fit” the experimental results and yield the asymptotic IT values of 3.8 and 3.1 bits that are given in the main body of the thesis.

B.3 Discussion

As mentioned in the beginning of Section 3.2, the amount of information transfer depends not only on the number of dimensions, but on how the different dimensions

1-D IDENTIFICATION ON LIO

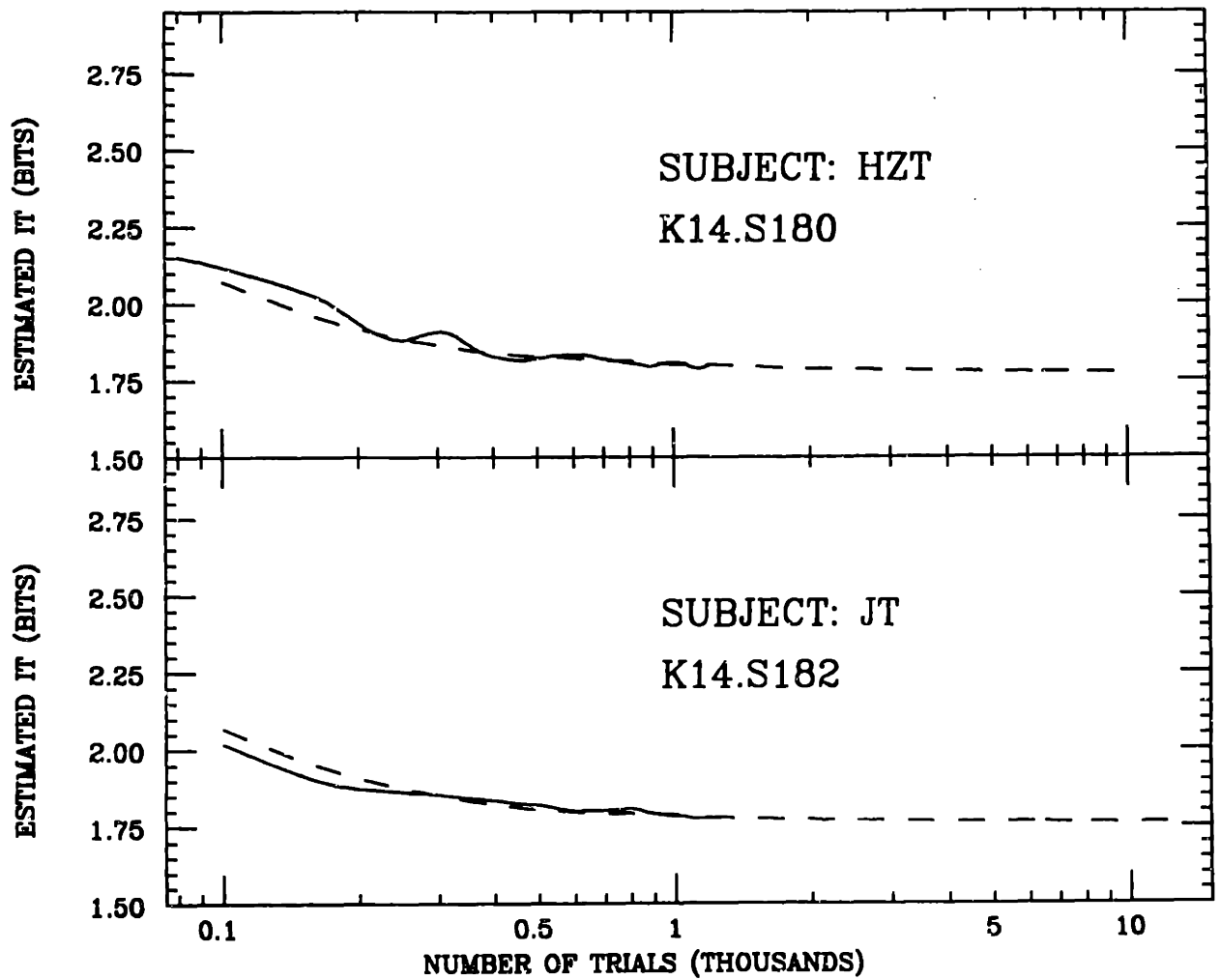


Figure B.2: Data processing for 1D identification tests on lio.

1-D IDENTIFICATION ON LUD

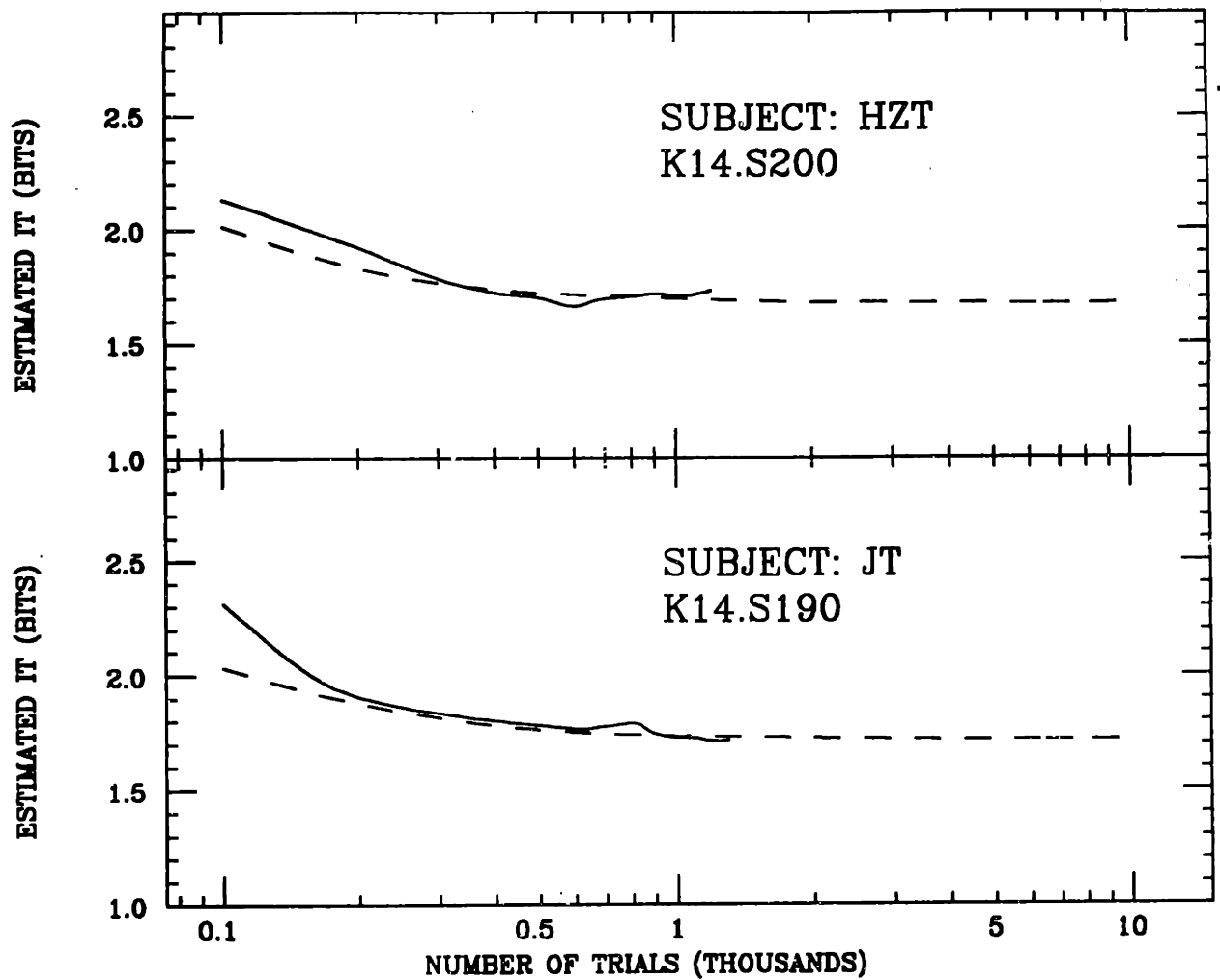


Figure B.3: Data processing for 1D identification tests on lud.

1-D IDENTIFICATION ON JUD

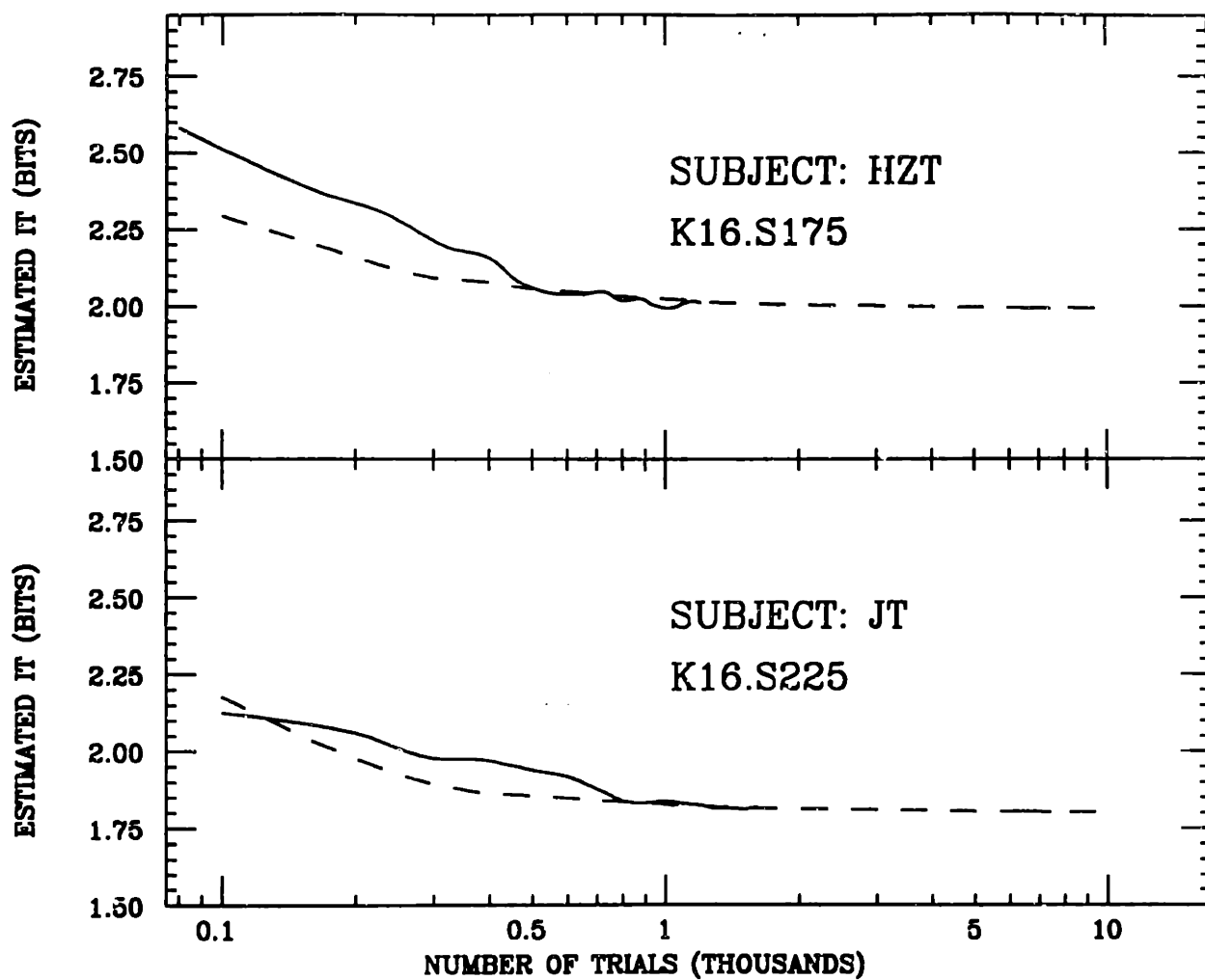


Figure B.4: Data processing for 1D identification tests on jud.

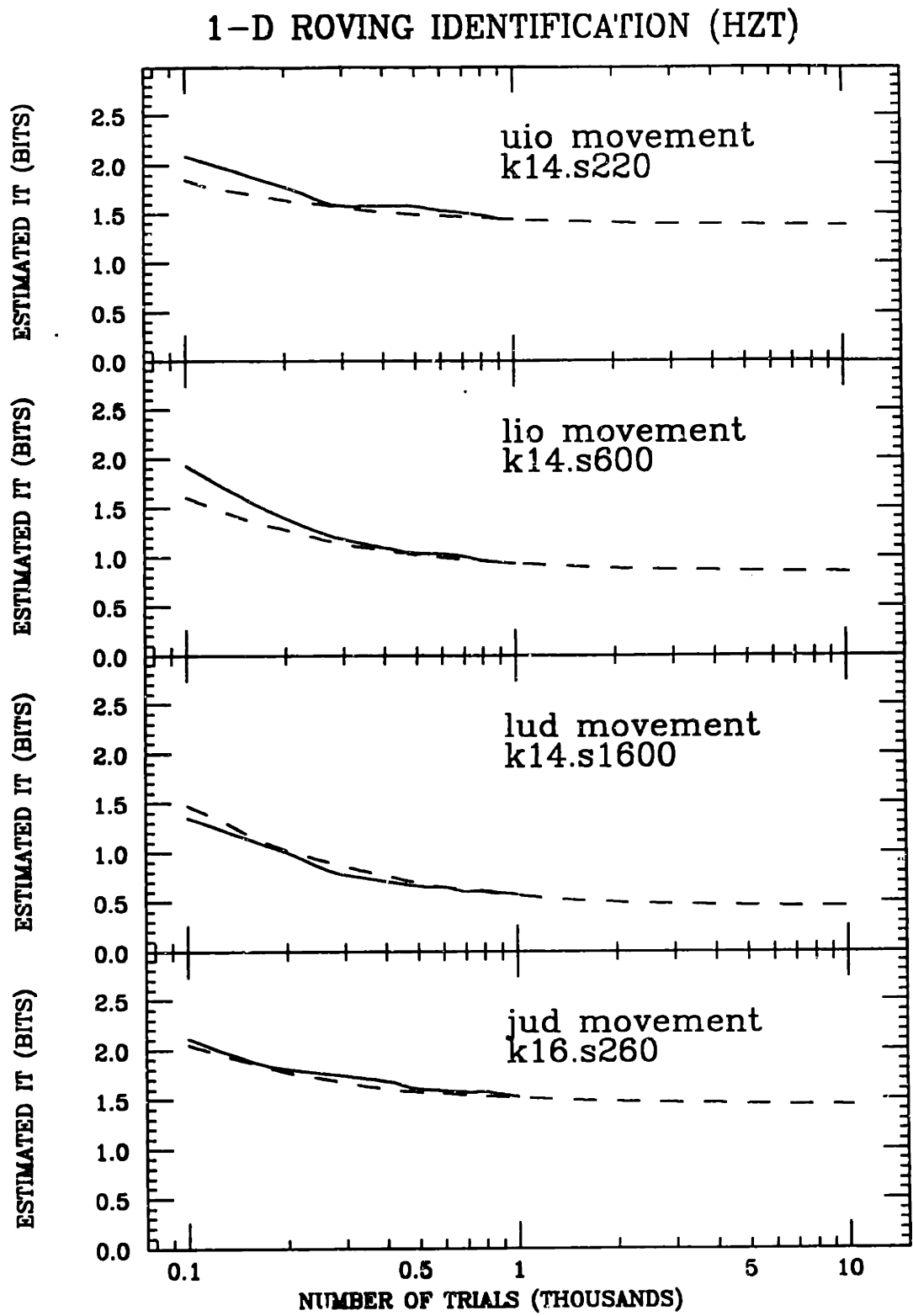


Figure B.5: Data processing for 1D roving identification tests (HZT).

1-D ROVING IDENTIFICATION (JT)

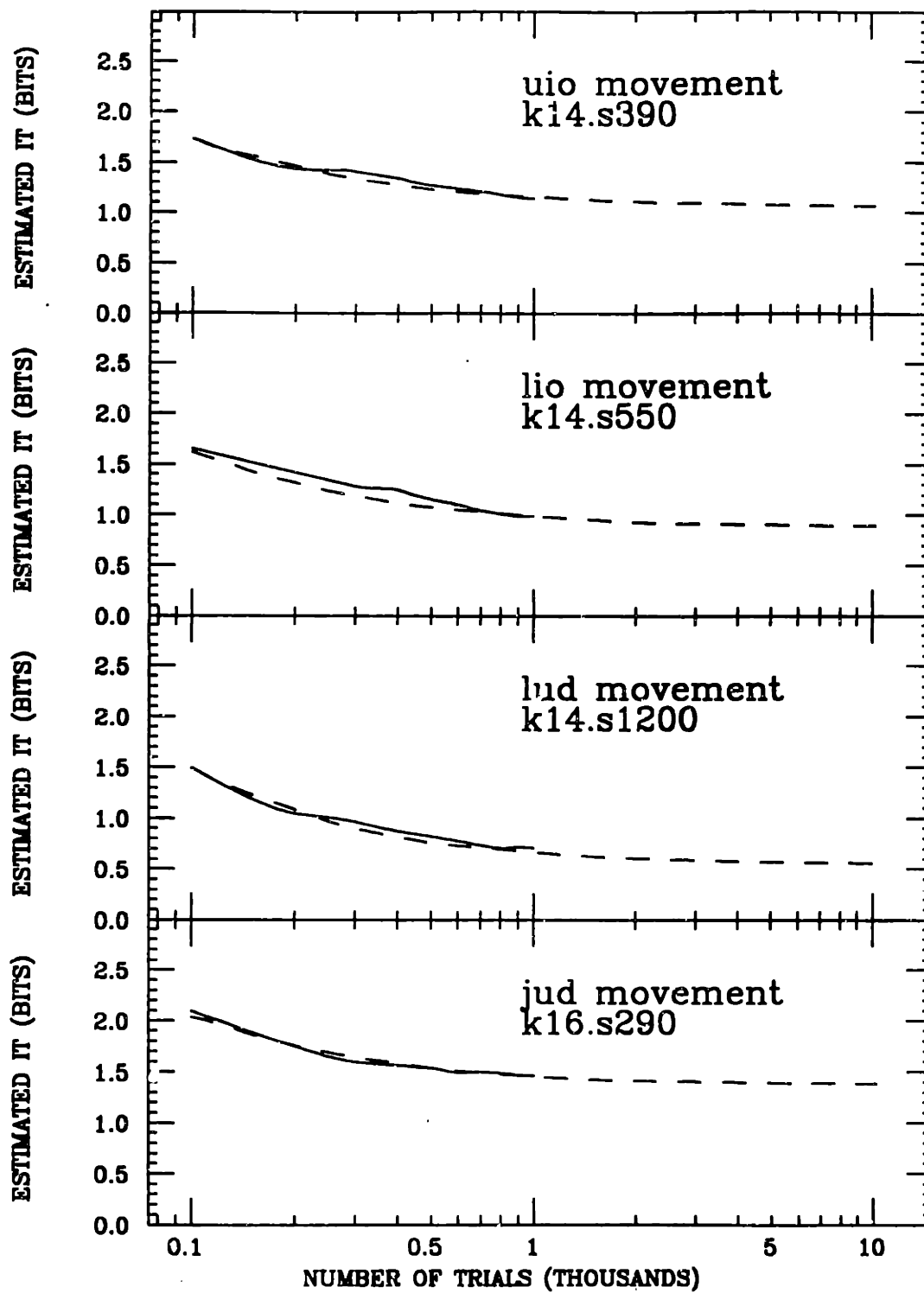


Figure B.6: Data processing for 1D roving identification tests (JT).

4-D IDENTIFICATION EXPERIMENTS

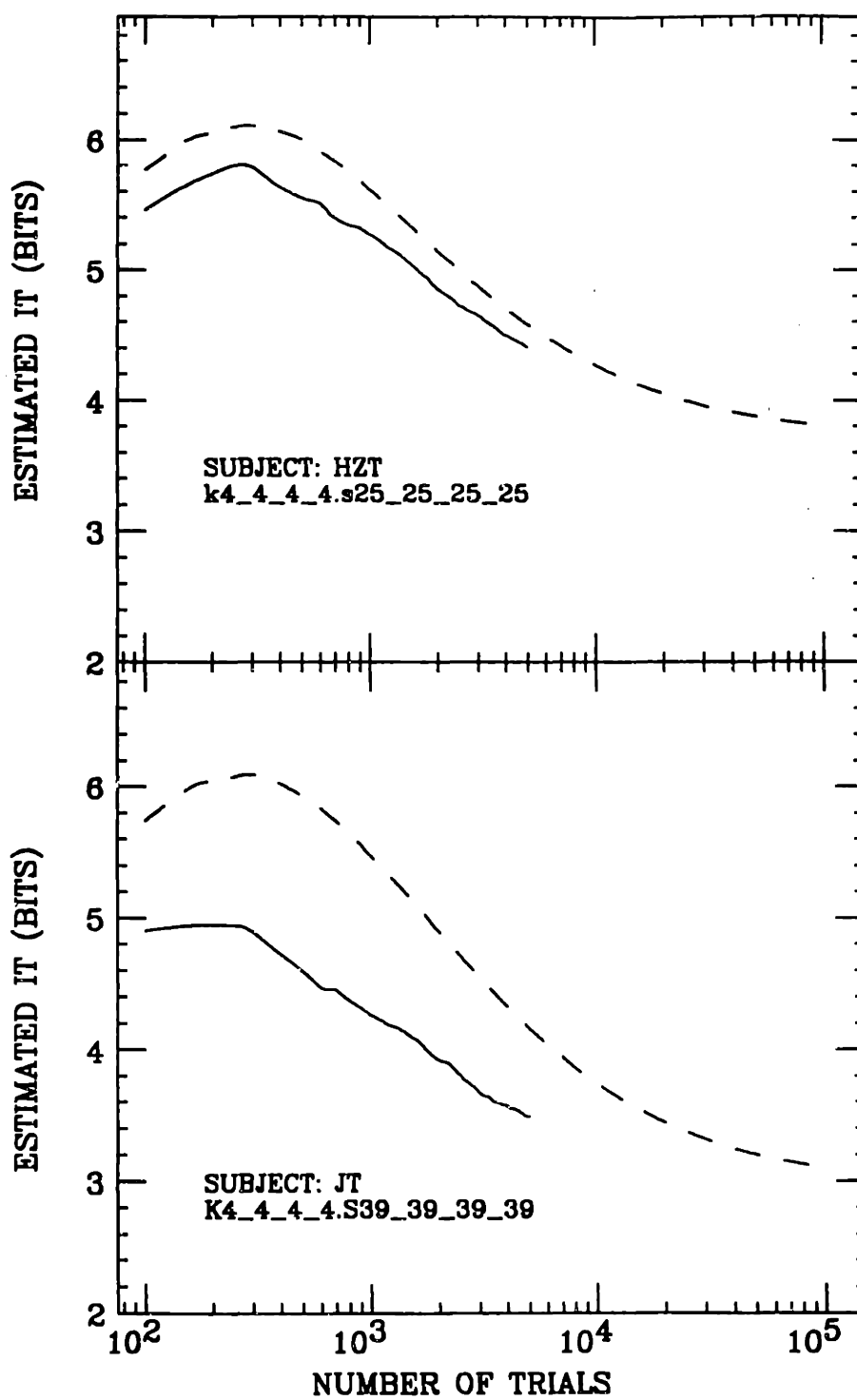


Figure B.7: Results of 4D identification experiments.

interact. When there are no interactions between different dimensions, these dimensions are said to be “orthogonal” and the ITs for each individual dimension will sum up to the IT for multidimensional identification experiments. In many cases, the “orthogonality” condition is not satisfied and multidimensional identification experiments must be performed. Depending on the number of dimensions and stimulus-response alternatives, multidimensional identification experiments often require that more than several thousands of trials be collected in order to obtain an unbiased estimate of IT.

It is interesting to consider whether some basic relations exist between 1D and 4D identification data. If so, we might then find a way to predict multidimensional identification IT from some 1D ITs which require much fewer trials for IT estimates. Some qualitative observations can be made regarding our data for 1D and 4D identification experiments.

For the four channels uio, lio, lud and jud, the sum of the ITs obtained from 1D identification experiments with a fixed background (7 bits, averaged over the two subjects) is greater than the bias-corrected IT obtained from 4D identification experiments (3.5 bits). However, the sum of the ITs obtained from 1D identification experiments with a roving background (4 bits) is close to the 4D IT. A further IT estimate for the 4D data can be obtained by transforming the 256×256 confusion matrix into 4×4 confusion matrices for each of the four channels by collapsing over the responses to the other three dimensions (see [14]). The sum of the ITs thus computed (2.6 bits) is smaller than the 4D IT. These results are summarized in Figure B.8 and lead us to suggest that 1D identification experiments with a roving background might predict 4D IT very closely. More work needs to be done before this conclusion can be generalized.

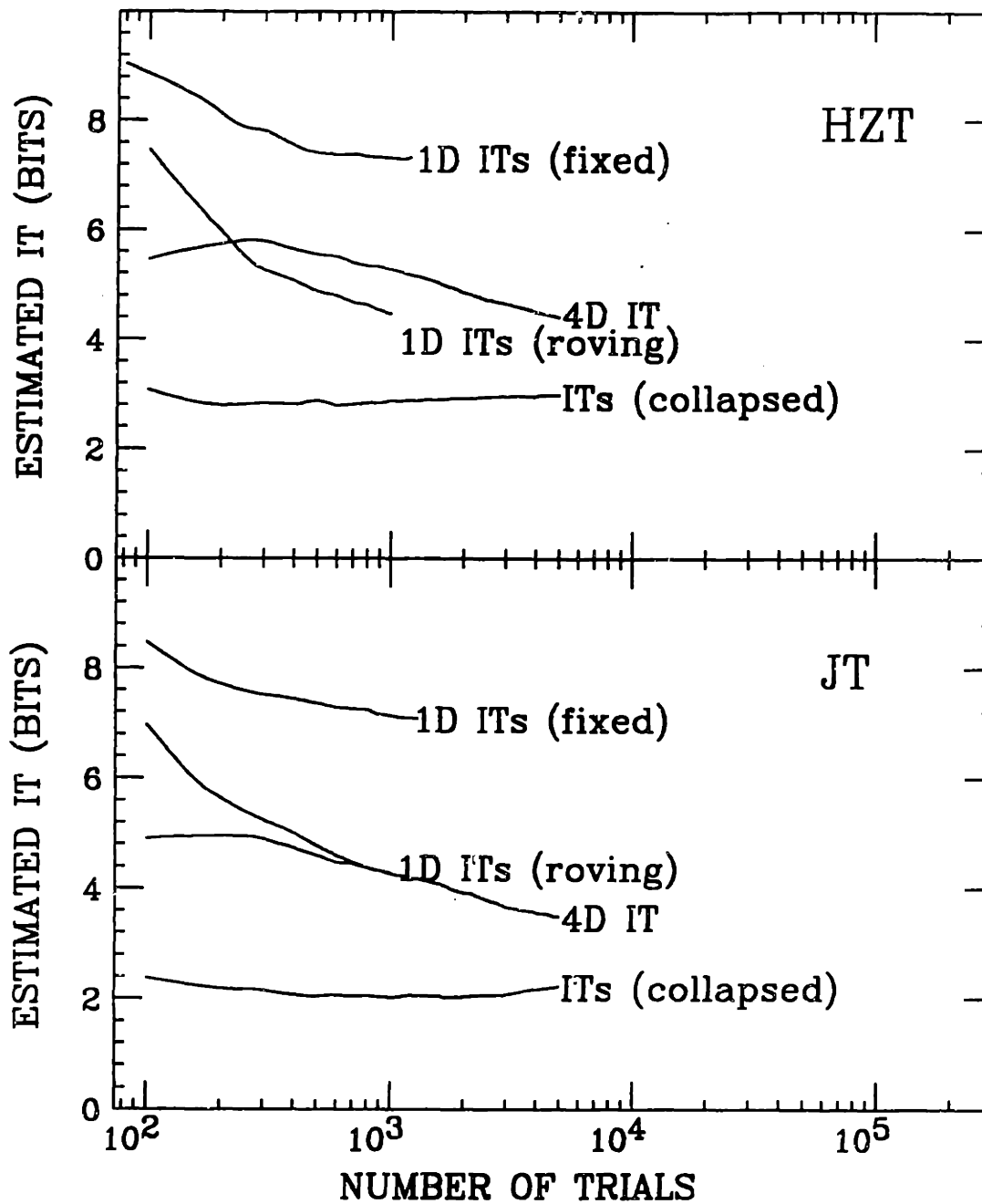


Figure B.8: Relation of ITs from 1D and 4D identification experiments.

Bibliography

- [1] Louis D. Braida and Nathaniel I. Durlach. Intensity perception. II. Resolution in one-interval paradigms. *The Journal of the Acoustical Society of America*, 51(2):483–502, 1972.
- [2] Nathaniel I. Durlach. A decision model for Psychophysics. (October 1968. available at CBG.
- [3] N.I. Durlach, L.A. Delhorne, A. Wong, W.Y. Ko, W.M. Rabinowitz, and J. Hollerbach. Manual discrimination and identification of length by the finger-span method. 1988. in press.
- [4] Paul M. Evans. Vibrotactile masking: Temporal integration, persistence, and strengths of representations. *Perception and Psychophysics*, 42(6):515–525, 1987.
- [5] Wendell R. Garner. *Uncertainty and Structure as Psychological Concepts*. John Wiley and Sons, Inc., New York, 1962. Chap.1–4.
- [6] Ira J. Hirsh and Carl E. Sherrick, Jr. Perceived order in different sense modalities. *Journal of Experimental Psychology*, 62(5):423–432, November 1961.
- [7] Adrianus J.M. Houtsma. Estimation of mutual information from limited experimental data. *The Journal of the Acoustical Society of America*, 74(5):1626–1629, November 1983.
- [8] Daniel Francis Leotta. *Synthetic Tadoma: Tactile Speech Reception from a Computer-Controlled Artificial Face*. Master's thesis, EECS, MIT, February 1985.
- [9] George A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63(2):81–97, March 1956.

- [10] Susan J. Norton, Martin C. Schultz, Charlotte M. Reed, Louis D. Braida, Nathaniel I. Durlach, William M. Rabinowitz, and Carol Chomsky. Analytic study of the Tadoma method: Background and preliminary results. *Journal of Speech and Hearing Research*, 20(3):574-595, September 1977.
- [11] Irwin Pollack. Multidimensional coding within the temporal microstructure of auditory displays. *The Journal of the Acoustical Society of America*, 53(2):420-428, 1973.
- [12] Irwin Pollack. Multidimensional encoding within the temporal microstructure of auditory displays II. *The Journal of the Acoustical Society of America*, 54(1):22-28, 1973.
- [13] Irwin Pollack and Lawrence Ficks. Information of elementary multidimensional auditory displays. *The Journal of the Acoustical Society of America*, 26(2):155-158, March 1954.
- [14] W.M. Rabinowitz, A.J.M. Houtsma, N.I. Durlach, and L.A. Delhorne. Multidimensional tactile displays: identification of vibratory intensity, frequency, and contactor area. *The Journal of the Acoustical Society of America*, 82(4):1243-1252, October 1987.
- [15] Charlotte M. Reed, Nathaniel I. Durlach, and Louis D. Braida. *Research on Tactile Communication of Speech: a Review*. Monograph 20, American Speech-Language-Hearing Association, May 1982.
- [16] Charlotte M. Reed, William M. Rabinowitz, Nathaniel I. Durlach, Louis D. Braida, S. Conway-Fithian, and M.C. Schultz. Research on the Tadoma method of speech communication. *The Journal of the Acoustical Society of America*, 77(1):247-257, January 1985.