For Robin, Marjorie, and Rebecca. Their support and love is cherished.

# Tactile Perception and the Information Processing Basis of Tactile Speech Prostheses for the Deaf.

*Patrick*

Douglas P. Mahar   BA(Hons)

Department of Psychology

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

*1991*
*conferred.*

**Statement**

    I hereby declare that the research reported here is my own original work which has not been submitted for any other degree or diploma. To the best of my knowledge all contributions to the body of this work which are not my own have been appropriately cited and acknowledged.

. . . . . . . . . . . . . . . . .

Douglas P. Mahar

## Abstract

The research reported in this thesis took as its starting point the question of whether touch has the information processing characteristics required to deal with speech transforms. A review of the current status of tactile speech prostheses and of the range of potential limitations to touch's ability to deal with speech transforms, lead to the identification of two specific research foci.

First, it was argued that an appropriate strategy for tactile aid development is to establish where tactile and auditory processes overlap and where they differ, then to take advantage of any similarities, or compensate for any differences, between these modalities in the design of the prosthesis. In line with this argument, a series of experiments was undertaken to test previous suggestions that there is an underlying similarity between auditory and tactile representations of stimuli. In support of these claims, it was found that auditory and tactile versions of patterns are easier to compare than are auditory and visual versions of those patterns. Subsequent research revealed that one aspect of this processing affinity between touch and hearing is that both modalities, unlike vision, process temporally distributed information more efficiently than spatially distributed information. This finding has broader theoretical significance in view of the current controversy regarding the division of senses according to a spatial vs temporal criterion.

The second research focus addressed was whether touch

has the spatial and temporal acuity required to deal with
speech transforms. It was argued that the limiting factors in
tactile spatial and temporal acuity were more likely to occur
at the higher level of touch's ability to deal with the strong
interactive effects between pattern elements, rather than at
the lower level of two-point thresholds in time and space. As
masking is a primary interactive force between tactile pattern
elements, an attempt was made to resolve the ongoing debate
regarding the extent to which tactile masking effects either
limit the perception of complex tactile patterns by obscuring
the identity of pattern elements or facilitate this task via a
process of perceptual integration.

This question was investigated by measuring the
discriminability of three-element tactile patterns as the
spatial and temporal separation, and hence the level of
masking, between pattern elements was varied. It was expected
that performance would be best at closer element spacings, due
to the greater opportunity for perceptual integration to occur.
Contrary to this prediction, it was found that the increasing
levels of masking induced by decreasing the spatial and
temporal separation between pattern elements caused a decrease
in the discriminability of the patterns.

One caveat to the acceptance of this result was the
possibility that training may be required before touch can take
advantage of any beneficial interactions between pattern
elements, a possibility supported by the anecdotal reports of
the subjects. Tentative support for this suggestion was

provided by a pattern learning experiment involving three subjects from the previous experiment. After brief experience with closely spaced tactile patterns, the subjects were able to discriminate these stimuli at least as well as widely spaced tactile patterns.

While failing to demonstrate the proposed beneficial effects of integration, these results did indicate that close spatial and temporal proximity between tactile pattern elements may not adversely affect the discriminability of those patterns. If subsequent research confirms this tentative finding, then the implication for tactile speech prostheses is that the display employed need not avoid the strong masking effects induced by close spatial and temporal proximity between speech pattern elements.

In summary, this thesis showed that there is a general similarity between auditory and tactile perceptual representations of patterns which may both assist in the implementation of tactile speech prostheses, and advantage touch over vision for this purpose. Second, it appears that although tactile pattern perception is initially impeded by the occurrence of masking effects between pattern elements, this performance deficit may be removed once the observers have sufficient experience with the stimuli. There is, however, a clear need for further research before this tentative conclusion can be confirmed.
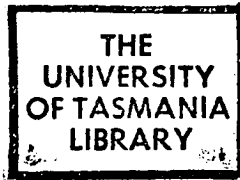
## Acknowledgement

Thanks to my supervisor of the past three years, Dr. Brian Mackenzie, to my initial supervisor, Professor Don McNicol, and to my associate supervisor, Dr. George Wilson, whose encouragement and assistance underlie this research. Thanks also to Mr. Peter Tattam for his assistance in the development of the computer software used in this research, as well as to the technical staff of the Psychology and Physics Department workshops for their assistance in the construction and maintenance of experimental apparatus. Finally, thanks to the many people who gave their time and enthusiasm as subjects.

# Table of Contents

Chapter 1: A Review of the Current Status of Tactile Speech
Prostheses.

The research reported here addresses several issues in the field of tactile information processing which are relevant to the development of tactile speech prostheses for the profoundly deaf. While the emphasis upon tactile prostheses guides the choice of research problems, this research maintains strong links with broader issues in the field of tactile perception.

This first chapter provides an overview of the rationale underlying tactile speech prostheses and a review of their current status. This review is intentionally selective, focusing on several of the most successful recent devices which characterize the various strategies which have been adopted. Following this review, several possible areas for further research are considered, with the conclusion that most emphasis should be placed on establishing whether touch has the information processing properties required to deal with transforms of speech stimuli.

Chapter 2 reviews this question at length, identifying two aspects of tactile information processing which must be examined on the path to the development of optimal tactile transforms of speech. The first of these is the extent to which auditory and tactile perceptual processes are alike or compatible, while the second is the capacity of touch to deal with the very complex patterns of stimulation associated with speech transforms.

Chapters 3 and 4 investigate the first of these, that is the extent to which auditory, visual, and tactile

representations of speech-like stimuli are alike, along with the related possibility that fundamental differences exist between these modalities in terms of the relative efficiency with which each processes spatially and temporally distributed information. Chapters 5 and 6 address the second aspect of tactile information processing raised in Chapter 2, that is the ability of touch to deal with complex patterns of stimulation. This issue is addressed with particular reference to the extent to which masking and integrative effects between tactile pattern elements either impede or facilitate the perception of those patterns.

### 1.1: The Principles Behind Tactile Aids.

Individuals with hearing losses greater than 90 dB effectively have no access to acoustic information. As a result, they have difficulties in the areas of speech perception, speech production, and the perception of acoustic events in the environment. Ideally, a remedial strategy should address all of these consequences of profound hearing loss.

Traditional approaches like air or bone conduction hearing aids, lipreading, and sign-language do not adequately meet these needs. First, as many as 60% of those suffering a profound hearing loss report that they receive minimal benefit from the fitment of conventional hearing aids (Lind, 1973; cited in Risberg, 1978). Second, contrary to popular perception, lipreading gives the user only limited access to the features necessary to understand speech. Particularly,

while lipreading provides information about the place of articulation of segmental speech features, it does not allow either prosodic features, voicing, or the manner of articulation to be identified (Risberg and Lubker, 1978; Woodward and Barber, 1960 ). Further, lipreading provides no cues to aid in either speech production or the perception of acoustic events in the environment. Finally, although sign-language allows relatively rapid rates of both receptive and expressive communication, it is only of use amongst the small population of trained users and does not assist the perception of acoustic environmental events.

Since the founding work of Gault (1924,1925,1930) the objective of tactile aid research has been to provide a device which, perhaps in conjunction with these traditional approaches, adequately meets the communicative and perceptual needs of the profoundly deaf. All tactile aids attempt to isolate salient features of the acoustic environment then to present an interpretable tactile analog of these features to the user. The acoustic signal is usually partitioned according to a spectral algorithm which extracts information about features such as the signal's amplitude envelope, its fundamental frequency, and the amplitude and frequency of some of its harmonic components. As the auditory system identifies sounds through just such a system of spectral analysis (Békésy, 1960), this approach can be seen as an attempt to mimic the

normal function of the ear on the surface of the skin[1]. Although the current generation of devices present this information through different types of transducers (for example either mechanical vibrators or electrodes), to different sites on the skin (for example to the fingers or forearm), most share this style of coding strategy and differ primarily in terms of the range of acoustic features encoded.

## 1.2: Varieties of Tactile Prostheses

In their review of tactile communication systems Weisenberger and Miller (1987) categorized auditory perceptual tasks along a continuum based upon the cumulative level of information required to perform each task. Their scale ranged from the simple detection of sounds, which requires only information about the presence or absence of acoustic energy, through the identification of environmental sounds and prosodic features of speech, which can be achieved with the addition of amplitude envelope information, to the complex tasks of word and connected discourse identification, which also require the addition of information about fine spectral characteristics of the signal. As there are marked differences amongst tactile aids in terms of the range of acoustic features encoded by the particular device, this hierarchy of perceptual tasks provides a useful framework from which to compare the current range of tactile communication systems.

---

1. A detailed analysis of the rationale underlying this strategy is presented in Section 2.3.

The simplest tactile devices only present information about the amplitude envelope of the acoustic signal via a single vibrator. These include devices such as the commercially available Minivib (AB Special Instruments) and Tactaid I (Audiological Engineering Corporation). Further up the scale of complexity are devices like the Minifonator (Siemens Hearing Instruments), Grant's electrotactile aid (Grant, Ardell, Kukl, and Sparks, 1986), and Plant's Sentiphone based aid (Plant, 1986). These intermediate level devices present either fundamental frequency information alone, or both amplitude envelope and fundamental frequency information, to the user. For example, Plant's (1986) Sentiphone based device modulates the amplitude and frequency of the signal presented to a single vibrator, while Grant's electrotactile aid (Grant et al., 1985, 1986) presents fundamental frequency information as the site of stimulation along a 10 element array of electrodes.

Attempts to present more complete speech-transforms to the skin have extended the strategy used in these simple aids so that information about each of a number of discrete bandwidths within the acoustic signal is presented to a specific transducer within a spatial array of transducers. In a sense, these multi-channel displays can be viewed as simple models of the basilar membrane, with particular frequency components being presented at specific sites on the skin.

Two current multi-channel devices are the Queen's University Tactile Vocoder (Brooks and Frost, 1983) and the Tickle Talker (Blamey and Clark, 1985; 1987). The Vocoder uses

6

an array of 1/3 octave band-pass filters to partition the acoustic signal into 18 channels with center frequencies ranging from 160 to 8000 Hz. The output of these channels is presented as an amplitude modulated 100 Hz signal to one of 16 vibrators (two of the 18 filter channels are combined) mounted along the forearm. These vibrators are arranged so that the output of low frequency channels is presented towards the wrist and that of high frequency channels towards the elbow.

The Tickle Talker differs from the Vocoder in that it employs electrotactile, rather than vibrotactile, transducers. These are affixed to the digital nerve bundles on the side of each finger of one hand. The coding strategy used with this device is derived from that successfully used with the Nucleus Ltd. multi-channel cochlear implant (Tong et al., 1983). The fundamental frequency of the signal is represented as the pulse rate of the electrical pulses presented to the 8 electrodes, the amplitude envelope of the signal is represented as the pulse width, while the frequency of the second formant is indicated by the particular electrode stimulated.

A final example of the multi-channel genre worthy of mention is the Optacon (Telesensory Systems Inc.). This device was originally designed for use as a reading aid for the blind (Bliss et al., 1970), but has subsequently been widely employed as a general purpose transducer in tactile perception research. The Optacon consists of a 24 by 6 matrix of "pins", each connected to a separate piezoelectric transducer. Spectral representations of spoken stimuli can be generated across this

matrix of pins by, for example, using each row of pins to represent a particular frequency bandwidth, with the number of pins activated within that row representing the amplitude of that bandwidth.

## 1.3: The Utility of Tactile Prostheses

As might be expected, the success which has been achieved on Weisenberger and Miller's (1987) hierarchy of auditory tasks with these various tactile aids correlates with the hierarchical level of information presented in each display. The following sub-sections review the performance achieved with some current prostheses at each level along this continuum.

### 1.3.1: Sound Detection and Identification

Success has been achieved with all levels of tactile device on lower level tasks such as the identification of environmental sounds and the perception of rhythmic properties of sounds. Participants in early trials of the Minivib reported substantial subjective improvement in their ability to perceive environmental sounds (Spens and Plant, 1983). In a more extensive investigation Weisenberger and Russell (1989) trained subjects to identify 20 common environmental sounds via the Minifonator. After 18 hours of training all subjects were able to identify the sounds more than ninety percent of the time. By comparison, Brooks and Frost (1986) trained a profoundly deaf subject on a list of 50 environmental sounds using their multi-channel Vocoder. After only 12 hours of training the

8

subject reached an overall criterion level of 80% correct.

Subjects using only the amplitude-envelope information available via the Minivib have succeeded in identifying the stressed syllable within two-syllable words at well above chance levels (Mahar, 1985). More recently, Bernstein, Eberhardt, and Demorest (1989) have also demonstrated high levels of stress and intonation identification performance using a fundamental frequency coding single channel device. Similarly, subjects using intermediate level devices have achieved accuracy scores in the 90% correct range on monosyllable, spondee, and trochee discrimination tests (Plant, 1983; Weisenberger, 1989), and have learnt to identify rhythmic changes within simple sentences (Grant et al., 1986).

The high levels of performance achieved on environmental sound identification and syllabic structure tasks with relatively simple amplitude envelope plus fundamental frequency displays suggests that these devices are adequate for the communication of information at this level. Further, it does not appear that the additional information available in multi-channel displays significantly improves performance on this type of task.

### 1.3.2: Perceiving Speech Features

At the highest levels of acoustic perception complexity, that is tasks ranging from the identification of segmental speech features to the comprehension of connected discourse, a distinction must be drawn between the ability of tactile devices to supplement lipreading performance and their capacity

to support the performance of these tasks in their own right. As was mentioned in Section 1.1, lipreading only provides information about the place of articulation of segmental speech features, while failing to distinguish either prosodic features, voicing, or manner of articulation. Clearly, in the case of prosodic speech features, even the low level amplitude envelope information provided by simple tactile devices could facilitate connected discourse perception via lipreading. However, this level of facility is far below that required of a stand-alone tactile communication system.

Current results with intermediate level devices confirm the utility of this level of device as a supplement to lipreading. Profoundly deaf subjects using Plant's Sentiphone based aid improved their consonant perception performance from an average of 47% correct using lipreading alone (L) to over 66% correct using both lipreading and the tactile aid (LT) (Plant, 1986). This improvement primarily resulted from improvements in the perception of manner and voicing features, characteristics not available via lipreading.

The benefits provided by the use of this level of device in conjunction with lipreading extend to the level of connected discourse perception. Plant (1986) found that profoundly deaf subjects' lipreading performance improved from a mean tracking rate (DeFilippo and Scott, 1978) of 33.3 words per minute (L) to 45.7 words per minute (LT). Equally impressive increases in tracking rate have also been reported for users of Grant's electrotactile aid (Grant et al., 1986), with an increase in

mean tracking rate from 49.9 words per minute (L) to 63.9 words per minute (LT).

Although the results achieved on segmental feature identification tasks with the intermediate level devices discussed above are impressive, even better results have been recorded with more complex multi-channel tactile aids. Profoundly deaf users of the Vocoder can, without the use of lipreading, classify CV and VC pairs into their phonemic category at accuracy levels of around 85% correct (Brooks, Frost, Mason, and Gibson, 1987). Further, one profoundly deaf user of this device has learnt to identify phonemes without the use of lipreading, reaching a criterion level of 80% correct with minimal training (Brooks and Frost, 1986). Promising results have also been achieved with other multi-channel devices, with high levels of vowel identification performance, though somewhat lower levels of consonant identification performance, being recorded by users of both the Tickle Talker (Blamey, Cowan, Alcantara, and Clark, 1988) and Optacon (Green, Craig, and Pisoni, 1983; Green, Craig, Wilson, and Pisoni, 1983).

The types of phonemic features identified with these three multi-channel displays included those voicing and manner of articulation characteristics which are not available via lipreading. As with the simpler envelope plus fundamental frequency displays discussed earlier, the availability of this type of information through these devices has led to significant improvements in segmental feature identification

when used in conjunction with lipreading. For example, subjects using the Tickle Talker in conjunction with lipreading recorded vowel and consonant identification performances approaching 100% correct (Cowan, Alcantara, Blamey, and Clark, 1988)

It should be noted that not all attempts to present phonemes via multichannel prostheses have met with this level of success. Carney (1988) presented phonemes via both single and multichannel prostheses under both T and LT conditions. He found that vowel identification performance peaked at about 40% correct in the T condition using either the single or multichannel device, with somewhat lower levels of consonant identification performance being recorded with both devices. Further, when used in conjunction with lipreading neither device facilitated performance above that observed in the lipreading condition.

Although Carney attributed the similarity in performance observed with the single and multichannel devices to limitations in the skin's capacity to perceive the more complex patterns generated by the multichannel device, it seems more likely that this was due to limitations in his multichannel display. Particularly, the transducers in this multichannel device vibrated at the unusually low (for this type of device) frequency of 60Hz. As is detailed in Section 2.2.1, this frequency is well below the region of maximum sensitivity of the skin, and may even stimulate a different receptor population to that activated by his single channel device and other current tactile prostheses (Bolanowski, Gescheider,

12

Verrillo, and Checkosky, 1988; Verrillo, 1985).

The most significant outcome of these studies with multi-feature tactile aids is that users appear to be able to identify many of the phonemic building blocks from which spoken words are constructed without recourse to lipreading, and thus may be able to learn to identify this type of material via these devices. This suggestion has received at least partial support in subsequent studies of word and connected discourse identification.

### 1.3.3: Perceiving Words and Sentences

Users of the Vocoder have been trained to identify up to 250 closed set words to a criterion level of 80% correct with 80.5 hours of training (Brooks and Frost, 1983; Brooks, Frost, Mason, and Chung, 1985; Brooks, Frost, Mason, and Gibson, 1987). In subsequent tests with randomly presented open set words the subject achieved an accuracy level of only 8.8% correct (Brooks, Frost, Mason, and Gibson, 1986a). This result is not as discouraging as it may appear because the errors made were often minor with the general form of the words being correctly identified (e.g. responding "cancel" to the target word "council").

More striking success on word identification tasks has been achieved with multi-channel tactile aid used in conjunction with lipreading. Brooks et al. (1986a) found that open set words identification accuracy improved from 39.4% (L) to 68.7% (LT). As in the Vocoder-alone condition, many of the

errors recorded were minor. Similar results have been achieved with the Tickle Talker. Cowan (Cowan et al., 1987, 1988) tested both hearing and hearing impaired subjects on an open set word identification task using either lipreading alone or lipreading plus the Tickle Talker. Average accuracy levels increased from 53.3% (L) to 69.2% (LT) correct. More recent work using the Tickle Talker in conjunction with both lipreading and limited auditory information yielded HG word identification scores as high as 83% correct (Cowan, Alcantara, Whitford, Blamey, and Clark, 1989).

Granted the levels of success achieved on segmental feature and word identification tasks with these multi-channel devices, it might be expected that they would be even more effective than intermediate level devices in facilitating the perception of connected discourse. Surprisingly, this expectation is not unambiguously supported by the results obtained.

When the Vocoder was used in conjunction with lipreading, open set sentence tracking rates improved from 15.3 words per minute (L) to 49.3 words per minute (LT) (Brooks, Frost, Mason, and Gibson, 1986b). These results have been replicated by a separate team using a simulation of the Vocoder (Engebretson and O'Connell, 1986). Comparable tests with the Tickle Talker have also revealed increases in lipreading performance, with mean tracking rates improving from 31.6 words per minute (L) to 48.5 words per minute (LT) (Cowan et al., 1987, 1988).

In terms of the absolute speech tracking rates achieved, these results are no better than those obtained with Plant's (1986) intermediate level device, and are actually worse than those achieved with Grant's electrotactile device (Grant et al., 1986). Of course, comparisons like this are confounded by variables such as the amount and type of training received with each device, the degree of difficulty of the test material used, and the level of lipreading proficiency of the subjects.

While there were marked differences in the extent of training received by participants in these studies, those studies which involved the most training did not always achieve the best results. For example, the subject in Brooks et al.'s (1986b) study had almost 200 hours experience with the vocoder, yet achieved a lower LT tracking rate (49.3) than did the subjects using Grant's electrotactile aid (Grant et al., 1986) after only 20 hours of training (63.9). As many of the studies cited failed to list the tracking rate test materials used, it is not possible to determine whether this also affected the results.

There is, however, strong evidence that the third possible contaminant, that is the subjects' initial level of lipreading proficiency, does nullify any direct comparisons between the results of these studies. In particular, the subject in the Vocoder study had a much lower lipreading alone tracking rate than the subjects in any of the other studies, while the subjects in Grant et al.'s, (1986) study had much higher lipreading alone tracking rates than those in the other

studies.

One way to counter these differences in lipreading proficiency between subjects is to the view the data in terms of the degree of improvement in tracking rate between the L and LT conditions. On this basis, the multi-channel devices appear superior to their less complex cousins, with mean improvements of 222.2% for the Vocoder (Brooks et al., 1986b) and 53.5% for the Tickle Talker (Cowan et al., 1987, 1988 ) compared with only 37.2% for Plant's Sentiphone (Plant, 1986) device and 28.0% for Grant's electrotactile aid (Grant et al., 1986). Of course, this type of correction ignores the possibility of a ceiling effect limiting the extent to which those subjects with superior lipreading alone tracking rates could benefit from the addition of tactile cues.

While no device currently available supports the fluent perception of speech, it does seem that both intermediate level and multi-channel aids can significantly facilitate lipreading performance at all levels up to the perception of connected speech. This improvement in performance seems to be due to the availability of information about features not available via lipreading which are identified primarily by fundamental frequency cues within the displays. Multi-channel devices appear superior to intermediate level devices on segmental feature and word identification tasks, and seem to provide greater proportional improvements in speech tracking performance than simpler systems. Finally, and most importantly, these more complex devices alone have proven

capable of supporting the perception of spoken words without recourse to lipreading assistance.

## 1.4 : Optimizing multi-channel prostheses

There is no evidence to suggest that the levels of performance currently being achieved with multi-channel tactile aids reflect the optimum level which can be achieved. For example, there was no indication that subjects' rate of word learning was approaching an asymptote at the end of the 85 hours of training undertaken (Brooks et al., 1985). Further, the high levels of performance achieved with direct haptic communication strategies like Tadoma (Norton et al., 1977, Reed, Doherty, Braida, and Durlach, 1982; Reed, et al., 1985) and Plant's "laryngeal vibration" method (Plant and Spens, 1986) also suggests that the tactile sense may have an even greater capacity than that exploited by the current generation of multi-channel tactile aids. These two methods involve the user placing their hands and fingers about the face and throat of the speaker in order to sense the patterns of articulation associated with the words spoken.

In his extensive review and analysis of tactile aid research, Sherrick (1984) identified seven issues which must be addressed on the path to an efficient tactile aid for the hearing impaired;

"(a) What are the processing capacities of the skin?
(b) What form of transducer system will provide a reliable and efficient display to the skin?
(c) Which dimensions of tactile experience can be mapped to the acoustic stream of events in order to "match" hearing and touch?
(d) If not all speech features can be handled by the substitute

17

channel, which are the important ones to emphasize?
(e) How should the target population be defined? What characteristics of the target population will determine how the aid will be used: eg., as a speech or sound reception aid, as a speech training aid, as a supplementary aid, etc.?
(f) What are the acceptable educational designs for promoting the acquisition of user skills and for showing a true information gain with the aid?
(g) What form of testing procedures would demonstrate the gains in aid use, while considering the issues posed by the two preceding questions?" (p.1327).

While this thesis directly addresses only the first three of these questions, it is not possible to clearly specify the goals of this research without first giving brief consideration to Sherrick's (1984) questions of the target population and functional role for a tactile prosthesis (Questions D and E). Questions F and G will not be discussed as they only become relevant once a viable tactile prosthesis has been developed. In spite of the favorable results obtained to date, all but the simplest current devices are primarily research tools, rather than fully functional prostheses.

It would be ill-advised to attempt the development of any tactile communication system without a clear specification of the target population, their needs, and the extent to which the device should meet those needs. As traditional remedial strategies such as hearing aids generally meet the needs of individuals suffering less than profound conductive hearing loss, it seems clear that the target populations for tactile aids are the profoundly deaf and those whose hearing loss is primarily characterized by dysacusis rather than sensitivity losses. As was explained in Section 1.1, these individuals require assistance with all those perceptual tasks normally

18

handled by the sense of hearing. The critical question then becomes which of these tasks the tactile aid designer should attempt to facilitate.

The data reviewed here indicates that current multi-channel aids can substantially facilitate the user's level of lipreading performance. While it would be a worthy goal to attempt to optimize tactile aids for this use, the results achieved on word and sentence identification tasks using only the tactile information provided by the Vocoder (Brooks et al., 1986a, 1986b) suggest that tactile aids may be capable of providing more than just a supplement to lipreading. If a tactile aid capable of allowing the perception of connected discourse were developed, there seems no reason why this device would not serve equally well as an aid in both environmental sound identification and speech production monitoring and control tasks. In effect, such a device would provide a replacement for the dysfunctional auditory system.

Thus, in answer to Sherrick's (1984) questions concerning the target population and functional goals (Questions E and D) for tactile prostheses, this thesis adopts the optimistic position that most effort should be directed towards establishing the viability of a complete tactile substitute for the dysfunctional hearing system of the profoundly deaf. Of course this stance does not imply that other roles for tactile prostheses should be ignored, nor that touch will inevitably prove suitable for this purpose.

With the adoption of this position, the question becomes

whether touch can serve this role as a replacement for hearing, and if so, how to implement such a prosthesis. Sherrick's (1984) first four questions reduce these general issues to specific questions concerning the perceptual capacities of the skin, the type of coding strategy most suited for presenting speech transforms to the skin, and the physical specifications of the tactile transducers employed. The remainder of this chapter looks briefly at some of the issues arising from these four questions.

Most of the coding strategies currently in use are based on spectral analyses of the acoustic signal. Devices like the Vocoder attempt to provide information about a large number of bandwidths within the acoustic signal, while the Tickle Talker and intermediate level devices focus upon smaller sets of the spectral features which distinguish important segmental speech features. Although the Vocoder appears to be the most efficient of these devices, particularly in terms of its ability to convey information without lipreading assistance, information is lacking on the extent to which a multiplicity of channels is required. To this end, it is necessary to establish which particular combination of spectral features are necessary to allow accurate discriminations between discrete auditory events to be made. A useful tool in this line of research may be acoustic models of the proposed coding schemes, a strategy which has already proven beneficial in assessing the efficacy of various coding strategies for cochlear implants (Blamey, Martin, and Clark, 1985; Grant, Ardell, Kuhl, and Sparks,

1985). The rationale of this approach is to present the speech features of interest to subjects as auditory signals. The accuracy with which these stimuli are identified then provides an index of the amount of information conveyed by that particular coding strategy.

This emphasis on spectral coding strategies does not imply that alternate, that is non-spectral, coding strategies should not be pursued. Indeed, the high levels of performance achieved with naturalistic articulatory based methods of communication like Tadoma reveal the promise of this type of strategy. Attempts to generate synthetic articulatory aids have not yet fulfilled this promise. Green et al. (1983) investigated one alternative strategy in which the momentary cross-sectional areas of various sites along the vocal tract were presented via an Optacon transducer. In theory, this information is sufficient to identify the majority of speech features. Although this "vocal tract location" scheme proved less efficient than its spectral alternatives, the results of a single experiment using a single example of this type of coding strategy are not sufficient to assess this family of coding methods.

Of course, it may turn out that spectral and non-spectral coding strategies, rather than being alternatives, should be combined, if Richardson and Frost's (1977) speculation that efficient uses of the tactile system requires the presentation of as many diverse and redundant types of information as possible is correct. Their speculations have

received empirical support through the demonstration of an additivity in tactile information transfer rates as the number of stimulus dimensions increases (Rabinowitz, Houstma, Durlach, and Delhorne, 1987) . Indeed, as Cowan et al. (1989) have shown that users of tactile prostheses can beneficially integrate tactile information with that provided by lipreading and residual hearing, perhaps we should not focus just on the richness of the tactile display, but on the richness of the complete perceptual environment of the hearing impaired individual.

An over-riding issue in the future of multi-channel aids is the provision of suitable transducers. Current technology allows the construction of pocket-sized speech processors, but has yet to provide suitable vibrators for use in multi-transducer systems. As Sherrick (1984) noted, the ideal transducer should offer compact dimensions, low power consumption, low radiation of acoustic energy, and high levels of fidelity. Indeed it was the lack of such a device which, in part, motivated the choice of an electrotactile transducer for the Tickle Talker (Blamey and Clark, 1987). This is not to say that electrotactile devices inevitably provide the only answer. As in the case of coding strategies, at this point in the development of tactile communication systems, as many diverse approaches as possible should be encouraged.

Although interesting in their own right, these questions of the optimum physical organization and specification of a tactile communication device are brought together in Sherrick's

(1984) Question C, which asks how we can best match the perceptual capacities of the skin with the information transfer requirements for an accurate transform of speech. This question provides the foundation of the research reported here.

In particular, this research focuses upon a specific aspects of this broader question; Does the skin have the processing characteristics necessary to perceive speech transforms? This question is, of course, not uniquely relevant to the field of tactile prostheses, because it requires the assessment of fundamental aspects of the processing capacities of the tactile modality. Chapter 2 begins the process of answering this question by reviewing current evidence on the extent to which the tactile modality has the information processing capacities required to deal with speech transforms.

**Chapter 2: A Review of Tactile Processing Abilities in the Context of Tactile Speech Prostheses.**

This chapter examines the extent to which the processing capacities of the tactile modality are compatible with the presentation of speech transforms. The suitability of touch for this purpose has been questioned on two grounds. First, it has been claimed that the tactile sense lacks the general processing capacities necessary to deal with such complex stimuli. Second, some theorists have maintained that the perception of speech stimuli requires the agency of a specialized decoder which is unique to the auditory modality, and consequently is unavailable to tactile input. These two issues are discussed in the following sections.

## 2.1: Is the Speech Code Special?

A longstanding argument against the presentation of transforms of speech via either touch or vision stems from Liberman's theory of speech perception (Liberman, 1970; Liberman et al., 1967; Liberman et al., 1968), which maintains that the complexity of speech stimuli demands the operation of a special decoder which is only available within the auditory modality. Briefly, Liberman argued that speech processing must involve a complex code, rather than a simple alphabet based on the phoneme, because the auditory system lacks the temporal resolution to process the number phonemes per unit time which our observed rate of speech perception demands.

Even without this temporal limitation, Liberman claimed that phonemes are not uniquely identified by invariant acoustic cues, instead being typically represented by different acoustic features in different contexts. Further, it appears that a

single acoustic cue, for example the second-formant transition, can carry information relating to more than one phoneme (Liberman et al., 1968), again suggesting that speech perception does not proceed on the basis of a phonemic alphabet, but on a complex code in which phonemic units are processed in parallel. Liberman believed that it is the lack of a one-to-one correspondence between the acoustic event and its phonemic content which necessitates the agency of a specialized decoder within the auditory system. Liberman (Liberman et al., 1968) explained the relative failure of (then) current tactile and visual direct transforms of speech as resulting from the lack of this special speech decoder within the tactile and visual modalities. In his view, the only practical way to present speech information via other modalities is to present that information in a "decoded" form, thus avoiding the need for this special decoder.

Liberman's thesis can be questioned on several grounds. First, regardless of the status of his theory of speech perception, his argument against the availability of a specialized speech decoder in other than the auditory modality is circular. He observes that tactile and visual transforms of speech are hard to "read", then proceeds to explain that this occurs because these modalities lack a specialized speech decoder, yet his evidence for this claim is his initial observation regarding the difficulty evident in deciphering non-auditory speech transforms.

Aside from this apparent circularity, Liberman's basic observation that non-auditory speech spectrum transforms are hard to read may now be questioned granted the considerable success recently achieved in identifying words using tactile spectral prostheses alone. Indeed, although far from being fully functional speech prostheses, there is growing evidence that we can learn to read visual representations of the speech spectrum (Cole, Radniky, Zue, and Reddy, 1979; Green, Pisoni, and Carrell, 1982). Finally, the high levels of phoneme identification performance achieved by users of a range of tactile prostheses described in Chapter 1 suggests that Liberman's specialized speech decoder either is available to modalities other than hearing, or that such a decoder is not essential for processing speech information.

These suggestions have received support from other sources. First, it has been demonstrated that variations in the spectral composition of auditory and tactile representations of vowels and consonants lead to equivalent changes in the identification and discrimination of those phonemes in either modality (Eilers, Ozdamar, Oller, Miskiel, and Urbano, 1988). In contrast to Liberman's model, Eilers et al. (1988) concluded that important aspects of speech perception appear to proceed on the basis of specific acoustic information contained in the speech signal which is not uniquely available to the auditory modality.

Kirman (1973) has also argued strongly against Liberman's phoneme based model of speech perception on three

27

grounds. First, he refutes Liberman's claim that speech perception could not proceed on the basis of directly identifying a very large number of extended units of information such as syllables, on the grounds that other senses, like vision, appear capable of identifying similarly large numbers of complex patterns holistically. Second, Kirman argued that phonemes, rather than being the primary unit of speech perception, are only identified after the perception of larger speech units, a position supported by the finding that it takes longer to identify phonemes within a syllable than to identify that syllable (Massaro, 1972; Savin and Bever, 1970; Warren, 1971). More recently, Warren (1976) has argued that the phonemic restoration effect, where-by listeners perceive illusory phonemes in words when the actual phonemes are either artificially removed and replaced by noise or masked by environmental noise, demonstrates that phonemes are not the basic unit of speech perception. Finally, Kirman rejected the notion that the inherently complex relationship between an acoustic speech event and its phonemic elements entails the operation of some unique decoder on the Gibsonian (Gibson, 1966) grounds that all perceptual processes involve a similar process of extracting higher order invariances from varying proximal stimuli.

If we are not constrained, as Liberman would have us believe, to using specially decoded transforms of speech stimuli when designing non-auditory prostheses for the deaf, does it then matter which coding scheme we apply? Houde (cited

in Sherrick, 1984) argued that any spectral transform of speech can be interpreted provided only that the stimuli are presented in a well organized manner with adequate levels of training being given.

What is entailed in ensuring that a spectral display is well organized? In the broadest terms, and in line with Kirman's (1973) claims about the extraction of higher level invariances from complex patterns of stimulation, it means that features which identify the original speech stimulus must be available in the perceptual representation of that stimulus processed by the receiving modality. Two criteria must be met in order for this to occur. First, that information must be preserved in the process of transduction from the acoustic signal to a tactile one. This issue is, of course, Sherrick's (1984) question of the appropriate coding strategy for a tactile prosthesis which was discussed in Chapter 1. Second, this information must be preserved both at the point of sensory encoding on the skin, and also in the subsequent higher level processing of that information by the tactile system. Whether this is possible granted the perceptual properties of the tactile system is the question posed at the end of Chapter 1, and pursued in the remainder of this thesis.

## 2.2: The Resolution of the Skin

The second general objection to the viability of tactile prostheses for the deaf is that touch is too limited in terms of its spatial, temporal, and frequency resolution to perceive the fine details of a spectral display of speech. For example,

Pickett and Pickett (1963) contended that; "It is probable though, that the pattern-resolving power of the skin, in terms of spatial locus, intensity, and frequency, will prove to have severe limitations that cannot be overcome by intensive training." (p.219)

A sensible way to assess this criticism is to compare the resolving powers of the skin with the demands imposed by the information content of the speech signal. The auditory properties of spoken words are defined by rapid changes in the frequency, amplitude and phase of the spectral components of the acoustic signal. This section reviews the processing capacities of the skin in this context (see Geldard, 1960; Kirman, 1973; Loomis, 1981; Richardson & Frost, 1977 for other reviews).

### 2.2.1: Frequency Specific Tactile Processing Channels

Before comparing the frequency resolving capacity of the skin with the demands imposed by speech transforms, it is necessary to discuss some psychophysical aspects of frequency processing in the tactile modality. A large number of perceptual phenomena are subsumed under the umbrella of the tactile modality, including sensations as diverse as pain, pressure, and temperature. It appears that these different sensations are generally elicited via the stimulation of different types of receptors in the skin (Iggo, 1976). Likewise, it appears that the process of cutaneous mechanoreception involves the action of a number of functionally different fiber populations.

The initial evidence of differences in the way various tactile receptors deal with mechanical vibrations came from thresholds studies, which revealed a bimodality in the shape of the threshold function depending upon the frequency of the stimulus. Verrillo (1963, 1985) found that below 100Hz thresholds were relatively low and insensitive to changes in frequency, while above that frequency thresholds varied with frequency in accordance with a U-shaped function centered about 300Hz.

This duplex model of mechanoreception received further support from studies of the effects of changes in the spatial and temporal characteristics of the stimulus. Briefly, above about 100Hz thresholds decrease with increasing stimulus size (Craig, 1968; Verrillo, 1963) and duration (Verrillo, 1965), while these manipulations have minimal effect on thresholds below this frequency. Thus it appears that only the high-frequency channel is capable of either spatial or temporal summation.

Subsequent studies have revealed that, although perceptual phenomena like adaptation, masking, and enhancement can be observed within each of these two frequency ranges, these effects do not occur when the respective stimulus frequencies lie on opposite sides of the 100Hz boundary (Hamer, Verrillo, and Zwislocki, 1983; Verrillo, 1985; Verrillo and Gescheider, 1977). Again, these results are consistent with a duplex model of tactile vibratory perception.

Early studies established that Pacinian corpuscles are

responsible for activity in the high-frequency tactile channel, with activity in the low-frequency channel resulting from stimulation of a number of receptors generically labeled the Non-Pacinian system (Verrillo, 1966a, 1966b). More recently, Bolanowski et al. (1988) have presented psychophysical data showing that this Non-Pacinian system consists of three separate channels, each responsible for different, but overlapping, components of the tactile threshold function. As with the general distinction between the Pacinian and Non-Pacinian systems, these three channels are characterized by the shape of the threshold function within the range of frequencies encoded by the channel, the presence or absence of summation phenomena within the channel, and the differential effects on thresholds of changes in skin temperature. Bolanowski et al. (1988) identified that these three Non-Pacinian channels (NP I, NP II, and NP III) are each served by separate types of neural fibers. Particularly, the NP I channel is associated with rapidly adapting (RA) fibers, the NP II channel with type II slow adapting (SA II) fibers, and the NP III channel with type I slow adapting (SA I) fibers.

To date, attempts to present speech transforms via the skin have worked primarily at frequencies within the range processed by the Pacinian system, principally because of the much greater sensitivity of that channel, and its primacy as the perceptual channel for the sensation of "vibration" (Talbot, Darian-Smith, Kornhuber, and Mountcastle, 1968). In spite of these advantages of the Pacinian channel, the

independence of these various tactile processing systems in terms of masking interactions between frequencies which fall within the domain of alternate channels suggest that it may be useful to investigate the extent to which this emphasis on stimulating only the Pacinian system is desirable. This suggestion is of particular significance considering the divergent views available on the effects of masking upon the perception of complex tactile patterns[1].

## 2.2.2: Tactile Frequency Resolution

Typically, the acoustic energy in the speech signal is contained within the 200 to 8000Hz frequency range. As the skin is responsive to frequencies in the range 0.4Hz to about 800Hz[2], and an even narrower range of 100 to 800Hz in the case of the high-sensitivity Pacinian channel (Goff, 1967), it is clear that a direct mapping of acoustic frequency to tactile frequency is not possible. This problem is compounded by the poor frequency discrimination powers of the skin (Goff, 1967) which further limits the number of discrete frequency intervals available within the tactile sense.

This is not as severe a limitation to the feasibility of tactile prostheses as it may at first appear. The critical-bands masking effect shows that the peripheral auditory system

---

1. See Chapter 5 for an review of this issue.

2. Although upper tactile frequency resolution values as high as 8192Hz have been reported (Geldard, 1940), Kenshalo (1978) has dismissed these reports on the grounds that tactile mechonoreceptors are not capable of responding at that rate. He suggested that the subjects were responding to subharmonics of the test frequency.

appears to act as a series of overlapping band-pass filters having relatively wide bandwidths (Hawkins and Stevens, 1950; Scharf and Meiselman, 1977; Zwicker and Fastl, 1972).

Briefly, if the bandwidth of a noise masker whose center frequency is the same as that of the target is progressively increased then the extent of masking increases in turn. However, beyond some critical mask bandwidth no additional masking is observed as the mask bandwidth is widened further. The width of the critical band varies with the frequency of the target, with higher frequency targets requiring wider bandwidth masks before maximum masking is observed. It has been argued that the existence of these critical bands demonstrates that the auditory system functions as a series of overlapping bandpass filters. These critical bands seem to correlate with small (1-2 mm) linear segment along the basilar membrane (Scharf, 1970), each responsive to the particular range of frequencies included within that critical band. This implies that auditory stimuli, including speech, can be characterized in terms of the amplitude envelope of each of these discrete bandwidths. Indeed, it has been demonstrated that subjects can accurately identify spoken material presented as the output of a series of bandpass filters mimicking this bandpass frequency analyzing action of the inner ear (Lebedev and Zagoruiko, 1985; Pols, 1975).

Even before these basic processing characteristics of the ear were understood, Gault devised a tactile prosthesis which implemented many of the features of this multiple

34

bandpass filter model of hearing. His "Teletactor" partitioned the speech signal into 5 bandwidths with the amplitude of each band being presented to a different fingertip (Gault and Crane, 1928). As was noted in Chapter 1, this technique of transposing auditory frequency to place of stimulation on the skin underlies the majority of current multichannel tactile prostheses. It overcomes the limited frequency resolution of the skin by removing the need to vary the frequency of the presented stimuli, and benefits from the fact that only a relatively small number of loci need be stimulated in order to cover the range of frequency bandwidths that appear necessary in order to preserve the identity of spoken words.

### 2.2.3: Tactile Spatial Resolution

With the locus of stimulation providing the tactual correlate of the frequency of auditory stimuli, the next question is whether the skin has sufficient spatial acuity to deal with an array of transducers spread across its surface. The skin is not a uniform receptive surface with respect to its ability to discriminate between two sites of stimulation (see Loomis (1981) for a review of the various neurological, mechanical, and psychophysical factors affecting tactile spatial resolution). Using traditional measurement methods, the two-point limen varies from as little as 2.5mm on the mid-fingertip to 47mm on the calf (Weinstein, 1968). In general, other tactile parameters, like the threshold for vibration and force, the density of innervation, and the size of the cortical projection area tend to vary across body sites in proportion

with the two-point limen (Békésy, 1959, Kenshalo, 1978).

What does this data tell us about the implementation of auditory frequency to locus on the skin transformation schemes? First, it seems reasonable to assume that the number of stimulated loci in the tactile display should be at least equal to the number of critical bands in the auditory system which are responsive to speech frequencies. Pols' (1975) data suggests that approximately 17 1/3 octave bandwidths are necessary to convey the speech signal, not surprisingly a number similar to the 18 channel analysis performed by the Vocoder, the most complex current multichannel tactile device (Brooks and Frost, 1983, 1986).

If the two-point threshold is taken as the limiting factor, then several restrictions apply to the spatial configuration of a tactile display of speech. First, a linear array of transducers, as foreshadowed in Békésy's (1959) vision of a model of the basilar membrane on the skin, cannot be placed on any reasonably contiguous body site (for example the forearm, thigh, or back) without adjoining vibrators lying within the two-point limen of that region. If necessary, this limitation can be overcome either by employing a two dimensional matrix of transducers, or by placing the transducers at diverse points about the skin surface. Second, compact two dimensional displays like that of the Optacon, regardless of their placement, inevitably stimulate sites within the two-point limen of the skin region used.

Must we avoid these limiting conditions in the design of

a tactile display of speech? Sherrick observed that, "..the cutaneous system cannot be characterized by a set of values on dimensions that are independent of one another" (1984, p. 1328), noting, in this context, that the two-point limen varies both with the temporal relationship between the presented stimuli and with the measurement method used. For example, Johnson and Phillips (1981) reported that two stimuli presented to the fingertip could be differentiated using a forced choice task even when both sites lay within 1mm, a value less than half the minimum fingertip two-point limen obtained by Weinstein (1968). Similarly, Guyot, Johnson, and Weaver (1981) found two-point limens as small as 2mm on the forearm using a two-point versus one-point of stimulation judgment task within a signal detection design, a value much below Weinstein's (1968) 38mm.

Likewise, the introduction of temporal and amplitude variations between stimuli may induce changes in the perceived spatial relationships between those stimuli, and can even lead to the generation of phantom percepts (Békésy, 1967; Geldard and Sherrick, 1972; Sherrick and Rogers, 1966). Clearly, the perceptual effects of stimulating multiple sites on the skin are more complex than can be predicted from a simple analysis of two-point limens, which are themselves largely determined by the type of task used to obtain them.

A final objection to the reliance of two-point limen data when prescribing the spatial configuration of a tactile display of speech comes from the work of Loomis (Loomis, 1981;

Loomis & Collins, 1978). He has described an example of "tactile hyperacuity" whereby the threshold for detecting a shift in the site of stimulation is 10 to 30 times less than the two-point threshold (Loomis & Collins, 1978). This apparent paradox has been explained in terms of differences in the way in which the neural responses of adjoining mechanoreceptors are analyzed when performing either two-point and shift-in-locus tasks (Loomis & Collins, 1978). More recently, Richardson and Wuillemin (1981) have reported tactile hyperacuity in the perception of orientation changes between stimuli. They also explained their effect in terms of response patterns of adjoining mechanoreceptors. As detecting changes in the location of stimulation is the primary spatial task when working under a frequency-to-locus scheme, Loomis and Collins (1978) data suggests that the two-point limen is not the limiting factor in the design of this type of tactile display.

It appears that the limit of tactile spatial acuity is difficult to isolate. If the two-point threshold is taken as the index of acuity, then the obtained value varies by up to a factor of 20 depending on the psychophysical method used. Likewise, other measures of acuity, like the threshold for detecting changes in locus, show much finer tactile spatial discrimination ability than is indicated by traditional two-point measures. Finally, it is clear that any measure of acuity taken in isolation is unlikely to fully characterize the tactile modality's ability to deal with the complex percepts which result from covarying a range of stimulus parameters. In

conclusion, if tactile spatial acuity is as limited as is suggested by two-point thresholds, then display configurations are available which overcome this limitation. On the other hand, if, as it appears, touch has finer spatial acuity than this, more diverse display designs may be possible.

### 2.2.4: <u>Tactile Temporal Resolution</u>

As speech is characterized by rapid changes in stimulus parameters, an obvious question is whether touch has the temporal acuity necessary to deal with speech transforms. One measure of temporal acuity is the interval between two stimuli required before both are detected as discrete stimuli (delta-t). Gescheider (1966, 1967) found that in both the auditory and tactile modalities delta-t decreases as a function of stimulus amplitude. However, delta-t in the tactile modality was at best 10ms, five times longer than in hearing.

This comparison may not, however, provide a true reflection of the respective abilities of touch and hearing to deal with rapid changes in stimuli. Hill and Bliss (1968) argued that a valid index of tactile temporal resolution must reflect the observer's capacity to identify the stimuli. Hirsh and Sherrick (1961) measured the threshold for detecting the temporal order in which two stimuli were presented. They found that in hearing, vision, and touch, two events had to be separated by approximately 20ms before they could be correctly ordered (using a 75% correct criterion). This constant held true in each modality regardless of the spatial location of each event. However, Hill and Bliss (1968) subsequently

2: Tactile Processing Abilities

demonstrated that the tactile limen for sequential recognition does vary with stimulus spacing, reporting an increase in temporal resolution as stimulus separation was decreased. Hence the invariance of Hirsh and Sherrick's (1961) 20ms temporal resolution constant must be questioned. The fact still remains that this more ecologically valid measure of temporal resolution shows that touch and hearing do not differ greatly in terms of their temporal acuity.

Even if the temporal resolution of touch is worse than that of hearing, the ability of touch to deal with rapid variations in speech derived stimuli may not be fully predicted by this index. At the very brief stimulus onset asynchronies (SOA) tested by Gescheider (1966, 1967) the limiting factors on both tactile and auditory temporal resolution are the strong masking effects which occur between successive stimuli. As is shown in Chapter 5, the perceptual consequences of these masking effects are currently unclear. Hence, as in the case of tactile spatial resolution, a simple analysis of tactile temporal resolution may not provide an accurate reflection of the skin's capacity to deal with brief temporal variations within a spectral representation of speech. Indeed, as is argued in Chapter 5, it may prove that speech perception does not require events on this time-scale to be perceived as discrete events.

### 2.2.5: Tactile Amplitude Acuity

A final parameter of tactile acuity relevant to the perception of speech transforms by the skin is the tactile

modality's ability to detect changes in stimulus amplitude. As with the other parameters discussed here, there are two main aspects of tactile amplitude perception worthy of consideration. First, the range of amplitudes to which the skin is responsive, and second, the ability of touch to discriminate between increments within this range.

The skin is not as sensitive to stimulation as are hearing and vision. Geldard (1972) concluded that the absolute threshold of the skin may be $10^7$-$10^9$ times that of hearing and vision. Obviously this shows that speech stimuli must be amplified before presentation to the skin.

As with spatial acuity, the sensitivity of the skin varies across its surface, in this case by several orders of magnitude. In general those sites with small two-point limens, high densities of mechanoreceptors and large cortical projection areas are the most sensitive (Békésy, 1959). However, the sensitivity ranking of regions does seem to vary depending upon the measure of sensitivity adopted. For example, Weinstein (1968) found that facial sites like the nose were most sensitive to static force, while Wilska (1954) found that finger and hand sites were most sensitive to vibration.

The significance of these differences in sensitivity across sites to the implementation of tactile speech prostheses is that the amplitude of vibrations presented to different sites may require adjustment to ensure that equally loud speech components presented to different sites yield equally intense tactile percepts. In the experiments reported in Chapter 6,

41

vibrations were presented to 3 sites within a 10cm region on the forearm. It was necessary to adjust the amplitude of these vibrations by up to 5dBm before the point of subjective equality was reached, a value similar to that previously reported for this forearm region (Békésy, 1959).

In the present context the most important characteristic of tactile amplitude coding is the extent to which we can discriminate changes in vibratory amplitude. Spector (cited in Geldard, 1957) found that the just noticeable difference (JND) on the chest ranged from 10μm at low amplitudes to 60μm at high amplitudes. Although Geldard (1957, 1960) detected up to 17 JND steps across a wide range of stimulus amplitudes, he claimed that the actual number of discrete amplitude steps that can be discriminated on an absolute recognition basis, rather than a JND basis, may be as low as three. Although three amplitude levels is insufficient for the purposes of perceiving tactile speech transforms, Geldard (1957, 1960) did note that this ability may improve with training. Indeed, it may also be that sites more sensitive than the chest allow a larger number of discrete steps to be recognized.

Studies of the utility of amplitude as an information carrier in tactile displays have also highlighted the limited intensity processing ability of the skin. Geldard (1960) reported that intensity is the least exploitable dimension in a tactile communication system, and Rabinowitz et al. (1987) have shown that the information transfer rate achieved using amplitude as the carrier is adversely effected by concurrent

changes across other display parameters.

### 2.2.5:  Conclusions

This section reviewed potential limitations to the use of touch for the presentation of speech transforms. First, it was shown that there is little support for claims that only the auditory modality can deal effectively with the speech signal. Second, although touch was found to lack the fine frequency discrimination capacities necessary to deal directly with the speech spectrum, a frequency-to-locus strategy seems capable of compensating for this limitation. Finally, while it may ultimately be found that touch lacks the spatial and temporal acuity to deal with speech transforms, the diverse and highly interactive nature of the tactile response to variations in stimulus dimensions suggests that this limitation will not be revealed by simple analyses of two-point thresholds in time and space.

The relevance of this complexity to the perception of tactile displays of speech has been noted by a number of reviewers (Kirman, 1973; Richardson and Frost, 1977, Sherrick, 1984), all of whom agree that the capacity of the skin to deal with speech transforms is dependent to a greater extent upon these effects, rather than on the simple resolution parameters of the skin. Indeed, some, like Kirman (1973) and Richardson and Frost (1977), go further, arguing that these interactive perceptual effects within the tactile modality are actually the vehicle which allows the perception of complex tactile spatiotemporal patterns. This issue is discussed at length in

43

Chapter 5. In conclusion, no clearly insurmountable limitations to the development of tactile speech prostheses have yet been identified.

### 2.3 : Similarities Between Touch and Hearing

The previous section compared the processing capacities of touch with those required to deal with tactile transforms of the speech spectrum. This process frequently involved comparing the processing capacities of touch and hearing. When differences in the processing abilities of these two modalities were identified, it was often found that strategies to compensate for these differences are available. An example of compensatory changes in the design of tactile prostheses flowing from this process of contrasting the capabilities of touch with those of hearing, is the "frequency-to-locus" coding strategy which accommodates differences in the frequency resolving powers of the two modalities. On the other hand, Eilers et al.'s. (1988) demonstration that changes in the spectral composition of acoustic stimuli lead to similar perceptual effects in both modalities is an example of a comparison which shows an area where minimal compensation in the design of the tactile prosthesis is required. Particularly, their data suggests that touch may respond well to a direct representation of acoustic spectral characteristics within its processing range.

These two examples highlight the advantages for the development of tactile speech prostheses which can flow from comparisons between auditory and tactile processes. This

44

section examines the extent to which tactile and auditory representations of stimuli are similar with the intention of clarifying the extent to which direct transformations between these two modalities are likely to yield similar percepts.

2.3.1: Auditory and Tactile Responses to Vibrations

Both the skin and the ear respond to mechanical vibrations. As vibratory stimuli are primarily characterized by their amplitude and frequency, an obvious question is whether touch and hearing deal with changes in these parameters of vibratory stimuli in perceptually and neurologically similar ways. This sub-section briefly reviews this question.

Békésy (1955,1959) has shown that the presentation of vibrations to either the skin or the ear leads to the generation of traveling waves across the respective sensory surfaces (the basilar membrane in the case of the ear). In spite of this spread of the stimulation, Békésy (1955, 1959) found that the perceived locus of stimulation was small in both modalities. He argued that in both cases a process of lateral inhibition between adjoining mechanoreceptors was responsible for this dampening of the spread of stimulation.

Both senses can detect changes in the frequency of vibrations, with the sensation of pitch resulting in either case. However, Békésy (1957a) noted a significant difference in the mechanisms underlying pitch perception in either modality. The perception of pitch on the skin is entirely dependent upon the rate of neural firing elicited by the vibration. In contrast, the place of maximum displacement of the basilar

45

membrane is the primary cue to auditory pitch, with neural firing rates playing only a minor role.

One perceptual consequence of this difference in frequency coding mechanisms is the frequency bandwidth of each modality. Due to the limited rate at which tactile receptors can discharge, only frequencies below about 800Hz can be readily detected by the skin (Goff, 1967). On the other hand, the dual frequency coding mechanisms of the ear (place of stimulation and rate of neural discharge) allow the aural perception of frequencies between about 20Hz and 15kHz[3] (Licklider, 1951).

Current models propose that the rate of neural discharge is the main mechanism of aural pitch perception at frequencies below about 100Hz (Békésy, 1956; Gulick, 1971). Granted that the rate of neural discharge is the only mechanism of pitch perception in touch, it is not surprising that it is only at frequencies below this level that tactile and auditory frequency acuity is comparable. For example, Goff (1967) found that while the tactile threshold for detecting changes in frequency is generally much worse than that of the ear, thresholds were similar for both senses below about 50Hz.

As with frequency, both the ear and the skin respond to changes in the amplitude of vibrations. Békésy (1958) found that increasing the amplitude of vibrations presented to the ear or the skin not only leads to a growth of in the magnitude

--------------------------------------------------

3. This frequency range is typical of a healthy adult. Children exhibit a higher maximum frequency detection limit.

of the resulting percept in each case, but also that the rate
of growth is similar in both modalities. He found the greatest
level of similarity to occur in the case of skin regions with
high densities of innervation. At other skin sites, the rate of
growth in sensation magnitude with amplitude has been found to
be steeper than for the ear (Békésy, 1958; Stevens, 1959a).
Perhaps, as Stevens (1959a) has suggested, this effect is
comparable to the loudness recruitment phenomena (Sanders,
1984) observed in patients suffering hearing loss due to
atrophy of auditory receptors.

Finally, Stevens (1959b) demonstrated that subjects
could readily match the amplitudes of vibrations presented to
the ear and the skin. Indeed, he found that the cross modal
equal-sensation function was linear and a grew with stimulus
amplitude in direct proportion to the ratio of the slopes of
the amplitude-sensation level functions in each modality.

2.3.2: Perceptual Similarities Between Hearing and Touch

Granted the similarities between hearing and touch in
terms of the way in which they process vibratory stimuli, it
might be expected that these stimuli are represented and
processed in similar ways at the perceptual and cognitive
levels. Handel and Buffardi (1968) have shown that there may
indeed be some fundamental similarity between auditory and
tactile percepts. They presented subjects with eight element
sequential patterns, with each element being identified by the
modality through which it was presented. When the respective
pattern elements were either tactile and visual or visual and

47

auditory, the subjects learnt to identify the pattern more rapidly than when the respective elements were auditory and tactile. This difference was independent of subjects' ability to learn similar patterns presented to each modality alone.

This finding indicates that the subjects had difficulty differentiating between auditory and tactile modes of stimulation, but not between either auditory and visual or tactile and visual inputs. Indeed, the subjects reported a sensation of "snapping back and forth between modalities" (Handel & Buffardi, 1968, p. 1028) in the latter cases, but not when auditory and tactile stimuli were paired. Subjects in our laboratory consistently report a similar effect. When presented with a vibrotactile stimulus, whose associated acoustic signal is completely masked by aurally presented white noise, most subjects report that the tactile stimulus "feels like a sound".

While these observations do show that some vibrotactile and auditory stimuli yield similar percepts, which differ from those arising from some visual stimuli, the generality of this observation is not clear. For example, would Handel and Buffardi (1968) have achieved the same pattern of results if both the visual and tactile stimuli were geometric shapes, rather than coloured lights and vibrations respectively? Perhaps these conditions would show the tactile and visual representations of the stimuli to be most alike.

Of particular relevance here is whether tactile and auditory representations of speech derived stimuli generate similar percepts. Eilers et al.'s (1988) demonstration of

48

equivalent perceptual changes in either modality in response to changes in spectral characteristics of speech stimuli does suggest that such perceptual similarities do exist. The following chapter describes a series of experiment which took as their starting point this question of the perceptual equivalence of auditory and tactile representations of speech stimuli.

### 2.4: Summary

This chapter reviewed some of the preliminary issues which must be addressed in deciding whether touch has the information processing properties necessary to deal with speech transforms. Section 2.1 demonstrated the flaws in Liberman's claim that only the auditory modality is equipped to deal with the particular processing demands imposed by the speech signal. Section 2.2 dealt with the broad issue of comparing the acuity of the tactile modality with the demands imposed by speech transforms. It was concluded that the available evidence is inadequate for a proper assessment of this issue. Most previous studies have explored tactile thresholds in time and space, while the important question in the present context is the ability of the perceiver to organize and differentiate between complex patterns of stimulation. This issue is addressed in Chapters 5 and 6. Finally, this chapter raised the need to compare and contrast auditory and tactile perceptual and information processing processes, arguing that this approach may facilitate the development of tactile prostheses. This last issue is pursued in the following chapter.

**Chapter 3: Cross-Modal Comparisons Between Auditory and Tactile and Auditory and Visual Representations of Spoken Words.**

The tactile modality is not the only one through which transforms of speech may be presented. A number of attempts have also been made to develop visual transforms of speech, most notably by Potter and his associates at the Bell Laboratories, who devised a system in which the speech spectrum was presented on a phosphorescent display (Potter, 1945; Potter, Kopp, and Green, 1947). Mahar (1985) asked whether there was any reason to favour tactile prostheses over these visual alternatives. He argued that if, as Handel and Buffardi (1968) claimed, there are greater similarities between the auditory and tactile perceptual representations of patterns than between these and visual representations, then this similarity in the way touch and hearing process patterns might make tactile speech transforms easier to interpret than their visual alternatives. The following section describes an experiment he undertook to confirm and extend Handel and Buffardi's (1968) findings.

### 3.1: Comparisons Between Auditory, Visual, and Tactile Representation of Spoken Words.

Although Handel and Buffardi (1968) found similarities between auditory and tactile perceptual representations of patterns, and differences between these and visual representations, they did not use speech stimuli. While it is true that Eilers et al.'s (1988) data indicates that tactile and auditory representations are similar when speech stimuli are used, they did not include a visual transform condition. Hence the degree of similarity between tactile, visual, and auditory representation of speech derived stimuli is unclear.

Mahar (1985) set out to clarify this issue by measuring the degree of similarity between both auditory and tactile and auditory and visual representations of spoken words. Subjects compared the pattern of stress in an auditory representation of a two-syllable word with that in subsequently presented tactile or visual transforms of a two-syllable spoken word. The rationale behind this task was that comparisons between similar types of perceptual representation should be performed both faster and more accurately than comparisons between less similar representations. Support for this design comes from cross-modal and intra-modal matching studies which show that the process of translation from one modality to the other imposes a penalty in terms of speed and accuracy (Bjorkman, 1967; Ittyerah & Broota, 1983). If auditory and tactile representations of speech derived stimuli are more alike than auditory and visual ones then the process of translation between auditory and tactile representations should impose a smaller performance penalty than applies in the case of auditory to visual translations.

The first word in each pair was presented to the subject via headphones as a 1 kHz pure tone modulated in amplitude to mirror changes in the speech-envelope of that word, then the second word was presented either as a tactile or visual transform. The tactile stimuli were 250 Hz square wave vibrotactile signals which varied in amplitude to mirror the momentary changes in loudness in the spoken word. These tactile stimuli were presented via a Minivib I single channel tactile

aid with the transducer affixed to the wrist of the subjects' preferred arm. The acoustic output of the transducer was controlled by placing the subject's arm in an acoustically damped box.

Two types of visual transform were used. The visual-spatiotemporal display displayed a graphical representation of amplitude changes within the spoken word on a cathode ray oscilloscope. These stimuli were generated in real-time across the screen so that by the end of each trial the entire representation of the word was visible. A schematic representation of the formation of this stimulus across time is presented in Figure 3.1. The visual-temporal display consisted of a row of LED's which were illuminated to represent momentary changes in stimulus amplitude. As stimulus amplitude increased, so too did the number of LED's illuminated.

The rationale behind the selection of these two modes of visual display was as follows. The visual spatiotemporal display was intended to reflect a simple version of the type of display used in previous attempts to convey speech transforms visually (eg. Potter, 1945; Potter et al., 1947). The visual temporal display was included in an attempt to prevent any effect due to modality being confounded with effects due to information distribution style. Many commentators have observed that cross-modal judgment studies have frequently failed to control for differences in the way the stimuli were administered to the subjects. In particular, subjects have frequently been asked to compare spatial arrays of visual

Time
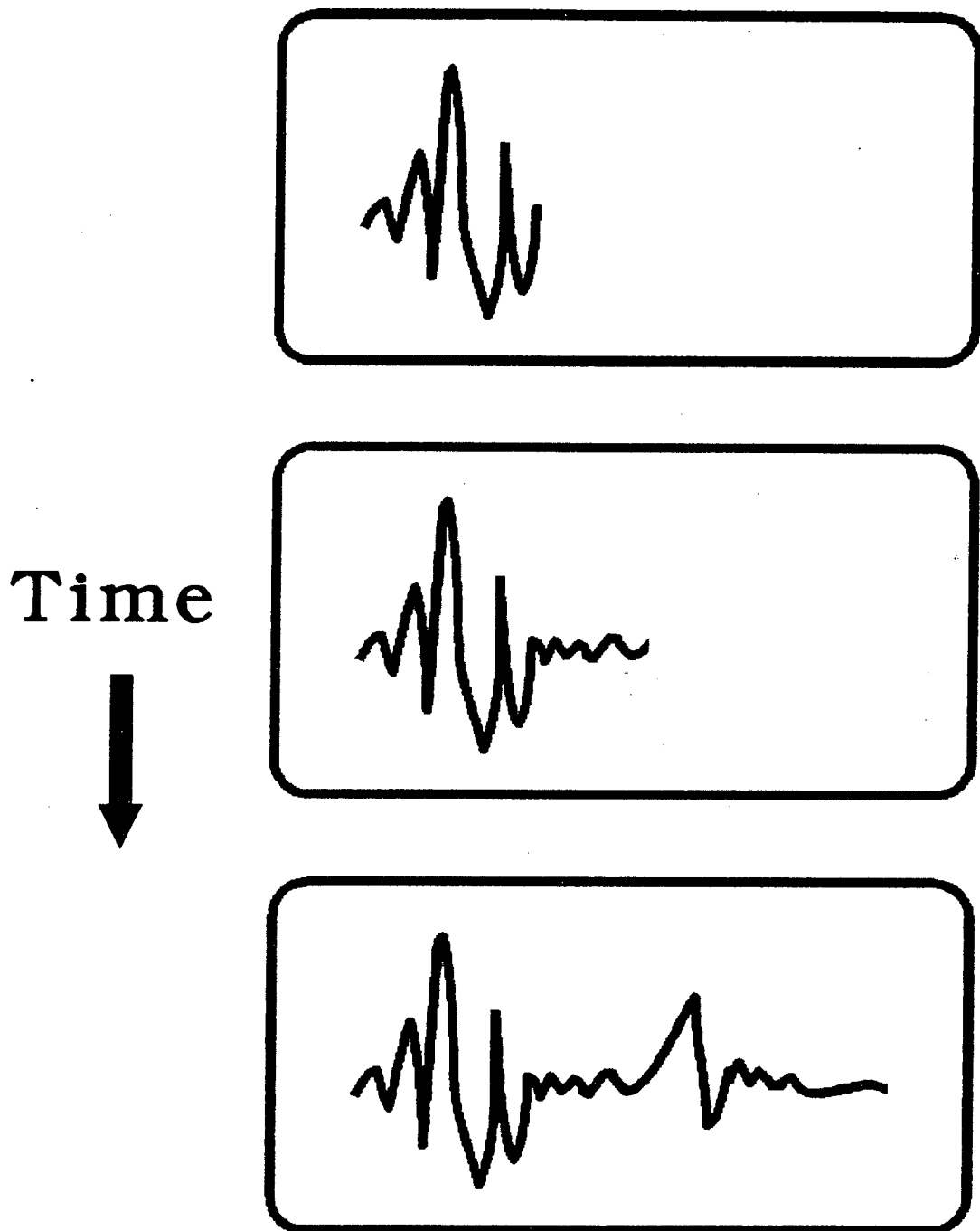
Figure 3.1: The progressive generation of a visual-spatiotemporal representation of a two-syllable spoken word stressed on the first syllable. The peaks and troughs represent momentary changes in the amplitude envelope of the spoken word.

stimuli with temporal arrays of auditory stimuli, in which case any cross-modal effects are confounded with cross-information distribution style effects (Friedes, 1974; Sawada & Jarman, 1982; Sterritt & Rudnick, 1966). Bryden (1972) found that comparisons across both modality and information distribution style are more difficult to perform than are across modality but within information distribution style ones. As the auditory and tactile stimuli used here were distributed purely across time, while the visual spatiotemporal stimuli varied across both time and space, it was hoped that inclusion of the visual temporal display would help clarify the cause of any differences in performance between the auditory to tactile and auditory to visual comparison tasks.

The results of this experiment are summarized in Table 3.1, which presents mean reactions times for correct responses and mean dprimes[1] for each type of transform. As can be seen, the auditory to vibrotactile comparisons were performed both faster and more accurately than were the auditory to visual-spatiotemporal comparisons. Although the auditory to visual-temporal comparison task was performed as rapidly as the auditory to tactile-temporal task, the near chance dprime scores recorded in the visual-temporal condition show that the subjects were effectively unable to compare the pattern of stress in the words presented in this way. In this case, it can only be concluded that the visual-spatiotemporal task,

---

1. Dprime is used as a measure of accuracy throughout this thesis. A summary of the derivation and usage of dprime is presented in Appendix A.

although performed more slowly than the visual-temporal comparison task, was performed more efficiently than its visual-temporal equivalent.

Table 3.1: Mean reaction time (sec.) for correct responses and mean accuracy (d') levels from Mahar (1985) as a function of transform modality and information distribution style.

|                  | Tactile Temporal | Visual Spatiotemporal | Visual Temporal |
|------------------|------------------|------------------------|-----------------|
| Accuracy (d')    | 1.27             | 0.83                   | 0.25            |
| Reaction Time (s)| 1.32             | 2.26                   | 1.39            |

Of course these results do not show that tactile transforms of speech are unequivocally easier to compare with auditory stimuli than are their visual equivalents. Other tactile transforms may be less efficient, with other visual ones being more efficient, than those used in this experiment. All that can be said with certainty is that this particular tactile display was easier to compare with the auditory reference stimulus than were these specific visual displays. Although these differences in performance may be due to the predicted greater similarity between auditory and tactile perceptual representations of patterns, another explanation is that the tactile display allowed the pattern of stress in the words to be discriminated more readily than did either of the visual displays.

### 3.2: Experiment 1: Discriminability of the Tactile and Visual Displays used in Mahar's (1985) Study

If the visual displays used in Mahar's (1985) cross-modal comparison experiment did not readily provide the subjects with the information needed to perform those comparisons then his results would tell us nothing about the relative ease with which tactile or visual representations of spoken words can be compared with auditory representations of words. On the other hand, if the visual displays were at least as discriminable as the tactile ones then the difference in the ease with which auditory to tactile and auditory to visual comparisons were made may be attributed to differences in the degree of compatibility between auditory, tactile, and visual representations of spoken words. This section describes an experiment conducted to test these alternative explanations by measuring the ease with which the stress patterns of words could be extracted from the various tactile and visual transforms of spoken words used in Mahar's (1985) experiment[2].

3.2.1:  Method

Fifteen university undergraduates with normal or corrected to normal vision participated in this experiment. They were presented with transforms of two-syllable spoken words via the tactile, visual-spatiotemporal, and visual-temporal displays described in Section 3.1. Half the test words were stressed on the first syllable, while the other half were stressed on the second syllable. The subjects task was to

-----------------------------------------------------------------

2.This experiment is the first conducted towards this thesis

identify the stressed syllable in each word via two-choice button-press. Each subject undertook 12 practice and 48 experimental trials in one of the three conditions. As the intention of this experiment was to establish whether subjects could determine the stressed syllable in the words presented, only accuracy data was gathered in this case.

3.2.2: Results and Discussion

Mean dprime scores from this display discrimination experiment are presented in Table 3.2[3]. Dprimes were calculated by counting words stressed on the first syllable which were correctly identified as hits, and words stressed on the second syllable which were incorrectly classified as false alarms. Table 3.2 shows that subjects could identify the stressed syllable in each word most accurately when the visual-spatiotemporal display was used. Accuracy levels in the tactile-temporal and visual-temporal conditions were similar. A series of planned comparisons confirmed these descriptive observations, revealing a significant difference in performance between the visual-spatiotemporal scores and those recorded in either the tactile-temporal, $F(1,12)=21.58$, $p<.001$, or visual-temporal, $F(1,12)=35.24$, $p<.001$, conditions. Accuracy levels in the tactile-temporal and visual-temporal conditions did not differ significantly, $F(1,12)=1.67$, $p>.05$.

As the two visual displays were at least as

_____

3. The mean scores recorded by each subject in this experiment are given in Appendix B1.

discriminable as the tactile one, these results show that the difficulty in performing the auditory to visual comparisons described in Section 3.1 was not due to any limitations in the design of these displays. Put simply, the stressed syllable in the visual representations of the words was clear to the subjects. Thus these difficulties must have arisen at the time that the cross modal comparisons were made, rather than when the subjects were extracting information about each word's stress patterns from the individual displays. Conversely, although the tactile display proved to be no more discriminable than either of the visual ones, the auditory to tactile comparison was performed better than either of the visual equivalents. In general, these result show that the ease with which a transforms of an auditory stimuli presented to touch or vision can be compared with an auditory representation of those stimuli is not a direct function of the ease with which that transform can be discriminated alone.

Table 3.2: Mean accuracy (d') levels from Experiment 1 as a function of transform modality and display type. Standard deviations are given in parentheses.

|  | Tactile Temporal | Visual Spatiotemporal | Visual Temporal |
|---|---|---|---|
| Accuracy (d') | 1.14 (0.15) | 1.98 (0.32) | 0.91 (0.27) |

An obvious explanation of these results is that they reflect some fundamental similarity between the auditory and tactile representations of a pattern, and a fundamental incompatibility between auditory and visual representations of that stimulus. While it is also possible that some aspect of

59

the particular visual displays used made the auditory to visual comparisons difficult to perform, this explanation awaits the demonstration of superior performance on an auditory to visual comparison task compared with that achieved on a similar auditory to tactile task, with the rider that the tactile display used is no less discriminable than the visual one.

Unlike these between modality comparisons, a comparison of the visual-temporal and visual-spatiotemporal results from these two experiments suggest that within the visual modality the discriminability of the particular transform used is a strong determinant of the relative ease with which that transform can be compared with an auditory representation of the same stimulus. This relationship does not seem to be a linear one. Although the visual-temporal display could be discriminated at well above chance levels, subjects found it almost impossible to compare this display with an auditory one.

It is tempting to conclude that these results tell us something about the way vision deals with temporally and spatiotemporally distributed displays. The amplitude information characterizing the stress pattern in the word was distributed purely across time in the visual-temporal display, while a spatiotemporal representation of this information was presented in the visual-spatiotemporal display. Perhaps it was this difference in information distribution style which led to the observed differences with which these two types of display could be either discriminated or compared with an auditory reference.

60

The main obstacle to assessing this explanation is that the two visual displays differed in many ways other than this information distribution style parameter. For example, the two displays differed in terms of the retinal size of the presented images and the luminance of the display. The following section describes an experiment which attempted to replicate the present results while ensuring that there was minimum uncontrolled variation between the characteristics of the various displays used. In addition, the range of information distribution styles used in the visual and tactile displays was extended to include both temporally, spatially, and spatiotemporally distributed variants within each modality.

### 3.3: Experiment 2: Cross-modal Comparisons and Information Distribution Style

The following experiment was modeled on the general design of the cross-modal comparison experiment described in Section 3.1, except that synthetic representations of two-syllable words were used instead of direct transforms of actual words. The visual stimuli now consisted of pairs of bars differing in luminance, while the tactile and auditory stimuli consisted of vibrations varying in amplitude. These systematic variations in luminance or amplitude provided an analog of the amplitude differences between the two syllables of the words used in the previous experiments. In addition to this change, all possible attempts were made to ensure that the temporally, spatially, and spatiotemporally distributed displays used in each modality differed only in terms of those characteristics

entailed in that style of display. It was hoped that the addition of these extra controls over the nature of the stimuli would allow a clearer assessment of the reasons for any within or between modality differences in performance.

### 3.3.1: Method

Subjects. Forty university undergraduates with normal or corrected to normal hearing and vision participated in this experiment.

Apparatus and Procedure. The subjects' task was to compare either a visual or tactile target stimulus with a subsequently presented auditory reference stimulus. The reference stimuli used were pairs of sequentially presented auditory pulses. Each pulse was a 550ms duration 1kHz square wave tone with a 300ms ISI between the two pulses in each pair. The amplitude of the two pulses in each pair was manipulated to provide a simple analog of the amplitude envelope representations of two syllable words presented in Experiment 1. The amplitude of the louder pulse in each pair was either 33, 30, 27, or 24 dBm[4], with that of the quieter pulse being set 6 dBm lower. The pulses were generated by an Applied Engineering Super Music Synthesizer controlled by a microcomputer and were presented binaurally via Senheiser HD22 headphones.

The tactile target stimuli consisted of pairs of 250Hz ramp wave vibrations presented via Oticon bone-conduction

---

4. 0dBm = 1mW into a 600Ω resistive load.

hearing aids placed under the index and third finger-tips of the subjects' preferred hand. These signals were generated by an Applied Engineering Super Music Synthesizer controlled by microcomputer. The amplitude of the two vibrations in each pair was varied in the same way as in the case of the auditory pairs, with the exception that a difference of 9dBm between the strong and weak pulses in each pair was employed. As the vibrators generated discernible auditory signals, they were mounted inside an acoustically dampened box into which the subject's hand was placed.

There were four tactile to auditory comparison conditions. In the tactile-temporal condition the first vibration was presented simultaneously to both finger-tips for 550 ms, then, 300 ms after its termination, the second vibration was also presented simultaneously to both finger-tips for 550 ms. In the tactile-spatial condition both vibrations were presented simultaneously for 550 ms to different finger-tips (the first and third fingers respectively). In the tactile-spatiotemporal condition the first vibration was presented for 550 ms to the subjects' index finger-tip. Three-hundred milliseconds after the termination of the first vibration, the second vibration was presented for 550 ms to the third finger-tip.

Due to the sequential presentation method used in the tactile-temporal and tactile-spatiotemporal conditions, it was possible for the subjects to process the first vibration in each pair before the second one was presented. As the

simultaneous presentation method used in the tactile-spatial condition precluded the subjects from preprocessing the first vibration, it was feared that this would impede performance on this task relative to those where the first vibration could be preprocessed. To test for this preprocessing effect, a second tactile-spatial condition was included in which an additional delay of 850ms (equivalent to the duration between the onset of the two vibrations in the tactile-temporal and tactile-spatiotemporal conditions) was included after the termination of the display before the auditory reference stimulus was presented. It was expected that this extra processing time available in the tactile-spatial-delayed condition would compensate for any preprocessing advantage inherent in the other two conditions.

The visual target stimuli were pairs of luminous bars generated by a Sprite Graphics card controlled by a microcomputer and displayed on a Sony CVM110 VDU. Two bars were varied in luminance to provide a visual analog of the auditory reference stimuli. The luminance of the high-intensity bar in each pair was set at either 125, 116, 98, or 67 cd/m$^2$, with the corresponding low-intensity bar being set at either 116, 98, 67, or 51 cd/m$^2$ respectively. Each bar extended across 3.30 degrees of visual field with a height of 0.75 degrees of visual field.

As with the tactile stimuli, four different presentation methods were used with these visual stimuli. In the visual-temporal condition, the two bars were presented sequentially at

the center of the display. Each bar was displayed for 550 ms with a 300 ms ISI. In the visual-spatial condition both bars were displayed side-by-side on the screen for 550 ms. The visual-spatial-delayed condition was identical to the visual-spatial condition except that an additional delay of 850ms was included after the termination of the display before the presentation of the auditory reference stimulus. As with the tactile-spatial-delayed condition, this visual condition was included to test for the occurrence of a preprocessing advantage in the visual-temporal and visual-spatiotemporal conditions. Finally, in the visual-spatiotemporal condition the first bar was presented for the duration of the display on the left-hand side of the display. The second bar was then presented for 550 ms on the right-hand side of the display 850 ms after the onset of the first bar.

Each subject participated in one of the eight cross modal comparison tasks, undertaking 12 practice and 48 experimental trials on that task. On each trial they were presented with the tactile or visual target stimulus then, after a 100ms delay (950ms in the cases of the tactile and visual spatial-delayed conditions) with the auditory reference stimulus. Their task was to identify whether the same pulse was most intense in both the target and reference displays, then respond by two-choice button press. They were instructed to respond as quickly as they could, using only their non-preferred hand to operate the button-press.

Both reaction time and accuracy data were recorded, with

reaction timing commencing at the onset of the auditory reference display. This ensured that there was a constant delay between the onset of timing and the termination of the trial regardless of the target stimulus presentation method.

### 3.3.2: Results and Discussion

Table 3.3 presents the mean reaction times for correct responses and mean dprime scores for each level of each factor in the experiment[5]. Dprimes were calculated counting same intensity pattern pairs which were correctly identified as hits and different intensity pattern pairs which were incorrectly identified as false alarms.

Table 3.3: Mean reaction time (sec.) for correct responses and mean accuracy (d') scores from Experiment 2 as a function of modality and information distribution style. Standard deviations are given in parentheses.

| | Target Stimulus Information Distribution Style | | | | |
|---|---|---|---|---|---|
| | Temporal | Spatial | Spatial-Delayed | Spatio-Temporal | Mean |
| | Touch | | | | |
| Reaction Time (s) | 1.57 (0.11) | 1.70 (0.16) | 1.87 (0.15) | 1.46 (0.08) | 1.65 (0.20) |
| Accuracy (d') | 2.24 (0.36) | 1.84 (0.73) | 2.12 (0.68) | 1.59 (0.61) | 1.95 (0.65) |
| | Vision | | | | |
| Reaction Time (s) | 1.71 (0.08) | 1.74 (0.17) | 1.81 (0.16) | 1.80 (0.15) | 1.77 (0.15) |
| Accuracy (d') | 1.13 (0.53) | 1.31 (0.43) | 1.36 (0.36) | 1.71 (0.73) | 1.38 (0.57) |

---

5. Individual subject means are presented in Appendix B2.

As can be seen, on average the tactile to auditory comparisons were performed both faster and more accurately than were the visual to auditory comparisons. These descriptive trends were confirmed by ANOVA, with a significant effect due to modality being evident in both the accuracy, $F(1,32)=7.94$, $p<.01$, and reaction time, $F(1,32)=5.93$, $p<.05$, data.

As further support for the general superiority of tactile to auditory comparisons over visual to auditory ones it should be noted that only the tactile-spatiotemporal and tactile-spatial-delayed to auditory comparisons yielded worse scores on either of the performance indices than were recorded in the best performed of the four visual to auditory comparisons. There is, however, evidence that both these results were influenced by the subjects in those conditions employing marked speed/accuracy tradeoffs. First, although yielding relatively low accuracy scores, the tactile-spatiotemporal to auditory comparison task also yielded the fastest response times of all conditions. Likewise, although yielding the slowest response latencies of any condition, the tactile-spatial-delayed to auditory comparison task was the second best performed of all in terms of accuracy. Hence, these observations make it difficult to conclude that either task was performed less well than any of the visual to auditory comparison tasks.

Table 3.3 also shows differences in reaction time and accuracy between the four styles of information distribution used within each modality. While this trend did prove

significant in the case of reaction time, $F(3,32)=4.25$, $p<.05$, it did not reach significance in the case of accuracy, $F(3,32)=0.11$, $p>.05$. Finally, although no significant interaction between target stimulus presentation modality and display style was found in the case of accuracy, $F(3,32)=1.65$, $p>.05$, a significant interaction between these two factors was evident in the case of response speed, $F(3,32)=3.09$, $p<.05$.

In an attempt to identify the source of the significant interaction observed in the reaction time data post hoc comparisons were performed between each presentation method within each modality using Fisher's Least Significant Difference Test[6] (Kaplan, 1987). Table 3.4 summarizes these comparisons, showing that only the differences in speed between the tactile-temporal and tactile-spatial-delayed, $t(32)=3.19$, $p<.01$, tactile-spatial and tactile-spatiotemporal, $t(32)=2.48$, $p<.05$, and tactile-spatial-delayed and tactile-spatiotemporal, $t(32)=4.27$, $p<.001$, conditions reached significance. These results suggest that the significant interaction between presentation modality and information distribution style involved a response speed deficit between the purely tactile-spatial tasks and the tactile tasks involving temporally distributed information which was not evident between those two types of visual task.

The results of these post-hoc comparisons between reaction time means, taken with the lack of a significant

---

6. This test is also known as Fisher's Protected t-test.

effect on response accuracy due to information distribution style, show that the spatial-delayed tasks in both modalities were performed no better than the spatial tasks. This suggests that the other tasks did not have an inherent preprocessing advantage over the spatial ones.

Table 3.4: Summary of post hoc comparisons between information distribution style reaction time means within the tactile and visual conditions in Experiment 2. (df = 32 in all cases)

| Comparison | Tactile | | Visual | |
|---|---|---|---|---|
| | t | p | t | p |
| Temporal vs Spatial | 1.39 | 0.17 | 0.23 | 0.80 |
| Temporal vs Spatial-Delayed | 3.19 | 0.003* | 1.08 | 0.29 |
| Temporal vs Spatiotemporal | 1.08 | 0.29 | 0.89 | 0.38 |
| Spatial vs Spatial-Delayed | 1.79 | 0.08 | 0.85 | 0.40 |
| Spatial vs Spatiotemporal | 2.48 | 0.02* | 0.67 | 0.52 |
| Spatial-Delayed vs Spatiotemporal | 4.27 | 0.0003* | 0.19 | 0.83 |

In general, these results show that the tactile to auditory comparisons were performed both faster and more accurately than were the visual to auditory comparison tasks. This result is consistent with the findings reported in Section 3.1, and hence is consistent with the hypothesis that tactile and auditory representations of patterns are more alike than are auditory and visual representations of the same patterns. Second, the style of information distribution used in the target displays was found to affect only the speed with which tactile to auditory comparisons were performed. In particular, responding was slower in the purely spatial tactile tasks than

in those tactile tasks involving temporally distributed information. Finally, as the tactile-spatial-delayed task was performed no better than the tactile-spatial one, it appears that this response speed deficit was not due to any preprocessing advantage inherent in the other tactile tasks.

As with the first cross modal comparison experiment, an obvious question is whether the observed difference in the ease with which the tactile to auditory and visual to auditory comparisons were performed was due to differences in the discriminability of the various tactile and visual displays used. This question is addressed in the following section.

### 3.4: Experiment 3: Display Discriminability and Information Distribution Style

This experiment was conducted to assess the ease with which the relative intensity of the two pulses presented in each target pair in Experiment 2 could be identified via the various types of target display employed in that experiment. If the superior performance observed in the tactile to auditory comparison tasks was due to those tactile displays being more discriminable than their visual equivalents then it was expected that the more intense pulse in each pair would be identified most accurately when presented via the tactile displays.

### 3.4.1: Method

Subjects. Twelve subjects participated in this experiment. All were students at the University of Tasmania with normal or corrected to normal vision.

70

Apparatus and Procedure. The stimuli used in this experiment were the same as the target stimuli used in Experiment 2 and were generated and displayed in the same way. The only modification was that the tactile and visual spatial-delayed displayed were not used in this experiment. The procedure followed was essentially the same as that described in Section 3.2. Each subject undertook 12 practice and 48 experimental trials with each of the six types of display. In order to control for practice and fatigue effects, the order in which the subjects undertook each task was counterbalanced. On each trial the subject was presented with the pulse-pair via the prescribed method, then identified which pulse was most intense via two-choice button press. Reaction timing was commenced at the onset of the display in the case of the spatial tasks and at the onset of the second pulse in the other tasks. This arrangement ensured that there was a constant delay of 550ms between the commencement of timing and the termination of the display regardless of presentation method.

3.4.2: Results and Discussion

Mean reaction times for correct responses and mean accuracy levels ($d'$) for each presentation method are shown in Table 3.5[7]. Dprimes were calculated in the manner described in Section 3.2.

Table 3.5 shows that while the tactile displays were discriminated more accurately than the visual ones, the reverse

------------------------------------------------------------

7. Individual subject means are presented in Appendix B3.

71

was true in the case of response speed. These descriptive observations were confirmed by ANOVA, which revealed a significant effect of presentation modality on both the speed, $F(1,11)=8.12$, $p<.05$, and accuracy, $F(1,11)=6.03$, $p<.05$, of subjects' responses. This suggests that the subjects applied different speed/accuracy tradeoff criteria in the visual and tactile conditions. The occurrence of this effect makes it difficult to assess which, if any, modality of presentation led to better performance on the task.

Table 3.5: Mean reaction times (sec.) for correct responses and mean accuracy (d') scores from Experiment 3 as a function of modality and information distribution style. Standard deviations are given in parentheses.

| | Information Distribution Style | | | |
| | Temporal | Spatial | Spatiotemporal | Mean |
|---|---|---|---|---|
| | Touch | | | |
| Reaction Time (s) | 0.66 (0.14) | 0.84 (0.18) | 0.72 (0.24) | 0.74 (0.21) |
| Accuracy (d') | 3.00 (0.79) | 1.79 (0.57) | 2.64 (0.67) | 2.48 (0.86) |
| | Vision | | | |
| Reaction Time (s) | 0.75 (0.24) | 0.59 (0.15) | 0.58 (0.13) | 0.64 (0.20) |
| Accuracy (d') | 1.86 (0.40) | 2.25 (0.36) | 2.25 (0.63) | 2.06 (0.52) |

It is also clear from Table 3.5 that subjects performance varied with the style of information distribution employed in the display. This trend was also confirmed by ANOVA, which revealed a significant effect on response accuracy

due to information distribution style, $F(2,22)=5.26$, $p<.05$. The effect of this factor on response speed did not, however, reach significance, $F(2,22)=2.67$, $p>.05$.

Table 3.5 shows that the effects of information distribution style on the speed and accuracy of subjects' responses differed between presentation modalities. In the case of touch, the temporal and spatiotemporal tasks were both performed much more efficiently than was the spatial task, while in the case of vision the spatial and spatiotemporal tasks were performed best. These descriptive observations were confirmed by the occurrence of a significant interaction between presentation modality and information distribution style evident in both the accuracy, $F(2,22)=15.76$, $p<.001$, and speed, $F(2,22)=14.89$, $p<.001$, data.

Post hoc analyses were conducted to clarify the nature of this interaction. It is clear that, for touch, temporal distribution of the stimuli led to better performance than spatial distribution. This was true both for speed, $F(1,11)=29.87$, $p<.001$, and for accuracy, $F(1,11)=139.0$, $p<.0001$. That is, responding was both faster and more accurate with temporal distribution.

For vision, the opposite results were obtained. Spatial distribution of the stimuli led to better performance than temporal distribution. This was true both for speed, $F(1,11)=9.08$, $p<.05$, and for accuracy, $F(1,11)=5.81$, $p<.05$. That is, responding was both faster and more accurate with spatial distribution.

In both modalities, spatiotemporal distribution led to results very similar to those found with the preferred distribution method. That is, for touch the reaction times for spatiotemporal distribution were not significantly different from those for temporal distribution, $F(1,11)=0.85$, $p>.05$, n.s., but were significantly shorter than those for spatial distribution, $F(1,11)=7.76$, $p<.05$. For vision, the reaction times for spatiotemporal distribution were not significantly different from those for spatial distribution, $F(1,11)=0.18$, $p>.05$, but were significantly shorter than those for temporal distribution, $F(1,11)=5.44$, $p<.05$. In each case, the results for accuracy mirrored those for speed.

As there was no evidence of subjects employing speed-accuracy tradeoffs within each modality, these results must reflect some differences in task difficulty between conditions. These differences in task difficulty may be attributed to variations in the coding strategy used by each modality, rather than to any procedural differences between presentation methods. The spatial and temporal tasks in each modality differed only in terms of those features which characterized them as being either spatial or temporal in nature, that is being presented at different locations or at different points in time. All other stimulus parameters, such as duration and intensity, were constant across conditions.

Further, these features which distinguished the spatial and temporal tasks were combined in the spatiotemporal tasks, which were performed as well as the 'preferred' tasks in each

modality. The presence of 'non-preferred' information (spatial for touch, temporal for vision) thus did not disrupt or prevent efficient processing of the 'preferred' information, but merely led to less efficient processing when it was presented alone.

Several conclusions can be drawn from this experiment. As subjects applied different speed/accuracy tradeoff criteria in the tactile and visual tasks, resulting in superior speed in the visual ones and superior accuracy in the tactile tasks, there is no compelling evidence that either mode of display was superior. In the absence of strong evidence for the tactile displays being more discriminable than the visual ones it cannot be concluded that the superior performance overall on the tactile to auditory comparison tasks observed in Experiment 2 was due to the subjects in the visual to auditory comparison tasks being unable to efficiently extract the required information from the visual displays. Thus the hypothesis that there is some fundamental incompatibility between the auditory and visual representations of a pattern, which does not exist between auditory and tactile versions of that pattern, remains the best account of those results.

While Experiment 1 merely hinted that there may be systematic variations in the discriminability of visual patterns depending upon the style of information distribution employed, the present study strongly confirmed the existence of such differences. Vision dealt with these patterns best when they were distributed either spatially or spatiotemporally, with much poorer performance being observed when the patterns

were distributed temporally. Differences in performance were also observed between the three styles of tactile display, but the relationship between information distribution style and performance was different from that in the visual modality. In this case it was the spatially distributed display which was poorly discriminated.

A cavil on the acceptance of the finding that touch processed the temporally distributed displays more efficiently than the spatially distributed ones it the possibility that these differences in performance were artifacts of the particular ways in which the temporally and spatially distributed displays were presented to the subjects. Each pulse in the tactile-temporal display was presented simultaneously to the subjects' first and third finger-tips, while each pulse in the tactile-spatial display was presented to only one finger-tip. Perhaps this arrangement favoured the temporally distributed display via a process of spatial summation (Békésy, 1959) between the concurrent presentations of each pulse in the temporally distributed displays. Such a process may have amplified the percept generated by each pulse, thus sharpening the difference in amplitude between these two sequentially presented pulses. On the other hand, both pulses in the tactile-spatial display were presented simultaneously to different finger-tips, and thus could not have benefited from this type of summation effect.

Second, the choice of the first and third finger-tips as the presentation sites for the tactile-spatial displays may

have provided the subjects with a particularly difficult comparison. Perhaps comparisons between a different pair of fingers would show that performance on a tactile-spatial task is not inevitably inferior to that on the tactile-temporal task. For example, the sites stimulated in this tactile-spatial display lay within the same dermatome[8] (Kalat, 1984). Perhaps this lead to detrimental neural interactions between the two stimuli which would not arise between pairs of sites selected from different dermatomes. The following experiment tested whether either of these alternative explanations could account for the present results.

### 3.5: Experiment 4: A Confirmation and Extension of the Tactile Data from Experiment 3

The purpose of this experiment was to test whether the superior performance observed on the tactile-temporal task compared to that observed in the tactile-spatial one in Experiment 3 was due to the particular choice of spatial and temporal presentation methods and sites used in that study. In particular, this experiment attempted to test whether the high level of performance observed in the tactile-temporal condition in Experiment 3 was due to a beneficial process of spatial summation, and whether the poor performance observed in the tactile-spatial task was due to a site-specific deficit, perhaps resulting from the stimulation of two sites within the same dermatome

---

8. A dermatome is an area of skin which is supplied with afferent nerve fibers by a single spinal nerve.

### 3.5.1:  Method

Subjects.  A total of 17 undergraduate volunteers from
the University of Tasmania participated in this experiment.
Eight of these undertook the tactile-spatial tasks, while nine
of them undertook the tactile-temporal tasks.

Apparatus and Procedure.  The stimuli used in this
experiment were identical to those used in the tactile-temporal
and tactile-spatial conditions in Experiment 3. The manner in
which they were present to the subjects was, however, varied
from that used in Experiments 2 and 3.

Three different methods of tactile-temporal presentation
were used. In all cases the two pulses in each pulse-pair were
presented sequentially. In the tactile-temporal-both task,
each pulse was presented simultaneously to both the subject's
first and third fingertips, just as in Experiments 2 and 3.  In
the tactile-temporal-first task each pulse was presented in
turn to the subject's index fingertip, while each pulse was
presented in turn to the subject's third fingertip in the
tactile-temporal-third task. As the temporal-first and
temporal-third conditions did not involve simultaneous
presentation of each pulse to two fingertips, neither of these
conditions could support the potentially beneficial spatial
summation effect which may have occurred in the tactile-
temporal condition in Experiment 3.

Two different methods of tactile-spatial presentation
were used. In the tactile-spatial-13 task both pulses were
presented simultaneously, one to the subject's first fingertip,

the other to the subject's third fingertip. A similar procedure was followed in the tactile-spatial-14 task except that the two pulses were presented simultaneously to the subjects first and fourth fingertips respectively. Aside from the obvious difference between these two methods in terms of the sites stimulated, the spatial-14 condition differed from the spatial-13 condition in that it did not involve stimulating sites which lay within the same dermatome.

Each subject undertook either all three of the tactile-temporal tasks or both the tactile-spatial tasks. The order in which they undertook these tasks was counterbalanced across subjects to control for practice and fatigue effects. The procedure followed during the experiment was identical to that used in Experiment 3.

### 3.5.2: Results and Discussion

Mean reaction times for correct responses and mean accuracy levels ($d'$) for each presentation method are shown in Table 3.6[9]. Dprimes were calculated in the manner described in Section 3.2.

As can be seen, on average the tactile-temporal tasks were performed more efficiently than were the tactile-spatial tasks, just as in Experiment 3. Inferential analysis via independent sample t-tests confirmed this descriptive observation, revealing that the temporal tasks were performed significantly faster, $t(15)=3.69$, $p=.001$, and significantly

---

9. Individual subject means are presented in Appendix B4.

more accurately, $\underline{t}(15)=2.36$, $\underline{p}<.05$, then were the spatial ones.

Table 3.6 also shows that, while the accuracy levels recorded in the three temporal conditions were similar, response speed did appear to differ between conditions. In particular, the temporal-third task was performed more slowly than the temporal-first task, with an intermediate level of response speed being evident on the tactile-both task. These observations were confirmed by repeated measures ANOVA which revealed a significant effect of temporal presentation method on response speed, $\underline{F}(2,16)=5.29$, $\underline{p}<.05$, but no significant effect of this manipulation on response accuracy, $\underline{F}(2,16)=0.39$, $\underline{p}>.05$. A series of post-hoc contrasts confirmed the descriptive observation that this effect was primarily due to the slow response speed evident on the temporal-third task. The temporal-third task was performed significantly slower than the temporal-first task, $\underline{F}(1,8)=10.37$, $\underline{p}<.05$, while no significant difference in speed was evident either between the temporal-both and temporal-first tasks, $\underline{F}(1,8)=1.65$, $\underline{p}>.05$, or between the temporal-both and temporal-third tasks, $\underline{F}(1,8)=4.89$, $\underline{p}>.05$. Similar post hoc comparisons revealed no significant differences in accuracy between either the temporal-both and temporal-first conditions, $\underline{F}(1,8)=0.40$, $\underline{p}>.05$, the temporal-both and temporal-third conditions, $\underline{F}(1,8)=0.05$, $\underline{p}>.05$, or the temporal-first and temporal-third conditions, $\underline{F}(1,8)=0.66$, $\underline{p}>.05$.

These results show that the tactile-temporal presentation method used in Experiment 3 (the tactile-both

method) did not lead to superior performance, either in terms of speed or accuracy, than was achieved with the two alternate temporal presentation methods used here. As neither of these tasks could support the type of beneficial spatial summation effect which was proposed to account for the superior performance observed in the tactile-temporal condition in Experiment 3, these results are inconsistent with that explanation. In more general terms, these results show that performance on tactile-temporal tasks is consistently better than on tactile-spatial tasks across a range of presentation methods.

Table 3.6: Mean reaction times (sec.) for correct responses and mean accuracy (d') scores from Experiment 4 as a function of information distribution style and presentation method. Standard deviations are given in parentheses.

| | Temporal Distribution | | | |
| --- | --- | --- | --- | --- |
| | Both | First | Third | Mean |
| Reaction Time (s) | 0.59 (0.11) | 0.53 (0.09) | 0.64 (0.11) | 0.59 (0.11) |
| Accuracy (d') | 2.42 (0.81) | 2.55 (1.13) | 2.37 (0.94) | 2.44 (0.97) |
| | Spatial Distribution | | | |
| | 1-3 | 1-4 | | Mean |
| Reaction Time (s) | 0.70 (0.09) | 0.75 (0.11) | | 0.72 (0.11) |
| Accuracy (d') | 2.01 (0.70) | 1.55 (0.54) | | 1.78 (0.67) |

Finally, Table 3.6 indicates that the spatial-13 task was performed both faster and more accurately than was the spatial-14 task. While analysis by ANOVA confirmed that a

significant difference existed between these two tasks in terms of response speed, $\underline{F}(1,7)=15.84$, $\underline{p}<.01$, no significant difference in terms of accuracy was found, $\underline{F}(1,7)=3.26$, $\underline{p}>.05$.

As the spatial-13 task was performed at least as well as was the spatial-14 task, it does not appear likely that the poor performance observed in the tactile-spatial condition in Experiment 3, which employed the spatial-13 task presentation method, was due to detrimental effects arising from the stimulation of sites from within the same dermatome. In conclusion, while other combinations of stimulated sites might lead to better levels of performance on a tactile-spatial task than were observed in Experiments 3 and 4, neither of these experiments provide any evidence that tactile-spatial tasks can be performed more efficiently than tactile-temporal ones.

### 3.6: Summary and Conclusions

In summary, this series of experiments has shown that comparisons between auditory and tactile stimuli are easier to perform than are comparisons between auditory and visual stimuli. This was true both in the case of speech derived and synthetic patterns, and did not appear to depend upon the format in which the stimuli were presented to the tactile and visual modalities. This latter observation is inconsistent with the results obtained by Bryden (1972). Further, the superiority of auditory to tactile comparisons does not seem to result from differences in the discriminability of the stimuli presented via the various types of visual and tactile displays employed.

Taken together, these results indicate that the observed

differences in the ease with which visual and tactile stimuli can be compared with auditory ones must arise at the time when the perceptual representations generated by the different modalities are being compared. The most parsimonious explanation of why the auditory and visual perceptual representations are more difficult to compare is that these representations are less alike than are auditory and tactile ones. In this case the relatively poor speed and accuracy with which the auditory and visual stimuli were compared is consistent with previous demonstrations that comparisons between different modes of perceptual representation impose a penalty in terms of both speed and accuracy of judgment (Bjorkman, 1967; Ittyerah & Broota, 1983).

The present results also suggest one way in which auditory and tactile representations may differ from visual ones. Experiments 1 and 3 showed that vision processes spatially and spatiotemporally distributed displays more efficiently than temporally distributed ones. In contrast, Experiment 3 and 4 demonstrated that touch may process spatially distributed displays less efficiently than these other kinds. If hearing, like touch, also excels in processing temporally distributed information, then this may contribute to the greater similarity evident between auditory and tactile perceptual representations.

With this hypothesis in mind, the following chapter reports a series of experiments which attempted to confirm the existence of the within modality differences in information

processing efficiency found here, and to relate them to the processing of spatially and temporally distributed information in the auditory modality. Further, these new experiments attempted to broaden the theoretical significance of these findings beyond the issue of the suitability of touch and vision for the presentation of speech transforms.

**Chapter 4: The Processing of Spatially and Temporally Distributed Information in Different Modalities.**

In the previous chapter it was suggested that one reason why auditory and tactile perceptual representations are more alike than auditory and visual representations may be that both hearing and touch, in contrast to vision, deal more efficiently with temporally distributed information than with spatially distributed information. This chapter investigates the extent to which the tactile and auditory modalities can be characterized as primarily temporal domains, and the extent to which the visual modality can be described as a primarily spatial domain. These questions are also relevant to the resolution of a currently controversial issue, the so-called Space:Time::Vision:Audition analogy.

### 4.1: The Space:Time::Vision:Audition Analogy

It has traditionally been argued that hearing is predominantly a temporal domain while vision is primarily a spatial sense, a view which has recently been challenged by Handel (1988a, 1988b). He argued that such a distinction between senses ignores the inherently spatiotemporal nature of both auditory and visual events. Kubovy (1988) has argued in turn that the spatiality of auditory stimuli, and the temporality of visual ones, are minimal, restricted, and largely absent from the perceptual experience of auditory and visual events.

As Kubovy (1988) noted, Handel's contention that atemporal visual events and nonspatial visual ones do not exist is trivial because the concept of an "event" implies spatiotemporal change. This does not, however, mean that

86

hearing and vision treat these spatiotemporal events in the same way.

While auditory events, such as spoken words, have spatial coordinates, and visual events have temporal ones, these features seem largely irrelevant in the recognition and identification of auditory and visual stimuli. Spoken words are primarily characterized by changes across time in the amplitude and frequency of the acoustic signal, rather than by variations in these parameters across spatial coordinates. Likewise, although less strongly, variations in characteristics like luminance and hue across space, rather than similar changes across time, are the primary features which uniquely identify a visual stimulus as a particular object. Simply put, although auditory stimuli may have spatially distributed aspects, and visual stimuli temporally distributed ones, these are not the primary cues used by each modality to identify stimuli. Most often, speech sounds are identified via their temporally distributed aspects, while visual stimuli are primarily identified via spatially distributed cues.

The intent of the preceding discussion was to make it clear that if vision is mostly concerned with changes across space, and audition with changes across time, then these different emphases are properties of the respective processing systems rather than of the auditory or visual events themselves. To say that vision is a spatial sense is to say that the visual system is more sensitive to, or more attuned to, or more efficient at the processing of, information which

is distributed across space rather than across time. Likewise, to say that audition is a temporal sense is to say that the auditory system is more sensitive to, or more attuned to, or more efficient at the processing of, information which is distributed across time rather than across space.

With this reformulation, the question of whether, or to what extent, vision and hearing are spatial and/or temporal senses stops being a question about an analogy (Handel, 1988a) or about "folk psychology" (Kubovy, 1988). Instead, it becomes an empirical question about the relative efficiency of each sense modality in processing information which is distributed across either time or space. This chapter addresses the issue of relative processing efficiency and extends it beyond vision and hearing into the tactile modality.

## 4.2: Spatial and Temporal Processing in Hearing and Vision

One way to assess the efficiency with which spatially and temporally distributed information is processed by hearing and vision is by comparing spatial and temporal two-point limens across modalities. As Handel (1988a) and Julesz and Hirsh (1972) have noted, such a comparison does show that hearing has the finer temporal, and vision the finer spatial, resolution.

Handel (1988a) argued that this observation has only marginal significance in an assessment of the Space:Time::Vision:Audition analogy because the true test of this analogy lies in the spatial and temporal nature of auditory and visual stimuli, rather than in the efficiency with

88

which each modality deals with these spatial and temporal changes. However, granted the preceding arguments showing that it is indeed the processing capacities of the auditory and visual modalities, rather than the properties of auditory and visual events, which determine the validity of the Space:Time::Vision:Audition analogy, this argument against the evidence from auditory and visual threshold studies cannot be accepted.

The rejection of Handel's (1988a) argument against the use of threshold data when assessing this analogy does not mean that we should rely solely on this source of evidence. In Chapter 2 it was argued that an analysis of threshold data may not be the best way to assess the information processing capacity of the tactile modality due both to the strong dependence of the obtained threshold values on the particular procedures and/or stimuli employed, and the limited extent to which these values may be predictive of performance on more complex perceptual and cognitive tasks. These factors may also limit the reliability of an assessment of the Space:Time::Vision:Audition analogy which is based purely on threshold data, in which case an alternate and/or complementary source of evidence is desirable.

One alternate way to test this analogy is through a comparisons of auditory and visual performance on supra-threshold tasks. In this case, both within and between modality contrasts are possible. Within modality contrasts have an advantage over between modality contrasts in that they do not

require the stimuli presented to each modality to be matched. Any differences revealed by between modality comparisons can be interpreted as reflecting differences in general processing ability only if it is assumed that the stimuli presented to each modality are equivalent in terms of their sensory attributes. When within modality contrasts are used, the relevant question becomes the relative efficiency with which each modality processes each style of information, rather than whether one modality processes a particular style of information more efficiently than does the other modality. Indeed, the Space:Time::Vision:Audition analogy would be supported by a demonstration of differences in the relative efficiency with which spatially and temporally distributed information are processed by hearing and vision, even if one modality exhibited superior absolute levels of performance on both types of task. Such a demonstration would still imply that a differences in information processing style specialization exists between these two modalities.

Aside from the experiments reported in the previous chapter, a number of studies have demonstrated within-modality differences in the processing of spatially and temporally distributed information using supra-threshold tasks. Metcalfe, Glavanov and Murdock (1981) investigated the processing of spatially and temporally distributed information in hearing and vision. They presented subjects with either auditory or visual sets of words in a spatially and temporally distinct order, then asked the subjects to recall either the spatial or

temporal order in which the lists were presented. Temporal order was recalled better than spatial order in the auditory condition, while spatial order was recalled better than temporal order in the visual condition.

Studies of the retention of spatial aspects of spoken words have yielded results in the auditory domain compatible with those of Metcalfe et al. (1981). Haberlandt and Baillet (1977) found that subjects could only accurately recall the location of spoken words at the expense of word identification accuracy. Similar results have been obtained by Geiselman and Bellezza (1976) using spoken sentences as the targets. Indeed, they found that subjects could not accurately recall the location of spoken sentences unless previously instructed to attend to this aspect of the stimuli. While not proving that hearing and vision differ in terms of the efficiency with which spatially and temporally distributed information is processed, these results are consistent with the earlier assertion that spatially distributed aspects of spoken words are not normally a significant factor in speech perception.

O'Connor and Hermelin (1972) also found differences in the processing of spatially and temporally distributed aspects of visual and auditory stimuli. They presented subjects with sets of three auditory or visual stimuli in which the temporal and spatial order of presentation were manipulated so that the middle stimulus in temporal terms never corresponded with the spatially central stimulus. When asked to recall the middle stimulus, subjects were more likely to identify the temporal

median in the auditory condition and the spatial median in the visual condition. They concluded that auditory presentation encourages temporal processing, while visual presentation promotes spatial processing. An alternative explanation is that there may have been some bias either in the instructions given to the subjects, or in the subjects' understanding of the term "middle", which lead them to interpret middle to have a spatial referent for visual stimuli and a temporal referent for auditory stimuli.

A number of other studies have shown differences in the efficiency with which temporally distributed information is processed by hearing and vision. Gault and Goodfellow (1938) employed a same/different task similar to that used in the cross-modal comparison experiments described in Chapter 3. They presented subjects with pairs of either auditory or visual temporally distributed patterns and asked the subjects to compare each pair of patterns on the basis of their rhythm. The subjects averaged 85% correct on the auditory task but only 75% correct on the visual task. This difference did not decrease with training. Similar results have been reported in subsequent studies (Garner & Gottwald, 1968; Rubinstein & Gruenberg, 1971).

In conclusion, it is clear that temporally distributed information is processed better by the sense of hearing than by vision both in threshold and suprathreshold tasks. Second, vision processes spatially distributed information better than hearing in threshold tasks. Finally, it appears that hearing

92

processes temporally distributed information more efficiently than spatially distributed information in both threshold and suprathreshold tasks, with the reverse being true in the case of vision. Clearly, all these observations are consistent with an efficiency of processing formulation of the Space:Time::Vision:Hearing analogy.

### 4.3: The Tactile Modality.

While threshold studies show that the spatial and temporal resolution of touch lies between that of hearing and vision (Kirman, 1973), no previous studies have directly investigated the relative efficiency with which spatially and temporally distributed information are processed by the sense of touch. It has been found that tracing the elements of alphabetic characters out via Optacon type displays leads to superior levels of performance than presenting the whole representation at once (Beauchamp, Matheson, and Scadden, 1971; Saida et al., cited in Loomis, 1981). In the case of Saida et al.'s study, accuracy improved from 25-30% correct in static mode to 90% correct in traced mode. These results clearly suggest that tactile perception is impeded if information is presented in a purely spatially distributed form.

Although other studies have suggested that similarities exist between auditory and tactile processes in general (Békésy, 1955, 1957a, 1957b, 1959; Eilers et al., 1988; Freides, 1974; Handel and Buffardi, 1968), and in the way each modality deals with temporal variations within patterns (Marks, 1987a), these studies did not consider the efficiency with

93

which spatially and temporally distributed information are processed within these two modalities.

In summary, the results reported in Chapter 3 suggest that touch is better suited to the processing of temporally distributed information, while vision excels in the processing of spatially distributed information. In the case of vision, other studies have supported these results, while, in the case of audition, it appears that temporally distributed information is processed more efficiently than spatially distributed information. These results show one way in which touch is like hearing, and, conversely, one way in which vision differs from both touch and hearing. Finally, within modality differences in the relative efficiency with which spatially and temporally distributed information are processed, like those found in Experiment 3, seem central to the resolution of the dispute regarding the Space:Time::Vision:Audition analogy. The experiments reported in the remainder of this chapter were conducted to confirm the results of Experiment 3, and to extend those results so that the theoretical issues discussed above might be clarified.

## 4.4: Experiment 5: A Replication and Extension of Experiment 3

This experiment was essentially a replication of Experiment 3. The subjects were again presented with pairs of tactile or visual pulses, although the two tactile pulses now differed in frequency rather than in amplitude, while the visual stimuli now differed in length rather than in luminance. The subjects' task was to identify which tactile pulse was

higher in frequency or which visual bar was longer. Although these modifications may weaken the analogy between the stimuli and amplitude variations in two syllable words, it did allow the generality of the effects found in Experiment 3 to be assessed. If those effects represent general properties of the tactile and visual modalities, then it was expected that this task would be performed worst in the tactile modality when the pulses were presented spatially, and worst in the visual modality when they were presented temporally.

4.4.1: Method

Subjects. Twelve undergraduate volunteers participated in this experiment. All had normal or corrected to normal vision.

Apparatus and Procedure. As in Experiment 3, the stimuli were 48 pairs of tactile or visual pulses distributed either spatially, temporally, or spatiotemporally. The tactile pulses were square-wave vibrations differing in temporal frequency so that one pulse in each pair was always higher in frequency than the other. The frequency of the high-frequency pulse in each pair was set at either 300, 500, 700, or 900 Hz with the low-frequency pulse always being set 200 Hz lower. The amplitude of each pulse was adjusted to compensate for the differences in sensitivity of the skin to signals of different frequency.

With two exceptions, these stimuli were generated and presented in the same three ways as in Experiment 3. First, different precautions were taken against auditory leakage from the transducers. In this case the acoustic output of the

vibrators was masked by 60 dB narrow-band white noise presented for the duration of each experimental session through Senheiser HD22 headphones. Second, the time-course of the tactile-spatiotemporal display was altered so that it was identical to that used in the visual-spatiotemporal condition. In Experiment 3 the second pulse in this tactile display was presented 300 ms after the termination of the first pulse. Now the first pulse was presented for the duration of the display with the second pulse being presented 850 ms after the onset of the first pulse.

The visual stimuli were pairs of luminous bars differing in length. The longer bar in each pair extended across either 2.0, 2.3, 2.6, or 2.9 degrees of visual field while the shorter bar always extended across 0.3 degrees of visual field less than the longer bar. Aside from this modification, these visual stimuli were presented in the same three ways used in Experiment 3.

Each subject undertook 16 practice and 48 experimental trials in each of the six conditions. The procedure followed was the same as that described in Section 3.4.

## 4.4.2: Results and Discussion

Reaction time and accuracy (d') means for each task are presented in Table 4.1[1]. As can be seen, all the tactile tasks in this experiment were performed less accurately, and no more rapidly, than any of the tactile tasks in Experiment 3. This

---

1. Individual subject means are presented in Appendix B5.

suggests that the frequency comparison task used here was more difficult than the amplitude comparison task used in Experiment 3. In spite of this difference, the pattern of performance across the three tactile tasks is the same in both cases, particularly with respect to the relatively poor performance achieved on the tactile-spatial task in both experiments. The tactile-temporal task was performed both faster, $F(1,11)=12.1$, $p<.01$, and more accurately, $F(1,11)=27.6$, $p<.001$, than the tactile-spatial task[2]. On the other hand, performance on the tactile-spatiotemporal task was not significantly different from that on the tactile-temporal task either in terms of speed, $F(1,11)=2.3$, $p>.05$, or accuracy, $F(1,11)=1.3$, $p>.05$.

As in Experiment 3, these results support the hypotheses that temporally distributed information is processed more efficiently than spatially distributed information when the input modality is touch. It should also be noted that the changes in the time-course of the spatiotemporal display introduced in this experiment did not change the pattern of results.

The results obtained in the three visual conditions differed somewhat from those obtained in Experiment 3. The visual-spatial task was performed faster, $F(1,11)=10.7$, $p<0.01$, although no more accurately, $F(1,11)=0.25$, $p>.05$, than

---

2. As the hypotheses of this experiment predicted specific differences between levels of the information distribution style factor at each level of the modality of presentation factor, these predictions were tested with planned comparisons rather than testing for a significant interaction effect then applying post-hoc contrasts.

the visual-temporal task. However, the visual-spatiotemporal task was performed more rapidly than either the visual-temporal task, $F(1,11)=27.76$, $p<.001$, or the visual-spatial task, $F(1,11)=9.07$, $p<.05$. This task was not, however, performed significantly more accurately than either the visual-temporal, $F(1,11)=2.0$, $p>.05$, or visual-spatial, $F(1,11)=3.56$, $p>.05$, tasks.

Table 4.1: Mean reaction times (sec.) for correct responses and mean accuracy (d') scores for each visual and tactile task in Experiment 5. Standard deviations are given in parentheses.

| | Information Distribution Style | | |
|---|---|---|---|
| | Temporal | Spatial | Spatiotemporal |
| | Touch | | |
| Reaction Time (s) | 0.90 (0.25) | 1.19 (0.32) | 0.81 (0.28) |
| Accuracy (d') | 1.63 (0.74) | 0.16 (0.35) | 1.33 (0.49) |
| | Vision | | |
| Reaction Time (s) | 0.92 (0.30) | 0.76 (0.26) | 0.64 (0.18) |
| Accuracy (d') | 1.20 (0.46) | 1.12 (0.59) | 1.44 (0.70) |

These findings again show that spatially distributed information is processed more efficiently than temporally distributed information by vision, in that the spatial task was performed faster than the temporal task with no concurrent penalty in terms of accuracy. On the other hand, performance on the visual-spatiotemporal task differed from that observed in Experiment 3. In that case performance on the visual-spatial

98

and visual-spatiotemporal tasks was similar, while in the present experiment the visual-spatiotemporal task was performed faster than the visual-spatial one with no concurrent penalty in terms of accuracy.

The similarity in performance between the visual-spatial and visual-spatiotemporal tasks and between the tactile-temporal and tactile-spatiotemporal tasks in Experiment 3 suggested that the presence of non-preferred information (spatially distributed for touch, temporally distributed for vision) information did not affect the processing of preferred information (temporally distributed for touch, spatially distributed for vision) in either modality. The finding of superior visual processing of spatiotemporally distributed displays in this experiment shows that this effect does not always occur, at least in the case of vision. Instead, it appears that, depending upon the type of information to be compared, vision may attend to both spatially and temporally distributed aspects of a display gaining a processing advantage, perhaps due to the availability of redundant information, as a result.

## 4.5: Experiment 6: A Confirmation of the Visual Data from Experiment 5

Experiments 3 and 5 supported a distinction between touch and vision based upon differences in the relative efficiency with which spatially and temporally distributed information are processed within each modality. While the results of Experiment 3 and the tactile data from Experiment 5

99

indicated the non-preferred information available in spatiotemporally distributed displays did not affect the processing of the concurrently available preferred style of information, this effect was not found in Experiment 5 in the case of vision. As the visual-spatiotemporal tasks in each experiment differed only in terms of the particular stimulus dimension which was to be compared, that is either luminance or length, these conflicting results suggest that it may be this dimensional factor which determines the particular visual coding strategy applied when processing the display.

This experiment was conducted to again confirm that spatially distributed information is processed more efficiently than temporally distributed information by vision, as well as to clarify whether vision usually takes advantage of both spatially and temporally distributed aspects of spatiotemporally distributed displays. With these aims, the three visual tasks used in the previous two studies were rerun with the modification that the stimuli differed in spatial frequency rather than in luminance or length.

### 4.5.1: Method

Subjects. Eighteen subjects participated in this experiment. All had normal or corrected to normal visual acuity.

Apparatus and Procedure. The equipment and procedure used in this experiment were similar to those used in the visual tasks in Experiments 3 and 5 with the exception that subjects were now presented with pairs of square-wave gratings

differing in spatial frequency. Their task was to identify which grating was the lower in spatial frequency.

Each grating extended across 2.8 degrees of visual field with a height of 3.0 degrees of visual field. In the spatial and spatiotemporal distribution conditions the two gratings were separated by 1.0 degrees of visual field. The low frequency grating in each pair had a spatial frequency of either 1.96, 3.03, 3.40, or 5.17 c/deg, while the corresponding high frequency grating had a spatial frequency of either 3.03, 3.40, 5.17, or 10.53 c/deg respectively. Due to the way in which these displays were generated it was no longer possible to present the subjects with either trial-warning or feedback messages via the VDU. Instead, the subjects were presented with a warning tone via headphones to indicate the onset of each trial. No feedback was given.

4.5.2: Results and Discussion

Mean reaction times and dprime scores for each task are presented in Table 4.3[3]. As can be seen, these results were similar to those obtained in Experiment 5.

The visual-spatial task was performed both faster, $F(1,17)=4.79$, $p<.05$, and more accurately, $F(1,17)=9.55$, $p<.01$, than the visual-temporal task[4]. This finding again confirms the superiority of spatial distribution over temporal distribution in visual processing.

---

3. Individual subject means are presented in Appendix B6.

4. As in Experiment 4.4 inferential analysis of the data was conducted using a series of planned comparisons.

**Table 4.2: Mean reaction times (sec) for correct responses and mean accuracy (d') scores for each visual task in Experiment 6. Standard deviations are given in parentheses.**

|  | Information Distribution Style | | |
|---|---|---|---|
|  | Temporal | Spatial | Spatiotemporal |
|  | Vision | | |
| Reaction Time (s) | 0.60 (0.20) | 0.50 (0.07) | 0.43 (0.05) |
| Accuracy (d') | 3.06 (0.57) | 3.48 (0.46) | 3.59 (0.54) |

The visual-spatiotemporal task was performed significantly faster, $F(1,17)=29.31$, $p<.001$, although not significantly more accurately, $F(1,17)=0.63$, $p>.05$, than the visual-spatial task. Finally, the visual-spatiotemporal task was performed both faster, $F(1,17)=13.48$, $p<.01$, and more accurately, $F(1,17)=11.58$, $p<.01$, than the visual-temporal task. The consistency of these results with those obtained in Experiment 5 suggests vision may frequently take advantage of both spatially and temporally distributed aspects of spatiotemporally distributed displays.

### 4.6: General Discussion

Several consistent effects emerged from this series of experiments. First, vision and touch were found to differ in terms of the style of representation (that is spatially or temporally distributed) which they process most efficiently. Experiments 3, 4, and 5 demonstrated that temporally distributed information is processed more efficiently than spatially distributed information when the input modality is

touch. Experiments 3, 5, and 6 demonstrated that the reverse is true in the case of vision, with spatially distributed displays being processed more efficiently than temporally distributed displays.

Spatial distribution proved superior to temporal distribution in the visual modality when either spatial characteristics of the display, like length or spatial frequency, or non-spatial characteristics of the display, like luminance, were to be compared. Likewise, temporal distribution of information proved superior to spatial distribution in the tactile modality regardless of whether temporal features, like temporal frequency, or non-temporal features, like amplitude, were to be compared. As these modality specific differences in the efficiency with which spatially and temporally distributed information are processed were consistent across a range of stimulus dimensions in each modality, it can be concluded that they were not artifacts of the particular type of information presented.

In the case of vision, these findings are consistent with the results previously obtained by Metcalfe et al. (1981). The similarity in the findings of these two studies is noteworthy due to the different type of tasks used in each case. Metcalfe et al. (1981) demonstrated superior processing of spatially distributed information in vision using a short-term memory task, while the present experiments revealed the same effect using a speed and accuracy of discrimination task. This is further support for the observed effects reflecting a

general property of visual processing, rather than a special case limited to a single type of task and/or stimulus.

A similar conclusion cannot be drawn in the case of touch because no previous study has investigated this aspect of tactile processing. However, by comparing the tactile results of this study with the auditory task results of Metcalfe et al. (1981) it can be concluded that one similarity between touch and hearing is that both modalities, unlike vision, process temporally distributed information more efficiently than spatially distributed information.

While it appears that visual processing may sometimes be enhanced or facilitated by the presence of both spatially and temporally distributed information within a display, no evidence of such a beneficial interaction was found in the case of touch. As the occurrence of this facilitative interaction in the case of vision appeared to depend upon the particular stimulus characteristic to be processed, the possibility that tactile processing may also sometimes be facilitated by the presence of both these styles of information cannot be discounted.

Returning to the issues raised at the beginning of this chapter, several conclusions can be drawn. First, it appears that Handel's (1988a, 1988b) attempted refutation of the traditional Space:Time::Vision:Audition analogy is mistaken when the efficiency with which these types of information are processed is considered. Taking the present results with those of previous studies, it appears that both hearing and touch can

justifiably be characterized as primarily temporal domains with vision being primarily a spatial domain. Second, as touch and hearing both seem to excel in the processing of temporally distributed information, these results demonstrate one way in which touch and hearing are alike.

Of course, these results do not show that the auditory, tactile, and visual modalities are incapable of processing displays where the information is distributed in a "non-preferred" form. Nor do these results conflict with the "unity of the senses" position advocated by Marks (Marks, 1987a, 1987b). He cited evidence that the various senses show, ".. a number of deep and significant communalities" (Marks, 1987b, p. 384), and demonstrated that the mechanisms underlying temporal pattern processing are similar in hearing, vision, and touch (Marks, 1987a). What the present data shows is that these various senses are differentially effective in their use of any shared or similar temporal processing mechanisms.

In Chapter 2 it was argued that the demonstration of similarities between hearing and touch may allow us to determine those areas in which minimal recoding of the auditory signal is required before presentation to the skin. As the defining characteristics of spoken words, that is changes in amplitude and frequency, are typically distributed across time, the present results suggest that tactile displays of speech should preserve this temporally distributed nature of the speech signal. Conversely, these results suggest that a visual display of speech may require some recoding of these temporally

105

distributed features into the spatially distributed order which it prefers.

It was established in Chapter 2 that one of the areas in which the speech signal must be recoded for presentation to the skin is in the mapping of auditory frequency into location on the skin. In this case, a display which preserves the temporally distributed nature of the speech signal, and implements a frequency to location transformation, must result in a spatiotemporally distributed, rather than temporally distributed, display. How would the skin deal with this style of display? Experiments 3 included a tactile spatiotemporal condition which presented vibrations differing in amplitude to different skin sites, thus representing a simple version of this type of display. The subjects' performance in discriminating this display was no worse than that observed with a purely temporal display. A similar result was achieved in Experiment 5 when different frequency vibrations, rather than different amplitude vibrations, were presented to the two skin sites. Thus, the present results reveal no difficulties with this type of tactile display.

Of course, this does not show that the skin is equipped to deal with the far more complex spatiotemporally distributed patterns necessary to accurately represent spoken words. The next two chapters of this thesis deal with a possible determinant of the tactile modality's ability to process these complex patterns, that is the effect of integrative and masking effects on tactile information processing efficiency.

**Chapter 5: The Role of Masking and Integration in the Processing of Tactile Patterns.**

# 5: Tactile Masking and Information Integration

Chapter 3 demonstrated that touch and hearing have an information processing affinity which may enhance the workability of tactile aids. Chapter 4 demonstrated that one aspect of this affinity is that both touch and hearing, unlike vision, process temporally distributed information more efficiently than spatially distributed information. Of course, the existence of these similarities alone dose not mean that touch has all the processing capacities required to deal with speech transforms. As was discussed in Chapter 2, another major factor in determining the potential of tactile prostheses is the extent to which the resolution of the skin is sufficient to deal with the demands imposed by the complexity of transforms of the speech signal. It was argued that, due to the complex and highly interactive effects of covarying the characteristics of tactile patterns, an accurate assessment of the acuity of the skin in this context is not likely to come from a simple analysis of parameters like the skin's two point spatial or temporal resolution.

One of the causes of this complex relationship between the characteristics of tactile patterns and the resulting percept, and, arguably, one of the primary limiters of tactile pattern perception capacity, is the occurrence of masking effects within the tactile processing system. This chapter investigates the extent to which these masking phenomena influence the capacity of the skin to deal effectively with complex patterns of stimulation, and consequently with speech transforms. Particular emphasis is given to divergent theories

on the effects of spatial and temporal proximity between the elements of tactile patterns on the intelligibility of those patterns.

## 5.1: What is Masking?

Masking is a phenomena which may be observed in all sensory systems. In the simplest sense, masking may be defined as the mechanism whereby the processing of one stimulus is impeded by the presence of another stimulus. It is conventional to refer to the stimulus whose processing is impeded as the target, and the stimulus causing this interference in the processing of the target as the mask (Gelfand, 1981).

Masking may be observed when the mask and target are presented at the same time (simultaneous masking), when the mask is presented before the target (forward masking), and when the mask is presented after the target (backward masking). Masking can manifest itself in two different ways. First, the mask can cause a reduction in sensitivity to the target. Simply put, the target must be more intense in the presence of the mask than in the absence of the mask in order to be detected. Second, although the target may remain detectable, the mask may reduce the ease with which the target can be recognized.

Studies of visual, auditory, and tactile masking effects have suggested two separate processes which may underlie masking effects; integration and interruption (Breitmeyer, 1984; Breitmeyer and Ganz, 1976, Kirman, 1984, Massaro, 1972, Michaels and Turvey, 1979, Turvey, 1973). Briefly, the interruption or erasure process only applies under backward

masking conditions and involves the arrival of the mask interrupting the processing of the target. The integration process applies to both backward and forward masking conditions and involves the neural representations of the mask and target becoming integrated into a single percept. While most theorist agree that integration is one of the mechanisms underlying masking, some, like Schultz and Eriksen (1977) and Felsten and Wasserman (1980), argue that interruption or erasure effects are best explained as the result of integrative mechanisms. The process of integration in tactile masking is discussed at greater length latter in this chapter.

Békésy (1957b,1958,1959) undertook one of the first systematic studies of tactile simultaneous masking phenomena. In one experiment he progressively decreased the distance between two uniformly vibrating tactile stimuli. When these two stimuli were presented at a spatial separation exceeding 3.5cm two separate percepts were felt. As the distance between the stimuli was decreased, so too did the perceived amplitude of the stimuli. Békésy attributed this effect to a process of lateral inhibition, similar to that observed along the basilar membrane or across the retina in the Mach bands phenomena. In this case, each stimulus appeared to be inhibiting the sensory response to the other. When the stimulus spacing was set below 2.5cm, the two stimuli merged into a single percept whose perceived magnitude was greater than that of either of the component stimuli. Békésy explained this phenomena as resulting from a process of spatial summation whereby the tactile system

combined the energy of the two stimuli. In another experiment Békésy (1957a) presented vibrations of differing frequency to five skin sites each separated by 2cm. Only the middle stimulus was perceived, yet its perceived amplitude was enhanced. This so called funneling effect thus involves both inhibitory and summation components.

This data shows that, even when only a small number of stimuli and no temporal variations are involved, the perceptual results of masking are complex. The existence of this complex relationship between display parameters, masking, and the resulting tactile percept suggests that, as in the case of tactile spatial, temporal, and frequency resolution, an analysis of basic masking functions is unlikely to fully predict the perceptual nature of complex spatiotemporal patterns presented to the skin. Of course, this does not mean that we can ignore simple masking effects when considering the way in which touch deals with spatiotemporal patterns of stimulation. The following section presents a brief review of factors effecting tactile masking phenomena.

## 5.2: Tactile Masking Phenomena.

Tactile masking has been studied using a diverse range of stimuli and psychophysical methods. Both vibrotactile and electrotactile presentation methods have been employed (Gilson, 1969a; Schmid, 1961) , with the target being either a patterned or energy stimulus (Craig, 1985a; Evans and Craig, 1986, 1987; Gescheider, Bolanowski, & Verrillo, 1989; Gilson, 1969a). Likewise, both patterned and energy maskers have been

employed (Craig, 1982b; Evans and Craig, 1986, 1987 Gilson, 1969a,1969b; Gescheider et al., 1989; Lechelt, 1986). Finally, the response methods used have included both threshold measurement and forced-choice procedures (Evans & Craig, 1986, 1987; Gilson, 1969a; Kirman, 1974, 1976; Schmid, 1961). Although Lechelt (1986) has charged that this diversity in the procedures used to investigate masking makes it difficult to make comparisons between studies, these diverse methods have led to reasonably consistent results.

5.2.1: Temporal Characteristics of Tactile Masking.

As was mentioned earlier, masking effects exist along a temporal continuum based on the delay between the target and mask. It has been found in the auditory, visual, and tactile (Craig, 1983b) modalities that it is the stimulus onset asynchrony (SOA) rather than the inter stimulus interval (ISI) between target and mask which determines the extent of masking. SOA is the time between the onsets of the target and mask, while ISI is the delay between the offset of the first stimulus (either the target or the mask depending respectively upon whether backward or forward masking is being induced) and the onset of the second stimulus.

This relationship between the extent of masking in the tactile modality and the temporal asynchrony between target and mask has been widely explored using most of the stimulus and psychometric variables mentioned above (Craig, 1983a; Evans, 1987; Evans & Craig, 1986, 1987; Gescheider, et al., 1989; Gilson, 1969a; Kirman, 1984,1986; Lechelt, 1986; Schmid, 1961).

Although some of these studies manipulated ISI rather than SOA, this difference does not alter the general shape of the resulting time-masking function. So long as the target and mask durations are held constant then the relationship *SOA = ISI + First Stimulus Duration* holds true.

Figure 5.1 presents the masking functions obtained by Evans (1987) and Lechelt (1986). These functions were selected because, aside from being consistent with those reported in the majority of studies, they were obtained using different procedures and different time-scales and thus highlight the effects of these differences on the shape of the masking function. Briefly, Lechelt (1986) used energy targets and measured the detection rates of these stimuli at brief SOA's (in the range -125ms to 75ms), while Evans (1987) employed patterned masks and targets and measured the recognition rates of these patterned targets over a relatively wide range of SOA's (between -526ms and 526ms SOA).

Both functions show maximum masking occurring when the target and mask are presented in very close temporal proximity. The extent of masking then decreases rapidly in both the forward and backward masking conditions as the delay between target and mask is increased from this point. Further, Evans' (1987) function shows that this rapid decrease in the extent of masking levels off when the SOA between target and mask is in the range 100 to 150ms. Beyond this SOA, a more gradual decrease in masking is evident, with recognition performance returning to unmasked levels at some point beyond 500ms SOA.
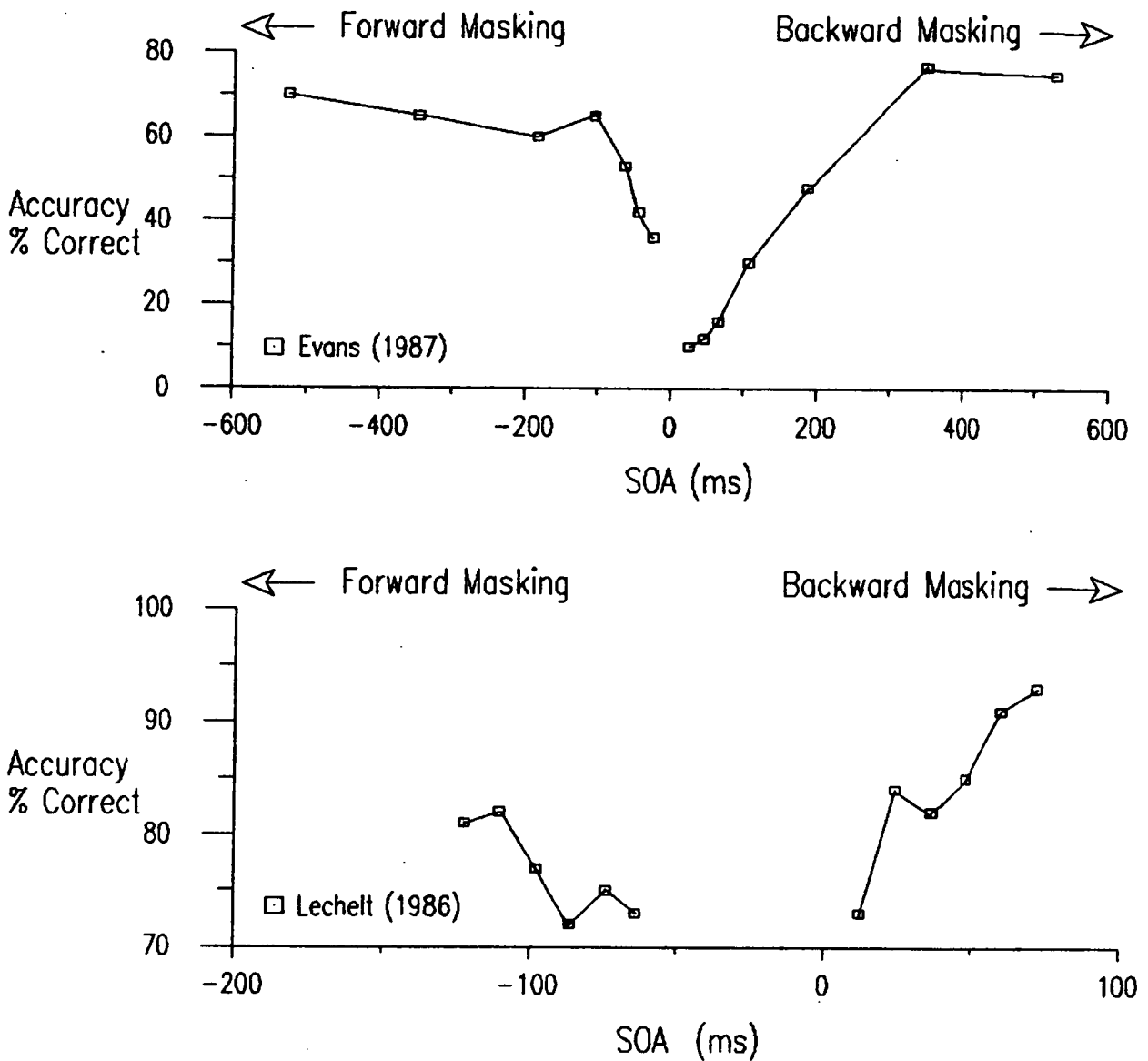
Figure 5.1: Tactile masking as a function of SOA. The upper graph is adapted from Evans (1987) and shows the recognition masking of a two-line pattern target by a single-line pattern mask. Masking is measured as the difference between target alone and target with mask recognition rates. The lower graph is adapted from Lechelt (1986) and shows the detection masking of an energy target by an energy mask. Masking is measured as the detection rate of the target. Lechelt's original ISI values have been converted to their SOA equivalents.

114

This leveling off in the rate of decline in the extent of masking is not evident in Lechelt's (1986) detection masking function due to the limited range of SOA's used. The actual point at which target detection/recognition performance returns its to unmasked level is unclear. Most studies have been restricted to SOA's less than 500ms, at which point performance was approaching unmasked levels. However, Evans and Craig (1987) and Gescheider et al. (1989) have reported evidence of masking extending beyond 1000ms SOA.

Where Evans' (1987) recognition masking function differs from Lechelt's (1986) detection masking function is in the relative extent of forward and backward masking evident. In particular, the recognition masking function shows more backward than forward masking occurring at SOA's below about 100ms, while the detection masking function shows the reverse, with more forward than backward masking being evident. As Gescheider et al. (1989) noted, backward masking is usually stronger than forward masking at brief SOA's when a recognition masking paradigm is used, with forward masking being at least as prominent as backward masking in detection masking paradigms.

Craig (1982a) suggested that this difference must reflect some fundamental difference in the processes involved in these two types of task. Kirman (1984) has argued that one difference between forward and backward masking is the stage in the tactile processing system at which the respective masking effects occur. He believes that forward masking is primarily a

peripheral interaction, while backward masking involves a strong central interaction component, a position consistent with findings in both the auditory and visual modalities (Massaro, 1972; Turvey, 1973). First, it appears that more forward than backward masking is found when the target and mask are presented to spatially proximate sites, with the reverse being true when the target and mask are presented to distant sites (Coquery and Amblard, 1973; Halliday and Mingay, 1961; Sherrick, 1964). Kirman (1984) noted that as there is greater opportunity for peripheral interaction between target and mask when they are presented close together on the skin, these demonstrations of optimal forward masking under such conditions are consistent with a peripheral explanation for this type of masking. Conversely, when target and mask are far apart, the opportunity for peripheral interaction is diminished, and thus the stronger masking force under these conditions, that is backward masking, must be primarily a central phenomenon.

## 5.2.2: Spatial Characteristics of Tactile Masking

The previous sub-section discussed the role of temporal proximity between target and mask in determining the extent of masking. This sub-section reviews the effects of spatial proximity on the level of masking observed.

Just as with temporal separation, it seems that the masking is a function of the spatial separation between target and mask both in the case of detection masking (Gescheider, Herman, and Phillips, 1970; Gilson, 1969a; Snyder, 1977) and recognition masking (Weisenberger, 1981 (cited in Craig,

1985a)). Gilson (1969a) measured the threshold of a 200ms 150Hz vibration presented to the subjects' thigh as the ipsilateral and contralateral longitudinal separation between the target and mask was varied from 0 to 75 cm. This involved progressively moving the mask from the thigh target site further up the trunk, with the most distant site used being on the arm near the shoulder. Figure 5.2 presents a summary of the results he obtained. In the ipsilateral condition, maximum masking was observed when the target and mask were presented to the same site, with a progressive decrease in masking as the mask was moved further away from the target site. A similar trend occurred in the contralateral condition, with the masking observed in this condition being similar to that in the ipsilateral condition.

The main finding of Gilson's (1969a) study, that is a decrease in masking with decreasing target-mask proximity, has been repeatedly confirmed (Gescheider et al., 1970; Snyder, 1977). However, his finding that contralateral and ipsilateral masks produce similar levels of masking has not received universal support. Both Gescheider et al. (1970) and Snyder (1977) have demonstrated that the type of task used to measure the extent of masking determines the relative strength of ipsilateral and contralateral masking. In particular, they found that tasks requiring the subject to detect the presence of the target, like that used by Gilson (1969a), result in more contralateral masking than do more complex tasks in which the subject attempts either to localize the target or compare
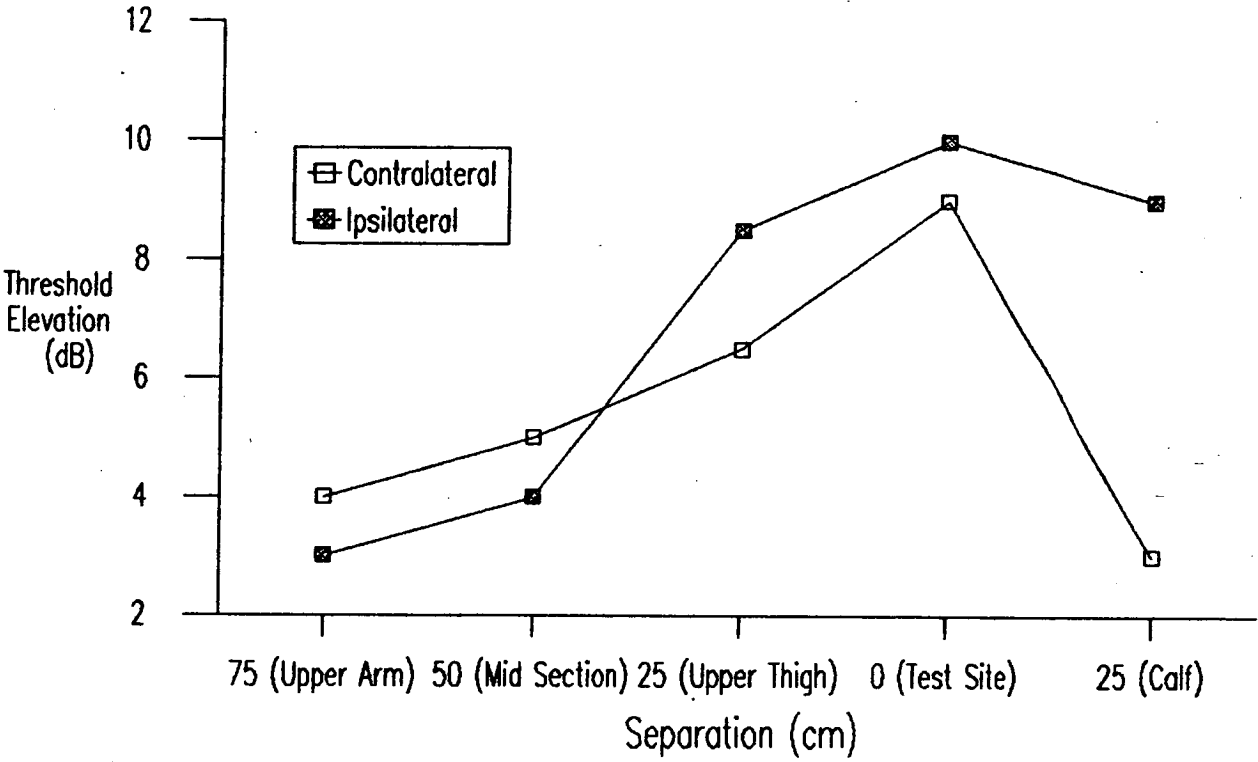
Figure 5.2: Tactile masking as a function of the spatial separation between the target and mask. This graph is derived from Gilson (1969a).

mask-alone and target-plus-mask displays. Neither study found differences in the extent of ipsilateral masking occurring with these two types of procedure.

Why do these two types of task lead to different patterns of ipsilateral and contralateral masking? Gescheider et al. (1970) suggested that simple detection tasks depend only upon our ability to detect the target, while the more complex procedures depend on our ability to discriminate changes in the general pattern of stimulation resulting from the target and mask. Gilson (1974) has expanded on this point, noting that certain masking effects, for example the apparent movement of the target towards the mask (Békésy, 1959), only occur under ipsilateral masking conditions. He argued that these effects make it more difficult to determine either whether the target was presented or where it was presented, thus decreasing performance on ipsilateral detection and localization tasks. However, the occurrence of these effects does not reduce the distinctiveness of the percept resulting from either target and mask or mask alone presentations. Thus performance on the more complex tasks used by Snyder (1977) and Gescheider et al. (1970) is not impeded under ipsilateral conditions relative to that observed in the absence of this type of interaction (that is under contralateral masking conditions).

5.2.3: Other Factors in Tactile Masking

The previous sub-sections have reviewed two of the main determinants of the strength of tactile masking, that is the spatial and temporal separation between target and mask. In

both cases, increasing the proximity of the target and mask has been shown to increase the extent of masking. Although this chapter is primarily concerned with the role of these factors in determining the intelligibility of tactile patterns, this sub-section briefly summarizes the masking effects of two other important target and mask characteristics; frequency and amplitude.

As was explained in sub-section 2.2.1, the tactile modality appears to contain four discrete channels for the processing of vibratory stimuli. Each channel is responsive to a particular range of frequencies, with the main distinction being drawn between the Pacinian channel, which is responsive to frequencies in the range 100 to 800Hz, and the Non-Pacinian channels, which respond to frequencies below 100Hz. It has been found that, while the masking functions obtained within each tactile channel are similar (Gescheider et al., 1989), masking generally occurs only when the target and mask frequencies activate the same processing channel (Bolanowski, et al., 1988; Gescheider, O'Malley, and Verrillo, 1983; Gescheider, Sklar, Van Doren, and Verrillo, 1985). However, Gescheider, Verrillo, and Van Doren (1982) have observed cross-channel masking at high mask amplitudes. They found that above a critical mask amplitude of 30 dB SL, the extent of cross-channel masking increases rapidly.

A final determinant of the extent of tactile masking is the relative intensity of the target and mask. It has consistently been found that the more intense the masker

relative to the target, the greater the extent of masking observed (Abramsky, Carmon, and Benton, 1971; Gescheider et al., 1970; Gescheider et al., 1982; Snyder, 1977). It appears that, provided the target and mask frequencies are selected to stimulate the same receptor populations, this increase in masking with mask amplitude is linear under simultaneous masking conditions (Gescheider et al., 1982; Snyder, 1977).

### 5.3: Masking and Pattern Perception : Competing Views

The previous section reviewed the main factors affecting the strength of tactile masking interactions. The remainder of this chapter investigates the extent to which these masking effects alter the intelligibility of complex tactile patterns.

In his seminal review of the perceptual basis of tactile communication systems, Kirman (1973) outlined_ two competing views on the role of masking phenomena in the process of tactile pattern perception. The first of these, here called the "Isolation Hypothesis", argues that the various tactile masking phenomena described in the preceding sections act to impede the accurate perception of tactile patterns. The Isolation Hypothesis provides both an explanation of the failure of various tactile communication systems and a recipe for minimizing these limiting factors. As Kirman (1973) explains, the poor performance then achieved with various tactile displays has commonly been attributed to the action of tactile masking effects. The corollary of this explanation is that these masking effects must be minimized before optimum tactile pattern perception performance can occur. Perhaps the most

emphatic expression of the Isolation Hypothesis is that offered by Brown, Nibarber, Ollie, and Solomon (1967). Their solution to the "problem" of interactions between stimuli in tactile patterns included the following elements; stimulating as few sites as possible, ensuring that these sites are widely distributed across the body surface, and stimulating only a single site at any given time. The goal of all these measures is to reduce masking between pattern elements by increasing the spatial and temporal separation between those elements.

The second position discussed by Kirman (1973), here called the "Integration Hypothesis", challenges this view, claiming instead that tactile masking phenomena reflect the operation of integrative information organization mechanisms which facilitate the perception of complex patterns. In Kirman's (1973) words; "Masking .. is but the negative manifestation of perceptual organization whose function is to detect relevant information in the environment, not to obscure it." (p. 64)

The core of this hypothesis is that the salient perceptual content of a pattern lies in the relationships between pattern elements. These relationships between pattern elements are represented at the sensory level by the interactions between these elements, rather than by the individual properties of the elements. The corollary of this is that sensory systems require the presence of these interactive or relational aspects of patterns if optimum pattern processing is to occur. This position is summarized by

Richardson and Frost (1977) who state;

"Like all perceptual systems, the cutaneous system responds best to relations between stimuli, not to individual point loci of stimulus energy. High acuity in vision and fine pitch perception are possible *because* of myriad interactions among receptors and not in spite of them." (p. 268)

What does this hypothesis imply about the role of masking in pattern discrimination, and consequently about the optimum spatial and temporal relationships between tactile pattern elements? When Kirman (1973) claimed that masking is the negative manifestation of this integrative process he was not, as it might appear, labeling the *process of masking* as simply an unfortunate by-product of perceptual integration. He believes that the process of masking is one of the means by which perceptual integration proceeds, with the disruptive nature of *observed masking phenomena* resulting from the types of task traditionally employed in studies of masking. In his words; "Most studies of masking do not allow the positive aspects of sensory organization to reveal themselves." (Kirman, 1974, p. 64)

If masking is one aspect of the process by which perceptual integration, and consequently pattern perception, occurs, then it follows that optimum pattern perception will not result when masking is minimized. On the other hand, this does not mean that the conditions which yield the maximum amount of masking are also those which lead to optimum pattern discrimination. For example, if we were presented with spoken

words at 10 times the normal rate then more masking than usual would occur between elements within the speech signal due to the closer temporal proximity of those elements. However, it is unlikely that these conditions would produce speech perception levels as great as those normally achieved. In this case the optimum level of masking for the purpose of pattern perception would seem to lie somewhere above the point where no masking occurs but below the point of maximum masking.

In summary, the Integration Hypothesis claims that masking represents the occurrence of a beneficial process of perceptual integration. The consequence of this in terms of the optimum level of spatial and temporal separation between elements in a tactile representation of speech is that these parameters should be set so that masking effects are strong, yet not necessarily at their maximum. In contrast, the Isolation Hypothesis argues that all masking effects are disruptive to pattern discrimination, and thus should be minimized in tactile communication systems, for example by decreasing the spatial and temporal proximity of pattern elements. The following section examines the evidence for and against the Isolation and Integration Hypotheses, and hence for and against their divergent prescriptions for the optimum spatial and temporal proximity of elements in tactile patterns.

### 5.4: The Case for the Isolation and Integration Hypotheses.

Neither the Integration or Isolation Hypothesis disputes that tactile masking does occur, nor that the extent of masking increases with the spatial and temporal proximity of the target

and mask. Where they differ is in terms of the extent to which the perception of complex patterns is either impeded or facilitated by masking. Consequently, this review will focus primarily upon that question.

5.4.1:    Masking and Integration

The central tenet of the Integration Hypothesis is that pattern perception proceeds on the basis of an integrative process of which masking is one component. Is there any evidence to suggest that masking results in or involves the generation of an integrated representation of the target and mask?

First, there is no doubt that the tactile modality is capable of some forms of integration. It has been demonstrated that the threshold of tactile stimuli decrease as a function of their duration, particularly at durations up to about 100ms, indicating the operation of a process of temporal summation (Gescheider & Joelson, 1983; Verrillo, 1965). Tactile summation has also been found between multiple stimuli. The spatial summation effect described by Békésy (1959), and discussed in Section 5.1, provides an example of the summation or integration of energy from two or more stimuli to produce a sensation whose magnitude is greater than that of any of its components.

Although there is some dispute concerning the mechanisms underlying masking, there is widespread agreement that temporal integration of the target and mask is one of the primary factors underlying masking effects in vision (Breitmeyer, 1984;

Breitmeyer & Ganz, 1976; Felsten & Wasserman, 1980; Massaro, 1975; Turvey, 1973). Integration is seen to occur between the target and mask in both backward and forward masking conditions, with the representations of these two stimuli merging at some point within the processing system to form a composite representation.

Craig (1980, 1981, 1982a) was amongst the first to propose that a process of integration also occurs in tactile masking. He worked with an Optacon based tactile display via which subjects received patterned stimuli, for example letters of the alphabet. He displayed the patterns in either static (that is all elements at once) or one of several sequential (that is elements or groups of elements displayed in turn) modes (Craig, 1981). At relatively long display times, he found that these different presentation methods led to divergent levels of pattern recognition accuracy. However, at very brief durations, performance was similar, both in terms of accuracy and error patterns, regardless of display mode. Craig (1981, 1982a) interpreted this as showing that the sequentially presented pattern elements were integrated together at brief display durations yielding a representation which was identical to that produced in the static mode. As in the case of temporal summation in detection tasks, Craig (1982a) found that this process of integration between pattern elements was primarily restricted to a temporal window of approximately 100ms duration.

Subsequent research has confirmed the role of

integration in tactile masking. Evans and Craig (1986) presented subjects with tactile patterns consisting of various one, two, or three line segments followed by a masking stimulus covering all the points included in any of the patterns. They found that subjects reported the presence of more lines than were actually presented, suggesting that elements of the mask were being integrated into the representation of the target. As in the previous demonstrations of tactile information integration, this effect ceased at SOA's greater than 100 ms.

Beyond this critical 100ms duration Evans and Craig (1986) still found evidence of masking. The difference at these longer SOA's was that the subjects' errors now reflected a failure to discriminate the target from other patterns containing the same number of lines, rather than misjudgements concerning the number of lines in the patterns. This finding suggests that some process other than integration is responsible for tactile backward masking at SOA's greater than 100ms. Evans and Craig (1986) attributed this long SOA masking effect to a process where-by the mask interferes in the extraction of higher-order relational details from the target. They found no evidence of interruption effects like those reported in some visual masking studies.

Evans and Craig (1987) and Evans (1987) reported similar effects in a forward masking design to those found in the above mentioned backward masking experiment. The primary difference between the results of these two studies was that masking was found to extend over a longer time period in the case of

forward masking.

An obvious question is what form the integrated representation of the target and mask takes. For example, does this representation consist of some aspects of the target mixed in with some aspects of the mask? Evans and Craig (1987) demonstrated that minimal masking occurs when the target and mask patterns are the same. This suggests that the integrated representation of the target and mask consists of an overlay of the features of the two stimuli, rather than a random mixture of features from each. Evans (1987) has supported this conclusion. He presented target and mask pairings such that if all features of both stimuli were preserved in the integrated representation then that representation would correspond with one of the forced choice responses available to the subjects. Under these conditions the subjects frequently identified the target as being the one which would result from such an integrated overlay of the target and mask. Clearly either a random or incomplete mixture of features from the target and mask is unlikely to result in a percept which is readily identifiable as that which would result from an "exact" overlay of all target and mask features.

While this data suggests that the integrated representation of the target and mask is a faithful overlay of the two stimuli, Evans (1987) has shown that the relative strength of the features contributed by each stimulus vary depending upon the order in which they are presented. In a backward masking situation subjects frequently gave the mask

pattern as their response to the identity of the target. This trend was absent in a forward masking situation, and was diminished in the backward masking condition when the relative amplitude of the target was increased. Evans (1987) interpreted these results as showing that the features of the temporally trailing stimulus were more strongly represented in the integrated percept than were those of the first stimulus.

In conclusion, it appears that at SOA's less than approximately 100ms the tactile system is capable of integrating information from successively presented patterns into a single perceptual representation which seems to preserve the features of both the target and the mask. Beyond this critical duration, tactile masking seems to be due to a higher level process of interference between target and mask. Finally, the observation that the perceptual representation of the target and mask consists of an integrated representation of the features of each stimulus is consistent with the Integration Hypothesis' claim that masking is one component in a general process of perceptual integration.

5.4.2:  Testing Whether Integration is Beneficial

The previous sub-section presented evidence showing that touch is capable of perceptual integration, particularly over brief time periods, and that this process is intimately associated with the occurrence of masking. Both these findings are consistent with the Integration Hypothesis. The pivotal question in assessing the Integration and Isolation Hypotheses thus remains whether this process of integration is either

beneficial or disruptive to the processing of complex tactile patterns. This sub-section primarily addresses the issue of how we should go about testing this question.

Craig (1985a) attempted to assess the effects of integration on the processing of tactile patterns. He presented random pairings of the letters "X" and "O" sequentially to subjects via an Optacon based display, and asked two types of question. First, the subjects were required to judge whether the two letters in a pair were the same or different. Second, the subjects were asked to identify the pair of letters as a whole, for example whether the pair was "O-O" or "O-X". Craig reasoned that the first task would require the subjects to focus upon aspects of the individual letters, a process which might be impeded by masking. On the other hand, he thought that the second task, might force the subjects to integrate the two stimuli. Performance on both tasks was found to be similarly poor.

To temper his failure to demonstrate beneficial tactile integration, Craig (1985a) noted that the types of patterns which he used may not have been amenable to constructive integration. For example, the letters "X" and "O" if integrated would yield, in Craig's words, an "indistinct blob" (1985a, p. 245), rather than a meaningful pattern.

Although not showing integration to be beneficial to pattern perception, Craig's (1985a) study does highlight the issues which must be addressed in testing the Integration Hypothesis. First, it is necessary to present the subject with

a task which allows any positive effects of tactile information integration to be revealed. Second, it is necessary to ensure that information is presented to the skin in a way which allows beneficial integrative effects to operate. Both these points are consistent with Kirman's (1973) criticism of attempts to extrapolate pattern perception ability from the results of traditional masking studies.

What is wrong, in this context, with the tasks used in traditional masking studies? Consider the letter and line discrimination tasks used by Evans and Craig (Craig, 1981,1982a,1985a; Evans, 1987; Evans and Craig, 1986, 1987). In these studies the subjects were presented with a meaningful target, such as a letter of the alphabet, in conjunction with either a similarly meaningful mask or an energy mask. The subjects' task was to identify either the target or some aspect of the target. Performance was best when the two patterns were presented far apart in time, and worst when they were temporally proximate.

Does this show that the Isolation Hypothesis is correct in arguing that the integrated representation of the target and mask which is generated at close temporal separations carries little useful information? In support of the Integration Hypothesis it can be argued that the tasks used in conventional masking studies do not allow an accurate assessment of this question. The integrated representation of the target and mask does not directly tell the perceiver what the *target* feels like, rather, it tells the perceiver what the *target-and-mask*

*event* feels like. Under these conditions, observers do not respond randomly. Evans and Craig (Evans, 1987;, Evans & Craig, 1986, 1987) found that the subjects' responses frequently reflected many aspects of how this integrated percept should appear. In this case, it is clear that, far from being devoid of information, much information was available in these integrated percepts. Further, this information did accurately reflect many features of the actual stimuli. Of course we may ask what use a perceptual representation is if the perceiver cannot extract certain key items of information, in this case the identity of one of the stimulus letters, from it. There are three answers to this question.

First, subjects can extract the identity of a letter or simple pattern from an integrated tactile representation with minimal experience provided the display only includes that one stimulus. In their masking studies Evans and Craig frequently measured masking in terms of the change in letter or pattern identification rates between target alone and target plus mask conditions. A tactile display of a single alphabetic character presented by any means is still subject to masking forces. The individual stimuli (vibrating pins in Evans and Craig's case) from which the representation of the letter is constructed are themselves both targets and masks. Each point of stimulation must exert masking forces on adjoining points while at the same time being subject to masking effects due to those adjoining stimuli.

Under these conditions of simultaneous presentation and

close spatial proximity strong masking effects must occur, yet the subjects achieve relatively high letter identification rates (for example 75% correct in Craig's 1983a study). In contrast, Craig (1982a) has shown that individual letter identification accuracy falls as a function of increasing SOA if the letter is presented in two successive halves. As the level of masking between letter elements should decrease as the SOA between successive letter halves increases, this indicates that tactile pattern perception accuracy does not always decrease as the level of masking between pattern elements increases.

In summary, the reason that subject have difficulty identifying the target when successive letters are used as the target and mask may be that the integrated representation of the target and mask no longer directly tells the observer what each individual letter felt like. In contrast, when a single letter is presented, with the strong masking effects present being due only to the interactions between elements constituting that letter, the integrated representation does directly tell the observer about the form of the letter, thus accounting for the relatively ease with which these stimuli can be identified.

This argument suggests that we can identify some integrated tactile percepts, with difficulties arising either as the complexity of the integrated representation increases or when we are required to identify one of the components of the integrated representation. Of course it is possible that

performance on these types of task may improve with training. In all complex sensory tasks considerable experience and practice is required before the observer can make efficient use of the available information. For example, individuals who have visual defects rectified after long periods of dysfunction often do not achieve normal visual function, even though there is evidence that their peripheral visual mechanisms are functioning normally (Valvo, 1971). Loomis (1981) has interpreted this observation as showing that, " .. normal form perception depends upon processes of perceptual integration that either develop with experience or require continued stimulation for normal functioning" (p. 10). Clearly, it is unwise to expect that touch can innately perform tasks of a complexity which hearing and vision cannot solve without practice.

Third, the communicative value of speech is not dependent upon what a given speech event sounds like to the listener. All that matters is that the listener can consistently discriminate that sound from all other speech sounds. Spoken words are often discriminated as much by their context as by their individual physical properties. Indeed, the listener need not be directly aware of the component sounds, that is phonemes, from which larger speech units are constructed. For example, Savin and Bever (1970) and Warren (1971) have shown that listeners can identify syllables more readily than they can the phonemes from which the syllable is constructed. Put simply, the extraction of the component parts

of a pattern is not a prerequisite in the identification of that pattern. Consequently, the important issue in tactile pattern perception is the ease with which a given pattern can be uniquely discriminated from all other patterns, not the extent to which individual components of that pattern can be identified.

Aside from explaining the relative failure of observers to identify individual elements from within integrated tactile percepts, these three points show why conventional masking methodologies are not suitable for showing the communicative value of these integrated tactile percepts. To summarize, these do not present the subject with a task which we can reasonably expect them to perform without training, nor do they allow the subject to demonstrate their capacity to discriminate between the percepts resulting from different patterns of stimulation.

There are at least two types of task where these limitations do not apply, and thus which allow an assessment of the predictions made by the Isolation and Integration Hypotheses. First, we may look at element and pattern identification studies where the subject has the benefit of training. Obvious examples of this type are the numerous attempts to train individuals to interpret tactile speech transforms. While this type of study cannot reveal beneficial effects of masking, unless compared with an equivalent study in which only the extent of masking is varied, it can show whether or not masking, as the Isolation Hypothesis claims, leads to inadequate levels of pattern identification. The results of

some of these previous studies will be discussed in the following sub-section.

A second approach, which does not impose the training load of learning to extract component elements from integrated percepts, is suggested by the earlier argument that the important issue in pattern perception is the extent to which a given pattern is uniquely discriminable, rather than the discriminability of its component elements. In this case, the most appropriate task to assess tactile pattern processing ability is to ask the observer to compare discretely presented tactile patterns. If the spatial and temporal proximity of the elements of each discrete pattern are varied, then the Isolation and Integration Hypotheses offer different predictions concerning the ease with which this task can be performed at each level of pattern element proximity. As the proximity of the pattern elements is increased then the level of masking between pattern elements, and consequently the extent to which an integrated perceptual representation is generated, should also increase. If these conditions of close pattern element proximity are detrimental to the process of pattern perception then the ease with which the two patterns can be compared should decrease with element proximity. However, if, as the Integration Hypothesis maintains, the occurrence of perceptual integration is beneficial to the process of pattern perception, then the ease with which the patterns can be compared should increase as the proximity of the pattern elements is increased to some optimal level.

Unlike conventional masking studies, where the observer is required to identify the target from an integrated representation of the target and mask, this task requires the observer to discriminate the pattern of stimulation generated by the integration of a number of targets and masks (the mutually interacting elements of the pattern) from that generated by the integration of a different set of elements. The advantage of this approach is that it removes the need for the experimenter to make assumptions about what the stimulus will feel like to the subject, and consequently how the subject will describe the stimulus through his/her response. The level of performance achieved on this type of task simply reflects the extent to which the subject can discriminate between differing patterns of stimulation. As was suggested earlier, this ability seems to be central to the success of any language, be it an auditory, tactile, or visual one. An experiment testing the Integration Hypothesis via this method is described in the next chapter.

## 5.4.3: Integration and Tactile Transforms of Speech

It was argued in the previous sub-section that an analysis of attempts to train subjects to identify speech stimuli may provide clues to the role of integration in tactile perception. As was described in Chapter 1, multichannel tactile speech prostheses generally consist of an array of closely spaced transducers through which momentary changes in the spectrum of the speech signal are presented. Due to the close spatial proximity of these transducers in most systems, and the

rapid rate at which the pattern of stimulation must change in order to accurately depict the speech spectrum, this type of display should induce strong masking effects between pattern elements. As was shown in Section 5.3.1, the perceptual representation of the pattern generated under these conditions of close pattern element proximity consists of an integrated representation of those elements' features.

There is evidence that masking does impede the perception of speech derived stimuli even after repeated exposure to the stimuli. Green et al. (1983) presented subjects with tactile transforms of consonant-vowel pairs via an Optacon based display. During the course of testing, each subject received each stimulus 320 times. Although subjects could make gross differentiations between consonants, they had difficulty when the consonants presented differed only in place of articulation (simply put, these stimuli were more alike). However, performance on this task improved when the duration of the following vowel was reduced, suggesting that the vowel was causing sufficient backwards masking to impede the recognition of these similar consonants. Green et al. (1983) concluded that, although the integrative effects associated with masking at very brief durations may not impede performance, the non-integrative masking effects evident at longer durations may do so.

As was described in Chapter 1, users of the Queen's University Tactile Vocoder, arguably the most successful current tactile prosthesis, have acquired tactile vocabularies

138

of up to 250 words (Brooks & Frost, 1983; Brooks et al., 1985, 1987). Although these observers achieved low levels of word identification accuracy when presented with novel words, the general pattern of many of these words was correctly identified. For example, when presented with the test word "staff" the subject would respond "stuff" (Brooks et al., 1986a).

These results suggest that the subject in Brooks et al.'s (1986a) study had begun to learn the skills necessary to deal with the integrated representations resulting from the presentation of closely spaced tactile patterns. In this case, the poor results reported by Green et al. (1983) may reflect inadequate training and/or an inappropriately designed display. For example, the Optacon's transducer density of 144 per $cm^2$, while obviously exceeding the recommendations of the Isolation Hypothesis, may overstep the limits of the Integration Hypothesis's proposed beneficial integrative mechanism. Indeed, a two dimensional display originally designed to present the spatial form of alphabetic characters may not be the ideal style of display for the representation of temporal forms like the speech signal.

In more general terms, the results achieved with the Vocoder throw doubt on the central premise of the Isolation Hypothesis; that masking resulting from the close spatial and temporal proximity of pattern elements in tactile transforms of speech renders these displays difficult to interpret. Perhaps those earlier studies which gave rise to this observation also

failed to provide appropriate levels of training and/or used deficient displays. Of course the results achieved with the Vocoder do not show that integrative effects resulting from close spatial and temporal proximity between pattern elements facilitate performance. What they do suggest is that such effects do not excessively impede performance.

### 5.4.4: Masking in Other Modalities

The Integration Hypothesis maintains that information integration is central to the processing of patterns in any modality. This sub-section briefly reviews this assertion. As this thesis is principally concerned with tactile speech prostheses, this review will focus upon masking in the auditory modality.

Elliott (1962a, 1962b) and Massaro (1970) have reported strong forward and backward masking effects in the auditory modality. The general shape and timecourse of the auditory masking functions observed by these researchers is very similar to that observed in the tactile modality (Kirman, 1986). For example, maximum masking is observed in either modality at very brief SOA's, with a rapid decline in the extent of masking with increasing SOA. In both cases, significant levels of masking seems to occur up to SOA's of at least 250ms duration.

Just as in the tactile and visual modalities, it appears that integration is one of the processes underlying auditory masking phenomena (Massaro, 1972; Pastore, Harris, & Goldstein, 1980). Granted the similarity between both the timecourse and underlying mechanisms of tactile and auditory masking, it seems

reasonable to assume that the speech signal undergoes similar changes due to masking when presented to either modality. While it is unclear whether the processing of the speech signal is either enhanced, impeded, or unaffected by the effects of masking, it is clear that their presence does not prevent us from achieving a perfectly acceptable level of speech perception performance.

The question of the relationship between masking and the efficiency of speech processing has not been the subject of systematic study. However some general observations can be made concerning the role of interactions between stimuli in auditory processing.

First, Békésy (1959) has argued that simultaneous masking effects within the auditory system reflect a process of lateral inhibition which, far from impeding performance, serves to sharpen the frequency selectivity of the ear. The similar Mach Bands effect in vision also seems to facilitate the perception of form by enhancing contours (Sekuler & Blake, 1985).

On the other hand, Lebedev et al. (1985) have shown that some masking effects are inconsequential to speech perception. They selectively removed weaker spectral components from the speech signal based upon the extent to which they would be masked by stronger components. They found that this manipulation did not effect speech perception accuracy. Thus it seems that some auditory masking effects are of no functional significance to the processing of speech.

Finally, there is evidence that the auditory system includes mechanisms designed to counteract some masking effects. Warren (1970, 1976, 1983, 1984) has demonstrated two such effects, the Phonemic Restoration Illusion and Auditory Induction. The Phonemic restoration illusion is observed when phonemes within phrases are removed and replaced with noise. Under these conditions the subjects report the perceptual presence of the removed phoneme. As this effect does not occur when the removed phonemes is replaced with a silent interval, it suggests the operation of a mechanism designed to compensate for masking of speech components. Auditory induction is observed when a portion of a constant amplitude sound is removed and replaced by a louder segment. Under these conditions the subjects fail to perceive the change in the signal's amplitude and report that the sound is constant. In both cases, these phenomena seem to reflect the operation of mechanisms designed to help the auditory system cope with the masking effect of unwanted signals upon the signal of interest.

The preceding discussion demonstrates that no one role can be assigned to masking in the process of auditory perception. However, there is no strong evidence available to show whether masking and integration is central to the perception of speech stimuli. Hence all that can be said with certainty is that we can adequately perceive speech in the presence of masking effects similar to those experienced when presented with tactile speech transforms.

## 5.4.5: Learning to Deal With Masking

Throughout this chapter it has been argued that we can learn to deal with the integrated percepts resulting from masking. While it appears that speech perception proceeds effectively in the presence of masking effects, and some individuals have learnt to identify many spoken words via the Vocoder, more direct evidence of the level of masking decreasing with practice is needed.

There is a growing body of evidence from the visual modality that observers do improve their target identification performance with practice in backward pattern masking situations, but do not improve in backward energy masking tasks (Hertzog, Williams, and Walsh; Schiller, 1965; Schiller & Wiener, 1963; Wolford, Marchak, and Hughes, 1988). In Wolford et al.'s (1988) study, subjects were presented with consonant targets in the presence of non-alphabetic character masks (eg. "#"). Across a 45 day training period the subject at least doubled his target identification rate at all of the six SOA's used. In the best case (SOA=68ms) his performance improved from 17% correct to 94% correct. Wolford et al. (1988) did not find comparable practice effects in similar situations, like lateral masking and whole report tasks, which do not involve backward masking.

Why was this practice effect confined to backward pattern masking situations, and what mechanism underlies it? Wolford et al. (1988) proposed that this practice effect is due to an enhancement in central sensory processing efficiency,

143

perhaps due to improved alertness to the target. They argued that central sensory processing performance is not a significant factor in those tasks where the practice effect is absent. For example, the limiting factor in whole report tasks is short-term memory capacity, which does not improve with practice (Klemmer, 1964; Pollack & Johnson, 1965), rather than central sensory processing capacity.

A different interpretation of these results can be cast in light of the Integration Hypothesis. Practice effects do not occur in backward energy masking tasks because the integration of the pattern target and energy mask yields a representation in which all features of the target are obliterated by the mask. In contrast, the integrated percept generated in a pattern masking paradigm retains features of both the target and mask. In this case, practice results in finer attunement to these features, thus leading to better target identification rates.

Evidence that tactile pattern masking effects are also mutable comes from Craig's (1977) investigation of two individuals who exhibited extraordinary Optacon reading rates. These individuals achieved reading rates of 80 and 100 words per minute respectively following only several hours training. In comparison, the majority of blind users of the Optacon achieve reading rates of about 30 to 60 words per minute following several years experience.

The exceptional levels of performance achieved by these two observers did not seem to result from superior visual

reading skills, nor from enhanced tactile sensitivity, nor from abnormally fine tactile temporal resolution. Where these observers did differ from normal observers was in terms of their performance on tactile recognition masking tasks. Although both exhibited normal detection masking functions, they exhibited virtually no recognition masking under either forward, backward, or forward and backward masking conditions.

Again, these results can be interpreted in term of the Integration Hypothesis. Detection masking does not involve making judgments about the form of the perceptual representation. Subjects simply have to identify the presence of the stimulus. On the other hand, as has been argued throughout this chapter, recognition masking does depend upon ability to deal with the integrated perceptual representation of the target and mask. In this case, these two observers may have innately higher levels of skill at interpreting such integrated representations. The fact that many blind users of the Optacon ultimately achieve reading rates approaching those of these two observers suggests that these skill are not unique. Instead, it seems that normal observers require considerable experience before learning to deal with these integrated tactile percepts.

It would be of interest to discover whether experienced blind users of the Optacon also exhibit lower levels of backward and forward masking. Although Craig (1977) did present data on the levels of masking observed in a number of blind observers, he did not specify their level of expertise in

Optacon reading. Never-the-less the blind observers did exhibit less forward and backward recognition masking than normal observers.

## 5.5: Conclusions

This chapter has examined two competing views on the role of masking in tactile pattern perception. The Isolation Hypothesis argues that masking hinders pattern perception, and hence must be minimized using strategies like wide spatial and temporal separations between pattern elements. In contrast, Kirman's (1983) Integration Hypothesis argues that masking reflects the operation of beneficial information integration processes essential to efficient pattern perception. It was explained that this view does not imply that optimum pattern discrimination will occur under conditions of maximum masking. What it does suggest is that optimum performance will not occur when masking is minimized.

A review of the literature clarified several of the issues required to assess these competing views. First, as the Isolation Hypothesis claims, the extent of masking is dependent on the spatial and temporal separation between stimuli. Further, masking does involve a reduction in the accuracy with which the target stimuli can be detected and/or recognized. On the other hand, it is clear that an integrated perceptual representation of target and mask features is generated under conditions of close pattern element proximity, an observation consistent with the Integration Hypothesis' claim that masking is but one element in a general process of perceptual

integration. With these points decided, the key question became the extent to which tactile pattern perception is either impeded or facilitated by the occurrence of masking.

Although conventional masking studies show a decrease in pattern perception performance with increasing levels of masking, it was suggested that the type of tasks used prevented any beneficial effects of masking from emerging. In particular, previous studies either failed to provide adequate training to give subjects a reasonable chance of success on the tasks used, or failed to employ tasks which minimized the need for such training. It was concluded that the only available data which allowed a reasonable assessment of the capacity of the tactile modality to deal with complex patterns came from extended training studies, like those undertaken with various tactile prostheses.

The promising results currently being achieved with these devices suggested that masking may not be the limiting factor in tactile pattern perception that the Isolation Hypothesis claims. However, it was concluded that there is insufficient evidence in either the case of touch or hearing to determine whether the occurrence of masking is beneficial to pattern perception, or whether these modalities manage to process complex patterns in spite of masking effects. Indeed, there is some evidence that the auditory sense may possess mechanisms to circumvent the effects of masking. Finally, evidence was presented from both the visual and tactile modalities which indicates that we can learn to overcome the

initially poor levels of performance achieved on recognition masking tasks.

It can be concluded that the Isolation Hypothesis appears to be wrong in arguing that masking prevents the perception of speech derived stimuli. This does not, however, show that the Integration Hypothesis is correct in arguing that masking facilitates the perception of such complex stimuli via a process of perceptual integration. As was noted earlier, what is needed to resolve this question is an experiment which both presents the subject with a task in which any positive effects of masking can emerge, and allows an assessment of the level of pattern element proximity which yields optimum performance levels. The following chapter describes an experiment of this kind.

# Chapter 6: Testing the Integration Hypothesis.

This chapter presents an experiment which investigated the Integration and Isolation hypotheses discussed in the previous chapter. The central premise of the Integration hypothesis is that tactile masking effects, although reducing the discriminability of individual pattern elements, are but one aspect of a general process which acts to facilitate the perception of patterns through the integration of discrete pattern elements into coherent perceptual units. It was shown in the previous chapter that tactile masking effects increase as the temporal and spatial separation between target and mask is decreased. The Integration Hypothesis implies that, although the discriminability of individual pattern elements decreases with decreasing spatial and temporal separation, optimum pattern perception performance will occur at some level of element proximity less than that needed to minimize masking. These predictions were tested in a temporal masking paradigm by measuring the discriminability of three-element spatially and temporally distributed tactile patterns as the spatial distance and temporal separation between the pattern elements was varied.

It was argued in Chapter 5 that many previous investigations of tactile pattern perception have employed designs which prevented the detection of any positive effects associated with masking. In an attempt to counter these problems, the present experiment employed a same/different task whereby subjects compared two sequentially presented tactile patterns. As was explained in Sub-section 5.3.2, this design

allows an assessment of the discriminability of the patterns which is independent of any arbitrary decision about which features the subject should be perceiving.

In addition to this Pattern Comparison Condition, a second condition was included in which the subjects were required to make decisions about individual elements within the tactile patterns. This Element Identification Condition was included as a control condition to provide an index of the extent of masking induced by the manipulation of the spatial and temporal separation between pattern elements which was independent of the performance observed on the Pattern Comparison Condition.

## 6.1: Experiment 7: Element Proximity and Pattern Recognition

### 6.1.1: Method

Subjects. The 18 subjects tested in this experiment were paid undergraduate volunteers attending the University of Tasmania. None had extensive experience with vibrotactile stimuli.

Apparatus and Procedure. Two types of task were used in this experiment. The Pattern Comparison Condition employed a same-different task in which subjects compared sequentially presented tactile patterns, while the Element Identification Condition employed a three-alternative forced choice task requiring the subjects to identify the most intense element within a tactile pattern. These tactile patterns consisted of three sequential 90ms duration 250Hz vibratory pulses, each presented to a different site on the forearm. Each pattern was

151

distinguished by a unique variation in amplitude between these three elements. Although the absolute amplitude of each pattern element was varied in the range 23 to 40dBm from trial to trial, there was always a 6dBm difference in amplitude between each of the three amplitude levels constituting one pattern. Table 6.1 shows the general form of the six types of pattern presented.

TABLE 6.1: Amplitude variations across stimulated sites for each pattern type.

STIMULATED SITE

|  |  | 1 | 2 | 3 |
|---|---|---|---|---|
|  | Pattern 1 | LOW | MEDIUM | HIGH |
|  | Pattern 2 | LOW | HIGH | MEDIUM |
| PATTERN | Pattern 3 | MEDIUM | LOW | HIGH |
| TYPE | Pattern 4 | MEDIUM | HIGH | LOW |
|  | Pattern 5 | HIGH | LOW | MEDIUM |
|  | Pattern 6 | HIGH | MEDIUM | LOW |

These tactile stimuli were generated by an Applied Engineering Super Music Synthesizer under the control of a microcomputer and were presented to the skin via Senheiser tactile transducers. Each transducer had a plastic flange glued to its underside which was used to mount the devices in a neoprene lined clamp constructed from dense craftwood. This arrangement allowed accurate spatial separations to be maintained between the transducers while minimizing any coupling effects between vibrators. A diagram of this apparatus is presented in Figure 6.1.

Within each task, the spatial and temporal separation between pattern elements was varied in three steps. The
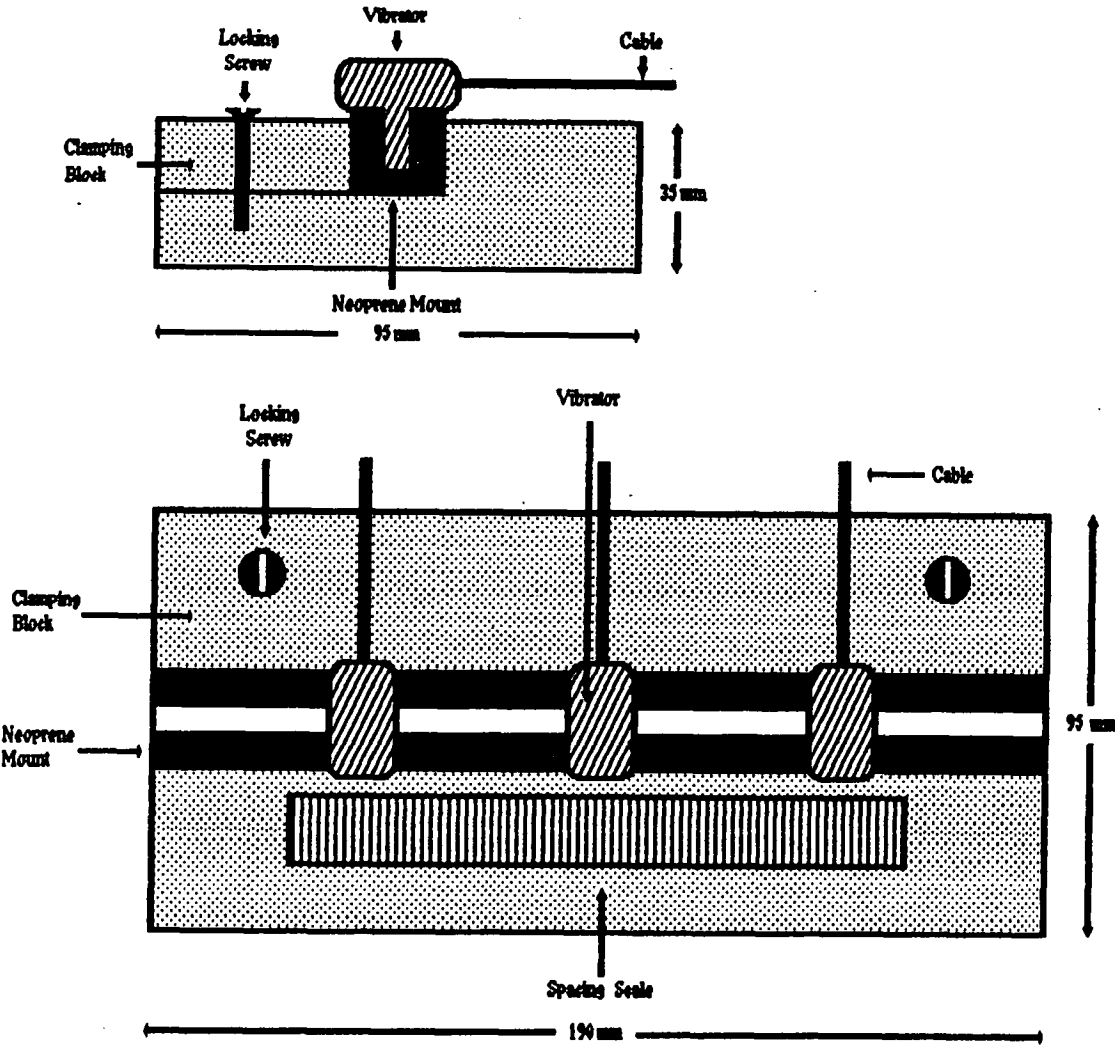
Figure 6.1: Schematic representation of the vibrator mounting device used in Experiment 7.

briefest SOA used was 94ms, a duration within Evans and Craig's (1986) 100ms masking by integration stage. The second SOA used was 200 ms, a duration within their higher-level interference stage. The final SOA used was 450ms, a duration outside the range in which masking effects are prominent.

The spatial separations used were constrained both by the size of the transducers used and by the physical dimensions of the forearm. The smallest center-to-center distance used between vibrators was 1.1cm, which left a 1mm gap between the sides of adjoining vibrators. This meant that adjoining stimulated sites lay completely within the two-point limen of the forearm region used (35-40mm per Békésy, 1959; Weinstein, 1968). The intermediate vibrator spacing used was 3cm which, taking the width of the vibrators into account, meant that adjoining stimulated sites lay partially within the two-point limen of the forearm region used. The widest center-to-center spatial separation used was 5cm, which left adjoining stimulated sites outside the two-point limen of the forearm region used.

As there is a difference of approximately 4-5dB in the point of subjective equality for amplitude across this range of sites (Békésy, 1959), it was necessary to adjust the amplitude of each pattern element depending upon the site to which it was presented. If this were not done, the perceived amplitude difference between pattern elements would have varied with stimulus spatial separation, thus confounding any masking effects due to spatial separation. Prior to testing, each

subject underwent a calibration session in which the point of subjective equality for amplitude between each stimulated site was established via the method of limits procedure. During piloting, it was established that a single calibration value for each site, obtained using the median amplitude level (32dBm) included in any pattern, was adequate to cover the whole range of stimulus amplitudes used in the experiment. The correction values obtained during this calibration session were applied to each stimulus as it was presented during each experimental trial.

Both the Pattern Comparison Condition and the Element Identification Condition employed a 3x3xSUBJECTS factorial design. The two independent variables were the temporal and spatial separation between pattern elements. As was described above, each factor had three levels. These two factor were fully crossed, resulting in nine experimental conditions. That is, each level of temporal separation was tested in conjunction with each level of spatial separation. The dependent variable in the Pattern Comparison Condition was the accuracy with which the two patterns presented on each trial were compared. The dependent variable in the Element Identification Condition was the accuracy with which the most intense element in each pattern could be identified.

The order in which the subjects undertook these nine experimental conditions was counterbalanced within each of the two task types in order to control for practice and fatigue effects. Similarly, the order in which subjects undertook the

two task types was counterbalanced.

Prior to undertaking both the Pattern Comparison and Element Identification Condition tasks the subjects underwent a screening session in which they were presented with blocks of 60 trials until they reach a criterion accuracy level on that block of dprime >= 1. Any subject who did not reach this level on either task after four blocks was excluded from the experiment. In most cases this criterion level was reached after only one block of trials, although one subject required four blocks to reach this level. Two potential subjects were excluded for failing to reach this criterion on either the Element Identification or Pattern Comparison Condition tasks. In the screening trials for the Element Comparison Condition the temporal and spatial separations between pattern elements were set at 450ms and 5cm respectively, these being the levels which both the Integration and Isolation Hypotheses predicted would produce the easiest discriminations. In the screening trials for the Pattern Comparison Condition the temporal and spatial separations between pattern elements were set at 250ms and 3cm respectively. As the Integration and Isolation hypotheses differed in terms of their predictions of which particular spacings should produce the best performance, these intermediate levels were chosen to avoid prejudging which separations would produce the easiest discriminations.

Immediately following the screening session each subject undertook 60 experimental trials in each of the nine spatial-temporal separation conditions for that task type. Including

the associated screening and calibration tasks, each task type session lasted about 2 hours, with the subjects receiving rest periods whenever they required. There was a delay ranging from 1 to 7 days between the two task type sessions for each subjects, dictated by the availability of the subject.

The subject was seated in front of a VDU with the vibrator mount placed on a waist-high table situated on the subject's dominant side. The subjects were instructed to ensure that the middle of the three vibrators was positioned under the middle of their forearm (the mid-point between the wrist and elbow) before each trial. This point was marked on their forearm to enable accurate placement. The acoustic output of the vibrators was masked by 400Hz 70dBA narrow-band white noise presented via Senheiser HD22 headphones.

Each trial was commenced by the subject pressing either of the response buttons. One second later the target pattern was presented to the subject. In the Pattern Comparison Condition, the reference pattern was then presented 500ms after the termination of the target. This ISI was chosen in order to balance the competing needs of minimizing masking effects between the two patterns and minimizing the memory load imposed by the task. The subjects then responded either by two-choice button-press in the Pattern Comparison Condition, or by three-choice button-press in the Element Identification Condition. They were instructed to respond as rapidly as they could while still maintaining a high level of accuracy, as indicated by feedback given following each response.

## 6.1.2: Results

Element Identification Condition. The raw data gathered in the Element Identification Condition consisted of the proportion of trials on which each subject correctly identified the strongest vibration in the target pattern. Dprimes were calculated for each subject in each condition using the multiple choice algorithm described by Green and Swets (1966). Mean dprime scores are presented in Table 6.2[1] as a function of the temporal and spatial separation between pattern elements and of the pattern element probed.

Table 6.2 shows that the level of spatial and temporal separation between pattern elements did influence the amount of masking observed. Performance increased with both the spatial and temporal separation between pattern elements. Further, it can be seen that Element 1 was identified most accurately, followed by Element 3 then Element 2. Figure 6.2 shows that this pattern of response accuracy across probed elements held true regardless of SOA or stimulus spacing. These trends were confirmed by ANOVA, which revealed a significant effect due to temporal separation, $F(2,34)=37.11$, $p<.0001$, a significant effect due to spatial separation, $F(2,34)=4.05$, $p<.05$, and a significant effect due to probed element, $F(2,34)=8.67$, $p<.01$. Neither the two-way nor three-way interactions between these factors approached significance.

Post-hoc tests were conducted to identify the source of these three significant effects. It was found that performance

---

1. Individual subject means are presented in Appendix B7.

increased between each of the three levels of temporal separation used in the experiment, with significant increases evident both between the 94ms and 250ms SOA levels, $F(1,17)=20.41$, $p<.001$, as well as between the 250ms and 450ms SOA levels, $F(1,17)=28.4$, $p<.001$. However, similar comparisons between the three levels of spatial separation used indicated that the increase in performance with increasing spatial separation was primarily evident between the extreme 1cm and 5cm vibrator spacings, $F(1,17)=11.53$, $p<.01$, rather than between either the 1cm and 3cm, $F(1,17)=1.52$, $p>.05$, or 3cm and 5cm, $F(1,17)=1.97$, $p>.05$, spacings. Finally, while Element 1 was identified significantly more accurately than either Element 2, $F(1,17)=19.07$, $p<.001$, or Element 3, $F(1,17)=7.05$, $p<.05$, no significant difference in identification accuracy was evident between Elements 2 and 3, $F(1,17)=2.3$, $p>.05$.

TABLE 6.2: Mean accuracy scores (d') as a function of SOA, element spacing, and probed element from the Element Identification Condition in Experiment 7. Standard deviations are given in parentheses.

| SOA (ms) | d' | Spacing (cm) | d' | Probe Element | d' |
|---|---|---|---|---|---|
| 94 | 0.62 (0.70) | 1 | 1.18 (1.49) | 1 | 1.79 (1.82) |
| 250 | 1.18 (1.30) | 3 | 1.33 (1.69) | 2 | 0.96 (1.26) |
| 450 | 2.25 (2.04) | 5 | 1.54 (1.62) | 3 | 1.30 (1.30) |

The purpose of including the Element Identification Condition in the experiment was to confirm that the changes in the SOA and spatial separation between patterns elements in the patterns to be used in the Pattern Comparison Condition did
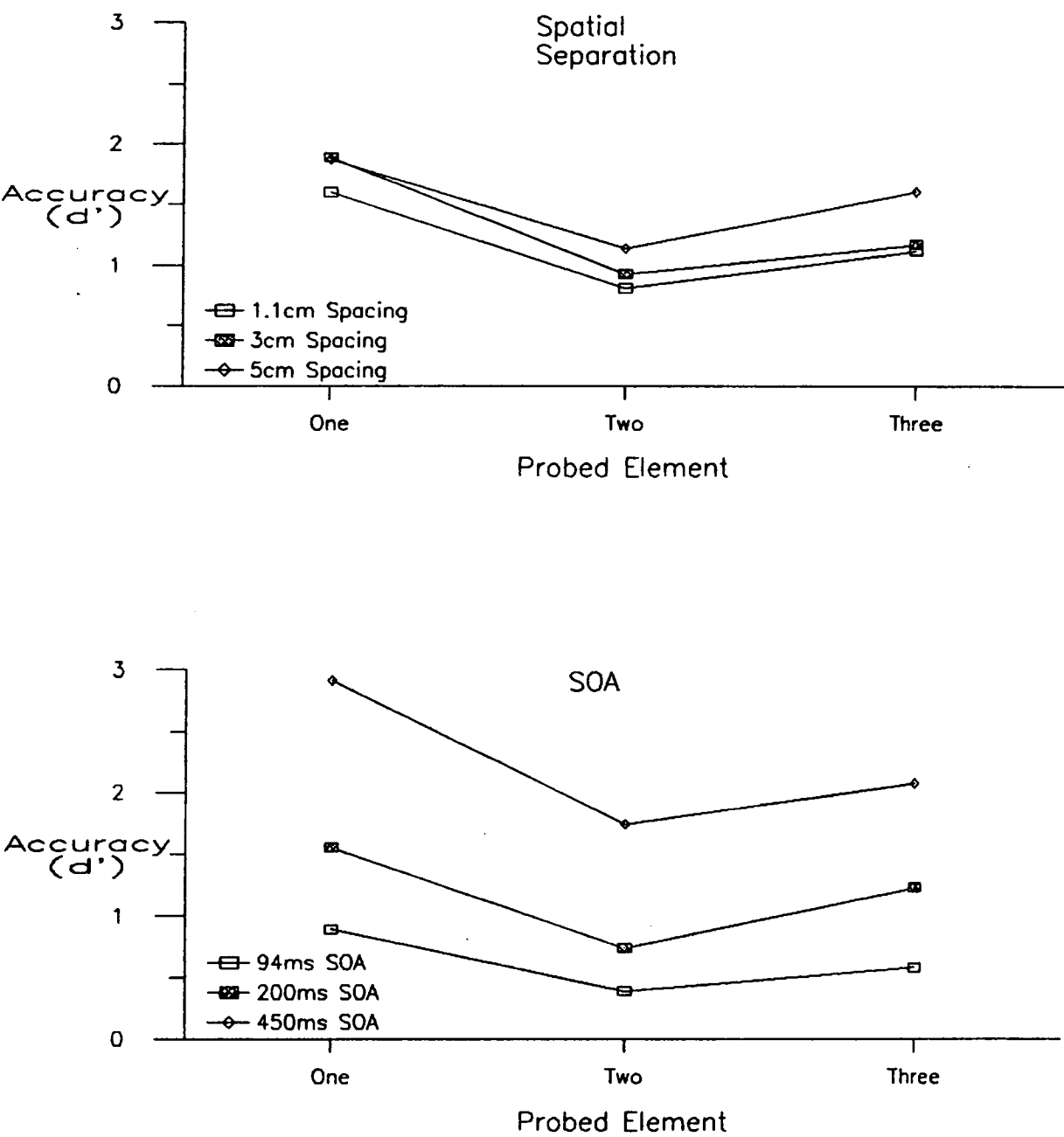
Figure 6.2: Accuracy as a function of probed element, SOA, and spatial separation from the Element Identification condition in Experiment 7.

induce  the desired pattern of change in the amount of  masking

occurring between pattern elements. The results obtained

confirmed  that these manipulations did lead to changes in  the

extent of both backward  and forward temporal masking.

Pattern Comparison Condition.  The raw data gathered  in

the  Pattern  Comparison Condition consisted of  the  frequency

with  which  each  subject correctly compared  the  target  and

reference patterns. Dprimes were calculated for each subject in

each  condition  counting  same-pattern  trials  correctly

identified  as hits and  different-pattern  trials  incorrectly

identified  as  false alarms. Mean dprimes for each  level  of

spatial and temporal separation are presented in Table 6.3.

TABLE  6.3  : Accuracy (d') as a function of  SOA  and  element
spacing  from  Pattern Comparison Condition  in  Experiment  7.
Standard deviations are given in parentheses.

| SOA (ms) | d' | Element Spacing (cm) | d' |
|---|---|---|---|
| 94 | 0.73 (0.38) | 1 | 0.83 (0.48) |
| 250 | 0.99 (0.40) | 2 | 0.91 (0.46) |
| 450 | 0.98 (0.54) | 3 | 0.96 (0.45) |

Table  6.3  shows  that  pattern  comparison  accuracy

initially  increased  with  increasing  SOA  before  stabilizing

between 250 and 450 SOA's. On the other hand, a trend towards a

progressive  increase  in  accuracy  with  increasing  spatial

separation between pattern elements was observed. Analysis  via

ANOVA  revealed that, while the trend for accuracy to  increase

with SOA was significant, $F(2,34)=5.07$, $p<.05$, the trend for accuracy to increase with spatial separation did not reach significance, $F(2,34)=1.40$, $p>.05$. Finally, this analysis revealed no significant interaction between these two factors, $F(2,34)=1.78$, $p>.05$.

Post hoc analysis confirmed the descriptive observation that the increase in accuracy with SOA was restricted to the 94ms to 250ms SOA interval, as a significant difference in accuracy was revealed between these two levels, $F(1,17)=10.78$, $p<.01$, but not between the 250ms and 450ms SOA levels, $F(1,17)=0.02$, $p>.05$. Similar comparisons revealed no significant differences in accuracy between either the 1cm and 3cm spatial separation levels, $F(1,17)=1.04$, $p>.05$, the 1cm and 5cm levels, $F(1,17)=2.7$, $p>.05$, or the 2cm and 5cm levels, $F(1,17)=0.43$, $p>.05$.

### 6.1.3: General Discussion

The Element Identification Condition results showed that masking between pattern elements increased as the spatial and temporal separation between elements was decreased. In the case of SOA, a significant decrease in performance was observed with each decrease in SOA. In the case of spatial separation, there was also a decrease in performance with decreasing spacing, although this effect was significant only between the extreme 1cm and 5cm element spacings.

Analysis of accuracy levels as a function of probed element revealed that Element 1, that is the first element activated, was identified more accurately than the equally well

identified Elements 2 or 3. This relationship held true regardless of SOA or spatial separation, suggesting that the same masking processes were operating at each level of spatial and temporal separation.

An obvious question is to what extent the observed levels of masking were due to the action of forward and backward masking components. If both types of masking acted equally, then Elements 1 and 3 should have been identified equally well, while Element 3 should have been identified best if backward masking effects predominated. Thus the only arrangement which can account for the observed pattern of results is that the forward masking effect was the strongest. In this case Element 3 would experience a strong forward masking effect due to Element 2 and, because of the greater spatial and temporal separation involved, a weaker forward masking effect due to Element 1. Likewise, Element 2 would be subject to a strong forward masking effect due to Element 1 and, because of the reduced force of backward masking, to a weaker backward masking effect due to Element 3. On the other hand, Element 1 would experience only a weak backward masking effect due to Element 2, and an even weaker one due to Element 3. This explanation implies that this task was more like a simple detection task, where forward masking is usually predominant, than a recognition task, where backward masking is strongest.

The results obtained in the Pattern Comparison Condition seem clear-cut. In this condition the subjects' performance was

worst at the closest spatial and temporal element separations. In the case of SOA, performance improved between this closest separation (94ms) and the intermediate SOA (250ms). There was no evidence of further improvement at the longest SOA (450ms). In the case of spatial separation, a non-significant trend for performance to increase with spatial separation was observed.

.     As the results obtained in the Element Identification Condition showed that masking effects between pattern elements increased as both the SOA and spatial separation between elements were decreased, the changes in accuracy observed in the Pattern Comparison Condition with equivalent variations in these parameters may be attributed to changes in the amount of masking occurring between pattern elements. To this extent, the results of the present experiment are similar to those of reported by Evans and Craig (Craig, 1985a; Evans, 1987; Evans & Craig, 1987), who also found a decrease in the accuracy with which patterns could be identified as masking increased.

Clearly, these results are consistent with the Isolation Hypothesis, which predicts a decrease in tactile pattern discriminability with increasing levels of masking, rather than with the Integration Hypothesis, which predicts an increase in pattern discriminability as masking increases to some optimum level. There is, however, a need to clarify several issues before adopting the Isolation Hypothesis. First, it is possible that the beneficial integrative effects proposed by the Integration Hypothesis only occur at briefer temporal separations and/or narrower spatial separations between pattern

elements than were used in this study. Subsequent experiments could test this possibility by employing element SOA's and spatial separations below those used here.

A second prospect which must be addressed before accepting the Isolation Hypothesis is that the particular style of tactile pattern used in this experiment may not have been conducive to beneficial perceptual integration. For example, both Kirman (1973) and Richardson and Frost (1977) have argued that efficient pattern perception in any modality requires the presentation of rich patterns which include as many diverse and redundant features as possible. Subsequent studies could examine whether the failure to observe beneficial integrative effects in the present study was due to a lack of richness in the particular patterns used here. These studies might employ a greater number of pattern elements varying across a wider range of pattern element characteristics; for example amplitude, frequency, and duration.

Finally, although it was argued in Chapter 5 that the design of the present experiment would minimize the need for training, it may be that observers require more experience than was available here before learning to make use of the information available in any integrated percept arising from close pattern element separations. This possible cavil on the acceptance of the Isolation Hypothesis did receive tentative support from the anecdotal reports of the subjects in this experiment. All subjects were questioned about what the stimuli felt like and how they performed each task. The universal

response when the elements were widely spaced was that all that was perceived was a series of unassociated pulses, rather than a pattern, with the pattern comparison being performed on the basis of a pulse by pulse amplitude comparison. However, when the elements were closely spaced, the subjects reported feeling a unified pattern. A common analogy offered was that it felt like a "word" rather than a discrete series of buzzes. In this case the subjects reported that they performed the pattern comparison task by judging whether the patterns felt the same, rather than by directly comparing the individual elements.

These reports are consistent with the conclusions reached by Garner and Gottwald (1968) in their study of the perception and learning of auditory and visual temporal patterns. They proposed that;

".. the perception of temporal pattern which occurs at faster rates (of element presentation) is one of an integrated sequence, is phenomenally compelling and immediate, and is a relatively passive process for the observer. On the other hand, the learning of temporal pattern is of a succession of single elements, and is derived, recoded and intellectualized process in which the observer is much more active." (Garner & Gottwald, 1969, p.108).

The subjects' anecdotal evidence does suggest that an integrative process may have been at work when close spacings were used, but did not occur at wider spacings. The question then becomes whether subjects can learn to take advantage of any integrative processes occurring at close levels of pattern element proximity. The next section describes an experiment which investigated this question.

## 6.2: Experiment 8: Learning to Discriminate Tactile Patterns

Although Experiment 7 yielded strong support for the Isolation Hypothesis, it was suggested that several issues require clarification before that hypothesis can be accepted. One of these was the possibility that observers may require some training or experience before being able to deal effectively with the integrated percepts which may result from close spatial and temporal proximity of pattern elements. The following experiment attempted to test whether the superior discriminability of widely spaced compared to closely spaced tactile patterns observed in Experiment 7 was due to a need for additional experience with closely spaced tactile patterns. If lack of experience or training with closely spaced tactile patterns was responsible for the effects observed in Experiment 7 then it was expected that the discriminability of those patterns would improve at least to the level observed with widely spaced patterns following training.

### 6.2.1: Method

Subjects. Three subjects participated in this experiment. Subjects RK and FH were 23 year old females who had participated in Experiment 7. Neither of these subjects had extensive experience with this type of vibrotactile stimuli prior to that. Subject DM was a 31 year old male who had participated in piloting for Experiment 7, and had extensive experience with vibrotactile stimuli.

Apparatus and Procedure. The stimuli used in this experiment were the same as those used in the Pattern

Comparison Condition in Experiment 7, that is sequentially presented three-element vibrotactile patterns. These patterns were generated and displayed in the same manner as was described in Sub-section 6.1.1.

The subjects undertook 8 training sessions during the experiment, with successive session being at least 24 hours apart. These training session commenced approximately 3 weeks after the subjects had participated in Experiment 7. All subjects undertook each training session on the same day. During each session they received four blocks of 60 trials at each of two levels of spatial and temporal pattern element separation. In the Close Condition the pattern elements were presented with an SOA of 94ms and a center-to-center spatial separation of 1.1cm. In the Far Condition the pattern elements were separated by an SOA of 450ms and a center-to-center distance of 5cm. These spacings were the same as those used in the closest and furthest spatial and temporal separation conditions in Experiment 7. The subjects underwent all four training blocks in a given condition back-to-back, with the training blocks for the other condition following after a 10 minute rest period. The order in which the two conditions were undertaken was counterbalanced between training sessions.

Due to limitations in the availability of the subjects, the 8 training session were not all undertaken on successive days. Sessions 1 to 3 were undertaken on successive days, then, after a 3 day break, sessions 4 to 6 were undertaken on successive days. After a final break of 7 days sessions 7 and 8 were undertaken on successive days.

The general procedure followed was similar to that used in Experiment 7. On each trial, the subject was presented with the target pattern then, 500ms latter, with the reference pattern. The PSE calibration values obtained for each subject in Experiment 7 were again used to adjust the amplitude of each pattern element as it was presented. The task was to judge whether the two pattern were the same or different then respond via two-choice button-press. Feedback was given to the subjects on a VDU following each trial.

6.2.2: Results

Figure 6.3, 6.4, and 6.5 present the mean accuracy level achieved in the Close and Far Conditions by each respective subject[2]. The pre-training scores given for subjects FH and RK are the scores achieved by those subjects in the equivalent conditions. The pre-training scores given for subject DM were those achieved by that subject during a pilot session for Experiment 7. The procedure followed during this pilot session was identical to that followed during the actual experiment.

Figures 6.3, 6.4, and 6.5 show that the three subjects exhibited different patterns of performance on the Far Condition task across training sessions. Subject DM showed no evidence of learning on this task. Although his scores fluctuated up and down from session to session, his performance on the first and last days of training was similar. This

---

2. Individual training block means are presented in Appendix B8.

descriptive observation was supported by regression analysis[3] which revealed no significant change in his pattern of discrimination performance across training sessions, $r=-.52$, $p>.05$. Remarkably, at no point during training did he achieve an accuracy level approaching his pre-training score. On the other hand, subject FH showed a significant downward trend in performance with training, $r=-.79$, $p<.05$. In spite of this, her performance during training was always above her pre-training level. Only subject RK showed any evidence of overall improvement in performance during training. Even in this case, her final level of performance was not markedly different from her pre-training level, with the trend for performance to increase with training failing to reach significance, $r=.17$, $p>.05$. The daily fluctuation in performance on each task can probably be attributed to the counterbalancing of the order in which the two tasks were undertaken on successive days of training.

Similarly, Figures 6.3, 6.4, and 6.5 show no evidence of a consistent learning effect across training sessions on the Close Condition task. Subject DM did show a slight improvement in performance with training, but this trend did not reach significance, $r=.53$, $p>.05$. Likewise, subject RK showed a non-significant trend for performance to improve with training, $r=.43$, $p>.05$. On the other hand, subject FH's level of accuracy fell significantly with training, $r=-.80$, $p<.05$. Where the

---

3. These correlations show the relationship between pattern discriminability and amount of training.
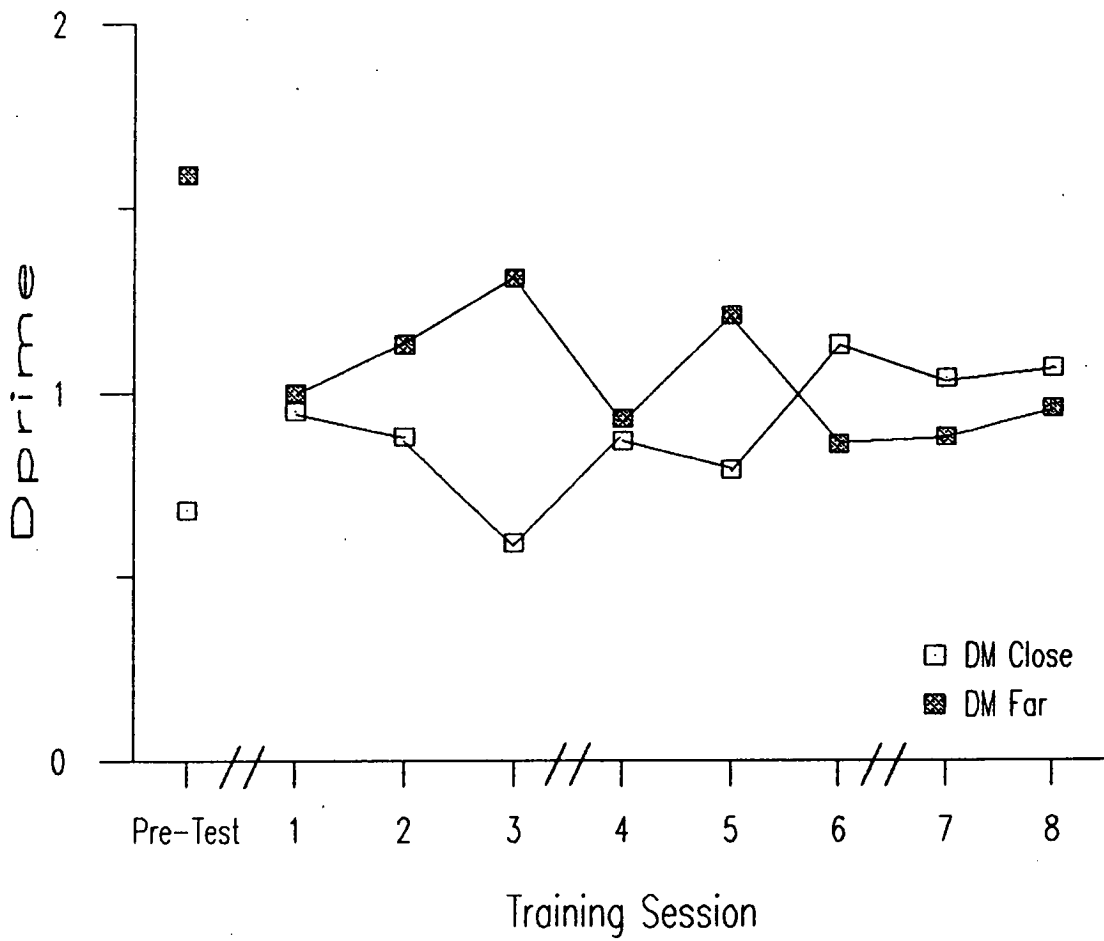
Figure 6.3: Accuracy levels a function of task and training session for subject DM.
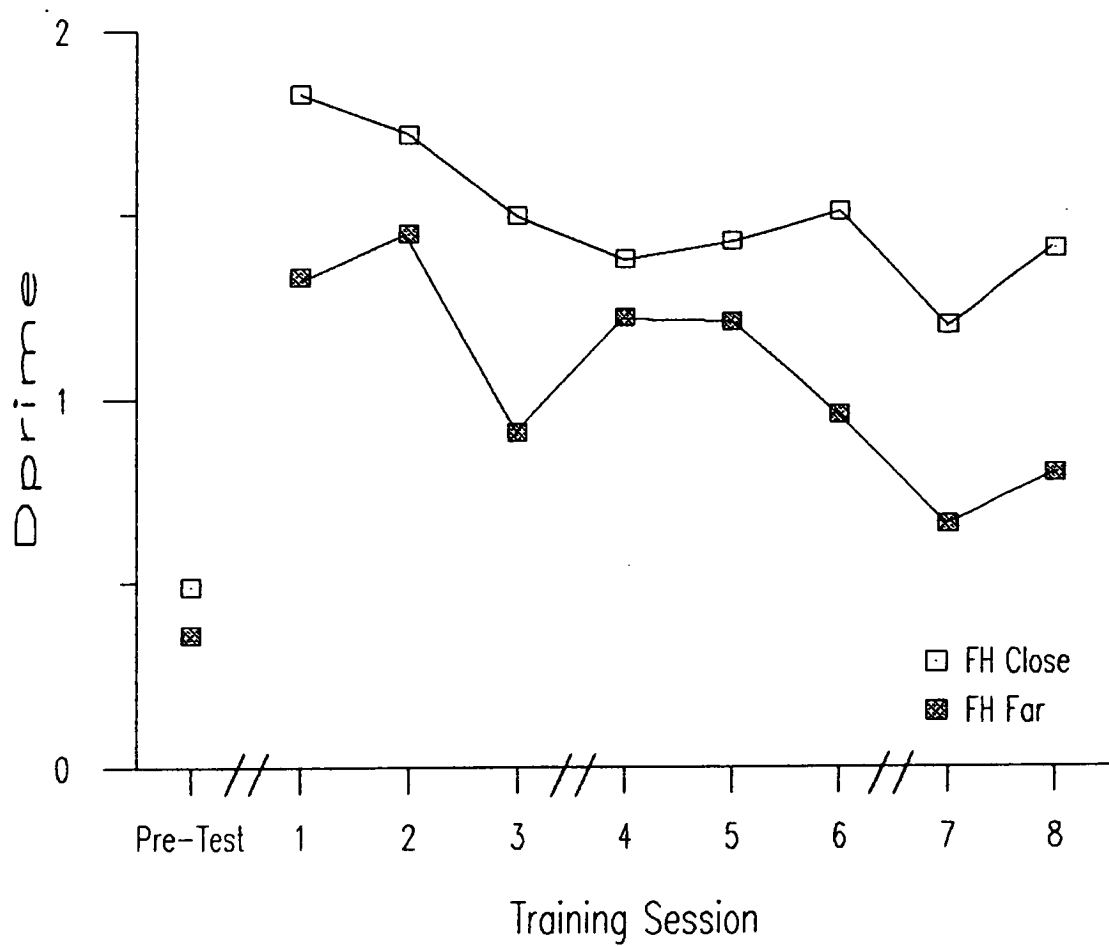
Figure 6.4: Accuracy levels as a function of task and training session for subject FH.
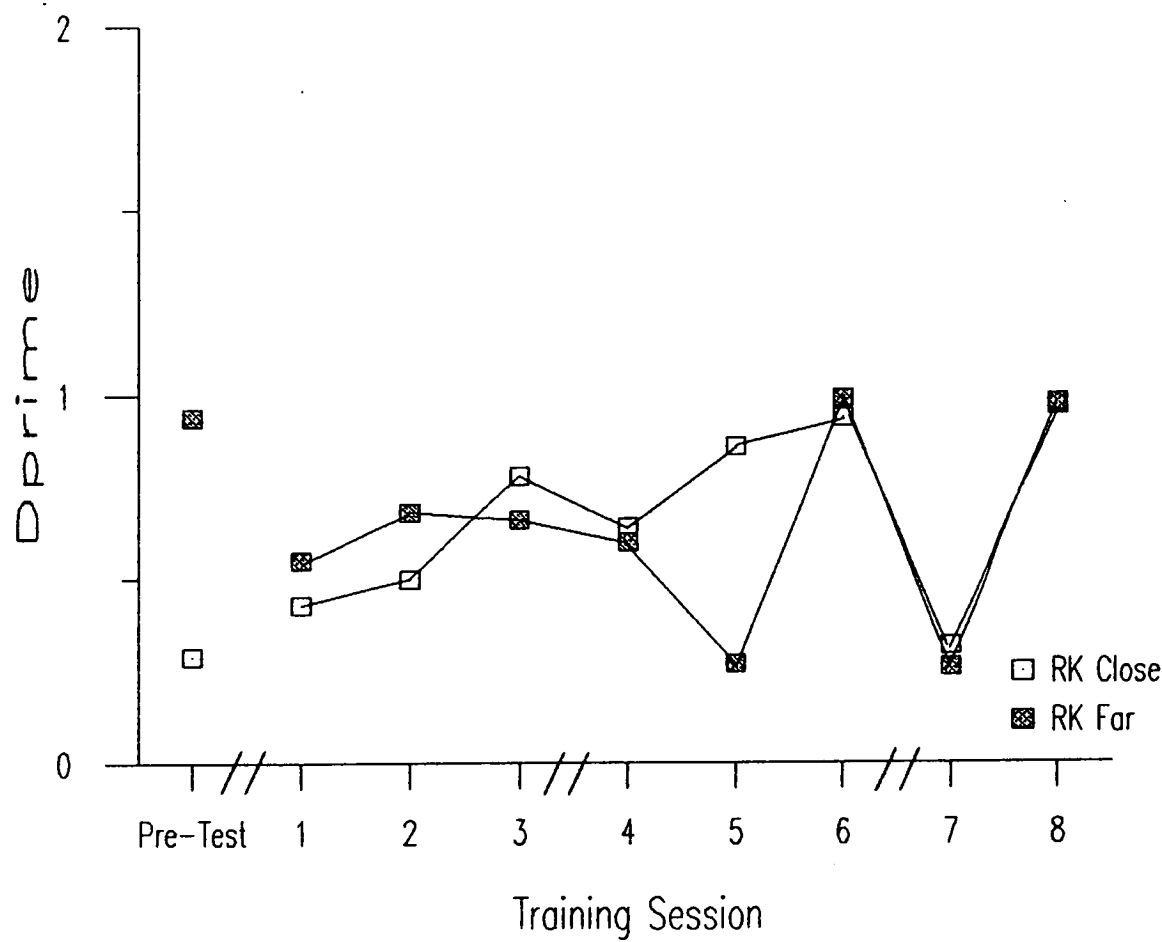
**Figure 6.5: Accuracy levels as a function of task and training session for subject RK.**

subjects' performance did differ from that observed in the Far conditions was in terms of their general improvement upon their pre-training scores. All subjects performed better on the first training session, and on all subsequent ones, than they did in the pretest session.

Perhaps the most important observation arises from a comparison of the subjects' performance on the Close and Far Condition tasks. At the completion of training, all subjects were performing the Close Condition task at least as well as the Far Condition task. Further, their level of performance on the Close Condition task at the completion of training was as good as, or better than, the best level of performance they recorded on any Far Condition training session. Finally, only in the case of DM was the best Far Condition score recorded during training better than the best Close Condition score.

6.2.3:  Discussion

This experiment did not provide any direct evidence that subjects' performance in either the Close Condition or the Far Condition improved consistently with training. The failure to observe general improvement in the subjects' performance in the Close Condition across training sessions does not support the hypothesis that the relatively poor discriminability of the closely spaced tactile patterns presented in Experiment 7 was due to the subjects lacking experience with that type of tactile display.

In spite of this, a comparison of the subjects' performance on the Close and Far Condition tasks in Experiment

7 with their performance on these tasks in the present experiment does highlight the need for further clarification of the results of Experiment 7. The subjects in this experiment consistently performed the Close Condition task better than they did in Experiment 7, and also consistently performed that task at least as well as the Far Condition task. A possible explanation of this is that, although the Far Condition task is initially easier to perform than the Close Condition, the limited amount of experience that the subjects had in performing the Close Condition task in Experiment 7 allowed them to improve their performance of that task during the present experiment to the level previously achieved only on the Far Condition task. If this trend persisted with a larger sample of subjects then it would provide support for the hypothesis that the strong evidence for the Isolation Hypothesis gathered in Experiment 7 was an artifact of the limited experience that the subjects had in discriminating closely spaced tactile patterns.

## 6.3: Conclusions

The research reported in this chapter set out to test the Integration and Isolation Hypotheses. The Integration Hypothesis maintains that masking effects act to facilitate the perception of complex patterns, and thus that optimum pattern discriminability occurs at some level of pattern element proximity above that where masking is minimized. On the other hand, the Isolation Hypothesis maintains that masking is always

destructive to the process of pattern perception, and thus should be minimized. An extensive review of these two hypotheses undertaken in Chapter 5 led to the development of an experimental design which it was expected would support the predictions of the Integration Hypothesis.

Experiment 7 tested these two hypotheses by measuring the accuracy with which pairs of tactile patterns could be compared as a function of the temporal and spatial proximity of the pattern elements. Contrary to expectations, it was found that performance on this task was best when masking was minimized, that is when wide temporal and spatial separations between pattern elements were employed.

Although this result is consistent with the Isolation Hypothesis, it is clear that several issues must be clarified before that hypothesis can be accepted. First, it is necessary to measure tactile pattern discrimination at closer levels of pattern element proximity than were employed here. If the Isolation Hypothesis is correct, then it would be expected that performance at these closer levels of pattern element proximity would remain below that observed when the pattern elements are far enough apart to minimize masking effects. Second, it is necessary to confirm that the present results in favour of the Isolation Hypothesis hold up as the level of tactile pattern complexity, or richness, is increased.

Finally, it is necessary to examine whether widely spaced tactile patterns are easier to discriminate than are closely spaced ones even after the observers have had

extensive experience with both these types of stimuli. The need for this type of research is supported by the anecdotal reports of the participants in Experiment 7, and by the similar levels of performance observed on both the Close and Far Condition tasks in Experiment 8.

In conclusion, these studies did not provide support for the Integration Hypothesis' claim that close proximity between tactile pattern elements enhances pattern perception via a process of perceptual integration. However, Experiment 8 did provide tentative evidence that the Isolation Hypothesis is wrong in arguing that performance at close element spacings must be worse than at wide element spacings. Regardless of the outcome of subsequent research on this question, the results will be of use in the development of tactile speech prostheses. In particular, studies such as these should allow the optimal level of pattern element proximity for use in complex tactile displays to be established.

**Chapter 7: Summary and Prospectus.**

This final chapter presents a brief summary of the present results, and suggests several issues which subsequent studies should investigate in order to clarify and extend the present findings.

## 7.1: Tactile Processing Abilities

Chapter 1 presented a review of the current status of tactile speech prostheses, concluding that, although these devices have great value as supplements to lipreading, the more ambitious goal of attempting to provide a tactile substitute for hearing should take priority. With this objective in mind, the immediate task became an assessment of whether touch is capable of dealing with stimuli as complex as representations of the speech spectrum.

Chapter 2 reviewed the question of tactile processing abilities in the context of the development of tactile speech prostheses, focusing first on Liberman's (Liberman, 1970; Liberman et al., 1967; Liberman et al., 1968) objection that only the auditory modality is equipped to deal with the speech signal. This position was rejected on the grounds that it is based on a false notion of the phoneme as the primary unit of speech perception, ignores the extent to which other modalities regularly deal with stimuli as complex as auditory ones, and includes an obvious circularity regarding the availability of the necessary processing resources in other modalities.

The second objection considered was that touch lacks the general sensory acuity to deal with complex patterns of stimulation. It was shown that some limitations, like the

narrow frequency bandwidth of the skin, can be overcome by appropriate modifications to the display design. However, in general, it was concluded that analyses of simple factors like spatial and temporal acuity do not allow an accurate assessment of the capacity of any modality to deal with complex patterns of stimulation. Instead, it is more likely that any limitations in tactile pattern processing ability occur at the higher level of touch's ability to deal with the complex percepts which arise from interactive effects between proximate stimuli.

Finally, Chapter 2 argued that the development of tactile speech prostheses may be facilitated by a process of comparison between the auditory and tactile modalities. When this process reveals differences between the two senses, the need to modify the tactile display to account for these differences arises. On the other hand, where similarities are found, the need for adjustment is minimized. This issue was pursued in Chapters 3 and 4, the results of which are summarized in the following section.

### 7.2: Comparisons Between Hearing and Touch

Previous studies indicated that there may be a greater similarity between auditory and tactile perceptual representations than between these and visual representations (Eilers et al., 1988; Handel & Buffardi, 1969; Mahar, 1985). The experiments reported in Chapters 3 and 4 confirmed and extended these earlier speculations.

Experiment 1 demonstrated that the visual displays used in Mahar's (1985) cross-modal comparison experiment were at

least as discriminable as the tactile one. This result confirmed that Mahar's (1985) demonstration that auditory and tactile representations of spoken words can be compared more readily than auditory and visual ones was not an artifact of variations in the discriminability of the tactile and visual displays he employed.

Experiment 2 extended the generality of Mahar's (1985) results by demonstrating that auditory and tactile representations of patterns can be compared more easily than auditory and visual ones regardless of the style of information distribution used in the tactile and visual displays. Experiment 3 demonstrated that the visual displays used in Experiment 2 could be processed at least as efficiently as the tactile ones, thus confirming that the superiority of tactile to auditory comparisons evident in Experiment 2 was not due to variations in the discriminability of the various displays used.

This result suggested that the difficulty which subjects exhibited in comparing auditory and visual patterns did not arise at the time that they were processing the visual displays, and thus must reflect higher level difficulties occurring at the time when the perceptual representations from each modality were compared. It was concluded that the most likely reason that these visual representations were more difficult to compare with auditory ones than were tactile representations was Handel and Buffardi's (1968) suggestion that auditory and tactile perceptual representations are more

alike than are auditory and visual ones.

Finally, Experiments 1, 3, and 4 provided tentative evidence that touch and vision may differ in terms of the relative efficiency with which each processes spatially and temporally distributed information. In particular it appeared that touch processes temporally distributed information more efficiently than spatially distribution, with the reverse being true in the case of vision.

Chapter 4 attempted to confirm these observation, and to extend them to the broader issue of the degree to which various modalities can be classified as spatial or temporal domains. A review of previous threshold and suprathreshold studies of the processing of spatially and temporally distributed information in hearing and vision suggested that these two modalities, like touch and vision, differ in terms of the relative efficiency with which each processes these two types of information. In particular, it appeared that hearing and touch share a common preference for temporally distributed information. Experiments 5 and 6 were conducted to confirm and extend these conclusions in the cases of touch and vision.

Experiment 5 confirmed the tactile data from Experiments 3 and 4, while Experiment 5 and 6 confirmed the visual data from Experiments 1 and 3. In all cases, temporally distributed information was processed more efficiently than spatially distributed information by touch, while spatially distributed information was processed with greater relative efficiency by vision. The generality of these findings was supported by the

observation that this pattern of results was obtained regardless of the particular tactile or visual stimulus characteristic manipulated in the display.

The visual and tactile data from Chapters 3 and 4, taken along with the results of previous studies of auditory and visual processing, demonstrate that both hearing and touch can be characterized as primarily temporal domains in terms of the efficiency with which each processes spatially and temporally distributed information. In contrast, it was shown that vision processes spatially distributed information more efficiently than temporally distributed information, and thus may be characterized as a primarily spatial domain.

It was concluded that this similarity between auditory and tactile perceptual processing may be one of the factors underlying the greater ease with which auditory to tactile comparisons can be performed relative to auditory to visual ones. Finally, it was argued that these results show that, while there may be a need to recode the primarily temporally structured speech signal into a spatially structured form for presentation via a visual display, there may be no need perform such a transformation in the case of tactile displays of speech.

Subsequent research should attempt to clarify the nature of and limits to the similarity between auditory and tactile perceptual and cognitive processes. One possible research direction is to determine whether the current demonstration that auditory and tactile patterns are relatively easy to

compare holds true when the patterns are modulated across stimulus dimensions other than amplitude. If the greater relative ease with which auditory to tactile comparisons are performed is, at least in part, due to the shared preference of these two modalities for temporally distributed information, then the present finding that this preference holds true in the case of touch regardless of the stimulus parameter manipulated suggests that auditory and tactile patterns should still be relatively easy to compare when the patterns are varied across other parameters (eg. frequency).

While further comparisons of this simple kind may be useful, Rabinowitz et al.'s (1987) study suggests that it is more likely that a multi-factor approach, that is one in which the tactile and auditory displays vary concurrently across a range of stimulus parameters, is required to accurately define the limits of the similarity between touch and hearing. They measured the IT rate of the tactile modality as additional display dimensions were included. One dimensional displays (using either frequency, amplitude, or duration as the information carrier) produced IT rates of 1 to 2 bits, while the richer three dimensional displays (using frequency, amplitude, and duration as concurrent information carriers) produced IT rates of 4 to 5 bits. These multi-dimensional IT rates did not appear to be a simple function of the IT rates obtained with each of the component dimensions. As tactile IT rates improve with added stimulus dimensions, and may not be predicted from uni-dimensional IT rates, it is clear that

questions like the degree of similarity between touch and hearing and the extent to which touch has the capacity to deal with transforms of the speech signal will only be answered by working with multi-dimensional displays.

One example of this type of research would be a continuation and extension of Eilers et al.'s (1988) study of the effects of changes in multiple stimulus parameters on the resulting tactile and auditory percepts. They found that the perceptual representations generated from auditory and tactile displays of speech segments varied in similar ways as the amplitude, frequency, and duration characteristics of the stimuli were manipulated. A sensible research strategy is to apply the cross modal comparison methodology used in Experiment 2 to the types of display used by Eilers et al. (1988). An analysis of variations in the ease with which these auditory and tactile patterns are compared as the multiple stimulus parameters are manipulated should allow a clearer specification of where and how auditory and perceptual processes are alike, and where and how they differ. Without this type of information we cannot complete a definitive specification of the ways in which the speech spectrum must be manipulated for optimum presentation to the skin.

### 7.3: Masking and Perceptual Integration

Chapters 5 and 6 took up this issue of the extent to which touch has the spatial and temporal resolution required to deal with speech transforms under the umbrella of two competing theories on the role of masking and perceptual integration in

tactile pattern perception. Briefly, the Isolation Hypothesis argues that the strong masking effects observed under conditions of close stimulus proximity impede the process of pattern perception in the tactile modality. Conversely, Kirman's (1973) Integration Hypothesis claims that masking is but one aspect of a beneficial process of perceptual integration which is necessary for efficient pattern processing in any modality. A review of the literature revealed that masking does appear to be associated with the generation of an integrated perceptual representation of target and mask features, but failed to find clear support for the Integration Hypothesis' contention that these integrative effects are beneficial to the process of pattern perception.

Experiment 7 was undertaken to test these two hypotheses in terms of their differing prediction about the effects of varying the proximity of pattern elements on the intelligibility of those patterns. It employed a novel design which was intended to circumvent the limited ability of traditional masking studies to show any positive aspects of this process. Contrary to the predictions of the Integration Hypothesis, it was found that the discriminability of both the patterns and of their elements fell as the proximity of the elements, and consequently the level of masking, was increased. Although this result strongly favoured the Isolation Hypothesis, it was noted that several issues required clarification before that hypothesis could be accepted.

Experiment 8 addressed one of these cavils, taking the

186

anecdotal reports of the subjects in Experiment 7 as its starting point, and investigating whether the proposed beneficial effects of perceptual integration would emerge with training. The results of this experiment were not clear-cut. While no consistent learning effect across training sessions was evident with either closely spaced or widely spaced patterns, it was found that both these types of patterns were discriminated equally well. A comparison of the results of Experiments 7 and 8 revealed that this effect was due to an improvement in pattern discrimination performance with closely spaced patterns rather than a decline in widely spaced pattern discrimination accuracy. It was suggested that this improvement may have been brought about by the experience, even if limited, which the subjects had in discriminating closely spaced tactile patterns during Experiment 7.

In summary, while Experiments 7 provided strong support for the Isolation Hypothesis, Experiment 8 provided tentative evidence that this result may have been an artifact of the limited experience that the subjects had in discriminating closely spaced tactile patterns prior to Experiment 7. Thus, while failing to support the Integration Hypothesis' claim that perceptual integration is beneficial to tactile pattern perception, these experiments did throw doubt on the Isolation Hypothesis' claim that tactile pattern perception is superior when masking effects are minimized.

What needs to be done to further clarify this issue? First, there is a clear need to replicate Experiment 7 using a

larger sample of subjects trained over a longer period of time. Such a study should include a wider range of spatial and temporal separation training conditions than was used in Experiment 8 in order to clearly identify the particular pattern element spacings which are likely to optimize the intelligibility of tactile displays of speech.

Further, it would be useful if such a study included regular measurements of the conventional recognition and detection masking functions of the subjects. Previous studies have shown greatly reduced levels of recognition masking, but normal levels of detection masking, in exceptional or highly trained observers of tactile and visual patterns (Craig, 1977; Hertzog et al., 1965; Schiller and Weiner, 1963; Wolford et al., 1988). Hence, this comparison should clarify whether improvements in tactile pattern processing capacity are associated with changes in the observer's recognition, but not detection, masking functions. Likewise, it may be useful to determine whether other highly experienced observers of tactile patterns, such as users of Braille, the Optacon, or Tadoma, also show significant reductions in their susceptibility to recognition masking.

Finally, future research on the role of perceptual integration and pattern element proximity on the intelligibility of tactile displays should employ more complex, multidimensional, tactile patterns. As was suggested in the case of subsequent research on the degree of similarity between hearing and touch, it is most likely that we must focus on this

type of display if we are to delimit the extent of tactile processing capacity. In particular, a full assessment of touch's capacity to deal with speech derived stimuli will not emerge until actual speech transforms are used. Speech stimuli have the advantage that the observer can ascribe meaning to the patterns. Once the patterns have meaning, the observer can begin to use context as a cue to aid in identifying the presented material. Surely it is this interaction between display features, meaning, and context which leads to high levels of communication in any modality.

If these subsequent studies show that the masking effects occurring at close levels of pattern element proximity do excessively impede the perception of tactile patterns then there are three means by which the extent of masking between pattern elements can be reduced; decreasing the temporal proximity of pattern elements, decreasing the spatial proximity of pattern elements, and varying the frequency of pattern elements. Unfortunately, the first of these alternatives carries a heavy penalty in terms of the speed of information transfer which can be achieved. Craig (1985a) explained that, because SOA is the temporal determinant of the strength of masking, any temporal modification which reduces masking must also increase the total time required to present the display. This fact imposes the limitation that speech could not be presented in real time.

Increasing the spatial separation is a more viable alternative. Indeed, Experiment 7 demonstrated that relatively

small changes in pattern element separation may lead to considerable reductions in the extent of masking. As Craig (1985b) has shown that the particular sites selected for stimulation effects the ease with which observers can attend to tactile stimuli, any attempt to implement a widely spaced tactile display should take care to ensure that attentional deficits are not induced.

A final way in which masking effects may be attenuated stems from the existence of a number of frequency specific channels within the tactile modality (Bolanowski et al., 1988). As was detailed in Chapter 2, one of the defining characteristics of these channels is the lack of masking interaction between them. Thus masking would be decreased if the vibrations presented to each transducer varied in frequency so that those occurring in close spatial and/or temporal proximity to one-another stimulate different channels. The major limitations of this scheme are the relatively small number of non-masking channels available, and the very low sensitivity of the Non-Pacinian channels.

In summary, Experiments 7 and 8 did not yield results which allow the Isolation vs Integration Hypothesis debate to be resolved. Never the less, these studies do point to further lines of research which should allow this debate, along with the issue of the optimum spatial and temporal parameters for use in tactile speech prostheses, to be settled.

## 7.4: Conclusions

This thesis adopted the approach of furthering the development of tactile speech prostheses by investigating the extent to which tactile information processes are compatible with, or can be adapted for, this purpose. Two particular foci were established; the extent to which tactile and auditory perceptual processes are alike, and the extent to which touch can deal with the complex patterns of stimulation required to represent the speech spectrum.

Investigation of the former issue did establish that touch and hearing share an underlying information processing affinity which may involve a shared processing preference for temporally, rather than spatially, distributed information. While investigation of the second focus in terms of establishing the extent to which masking and integrative effects either impede or facilitate the perception of complex tactile patterns proved inconclusive, that research did point to subsequent research directions which should resolve this issue.

In conclusion, while all these findings are of value, both in the design of tactile speech prostheses, and in terms of our general understanding of tactile information processing, they do not provide a definitive answer to the initial question of the extent to which touch has the processing capacities to deal with speech transforms. While subsequent research may ultimately show that touch is not able to support the fluent perception of speech transforms, in the process of discovering

this fact two benefits would emerge. First, we would be far better equipped to optimize the use of touch as a supplement to lipreading and other cues to speech perception. Second, our knowledge of tactile processes would be advanced to the point where Geldard's (1960) accusation that "We have not yet really begun to look carefully into the communication possibilities offered by the human integument or even into the bare facts that provide the possibilities." (p. 1588) might finally be dismissed.

# References

Abramsky, O., Carmon, A., & Benton, A.L. (1971). Masking of and by tactile pressure stimuli. Perception and Psychophysics, 10, 353-355.

Beauchamp, K.L., Matheson, D.W., & Scadden, L.A. (1971). Effects of stimulus-change method on tactile image recognition. Perceptual and Motor Skills, 33, 1067-1070.

Békésy, G, v. (1955). Human skin perception of traveling waves similar to those on the cochlea. Journal of the Acoustical Society of America, 27, 830-841.

Békésy, G, v. (1956). Current status of theories of hearing, Science, 123, 779-783.

Békésy, G, v. (1957a). Neural volleys and the similarity between some sensations produced by tones and by skin vibrations. Journal of the Acoustical Society of America, 29, 1059-1069.

Békésy, G, v. (1957b). Sensations on the skin similar to directional hearing, beats, and harmonics of the ear. Journal of the Acoustical Society of America, 29, 489-501.

Békésy, G, V. (1958). Funneling in the nervous system and its role in loudness and sensitivity on the skin. Journal of the Acoustical Society of America, 30, 399-412.

Békésy, G, v. (1959). Similarities between hearing and skin sensations. Psychological Review, 66, 1-22.

Békésy, G, v. (1960). Experiments in Hearing. New York: McGraw-Hill.

Békésy, G, v. (1967). Sensory Inhibition, Princeton: Princeton University Press.

Bernstein, L.E., Eberhardt, S.P., & Demorest, M.E. (1989). Single-channel vibrotactile supplement to visual perception of intonation and stress. Journal of the Acoustical Society of America, 85, 397-405.

Bjorkman, M. (1967). Relations between intra-modal and cross-modal matching. Scandinavian Journal of Psychology, 8, 65-76.

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

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

Blamey, P.J., Cowan, R.S.C., Alcantara, J.L., & Clark, G.M. (1988). Phonemic information transmitted by a multichannel electrotactile speech processor. _Journal of Speech and Hearing Research_, _31_, 620-629.

Blamey, P.J., Martin, L.F.A., & Clark, G.M. (1985). A comparison of three speech coding strategies using an acoustic model of a cochlear implant. _Journal of the Acoustical Society of America_, _74_, 209-217.

Bliss, J.C., Katchner, M.H., Rogers, C.H., & Shepard, R.P. (1970). Optical-to-tactile image conversion for the blind. _IEEE Transactions on Man-Machine Systems_, _MMS-11_, 58-63.

Bolanowski, S.L., Gescheider, G.A., Verrillo, R.T., & Checkosky, C.M. (1988). Four channels mediate the mechanical aspects of touch. _Journal of the Acoustical Society of America_, _84_, 1680-1694.

Breitmeyer, B.G. (1984). _Visual Masking: An Integrative Approach_. New York: Oxford University Press.

Breitmeyer, B.G. & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, sacadic suppression, and information processing. _Psychological Review_, _83_, 1-36.

Brooks, P.L. & Frost, B.J. (1983). Evaluation of a tactile vocoder for word recognition. _Journal of the Acoustical Society of America_, _74_, 34-39.

Brooks, P.L. & Frost, B.J. (1986). The development and evaluation of a tactile vocoder for the profoundly deaf. _Canadian Journal of Public Health_, _77_, 108-113.

Brooks, P.L., Frost, B.J., Mason, J.L., & Chung, K. (1985). Acquisition of a 250-word vocabulary through a tactile vocoder. _Journal of the Acoustical Society of America_, _77_, 1576-1579.

Brooks, P.L., Frost, B.J., Mason, J.L., & Gibson, D.M. (1986a). Continuing evaluation of the Queen's University Tactile Vocoder I: Identification of open set words. _Journal of Rehabilitation Research and Development_, _23_, 119-128.

Brooks, P.L., Frost, B.J., Mason, J.L., & Gibson, D.M. (1986b). Continuing evaluation of the Queen's University Tactile Vocoder II: Identification of open set sentences and tracking narrative. _Journal of Rehabilitation Research and Development_, _23_, 129-138.

Brooks, P.L., Frost, B.J., Mason, J.L., & Gibson, D.M. (1987). Word and feature identification by profoundly deaf teenagers using the Queen's University Tactile Vocoder. _Journal of Speech and Hearing Research_, _30_, 137-141.

Brown, R.L., Nibarber, D., Ollie, G., & Solomon, A. (1967). A differential comparison of two types of electropulse alphabets based on locus of stimulation. Perception and Motor Skills, 24, 1039-1044.

Bryden, M.P. (1972). Auditory-visual and sequential spatial matching in relation to reading ability. Child Development, 43, 824-832.

Carney, A.E. (1988). Vibrotactile perception of segmental features of speech: A comparison of single-channel and multichannel instruments. Journal of Speech and Hearing Research, 31, 438-448.

Cole, R., Radniky, A., Zue, V., & Reddy, D.R. (1979). Speech as patterns on paper, in R. Cole (Ed.) Perception and Production of Fluent Speech, Hillsdale: Erlbaum, (pp. 3-50).

Coquery, J. & Amblard, B. (1973). Backward and forward masking in the perception of cutaneous stimuli. Perception and Psychophysics, 13, 161-163.

Cowan, R.S.C., Alcantara, J.I., Blamey, P.J., & Clark, G.M. (1987). Interim results of open-set speech discrimination with a wearable multichannel electrotactile speech processor. Journal of the Acoustical Society of America, 82, 1456-1457.

Cowan, R.S.C., Alcantara, J.I., Blamey, P.J., & Clark, G.M. (1988). Preliminary evaluation of a multichannel electrotactile speech processor. Journal of the Acoustical Society of America, 83, 2328-2338.

Cowan, R.S.C., Alcantara, J.I., Whitford, L.A., Blamey, P.J., & Clark, G.M. (1989). Speech perception studies using a multichannel electrotactile speech processor, residual hearing, and lipreading. Journal of the Acoustical Society of America, 85, 2593-2607.

Craig, J.C. (1968). Vibrotactile spatial summation. Perception and Psychophysics, 4, 351-354.

Craig, J.C. (1977). Vibrotactile pattern perception: Extraordinary observers. Science, 196, 450-452.

Craig, J.C. (1980). Modes of vibrotactile pattern perception. Journal of Experimental Psychology: Human Perception and Performance, 6, 151-166.

Craig, J.C. (1981). Tactile letter recognition: Pattern duration and modes of pattern generation. Perception and Psychophysics, 30, 540-546.

Craig, J.C. (1982a). Temporal integration of vibrotactile patterns. Perception and Psychophysics, 32, 219-229.

Craig, J.C. (1982b). Vibrotactile masking: A comparison of energy and pattern maskers. Perception and Psychophysics, 31, 523-529.

Craig, J.C. (1983a). Some factors affecting tactile pattern recognition. International Journal of neuroscience, 19, 47-58.

Craig, J.C. (1983b). The role of onset in the perception of sequentially presented vibrotactile patterns. Perception and Psychophysics, 34, 421-432.

Craig, J.C. (1985a). Tactile pattern perception and its perturbations. Journal of the Acoustical Society of America, 77, 238-246.

Craig, J.C. (1985b). Attending to two fingers: Two hands are better than one. Perception and Psychophysics, 38, 496-511.

DeFilippo, C.L., & Scott, B.L. (1978). A method for training and evaluating the reception of ongoing speech. Journal of the Acoustical Society of America, 65, 1186-1191.

Eilers, R.E., Ozdamar, O., Oller, D.K., Miskiel, E., & Urbano, R. (1988). Similarities between tactual and auditory speech perception. Journal of Speech and Hearing research, 31, 124-131.

Elliott, L.L. (1962a). Backward masking: Monotonic and dichotic conditions. Journal of the Acoustical Society of America, 34, 1108-1115.

Elliott, L.L. (1962b). Backward and forward masking of probe tones of different frequencies. Journal of the Acoustical Society of America, 34, 1116-1117.

Engebretson, A.M., & O'Connell, M.P. (1986). Implementation of a microprocessor-based tactile hearing prosthesis. IEEE Transactions on Bio-Medical Engineering, BME-33, 712-716.

Evans, P.M. (1987). Vibrotactile masking: Temporal integration, persistence, and strengths of representations. Perception and Psychophysics, 42, 515-525.

Evans, P.M. & Craig, J.C. (1986). Temporal integration and vibrotactile backward masking. Journal of Experimental Psychology: Human Perception and Performance, 12(2), 160-168.

Evans, P.M. & Craig, J.C. (1987). Vibrotactile masking and the persistence of tactual features. Perception and Psychophysics, 42, 309-317.

Felsten, G., & Wasserman, S.S. (1980). Visual masking: Mechanisms and theories. Psychological Bulletin, 88, 329-353.

Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. Psychological Bulletin, 81, 284-310.

Garner, W.R. & Gottwald, R.L. (1965). The perception and learning of temporal patterns. The Quarterly Journal of Experimental Psychology, 20(2), 97-109.

Gault, R.H. (1924). Progress in experiments on tactual interpretation of oral speech. Journal of Abnormal and Social Psychology, 19, 155-159.

Gault, R.H. (1925). Progress in experiments on interpretation of speech by touch. Journal of Abnormal and Social Psychology, 20, 118-127.

Gault, R.H. (1930). On the effect of simultaneous tactual-visual stimulation in relation to the interpretation of speech. Journal of Abnormal and Social Psychology, 24, 498-517.

Gault, R.H. & Crane, G.W. (1928). Tactual patterns from certain vowel qualities instrumentally communicated from a speaker to a subject's fingers. Journal of General Psychology, 1, 353-359.

Gault, R.H., & Goodfellow, L.D. (1938). An empirical comparison of audition, vision, and touch in the discrimination of temporal patterns and ability to reproduce them. The Journal of General Psychology, 18, 41-47.

Geiselman, R.E. & Bellezza, F.S. (1976). Long-term memory for speaker's voice and source location. Memory and Cognition, 4, 483-489.

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

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

Geldard, F.A. (1972). The Human Senses. New York: Wiley.

Geldard, F.A. & Sherrick, C.E. (1972). The cutaneous 'rabbit': A perceptual illusion. Science, 178, 178-179.

Gelfand, S.A. (1981). Hearing: An Introduction to Psychological and Physiological Acoustics, New York: Marcel Dekker.

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

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

Gescheider, G.A., Bolanowski, S.J., & Verrillo, R.T. (1989). Vibrotactile masking: Effects of stimulus onset asynchrony and stimulus frequency. Journal of the Acoustical Society of America, 85, 2059-2064.

Gescheider, G.A., Herman, D.D., & Phillips, J.N. (1970). Criterion shifts in the measurement of tactile masking. Perception and Psychophysics, 8, 433-436.

Gescheider, G.A., & Joelson, J.M. (1983). Vibrotactile temporal summation for threshold and suprathreshold levels of stimulation. Perception and Psychophysics, 33, 156-162.

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

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

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

Gibson, J.J. (1966). The senses as considered perceptual systems, Boston: Houghton Mifflin.

Gilson, R.D. (1969a). Vibrotactile masking: Some spatial and temporal aspects. Perception and Psychophysics, 5(3), 176-180.

Gilson, R.D. (1969b). Vibrotactile masking: Effects of multiple maskers. Perception and Psychophysics, 5(3), 181-182.

Gilson, R.D. (1974). Vibrotactile masking. In F.A. Gelfand (Ed.), Cutaneous communication systems and devices (pp. 53-56). Austin: Psychonomic Society.

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

Grant, K.W., Ardell. L.H., Kuhl, P.K., & Sparks, D.W. (1985). The contribution of fundamental frequency, amplitude envelope, and voicing duration cues to speechreading in normal hearing subjects. Journal of the Acoustical Society of America, 77, 671-677.

Grant, K.W., Ardell. L.H., Kuhl, P.K., & Sparks, D.W. (1986). The transmission of prosodic information via an electrotactile speechreading aid. Ear and Hearing, 7(5), 328-335.

Green, B.G., Craig, J.C., & Pisoni, D.B. (1983). Vibrotactile communication of information about consonants: Vowels mask consonants. Perception and Psychophysics, 33, 507-515.

Green, B.G., Craig, J.C., Wilson, A.M., Pisoni, D.B., & Rhodes, R.P. (1983). Vibrotactile identification of vowels. Journal of the Acoustical Society of America, 73, 1766-1778.

Green, B.G., Pisoni, D.B., & Carrell, T.D. (1982). Learning to identify visual displays of speech: A first report. Journal of the Acoustical Society of America Supplement 1, 71, S96.

Green, D.M., & Swets, J.A. (1966). Signal Detection Theory and Psychophysics. New York: Wiley.

Gulick, W.L. (1971). Hearing: Physiology and Psychophysics, New York: Oxford University Press.

Guyot, G.W., Johnson, F.D., & Weaver, C. (1981). Two-point tactual discrimination: A signal detection approach. Perceptual and Motor Skills, 52, 433-434.

Haberlandt, K.F. & Baillet, S. (1977). Trade-off between remembering words and their source of location. Perceptual and Motor Skills, 45, 663-670.

Halliday, A.M. & Mingay, R. (1961). Retroactive raising of a sensory threshold by a contralateral stimulus. Quarterly Journal of Experimental Psychology, 8, 1-11.

Hamer, R.D., Verrillo, R.T., & Zwislocki, J.J. (1983). Vibrotactile masking of Pacinian and non-Pacinian channels. Journal of the Acoustical Society of America, 73, 1293-1303.

Handel, S. (1988a). Space is to time as vision is to audition: Seductive but misleading. Journal of Experimental Psychology: Human Perception and Performance, 14(2), 315-317.

Handel, S. (1988b). No one analogy is sufficient: Rejoinder to Kubovy. Journal of Experimental Psychology: Human Perception and Performance, 14(2), 321.

Handel, S., & Buffardi, L. (1968). Pattern perception: Integrating information presented in two modalities. Science, 162, 1026-1028.

Hawkins, J.E. & Stephens, S.S. (1950). The masking of pure tones and of speech by white noise. Journal of the Acoustical Society of America, 22, 6-13.

Hertzog, C.K., Williams, M.V., & Walsh, D.A. (1976). The effects of practice on age differences in central perceptual processing. Journal of Gerontology, 31, 428-433.

Hill, J.W. & Bliss, J.C. (1968). Perception of sequentially presented tactile point stimuli. Perception and Psychophysics, 4, 289-295.

Hirsh, I.J. & Sherrick, C.E. (1961). Perceived order in different sense modalities. The Journal of Experimental Psychology, 62, 423-432.

Iggo, A. (1976). Is the physiology of cutaneous receptors determined by morphology? In A. Iggo & O.B. Ilyinsky (Eds.), Somatosensory and visceral receptor mechanisms. (Progress in brain research, Vol.43), (pp. 15-31), Amsterdam: Elsevier.

Ittyerah, M. & Broota, K. (1983). Inter- and intra-modal processing of sensory-specific stimuli. Perceptual and Motor Skills, 56, 507-517.

Johnson, K.O. & Phillips, J.R. (1981). Tactile spatial resolution: I. Two-point discrimination, gap detection, grating resolution, and letter recognition. Journal of Neurophysiology, 46, 1177-1191.

Julesz, B., & Hirsh, I.J. (1972). Visual and auditory perception - An essay of comparison. In E.E. David & P. Denes (Eds.), Human Communication: A Unified View. (pp.283-340), New York: McGraw-Hill.

Kalat, J.W. (1984). Biological Psychology (2nd ed.). Belmont: Wadsworth Publishing Company.

Kaplan, R.M. (1987). Basic statistics for the behavioral sciences. Newton: Allyn and Bacon Inc.

Kenshalo, D.R. (1978). Biophysics and psychophysics of feeling. In E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception: VIB: Feeling and hurting. (pp. 29-74), New York: Academic Press.

Kirman, J.H. (1973). Tactile communication of speech: A review and analysis. Psychological Bulletin, 80(1), 54-74.

Kirman, J.H. (1984). Forward and backward tactile recognition masking. The Journal of General Psychology, 111, 83-89.

Kirman, J.H. (1986). Vibrotactile frequency recognition: Forward and backward masking effects. The Journal of General Psychology, 113(2), 147-158.

Klemmer, E.T. (1964). Does recoding from binary to octal improve the perception of binary patterns? Journal of Experimental Psychology, 67, 19-21.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? Journal of Experimental Psychology: Human Perception and Performance, 14(2), 318-320.

Lebedev, V.G. & Zagoruiko, N.G. (1985). Auditory perception and speech recognition. Speech Communication, 4, 97-103.

Lechelt, E.C. (1986). A comparison of local and remote masking on tactile pulse detection using different masking patterns. Perceptual and Motor Skills, 63, 343-351.

Liberman, A.M. (1970). The grammar of speech and language. Cognitive Psychology, 1, 301-323.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. Psychological Review, 74, 431-461.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1968). Why are speech spectrograms hard to read? American Annals of the Deaf, 113, 127-132.

Licklider, J.C.R. (1951). Basic correlates of the auditory stimulus. In S.S. Stevens (Ed.) Handbook of experimental Psychology, pp. 985-1039, New York: Wiley.

Loomis, J.M. (1981). Tactile pattern perception. Perception, 10, 5-27.

Loomis, J.M. & Collins, C.C. (1978). Sensitivity to shifts of a point stimulus: An instance of tactile hyperacuity. Perception and Psychophysics, 24, 487-492.

Mahar, D.P. (1985). Comparisons between auditory, vibrotactile, and visual representations of spoken words. Unpublished honors thesis, University of Tasmania, Hobart.

Marks, L.E. (1987a). On cross-modal similarity: Perceiving temporal patterns by hearing, touch, and vision. Perception and Psychophysics, 42, 250-256.

Marks, L.E. (1987b). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. Journal of Experimental Psychology: Human Perception and Performance. 13, 384-394.

Massaro, D.W. (1970). Preperceptual auditory images. Journal of Experimental Psychology, 85, 411-417.

Massaro, D.W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. Psychological Review, 79, 124-145.

Metcalfe, J., Glavanov, D., & Murdock, M. (1981). Spatial and temporal processing in the auditory and visual modalities. Memory and Cognition, 9, 351-359.

Michaels, C.F., & Turvey, M.T. (1979). Central sources of visual masking: Indexing structures supporting seeing at a single, brief glance. Psychological review, 41, 1-61.

Norton, S.J., Schultz, M.C., Reed, C.M., Braida, L.D., Durlach, N.I., Rabinowitz, W.M., & Chomsky, C. (1977). Analytic study of the tadoma method: Background and preliminary results. Journal of Speech and Hearing Research, 20, 574-595.

O'Connor, N., & Hermelin, B. (1972). Seeing and hearing and space and time. Perception and Psychophysics, 11, 46-48.

Pastore, R.E., Harris, L.B., & Goldstein, L. (1980). Auditory forward and backward masking interaction. Perception and Psychophysics, 28, 547-549.

Pickett, J.M. & Pickett, B.M. (1963). Communication of speech sounds by a tactual vocoder. Journal of Speech and Hearing Research, 6, 207-222.

Plant, G. (1983). The use of vibrotactile aids with profoundly deaf children. Speech Transmission Laboratory: Quarterly Progress and Status Report, 1/1986, 87-110.

Plant, G. (1986). A single-transducer vibrotactile aid to lipreading. Speech Transmission Laboratory - Quarterly Progress and Status Report, 1/1986, 41-63.

Plant, G., & Spens, K. (1986). An experienced user of tactile information as a supplement to lipreading. An evaluative study. Speech Transmission Laboratory: Quarterly Progress and Status Report, 1/1986, 87-110.

Pollack, I. & Johnson, L.B. (1965). Memory span with efficient coding procedures. American Journal of Psychology, 78, 609-614.

Pols, L.C.W. (1975). Analysis and synthesis of speech using broad-band spectral representations. In G. Fant & M.A.A. Tatham (Eds.), Auditory analysis and perception of speech, (pp. 23-36), London: Academic Press.

Potter, R.K. (1945). Visible patterns of sound. Science, 102, 463-470.

Potter, R.K., Kopp, G.A., & Green, H.C. (1947). Visible Speech, New York: Van Nostrand.

Rabinowitz, W.M., Houtsma, A.J.M., Durlach, N.I., & Delhorne, L.A. (1987). Multidimensional tactile displays: Identification of vibratory intensity, frequency, and contractor area. Journal of the Acoustical Society of America, 82, 1243-1252.

Reed, C.M., Doherty, M.J., Braida, L.D., & Durlach, N.I. (1982). Analytic study of the Tadoma method: Further experiments with inexperienced observers. Journal of Speech and Hearing research, 25, 216-223.

Reed, C.M., Rabinowitz, W.M., Durlach, N.I., Braida, L.D., Conway-Fithian, S., & Schultz, M. (1985). Research on the Tadoma method of speech communication. Journal of the Acoustical Society of America, 77, 247-257.

Richardson, B.L., & Frost, B.J. (1977). Sensory substitution and the design of an artificial ear. Journal of Psychology, 96, 259-285.

Richardson, B.L., & Frost, B.J. (1979). Tactile localization of the direction and distance of sounds. Perception and Psychophysics, 25, 336-344.

Richardson, B.L. & Wuillemin, D.B. (1981). Different orientations of sub-two-point threshold tactile stimuli can be discriminated. Bukketin of the Psychonomic Society, 18, 311-314.

Risberg, A. (1978). Requirements of speech processing aids for the profoundly deaf. Some preliminary results. Scandinavian Audiological Supplement, 6, 179-198.

Risberg, A., & Lubker, J.L. (1978). Prosody and speechreading. Speech Transmission Laboratory - Quarterly Progress and Status Report, 4/1978, 1-15.

Rubinstein, L. & Gruenberg, E.M. (1971). Intramodal and crossmodal sensory transfer of visual and auditory temporal patterns. Perception and psychophysics, 9(5), 385-390.

Sanders, J.W. (1984) Diagnostic Audiology. In J.L. Northern (Ed.) Hearing Disorders, pp. 25-39, Boston: Little, Brown and Company.

Savin, H.B., & Bever, T.G. (1970). The nonperceptual reality of the phoneme. Journal of verbal Learning and Verbal Behavior, 9, 295-302.

Scharf, B. (1970). Critical bands. In J.V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. 1, New York: Academic Press.

Scharf, B., & Meiselman, C.H. (1977). Critical bandwidth at high intensities. In E.F. Evans & J.P. Wilson (Eds.) Psychophysics and physiology of hearing, (pp. 221-235), London: Academic Press.

Schiller, P.H. (1965). Monoptic and dichoptic visual masking by patterns and flashes. Journal of Experimental Psychology, 69, 193-199.

Schiller, P.H., & Wiener, M. (1963). Monoptic and dichoptic visual masking. Journal of Experimental Psychology, 66, 386-393.

Schmid, E. (1961). Temporal aspects of cutaneous interaction with two-point electrical stimulation. Journal of Experimental Psychology, 61, 400-409.

Schultz, D.W. & Eriksen, C.W. (1977). Do noise masks terminate target processing? Memory and Cognition, 5, 90-96.

Sekuler, R. & Blake, R. (1985). Perception. New York: Alfred A. Knopf.

Sherrick, C.E. (1964). Effects of double simultaneous stimulation of the skin. American Journal of Psychology, 77, 42-53.

Sherrick, C.E. (1984). Basic and applied research on tactile aids for deaf people: Progress and prospects. Journal of the Acoustical Society of America, 75(5), 1325-1342.

Sherrick, C.E., & Rogers, R. (1966). Apparent haptic movement. Perception and Psychophysics, 1, 175-180.

Snyder, R.E. (1977). Vibrotactile masking: A comparison of psychophysical procedures. Perception and Psychophysics, 22, 471-475.

Spens, C.A., & Plant, G. (1983). A tactual "hearing" aid for the deaf. Speech Transmission Laboratory: Quarterly Progress and Status Report, 1/1983, 52-58.

Sterritt, G.M. & Rudnick, M. (1966). Auditory and visual rhythm perception in relation to reading ability in fourth grade boys. Perceptual and Motor Skills, 22, 859-864.

Stevens, S.S. (1959a). Tactile vibration: Dynamics of sensory intensity. Journal of Experimental Psychology, 57, 210-218.

Stevens, S.S. (1959b). Cross-modality validation of subjective scales for loudness, vibration, and electrical shock. Journal of Experimental Psychology, 57, 201-209.

Swets, J.A. (1964). Signal detection and recognition by human observers. New York: Wiley.

Talbot, W.H., Darian-Smith, I., Kornhuber, H.M., & Mountcastle, V.B. (1968). The sense of flutter-vibration: Comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. Journal of Neurophysiology, 31, 301-334.

Tong, Y.C., Blamey, P.J., Dowell, R.C., & Clark, G.M. (1983). Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant. Journal of the Acoustical Society of America, 74, 73-80.

Turvey, M.T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. Psychological Review, 80, 1-52.

Valvo, A. (1971). Sight Restoration After Long-Term Blindness: The Problems and Behavior Patterns of Visual Rehabilitation, New York: American Foundation for the Blind.

Verrillo, R.T. (1963). Effects of contractor area on vibrotactile threshold. Journal of the Acoustical Society of America, 35, 1962-1966.

Verrillo, R.T. (1965). Temporal summation in vibrotactile Journal of the Acoustical Society of America, 37, 834-846.

Verrillo, R.T. (1966a). Vibrotactile sensitivity and the frequency response of the Pacinian corpuscle. Psychonomic Science, 4, 135-136.

Verrillo, R.T. (1966b). Specificity of a cutaneous receptor. Perception and Psychophysics, 1, 149-153.

Verrillo, R.T. (1985). Psychophysics of vibrotactile stimulation. Journal of the Acoustical Society of America, 77, 225-232.

Verrillo, R. T., & Gescheider, G.A. (1977). Effects of prior stimulation on vibrotactile thresholds. Sensory Processes, 1, 292-300.

Warren, R.M. (1970). Perceptual restoration of missing speech sounds. Science, 167, 392-393.

Warren, R.M. (1971). Identification times for phonemic components of graded complexity and for spelling. Perception and psychophysics, 9, 345-349.

Warren, R.M. (1976). Auditory illusions and perceptual processes. in N.J. Lass (Ed.), Contemporary Issues in Experimental Phonetics. New York: Academic Press.

Warren. R.M. (1983). Auditory Illusions and their relation to mechanisms normally enhancing accuracy of perception. _Journal of the Audio Engineering Society_, _31_, 623-629.

Warren, R.M. (1984). Perceptual restoration of obliterated sounds. _Psychological Bulletin_, _96_, 371-383.

Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex and laterality. In D.R. Kenshalo (Ed.), _The skin senses_, (pp. 195-222), Springfield: Thomas.

Weisenberger, J.M. (1989). Evaluation of the Siemens Minifonator vibrotactile aid. _Journal of Speech and Hearing research_, _32_, 24-32.

Weisenberger, J.M., & Miller, J.D. (1987). The role of tactile aids in providing information about acoustic stimuli. _Journal of the Acoustical Society of America_, _82(2)_, 906-916.

Weisenberger, J.M. & Russell, A.F. (1989). Comparison of two single-channel vibrotactile aids for the hearing-impaired. _Journal of Speech and Hearing Research_, _32_, 83-92.

Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. _Acta Physiologica Scandinavia_, _31_, 285-289.

Wolford, G., Marchak, F., & Hughes, H. (1988). Practice effects in backward masking. _Journal of Experimental Psychology: Human Perception and performance_, _14_, 101-112.

Woodward, M.F., & Barber, C.G. (1960). Phoneme perception in lipreading. _Journal of Speech and Hearing Research_, _3_, 212-222.

Zwicker, E. & Fastl, H. (1972). On the development of the critical band. _Journal of the Acoustical Society of America_, _52_, 699-702.

# Appendices

## Appendix A: Signal Detection Theory

Signal Detection Theory (SDT) provides a means of measuring the accuracy with which perceptual and cognitive tasks are performed which is independent of the observer's response proclivity. Extensive discussion on the derivation and application of this theory may be found in Swets (1964) and Green and Swets (1966).

SDT was originally developed as an alternative to the traditional concept of absolute sensory thresholds, whereby a given signal only becomes detectable once its magnitude exceeds some threshold value. Consequently, SDT is most commonly discussed in terms of the detectability of such simple sensory stimuli. This discussion will, however, focus upon its application to higher level perceptual and cognitive tasks like those employed in the experiments reported in this thesis. To highlight the main points of SDT, this discussion will use the example of Mahar's (1985) experiment in which subjects were asked to determine whether two words presented sequentially as either auditory and tactile or auditory and visual representations were the same or different[1].

When two different words are presented to an observer the degree to which the cognitive representations of the two words differ will vary from trial to trial. For example, the words "cat" and "hat" share more features in common than the words "cat" and "dog". In addition to these innate variations between the stimuli, SDT proposes that other factors, like

---

1. See Section 3.1 for a description of this experiment.

momentary variations in the fidelity of the transducers used to present the stimuli and/or in the fidelity with which the stimuli are encoded by the respective modalities, will alter the extent to which the cognitive representations of the two different words appear alike. Likewise, these factors will cause the cognitive representations of two physically identical words to vary in terms of their degree of sameness from trial to trial. As the effects of these momentary factors should be weaker than that due to the physical identity of the words, it is reasonable to assume that the average degree of similarity between the cognitive representations of same word pairs will be higher than that of different word pairs. If we assume that, over a large sample of trials, the degree of similarity between both same and different word pairs will be normally distributed, then the degree of sameness of these two types of word pairs across this large number of trials can be represented by the two curves depicted in Figure A1.1.

SDT proposes that the observer determines whether the two words are the same or different by comparing the degree of similarity between their respective cognitive representations. When the degree of similarity between these two representations is above some criterion level set by the observer, the observer will judge that the two words were the same. Conversely, when the degree of similarity is below this criterion level, the observer will judge that the two words were different. This criterion level ($\beta$) is represented by the vertical division placed along the Decision Axis in Figure A1.1.
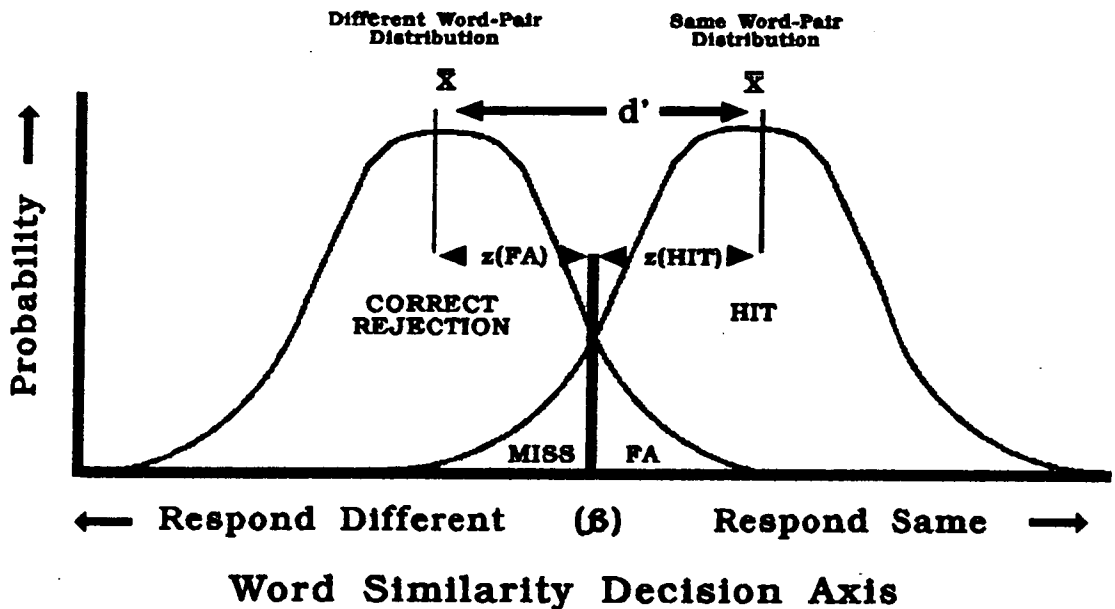
Figure A1.1: Hypothetical probability distribution of Same-Word pair and Different-Word pair trials as a function of the degree of similarity between each word. This figure also shows the possible outcomes of a "same" or "different" response based upon the position of the observer's decision criterion (β) along the decision axis.

It follows from Figure A1.1 that, provided sufficient trials are administered to allow the same and different word-pair representation distributions to approach normalcy, there will be occasions on which the degree of similarity between the cognitive representations of physically identical words will fall below β, and hence the observer will incorrectly judge that the two words were different. Likewise, there will be occasions on which the observer will incorrectly judge that two physically different words were the same. Of

course, as Figure A1.1 shows, the observer should also frequently correctly determine that the two words were either the same or different. These four possible outcomes of a given trial are summarized in Table A1.1.

Table A1.1: Stimulus-response table showing the four possible outcomes of any trial in an SDT experiment.

RESPONSE

|  |  | SAME | DIFFERENT |
|---|---|---|---|
| STIMULUS | SAME | HIT | MISS |
| | DIFFERENT | FALSE ALARM (FA) | CORRECT REJECTION (CR) |

As can be seen in Figure A1.1, these four response types represent areas under the same-word and different-word normal curves. In this case, the proportion of Hits recorded by a observer tells us the proportion of the same-word distribution lying above β, while the proportion of FA's recorded by the observer tells us the proportion of different-word distribution which also lies above β. By converting the proportion of Hits and False Alarms recorded by the observer into z-scores we can obtain a measure of the distance between β and the means of the respective trial type distributions. If we know the distance from β to the mean of each distribution, it follows that we can calculate the distance between the means of these two distributions by summing these two distances. SDT calls this

distance dprime (d'), and it may be calculated via the following formula;

$$d' = z(HIT) - z(FA)^2$$

SDT argues that the size of d' determines the accuracy with which the observer can discriminate same-word pairs from different-word pairs. As the cognitive representations of different-word and same-word pairs grow more alike, and consequently the distance between the same-word and different-word distributions decreases, then proportionally more same-word pair trials will lie below β and more different-word pair trials will lie above β. This will lead the observer to make more FA's and Misses, and less Hits and CR's. In simple terms, this means that as the cognitive representations of same-word and different-word pairs grow more alike, the observer will make more errors in judging whether the two words were the same or different. In summary, as d' increases from 0 (representing chance level performance), the accuracy with which the observer can make the required discrimination increases.

Finally, it should be noted that d' is independent of β. The observer's decision criterion only reflects the degree of similarity between the two presented words required before they will be judged to be identical. The placement of β along the decision axis is under the control of the observer, and may vary both between observers and between observations for the

---

2. Although dprime is the sum of the distances from β to the means of each distribution, the computational formula subtracts z(FA) from z(HIT) to account for the change of the sign of z from negative when probability (p) lies below 0.5 to positive when p lies above 0.5.

same observer. The only factor which the position of β determines is the proclivity of the observer to give either a "same" or "different" response. On the other hand, d' is not under the control of the observer, it is determined by an interaction between the degree of physical similarity between the stimuli and the various factors which may act to alter the degree of similarity between the cognitive representations of those stimuli. Thus, while β may be set by the observer at any point along the decision axis, this does not alter the relative positions of the two distributions, and thus does not alter d'.

Appendix B1: Mean accuracy (d') scores for each subject in Experiment1.

| Subject N° | Tactile Temporal | Visual Spatiotemporal | Visual Temporal |
|------------|------------------|-----------------------|-----------------|
| 1 | 1.13 | 1.69 | 0.66 |
| 2 | 1.20 | 1.57 | 1.36 |
| 3 | 0.92 | 2.23 | 0.62 |
| 4 | 1.07 | 2.02 | 0.87 |
| 5 | 1.39 | 2.41 | 1.03 |
| Mean | 1.14 | 1.98 | 0.91 |
| SD | 0.15 | 0.32 | 0.27 |

**Appendix B2: Mean accuracy (d') and reaction time (sec) scores for each subject in Experiment 2.**

| | Tactile Target | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Temporal | | Spatial | | Spatial-Delayed | | Spatiotemporal | |
| | d' | RT(s) | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| | 2.11 | 1.50 | 1.90 | 1.73 | 1.63 | 2.01 | 2.47 | 1.45 |
| | 2.58 | 1.41 | 2.79 | 1.61 | 1.06 | 1.79 | 1.82 | 1.47 |
| | 2.72 | 1.72 | 2.07 | 1.71 | 2.32 | 1.62 | 1.82 | 1.52 |
| | 1.75 | 1.62 | 1.90 | 1.49 | 2.88 | 1.92 | 1.14 | 1.55 |
| | 2.04 | 1.58 | 0.54 | 1.96 | 2.72 | 2.02 | 0.71 | 1.32 |
| Mean | 2.24 | 1.57 | 1.84 | 1.70 | 2.12 | 1.87 | 1.59 | 1.46 |
| SD | 0.36 | 0.11 | 0.73 | 0.16 | 0.68 | 0.15 | 0.61 | 0.08 |

| | Visual Target | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Temporal | | Spatial | | Spatial-Delayed | | Spatiotemporal | |
| | d' | RT(s) | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| | 1.42 | 1.75 | 1.28 | 1.56 | 0.90 | 1.64 | 1.63 | 1.79 |
| | 1.27 | 1.78 | 1.27 | 1.82 | 1.68 | 1.69 | 0.64 | 1.91 |
| | 0.71 | 1.58 | 1.39 | 2.02 | 1.88 | 2.02 | 2.59 | 1.82 |
| | 0.38 | 1.80 | 1.99 | 1.62 | 1.13 | 1.74 | 1.28 | 1.96 |
| | 1.87 | 1.66 | 0.64 | 1.66 | 1.19 | 2.00 | 2.43 | 1.52 |
| Mean | 1.13 | 1.71 | 1.31 | 1.74 | 1.36 | 1.82 | 1.71 | 1.80 |
| SD | 0.53 | 0.08 | 0.43 | 0.17 | 0.36 | 0.16 | 0.73 | 0.15 |

Appendix B3

Appendix B3: Mean accuracy (d') and reaction time (sec) scores for each subject in Experiment 3.

Touch

| Subject N | Temporal | | Spatial | | Spatiotemporal | |
|---|---|---|---|---|---|---|
| | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| 1 | 4.07 | 0.92 | 2.36 | 1.30 | 2.44 | 1.43 |
| 2 | 2.71 | 0.71 | 1.10 | 0.92 | 2.89 | 0.47 |
| 3 | 3.77 | 0.58 | 2.71 | 0.72 | 3.20 | 0.62 |
| 4 | 2.89 | 0.52 | 1.41 | 0.80 | 3.77 | 0.63 |
| 5 | 3.43 | 0.48 | 1.64 | 0.70 | 3.03 | 0.63 |
| 6 | 2.89 | 0.49 | 1.78 | 0.62 | 2.76 | 0.54 |
| 7 | 3.77 | 0.54 | 1.97 | 0.75 | 2.89 | 0.73 |
| 8 | 1.83 | 0.87 | 1.39 | 1.07 | 2.21 | 0.81 |
| 9 | 1.64 | 0.69 | 1.03 | 0.99 | 2.77 | 0.62 |
| 10 | 2.57 | 0.78 | 1.60 | 0.76 | 1.97 | 0.78 |
| 11 | 4.07 | 0.67 | 2.89 | 0.83 | 2.76 | 0.76 |
| 12 | 2.44 | 0.72 | 1.60 | 0.70 | 0.98 | 0.72 |
| Mean | 3.01 | 0.66 | 1.79 | 0.84 | 2.64 | 0.72 |
| SD | 0.79 | 0.14 | 0.57 | 0.18 | 0.67 | 0.24 |

Vision

| Subject N | Temporal | | Spatial | | Spatiotemporal | |
|---|---|---|---|---|---|---|
| | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| 13 | 1.95 | 1.35 | 2.71 | 1.06 | 1.78 | 0.97 |
| 14 | 2.07 | 1.07 | 2.21 | 0.53 | 2.16 | 0.55 |
| 15 | 2.02 | 0.69 | 1.94 | 0.60 | 3.77 | 0.58 |
| 16 | 1.14 | 0.54 | 2.11 | 0.53 | 1.64 | 0.55 |
| 17 | 2.21 | 0.68 | 1.92 | 0.46 | 2.89 | 0.43 |
| 18 | 0.99 | 0.95 | 1.83 | 0.54 | 1.84 | 0.65 |
| 19 | 2.25 | 0.51 | 3.03 | 0.46 | 2.25 | 0.58 |
| 20 | 1.92 | 0.66 | 2.44 | 0.65 | 1.71 | 0.49 |
| 21 | 2.11 | 0.61 | 2.32 | 0.62 | 2.21 | 0.53 |
| 22 | 1.95 | 0.66 | 1.74 | 0.55 | 2.71 | 0.60 |
| 23 | 2.21 | 0.78 | 2.32 | 0.59 | 2.54 | 0.56 |
| 24 | 1.54 | 0.50 | 2.44 | 0.53 | 1.44 | 0.50 |
| Mean | 1.86 | 0.75 | 2.25 | 0.59 | 2.25 | 0.58 |
| SD | 0.40 | 0.24 | 0.36 | 0.15 | 0.63 | 0.13 |

**Appendix B4: Mean accuracy (d') and reaction time (sec) scores for each subject in Experiment 4.**

| Subject | Temporal Distribution | | | | | |
|---|---|---|---|---|---|---|
| | Both | | First | | Third | |
| N° | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| 1 | 2.30 | 0.61 | 3.20 | 0.53 | 2.35 | 0.56 |
| 2 | 2.50 | 0.48 | 2.20 | 0.42 | 1.49 | 0.63 |
| 3 | 4.10 | 0.50 | 3.77 | 0.49 | 3.46 | 0.53 |
| 4 | 2.35 | 0.85 | 2.77 | 0.74 | 2.53 | 0.79 |
| 5 | 0.98 | 0.65 | 0.21 | 0.49 | 0.42 | 0.74 |
| 6 | 3.12 | 0.58 | 3.77 | 0.60 | 3.77 | 0.75 |
| 7 | 2.09 | 0.64 | 2.20 | 0.46 | 2.77 | 0.53 |
| 8 | 2.53 | 0.50 | 3.46 | 0.54 | 2.35 | 0.50 |
| 9 | 1.78 | 0.55 | 1.35 | 0.56 | 2.20 | 0.78 |
| Mean | 2.42 | 0.60 | 2.55 | 0.54 | 2.37 | 0.65 |
| SD | 0.81 | 0.11 | 1.13 | 0.09 | 0.94 | 0.11 |

| Subject | Spatial Distribution | | | |
|---|---|---|---|---|
| | First & Third | | First & Fourth | |
| N° | d' | RT(s) | d' | RT(s) |
| 10 | 2.05 | 0.78 | 1.28 | 0.81 |
| 11 | 1.10 | 0.86 | 0.54 | 0.97 |
| 12 | 1.40 | 0.58 | 1.70 | 0.61 |
| 13 | 3.42 | 0.67 | 1.34 | 0.75 |
| 14 | 1.61 | 0.66 | 1.28 | 0.71 |
| 15 | 1.62 | 0.79 | 1.70 | 0.83 |
| 16 | 2.53 | 0.61 | 2.19 | 0.61 |
| 17 | 2.41 | 0.62 | 2.35 | 0.73 |
| Mean | 2.02 | 0.70 | 1.55 | 0.75 |
| SD | 0.70 | 0.09 | 0.54 | 0.11 |

**Appendix B5: Mean accuracy (d') and reaction time (sec) scores for each subject in Experiment 5.**

| Subject N° | Touch | | | | | |
|---|---|---|---|---|---|---|
| | Temporal | | Spatial | | Spatiotemporal | |
| | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| 1 | 0.21 | 0.96 | 0.35 | 1.03 | 1.54 | 0.61 |
| 2 | 0.76 | 0.93 | 0.51 | 1.47 | 2.02 | 1.36 |
| 3 | 1.30 | 0.80 | 0.22 | 0.92 | 0.21 | 0.67 |
| 4 | 1.97 | 0.57 | -0.48 | 1.42 | 0.93 | 0.57 |
| 5 | 1.59 | 1.07 | 0.50 | 1.46 | 1.14 | 1.10 |
| 6 | 0.76 | 1.50 | 0.22 | 1.78 | 1.44 | 1.26 |
| 7 | 1.36 | 0.62 | 0.22 | 0.55 | 1.62 | 0.45 |
| 8 | 1.95 | 1.19 | 0.21 | 1.12 | 0.75 | 0.90 |
| 9 | 2.35 | 0.79 | 0.42 | 1.27 | 1.74 | 0.67 |
| 10 | 2.35 | 0.83 | 0.42 | 0.82 | 1.41 | 0.60 |
| 11 | 2.30 | 0.70 | -0.64 | 1.17 | 1.84 | 0.77 |
| 12 | 2.68 | 0.78 | -0.05 | 1.31 | 1.31 | 0.72 |
| Mean | 1.63 | 0.90 | 0.16 | 1.19 | 1.33 | 0.81 |
| SD | 0.74 | 0.25 | 0.35 | 0.32 | 0.49 | 0.28 |

| Subject N° | Vision | | | | | |
|---|---|---|---|---|---|---|
| | Temporal | | Spatial | | Spatiotemporal | |
| | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| 13 | 1.30 | 0.82 | 1.39 | 0.79 | 1.25 | 0.66 |
| 14 | 1.41 | 1.19 | 0.75 | 1.16 | 2.07 | 0.87 |
| 15 | 0.33 | 0.59 | 1.06 | 0.44 | 0.55 | 0.44 |
| 16 | 1.50 | 1.09 | 1.29 | 0.79 | 2.12 | 0.56 |
| 17 | 1.09 | 1.05 | 0.53 | 0.60 | 1.39 | 0.62 |
| 18 | 0.99 | 1.50 | 1.36 | 1.19 | 0.96 | 1.01 |
| 19 | 0.45 | 0.37 | -0.11 | 0.36 | 0.23 | 0.28 |
| 20 | 1.84 | 0.88 | 2.02 | 0.61 | 2.44 | 0.60 |
| 21 | 1.03 | 0.90 | 0.66 | 0.71 | 1.64 | 0.60 |
| 22 | 1.93 | 1.17 | 1.35 | 1.14 | 1.11 | 0.68 |
| 23 | 1.41 | 0.56 | 1.03 | 0.66 | 1.03 | 0.59 |
| 24 | 1.09 | 0.86 | 2.12 | 0.70 | 2.54 | 0.74 |
| Mean | 1.20 | 0.92 | 1.12 | 0.76 | 1.44 | 0.64 |
| SD | 0.46 | 0.30 | 0.59 | 0.26 | 0.70 | 0.18 |

**Appendix B6: Mean accuracy (d') and reaction time (sec) scores for each subject in Experiment 6.**

| Subject N° | Temporal | | Vision Spatial | | Spatiotemporal | |
|---|---|---|---|---|---|---|
| | d' | RT(s) | d' | RT(s) | d' | RT(s) |
| 1 | 3.44 | 0.43 | 3.09 | 0.41 | 3.41 | 0.37 |
| 2 | 3.43 | 0.57 | 4.07 | 0.53 | 4.07 | 0.47 |
| 3 | 2.89 | 0.86 | 3.77 | 0.47 | 4.07 | 0.45 |
| 4 | 2.71 | 0.48 | 2.48 | 0.49 | 3.46 | 0.42 |
| 5 | 3.12 | 0.40 | 3.12 | 0.66 | 2.53 | 0.45 |
| 6 | 2.07 | 0.43 | 3.43 | 0.46 | 4.07 | 0.47 |
| 7 | 3.46 | 0.57 | 3.20 | 0.46 | 4.07 | 0.43 |
| 8 | 3.02 | 0.54 | 3.43 | 0.51 | 4.07 | 0.43 |
| 9 | 3.46 | 0.84 | 4.07 | 0.47 | 3.43 | 0.40 |
| 10 | 2.89 | 0.79 | 3.12 | 0.54 | 4.07 | 0.41 |
| 11 | 3.77 | 0.38 | 3.77 | 0.46 | 4.07 | 0.39 |
| 12 | 2.71 | 0.41 | 3.46 | 0.44 | 3.12 | 0.46 |
| 13 | 3.20 | 0.46 | 3.12 | 0.42 | 3.77 | 0.37 |
| 14 | 2.57 | 0.72 | 4.07 | 0.51 | 3.77 | 0.42 |
| 15 | 1.64 | 0.99 | 2.76 | 0.60 | 2.20 | 0.47 |
| 16 | 3.12 | 0.97 | 3.77 | 0.65 | 3.46 | 0.59 |
| 17 | 4.07 | 0.52 | 4.07 | 0.48 | 3.77 | 0.42 |
| 17 | 3.46 | 0.37 | 3.77 | 0.48 | 3.12 | 0.39 |
| Mean | 3.06 | 0.60 | 3.48 | 0.50 | 3.59 | 0.43 |
| SD | 0.57 | 0.20 | 0.46 | 0.07 | 0.54 | 0.05 |

**Appendix B7: Mean accuracy (d') scores for each subject in Experiment 7.**

| Subject N° | Element Identification Task | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pattern Element Spacing | | | | | | | | |
| | 94ms 1cm | 94ms 3cm | 94ms 5cm | 250ms 1cm | 250ms 3cm | 250ms 5cm | 450ms 1cm | 450ms 3cm | 450ms 5cm |
| 1 | -0.06 | 0.29 | 0.67 | 0.72 | 1.12 | 0.67 | 1.00 | 0.89 | 1.12 |
| 2 | 0.34 | 0.23 | 0.94 | 0.50 | 0.67 | 0.29 | 0.45 | 0.94 | 1.30 |
| 3 | 0.17 | 1.12 | 0.89 | 0.34 | 0.61 | 0.72 | 0.61 | 0.83 | 1.00 |
| 4 | 0.00 | 0.00 | 0.89 | 0.50 | 0.67 | 0.72 | 1.82 | 0.72 | 1.82 |
| 5 | 0.56 | 0.42 | 0.50 | 1.43 | 1.12 | 1.30 | 2.22 | 1.49 | 2.00 |
| 6 | 1.30 | 0.78 | 1.37 | 2.22 | 2.00 | 1.65 | 2.11 | 2.94 | 2.94 |
| 7 | 0.94 | 0.72 | 1.90 | 1.58 | 2.11 | 1.58 | 1.82 | 7.05 | 3.32 |
| 8 | 0.02 | 0.08 | 0.29 | 0.17 | 0.50 | 0.29 | 0.89 | 1.73 | 1.24 |
| 9 | 0.40 | 0.94 | 1.00 | 1.37 | 0.89 | 1.00 | 0.78 | 2.11 | 1.43 |
| 10 | 0.29 | 0.40 | 0.45 | 1.18 | 1.43 | 1.50 | 1.65 | 1.82 | 2.00 |
| 11 | 0.06 | 0.23 | 0.78 | 0.40 | 0.56 | 0.67 | 1.73 | 1.90 | 1.58 |
| 12 | 0.56 | 0.50 | 0.72 | 1.43 | 1.58 | 1.24 | 1.58 | 1.43 | 1.90 |
| 13 | 0.56 | 0.72 | 0.78 | 1.30 | 1.24 | 1.12 | 2.94 | 2.51 | 2.11 |
| 14 | 0.00 | 0.12 | 0.83 | 0.20 | 0.50 | 0.83 | 0.67 | 0.61 | 0.67 |
| 15 | 0.72 | 1.06 | 1.06 | 1.82 | 1.58 | 1.06 | 1.73 | 2.22 | 2.51 |
| 16 | 0.56 | 0.67 | 0.83 | 1.00 | 1.82 | 1.50 | 2.36 | 2.11 | 1.90 |
| 17 | 1.06 | 0.78 | 0.94 | 1.65 | 1.73 | 1.82 | 1.12 | 1.73 | 1.90 |
| 18 | 0.50 | 0.34 | 0.89 | 0.45 | 0.83 | 1.06 | 1.65 | 1.18 | 2.00 |
| Mean | 0.44 | 0.52 | 0.87 | 1.02 | 1.16 | 1.06 | 1.51 | 1.90 | 1.82 |
| SD | 0.38 | 0.33 | 0.34 | 0.60 | 0.53 | 0.44 | 0.66 | 1.40 | 0.65 |

| Subject N° | Pattern Comparison Task | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pattern Element Spacing | | | | | | | | |
| | 94ms 1cm | 94ms 3cm | 94ms 5cm | 250ms 1cm | 250ms 3cm | 250ms 5cm | 450ms 1cm | 450ms 3cm | 450ms 5cm |
| 1 | 0.87 | 1.05 | 1.07 | 0.68 | 0.61 | 0.83 | 1.13 | 1.65 | 1.16 |
| 2 | 0.29 | 0.09 | 0.08 | 0.57 | 0.71 | 1.07 | 0.77 | 0.69 | 0.94 |
| 3 | 0.75 | 1.39 | 1.10 | 1.19 | 0.55 | 0.66 | 1.28 | 1.15 | 0.90 |
| 4 | 0.78 | 1.46 | 0.88 | 1.05 | 1.37 | 1.41 | 1.15 | 1.65 | 0.79 |
| 5 | 0.87 | 0.59 | 0.51 | 1.50 | 0.90 | 1.05 | 0.83 | 1.46 | 0.68 |
| 6 | 0.99 | 0.42 | 1.25 | 0.99 | 0.71 | 1.97 | 1.97 | 0.55 | 0.38 |
| 7 | 0.51 | 0.52 | 0.59 | 0.95 | 1.19 | 1.15 | 0.08 | 0.99 | 0.95 |
| 8 | 0.13 | 1.00 | 0.77 | 0.38 | 0.08 | 1.05 | -0.12 | 0.57 | 0.13 |
| 9 | 0.00 | 0.36 | 1.10 | 0.95 | 0.88 | 0.88 | 0.25 | 0.87 | 0.97 |
| 10 | 0.51 | 0.87 | 1.07 | 0.53 | 0.78 | 0.90 | 0.87 | 1.19 | 1.10 |
| 11 | 0.68 | 1.05 | 1.24 | 1.50 | 1.84 | 1.82 | 1.19 | 0.74 | 1.15 |
| 12 | 1.50 | 0.27 | 1.07 | 0.42 | 0.82 | 1.15 | 0.71 | 0.53 | -0.19 |
| 13 | 0.49 | 0.88 | 0.46 | 0.83 | 0.99 | 0.55 | 1.38 | 1.65 | 0.36 |
| 14 | 0.20 | 0.00 | 0.69 | 0.53 | 0.86 | 1.15 | 0.66 | 0.52 | 1.26 |
| 15 | 0.68 | 0.25 | 0.34 | 0.77 | 0.77 | 0.94 | 0.57 | 1.24 | 1.59 |
| 16 | 0.95 | 0.57 | 0.71 | 1.65 | 0.69 | 0.90 | 1.65 | 1.92 | 1.56 |
| 17 | 0.61 | 0.61 | 1.02 | 1.49 | 1.74 | 1.41 | 1.65 | 1.65 | 2.24 |
| 18 | 0.90 | 1.33 | 0.82 | 0.99 | 0.87 | 1.33 | 0.17 | 0.94 | 0.69 |
| Mean | 0.65 | 0.71 | 0.82 | 0.94 | 0.91 | 1.12 | 0.90 | 1.11 | 0.93 |
| SD | 0.35 | 0.43 | 0.32 | 0.39 | 0.40 | 0.35 | 0.56 | 0.45 | 0.55 |

**Appendix B8: Mean accuracy (d') scores for each subject in Experiment 8.**

| Subject & Condition | DAY 1 Session N° | | | | | | DAY 2 Session N° | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | Mean | SD | 1 | 2 | 3 | 4 | Mean | SD |
| FH Close | 1.57 | 1.57 | 1.92 | 2.25 | 1.83 | 0.28 | 1.19 | 1.80 | 2.20 | 1.70 | 1.72 | 0.36 |
| FH Far | 1.60 | 1.20 | 0.90 | 1.60 | 1.33 | 0.29 | 1.40 | 1.80 | 2.00 | 0.60 | 1.45 | 0.536 |
| DM Close | 0.70 | 1.00 | 1.20 | 0.90 | 0.95 | 0.18 | 0.60 | 0.90 | 0.80 | 1.20 | 0.88 | 0.217 |
| DM Far | 1.00 | 0.70 | 1.10 | 1.20 | 1.00 | 0.19 | 0.60 | 1.60 | 1.20 | 1.10 | 1.13 | 0.356 |
| RK Close | -0.20 | 0.40 | 0.40 | 1.10 | 0.43 | 0.46 | -0.20 | 0.60 | 0.70 | 0.90 | 0.50 | 0.418 |
| RK Far | 0.70 | 0.80 | 0.30 | 0.40 | 0.55 | 0.21 | 0.70 | 0.80 | 0.40 | 0.80 | 0.68 | 0.164 |

| Subject & Condition | DAY 3 Session N° | | | | | | DAY 4 Session N° | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | Session N° | 4 | | Mean | SD | 1 | 10 | 11 | 12 | Mean | SD |
| FH Close | 1.30 | 1.40 | 1.70 | 1.60 | 1.50 | 0.16 | 1.15 | 1.32 | 1.57 | 1.49 | 1.38 | 0.162 |
| FH Far | 0.68 | 0.61 | 1.57 | 0.79 | 0.91 | 0.38 | 1.05 | 1.50 | 0.74 | 1.57 | 1.22 | 0.339 |
| DM Close | 0.52 | 0.10 | 0.88 | 0.86 | 0.59 | 0.32 | 1.19 | 0.61 | 0.51 | 1.15 | 0.87 | 0.307 |
| DM Far | 1.82 | 1.49 | 1.10 | 0.82 | 1.31 | 0.38 | 0.63 | 1.02 | 1.15 | 0.90 | 0.93 | 0.192 |
| RK Close | 0.50 | 0.53 | 1.19 | 0.90 | 0.78 | 0.28 | 1.02 | 1.25 | 0.08 | 0.19 | 0.64 | 0.508 |
| RK Far | 1.35 | 0.69 | 0.36 | 0.25 | 0.66 | 0.43 | 0.13 | 0.51 | 0.77 | 0.99 | 0.60 | 0.32 |

## Appendix C: Publication List

### Journals

Mahar, D.P. (1988). Progress with tactile aids for the profoundly deaf. <u>Australian Hearing and Deafness Review</u>, <u>5(1)</u>, 4-8.

### Conferences

Mahar, D.P. (1986, May). <u>Comparisons between auditory, visual and vibrotactile representations of spoken words</u>. Paper presented at the 13$^{th}$ Experimental Psychology Conference, Monash University.

Mahar, D.P. and Mackenzie, B.D. (1987). Non-auditory aids for the hearing impaired: Modality-specific differences in the processing of spatially and temporally presented information. <u>Proceedings of the Australian Acoustical Society Annual Conference: Acoustics in the Eighties</u>, Australian Acoustical Society: Sydney, 139-148.

Mahar, D.P. and Mackenzie, B.D. (1989). Differences between modalities in the processing of spatially and temporally distributed information. Paper presented at the 1989 APS Annual Conference, Hobart.