

Somatic ABC's:  
A Theoretical Framework for Designing, Developing and Evaluating the Building Blocks  
of Touch-Based Information Delivery

by  
Troy Lee McDaniel

A Dissertation Presented in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

Approved April 2012 by the  
Graduate Supervisory Committee:

Sethuraman Panchanathan, Chair  
Hasan Davulcu  
Baoxin Li  
Marco Santello

ARIZONA STATE UNIVERSITY

May 2012

## ABSTRACT

Situations of sensory overload are steadily becoming more frequent as the ubiquity of technology approaches reality—particularly with the advent of socio-communicative smartphone applications, and pervasive, high speed wireless networks. Although the ease of accessing information has improved our communication effectiveness and efficiency, our visual and auditory modalities—those modalities that today’s computerized devices and displays largely engage—have become overloaded, creating possibilities for distractions, delays and high cognitive load; which in turn can lead to a loss of situational awareness, increasing chances for life threatening situations such as texting while driving. Surprisingly, alternative modalities for information delivery have seen little exploration. Touch, in particular, is a promising candidate given that it is our largest sensory organ with impressive spatial and temporal acuity. Although some approaches have been proposed for touch-based information delivery, they are not without limitations including high learning curves, limited applicability and/or limited expression. This is largely due to the lack of a versatile, comprehensive design theory—specifically, a theory that addresses the design of touch-based building blocks for expandable, efficient, rich and robust touch languages that are easy to learn and use. Moreover, beyond design, there is a lack of implementation and evaluation theories for such languages. To overcome these limitations, a unified, theoretical framework, inspired by natural, spoken language, is proposed called *Somatic ABC’s* for **A**rticulating (designing), **B**uilding (developing) and **C**onfirming (evaluating) touch-based languages. To evaluate the usefulness of *Somatic ABC’s*, its design, implementation and evaluation theories were applied to create communication languages for two very unique application areas: audio described movies and motor learning. These applications were chosen as they presented opportunities for complementing communication by offloading

information, typically conveyed visually and/or aurally, to the skin. For both studies, it was found that Somatic ABC's aided the design, development and evaluation of rich somatic languages with distinct and natural communication units.

## DEDICATION

To my parents, who taught me that with hard work, perseverance, and integrity, you are limited only by your imagination; and to Jennifer, whose encouragement and support gave me the confidence and motivation needed to achieve my dreams and aspirations.

## ACKNOWLEDGMENTS

I am sincerely grateful to Dr. Sethuraman (Panch) Panchanathan for his guidance toward completing this dissertation. During the time I have worked with Dr. Panch, I have grown and benefited from his mentorship. He has provided me with the independence I needed to explore and learn, while enabling my progress through thought-provoking discussions, project collaborations and funding opportunities. I would also like to thank my committee members including Dr. Marco Santello, Dr. Baoxin Li and Dr. Hasan Davulcu, for their helpful discussions and feedback.

I am very appreciative of Dr. Morris Goldberg for his help and advice with formulating the proposed theory and experimental procedures. Our insightful discussions helped coalesce the components of my dissertation. I would also like to thank Dr. Vineeth Balasubramanian, Dr. John Black and Dr. Terri Hedgpeth for their insightful feedback and discussions over the course of this dissertation.

I would like to thank Daniel Villanueva, Shantanu Bala and Bijan Fakhri for their help with developing the haptic suit. I would like to thank Lakshmi Narayan Viswanathan for his contribution to the haptic-audio descriptions project. I would also like to thank Jacob Rosenthal and Nathan Edwards for their advice and helpful suggestions over the years.

Finally, I would not have accomplished my research goals without the encouragement and teamwork of my fellow research colleagues at CUbiC. CUbiC cultivates a positive and synergistic learning and work environment that naturally attracts creativity—an environment that has been integral to the successful completion of my dissertation. For that, I am forever grateful.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	xii
LIST OF FIGURES.....	xiii
CHAPTER	
1 INTRODUCTION .....	1
2 BACKGROUND.....	6
Neurophysiology of Touch .....	6
The human brain and nerve cell.....	7
The human skin .....	10
Mechanoreceptors.....	13
Mechanoreceptors in glabrous skin.....	18
Mechanoreceptors in hairy skin.....	22
Mechanoreceptors in muscles, tendons and joints .....	23
Somatosensory cortex.....	25
Psychophysics and Perception of Vibrotactile Stimulation .....	29
Vibrotactile sensitivity and vibration frequency.....	29
Vibrotactile sensitivity .....	29
Dependencies between frequency, amplitude and pitch. 33	
Relative and absolute frequency discrimination .....	34
Vibration intensity .....	35
Dependencies between amplitude, frequency and sensation magnitude.....	36
Relative and absolute amplitude discrimination .....	38

CHAPTER	Page
Varying amplitude over time: Complex waveforms as a parameter.....	39
Vibrotactile timing and rhyhtm.....	43
Relative and absolute temporal discrimination.....	44
Tactile rhythm.....	45
Body site.....	47
Anatomical reference points.....	50
Endpoints.....	54
Odd sites.....	55
Spacing versus numbers.....	56
Vibrotactile sensitivity, age and other factors.....	58
Two-dimensional localization.....	59
Relative localization.....	63
Spatio-temporal patterns.....	65
3 RELATED WORK.....	74
Literal Translations via Touch.....	79
Human-to-human literal translation.....	80
Tadoma.....	80
Tactile sign language and fingerspelling.....	82
Social-haptic communication.....	85
Computerized systems for sensory substitution.....	87
Tactile-vision translation.....	87
Tactile perception of print.....	87
Tactile-vision sensory aids.....	90

CHAPTER	Page
Tactile-audio translation .....	93
Alphanumeric Information Delivery via Touch.....	95
Braille.....	96
Morse code.....	97
Vibratese .....	99
Automated Conceptual Information Delivery via Touch .....	102
Mediated haptic interpersonal communication technology .....	104
Tactile icons.....	108
Haptic glyphs .....	112
Haptic icons .....	116
4 SOMATIC ABC'S .....	121
Somatic Language Articulation .....	123
Somatic language.....	125
Somatic phonemes.....	126
Phonemes versus somatemes.....	126
Designing somatemes .....	127
Somatic words .....	128
Vocabulary .....	128
Using context to alter meaning.....	129
Simulation variations .....	129
Using stimulation variations to ensure successful delivery of a message .....	130
Using stimulation variations to alter the meaning of a message .....	131

CHAPTER	Page
Designing somatic words.....	132
Level of abstraction.....	132
Signal parameter association.....	133
Parameter value selection.....	134
Somatic sentences.....	135
Building a System for Somatic Language Communication.....	137
Functionality requirements.....	138
Performance requirements.....	139
Usability requirements .....	140
Confirm Somatic Language through Evaluation.....	141
Pilot and formal evaluations.....	142
Objective evaluation.....	143
Subjective evaluation.....	143
5 APPLICATION #1: AUDIO-HAPTIC DESCRIBED MOVIES .....	145
Related Work.....	149
Proposed Approach.....	154
Articulate .....	154
Build.....	159
Vibrotactile belt.....	160
Hardware description .....	160
Software description.....	163
Audio-haptic movie player .....	164
Authoring .....	165
Confirm.....	165

CHAPTER	Page
Subjects .....	165
Materials and apparatus .....	166
Procedure.....	168
Audio-only condition .....	168
Audio-haptic condition .....	169
Results .....	172
Discussion .....	176
Functionality and performance assessment ....	176
Subjective assessment .....	177
Objective assessment .....	179
Conclusion and Future Work.....	181
6 APPLICATION #2: VIBROTACTILE MOTOR INSTRUCTIONS AND	
FEEDBACK.....	184
Related Work.....	189
Virtual reality.....	189
Acoustic systems .....	191
Kinesthetic systems .....	192
Vibrotactile systems .....	193
Proposed Approach .....	201
Articulate: vibrotactile motor instructions.....	201
Build: vibrotactile motor instructions .....	209
Haptic suit for vibrotactile motor instructions .....	209
Hardware description .....	209
Firmware description .....	211

Vibrotactile motor instruction software .....	212
Confirm: vibrotactile motor instructions .....	213
Aim.....	214
Subjects .....	214
Procedure.....	214
Results .....	217
Discussion .....	221
Learning rate.....	221
Recognition accuracy .....	221
Response time.....	223
Posture-free vibrations .....	223
Subjective feedback .....	224
Conclusion and future work: vibrotactile motor instructions.....	225
Articulate: vibrotactile feedback .....	227
Feedback for positioning errors.....	227
Feedback for speed errors.....	231
Build: vibrotactile feedback .....	234
Haptic feedback suit.....	235
Hardware description .....	235
Firmware description .....	239
Software description.....	242
Confirm: vibrotactile feedback .....	244
Aim.....	245
Subjects .....	245

CHAPTER	Page
Apparatus .....	245
Procedure.....	247
Results .....	251
Discussion .....	259
Learning rate (position and speed).....	259
Recognition accuracy (position) .....	259
Recognition accuracy (speed) .....	263
Response time (position).....	265
Response time (speed).....	265
Post-experiment questionnaire.....	267
Conclusion and future work .....	268
7 CONCLUSION AND FUTURE WORK .....	270
REFERENCES .....	274

## LIST OF TABLES

Table		Page
1.	Summarization of Proposed Approaches in terms of Desired Design and Performance Criteria for Achieving Functional and Practical Somatic Languages .....	78
2.	Design of Haptic Icons for Application Sharing .....	118
3.	Functionality, Performance and Usability Requirements during the Construction of a Somatic Language Communication System .....	138
4.	List of Clips from Audio Described Movies for User Study .....	167
5.	Mean Responses to the Post-Experiment Questionnaire for Vibrotactile Motor Instructions .....	220
6.	Subjective Evaluation of Naturalness of Vibrotactile Motor Instructions....	221
7.	Mean Responses to General Usability Questions .....	256
8.	Mean Responses to Position Feedback Questions .....	257
9.	Mean Responses to Speed Feedback Questions.....	258

## LIST OF FIGURES

Figure	Page
1. Conceptual drawing of one hemisphere of the human brain .....	9
2. Anatomy of the skin including the epidermis, dermis and hypodermis .....	13
3. Adaptation characteristics for slowly adapting and rapidly adapting (or fast adapting) Type I and II receptors .....	16
4. Receptive field characteristics for Type I receptors .....	17
5. Receptive field characteristics for Type II receptors.....	18
6. Anatomical sketch of Meissner's corpuscle .....	19
7. Anatomical sketch of Pacinian corpuscle.....	21
8. Anatomical sketch of Ruffini ending.....	22
9. Anatomical sketch of a portion of a muscle spindle receptor .....	24
10. Anatomical sketch of a Golgi tendon organ .....	25
11. Surface sketch of the left cerebral hemisphere of the brain with the parietal lobe highlighted .....	26
12. Amplitude thresholds, measured in micrometers, for different regions of the body .....	31
13. Plot of amplitude threshold as a function of contactor area for various frequencies .....	32
14. Plot of amplitude threshold as a function of frequency for various contactor areas .....	33
15. Plot of equal sensation magnitude curves across frequencies and vibration amplitudes .....	37
16. Plot of intensity values versus size of just noticeable differences for the chest region .....	39

Figure	Page
17. Example of a “rough” waveform generated by multiplying a 250 Hz sinusoid with a 30 Hz sinusoid .....	40
18. Plot of burst durations versus size of just noticeable differences .....	44
19. Localization accuracies for seven piezoceramic tactors on the ventral side of the left forearm .....	52
20. Localization accuracies for twelve tactors around the lower or upper part of the abdomen and back .....	53
21. Localization accuracies for seven C2 tactors around the lower part of the abdomen and back .....	55
22. Conceptual drawing of the 4x4 tactor arrangement .....	62
23. Contour plot showing the interaction between timing parameters—burst duration (BD) and stimulus onset asynchrony (SOA)—and localization performance .....	64
24. Spatio-temporal patterns, A through H, for Piatetski and Jones’ pattern recognition experiment using the volar side of the forearm .....	67
25. Two sets of spatio-temporal patterns for Jones, Kunkel and Piatetski’s pattern recognition experiment using the volar side of the forearm .....	68
26. Spatio-temporal patterns, A through H, for Piatetski and Jones’ pattern recognition experiment using the torso (back) .....	69
27. Additional patterns for recognition, similar in timing and activation to those of figure 26 .....	70
28. Spatio-temporal patterns, A through F, for Jones and Ray’s pattern recognition experiment using a belt.....	71

Figure	Page
29. Intuitive spatio-temporal patterns, A through G, for Jones, Kunkel and Piatieski’s pattern recognition experiment using a torso-based tactile display for military applications .....	73
30. Visual demonstration of Tahoma method for haptic human-to-human communication by individuals who are deaf-blind .....	82
31. Visual demonstration of tactile sign language for haptic human-to-human communication by individuals who are deaf-blind .....	83
32. Visual demonstration of tactile fingerspelling for haptic human-to-human communication by individuals who are deaf-blind .....	84
33. System setup of the Optohapt .....	89
34. Original apparatus of TVSS with vibrotactile display .....	91
35. Tongue Display Unit (TDU) version of TVSS .....	92
36. Teletactor device for audio-to-tactile conversion .....	94
37. Example of Braille cells spelling the word ‘braille’ .....	96
38. International Morse Code .....	98
39. Vibratese language for alphanumeric communication through vibrations ..	101
40. Conceptual design of Shake2Talk .....	107
41. Two examples of haptic glyphs (hlyphs).....	113
42. Somatic ABC’s theoretical framework to support the creation and evaluation of functional and practical somatic languages, and their integration into larger systems. ....	122
43. General terminology, defined in terms of somatic (body related) stimulation, to accommodate any modality of touch, from tactile to kinesthetic .....	123

Figure	Page
44. Somatic language defined in terms of a metaphorical interpretation of natural language .....	124
45. Haptic remote control for adding haptic feedback and interactivity to video media .....	150
46. Movie enhancing tactile jacket .....	151
47. Tactile stimulation felt over time using a glove-based system for experiencing authored tactile video .....	153
48. Hybrid haptic device for experiencing haptic movies using a force-feedback device combined with a tactile display for the finger.....	154
49. Proposed vibrotactile words for communicating the distance of a character from the camera for use in audio-haptic descriptions .....	156
50. Division of screen into six regions of equal width for audio-haptic described movies .....	158
51. System setup for viewing audio-haptic described movies .....	160
52. System architecture of vibrotactile belt and command control .....	161
53. Vibrotactile belt implementation depicting factor modules, control module, communication bus and belt .....	162
54. Graphical user interface of vibrotactile belt command console on PDA (left), and tactile rhythm authoring tool (right).....	163
55. Graphical user interface of audio-haptic described movie player .....	164
56. Average localization accuracy with standard deviation (SD) across participants and rhythms for L3 through R3 (first part of audio-haptic condition).....	172

Figure	Page
57. Average rhythm recognition accuracy with standard deviation (SD) across participants and body sites for close, middle and far rhythm (first part of audio-haptic condition) .....	173
58. Likert ratings for participants' perception of their understanding of an audio-haptic described clip, their needed concentration and the overall complexity of a clip .....	174
59. View of Tai Chi system setup (left) and virtual scene (right) depicting the virtual avatars of the student (user) and instructor .....	191
60. MusicJacket system being used by a young violin student .....	194
61. System and experimental setup of the tactile motion instruction system ....	196
62. Tactile motion instruction system's placement of vibration motors configured into groups based on the movements they represent .....	197
63. TIKL marker placement for use with the Vicon system for capturing arm movements .....	198
64. Motor placement for TIKL .....	199
65. Anatomical planes of the human body. ....	203
66. Visual depiction of the right arm fundamental movements involved in this work .....	204
67. Depiction of conceptual mappings "follow me" and push/pull for vibrotactile instructions .....	205
68. Haptic suit being worn and demonstrated by a student researcher.....	210
69. Hardware details of haptic suit .....	211
70. Graphical user interface for triggering vibrotactile motor instructions and individual motors through the haptic suit .....	213

Figure	Page
71. Novel postures used in user study to assess if recognition performance of vibrotactile motor instructions is independent of postures.....	216
72. Summary of mean recognition accuracies averaged across participants (left) and misclassifications displayed using confusion matrices (right) for vibrotactile motor instructions .....	218
73. Mean response times averaged across participants for each vibrotactile motor instruction .....	219
74. Motor configurations and conceptual mappings used for vibrotactile positioning feedback .....	228
75. Visual representation of vibrotactile rhythms used in the proposed vibrotactile feedback for correcting speed errors .....	234
76. IMU with coordinate system overlaid and inset for detail .....	236
77. IMU placement on the haptic suit .....	238
78. Sensors and vectors involved in relative angle calculations for vibrotactile feedback .....	241
79. Graphical user interface for recording movements .....	244
80. Mean recognition accuracy per pattern (position experiment) .....	253
81. Mean response accuracy per pattern (position experiment).....	254
82. Mean recognition accuracy per pattern (speed experiment) .....	254
83. Mean response time per pattern for training and testing phase (position experiment) .....	255
84. Mean response time per pattern for training and testing phase (speed experiment) .....	255

## Chapter 1

### INTRODUCTION

In the 21st century, our daily lives are inundated with information given its ease of access, particularly in recent years due to the ubiquity of technology and advancements toward faster wireless networks, mobile web browsing and mobile tools and applications that enable users to stay well connected via social networking. The digital revolution that began in the 1980's and 1990's continues to build momentum: There are 232 million cell phone users in the United States with almost half being smart phone users where the majority of usage time is spent texting (14%), web browsing (10%) and using applications including Facebook and Twitter (53%) (State of the Media: Consumer Usage Report, 2011). Although the digital revolution and ubiquity of technology continues to afford improved connectivity, efficacy and efficiency through technological advances, the influx of information is creating increasing situations of information overload due to today's computerized devices and technologies that largely engage our visual and auditory modalities.

Problems of sensory overload can also be found in many areas of the workforce including aircraft operation and military roles where a myriad of audiovisual displays and controls require constant scanning and assessment while simultaneously maintaining situational awareness. For example, pilots rely on large information “dashboards” that are demanding of visual attention (Rupert A. H., 2000) (van Erp J. B., van Veen, Jansen, & Dobbins, 2005). By assuming that vision and/or hearing provides an optimal channel for information delivery, these senses have been overloaded, increasing distractions, cognitive load, and operation/decision-making delay—thereby increasing chances for life-threatening situations. For soldiers, these hindrances must be overcome as accurate, on-the-fly decision making equates not only to their own safety, but the safety of their

fellow warfighters and civilians. A recent example of how information overload can endanger lives is the rise of texting and cell phone use while driving. In 2009, almost 5,500 people were killed in car accidents involving a distracted driver—one in five of those deaths involved a cell phone (Distracted Driving 2009 NHTSA Traffic Safety Facts Research Note, 2010).

With the advent of touchscreens and gesture-based input, the human hand has become an effective and efficient means for directly operating computerized devices with touchscreens, such as smartphones. It is surprising, however, that the digital revolution has failed to seize the opportunity of our skin's ability to *receive* information. Therefore, as a receptive channel for computerized information delivery, our sense of touch is underutilized compared to vision or hearing (Tan & Pentland, 2001). This is unclear given that it is our largest and oldest sensory organ (Montagu, 1986) well equipped for rich spatial and temporal perception (Geldard F. A., 1960). One reason is that the field of *haptics*<sup>1</sup> is still in its infancy compared to vision and hearing research; but this is slowly changing: the number of researchers, engineers and hobbyists exploring haptic cyber-physical systems, multimodal immersive environments and human haptic perception, is steadily rising. If the rich, multimodal sensory capabilities of the skin, and the processing power of the somatosensory cortex within the parietal lobe of the human brain, can be effectively leveraged, this may pave the way for rich and efficient haptic communication systems. However, if alternative modalities for information delivery, such as touch, are ignored, this will only exasperate the problem of information overload as data becomes more accessible and ubiquitous.

Recent research is beginning to show promise in that it supports the feasibility and versatility of touch as a useful communication channel for augmenting visual and

---

<sup>1</sup> The word *haptic* means “of or relating to the sense of touch”, and comes from the Greek word *haptikos*.

auditory presentation to distribute data between modalities (Rupert A. H., 2000) (van Veen & van Erp, 2001) (van Erp J. B., van Veen, Jansen, & Dobbins, 2005). But for touch-based information delivery to receive more widespread use, a *versatile* and *rich* design theory is needed; that is, a design theory that could be effectively applied across diverse application areas where rich haptic communication might be useful. Among the several theories for touch-based information delivery that have been proposed, including *tactile icons* (Brewster & Brown, 2004), *haptic icons* (MacLean & Enriques, 2003) (Enriquez, MacLean, & Chita, 2006) and *vibratese* (Geldard F. A., 1957), such a design theory is still needed. Our approach begins by exploring the structure of *somatic*<sup>2</sup> *building blocks* for touch-based information delivery, toward the creation of *somatic languages* that are easy-to-learn, easy-to-use, versatile, expandable, efficient, rich and robust—attributes needed for practical, useful touch-based information delivery systems.

Our proposed design theory is inspired by natural, spoken language; particularly, how language’s metaphorical building blocks, *phonemes*, are combined to create words, of which a small vocabulary can be used to create unlimited, expressive sentences. In addition to a theory of design, both an implementation theory and evaluation theory are proposed. The implementation theory provides construction guidelines in terms of functionality, performance and usability requirements. The proposed evaluation theory presents guidelines for testing somatic information delivery systems with users; and key objective and subjective attributes that must be assessed including distinctness and naturalness of haptic stimulation as it relates to its associated meaning. These theories are combined into one unified theoretical framework called *Somatic ABC’s* for **A**rticulating (designing), **B**uilding (developing) and **C**onfirming (evaluating) somatic languages. The proposed framework is intended to overcome the aforementioned limitations of existing

---

<sup>2</sup> The word *somatic* means “of or relating to the body”, and comes from the Greek word *somatikos*, and the Greek word for body, *soma*.

design theories through enabling designers to create rich and versatile somatic languages, using natural, spoken language as a basis, across diverse applications.

The following sections of this dissertation are organized as follows. Chapter 2 presents background work beginning with a broad, but detailed, coverage of how biological touch works—specifically, the neurophysiology of touch from the functional characteristics of peripheral touch receptors, to where and how touch signals are processed in the brain. This is followed by an in-depth presentation of the psychophysics and perception of vibrotactile stimulation as both applications used to assess the proposed Somatic ABC’s framework utilize vibrations to convey information. Chapter 3 compares existing approaches for touch-based information delivery, deriving design and performance criteria for somatic languages. Existing approaches are divided into three categories: *literal translation*, *alphanumeric* and *conceptual*. Chapter 4 presents the proposed theoretical framework, Somatic ABC’s, detailing its theoretical components including design, implementation and evaluation theories. Chapter 5 and 6 each present the results from two different applications in which Somatic ABC’s was applied to design, develop and evaluate a language for conveying information through touch. The first application, haptic-audio described movies for individuals who are blind, is presented in Chapter 5. Somatic ABC’s was used to haptically augment audio descriptions (narrations describing visual content within a film) such that complementary information could be presented tactually. The second application, vibrotactile motor learning, is presented in Chapter 6. Somatic ABC’s was used to design, develop and evaluate vibrotactile motor instructions and feedback for augmenting traditional motor learning. In both applications, Somatic ABC’s was found to support the design, implementation and evaluation processes, effectively creating somatic languages that are rich, intuitive and practical. Lastly, Chapter 7 discusses possible directions for future

work that have the potential to open new vistas for research in haptics and haptic information delivery; these include the exploration of neurological bases for improving the distinctness and naturalness of communication units within somatic languages; multimodal design approaches for achieving high bandwidth touch-based communication; and novel applications areas that might benefit from offloading information to the skin for fast, parallel processing.

## Chapter 2

### **BACKGROUND**

In the first part of this section, the neurophysiology of touch is discussed, which provides an overview of both the peripheral and central mechanisms of touch; in particular, how receptors in the skin, muscles, joints and tendons mediate touch stimulations to the brain through nerve impulses, and how touch centers in the brain interpret these incoming haptic signals. Interactions between touch centers, as well as their interactions with other areas of the brain, will be described. A thorough understanding of the neurophysiology of touch is useful as a basis toward guiding the design of somatic information delivery systems.

In the second part of this section, human psychophysical and perceptual aspects of vibrotactile stimulation will be explored in preparation for discussion of the applications used to explore the effectiveness and usefulness of Somatic ABC's. As both applications utilize vibrotactile stimulation, an understanding of how to design vibratory signals to optimize human sensitivity, perceptual distinctness and naturalness, is critical. Toward this goal, this section explores the sensibility of individual vibrotactile dimensions (frequency, intensity, timing and location) and higher order dimensions (rhythms and spatio-temporal patterns) toward distinct and natural touch-based information delivery systems.

#### **Neurophysiology of Touch**

When an object is intriguing, the viewer will often pursue a “closer look” by actively exploring it with his or her hands, perceiving its texture, material and shape, among many other properties, through touch. Humans employ various exploratory procedures (Lederman & Klatzky, 1987) to extract different haptic properties of objects; for example, the lateral movements of the fingers across an object's surface pick up fine

textural details, as well as imperfections or irregularities; contours are followed for precise shapes; an unsupported hold helps estimate an object's weight; and moveable parts are located and engaged to predict function; among many other procedures for various object features. These processes of haptic feature recognition are mediated by the peripheral and central mechanisms of touch—in particular, the receptors in the skin, and their pathways leading to touch centers in the brain.

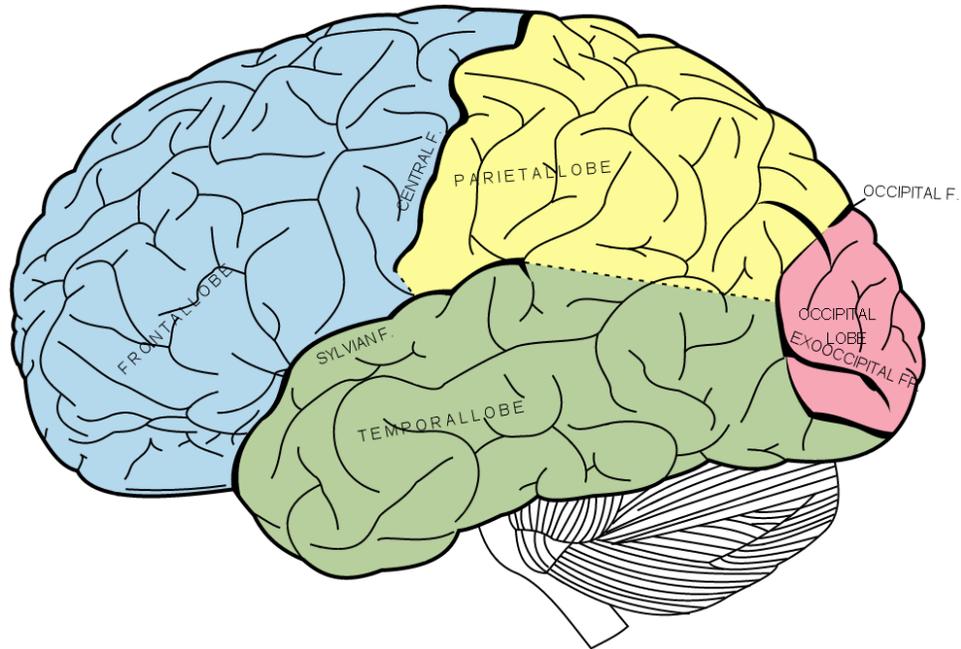
The skin acts as an interface to the environment, providing both a protective boundary and a rich sensory channel through its dense array of receptors. The sensations provided through touch are rich and engaging including light to heavy pressure, cold to warm temperature, pain, and kinesthesia. The modalities of touch are afforded by the various physiologies of touch receptors, which determine their sensitivity to external stimuli. In this chapter, an overview of the peripheral and central mechanisms of touch is presented related to sensation and perception of stimuli in contact with the skin.

**The human brain and nerve cell.** It is often said that the human brain is the most complex of machines, biological or man-made, and for justifiable reason: humans can master both physical and mental skills; invent devices and technologies; learn languages; hold conversations in varied topics; reason about and solve complex problems; and articulate and express emotions, thoughts and ideas. Although much is still not known about the inner workings of the brain, a solid understanding of many of its functions has been achieved through persistent research efforts throughout the nineteenth and twentieth century to the present.

The central nervous system (CNS) consists of the brain and spinal cord, and is protected by the skull and vertebrae, respectively. The brain may be further divided into the cerebral hemispheres, brain stem and cerebellum. The peripheral nervous system (PNS) consists of sensory cells and nerve fibers creating connectivity between receptors

and the CNS. The brain is divided into a left and right hemisphere, each of which process information from, and controls movement of, the contralateral side of the body. The outer layer of each hemisphere is called the cerebral cortex (gray matter), and provides much of the brain's processing capability through nerve cells and their interconnections. Below the outer layer is white matter, which largely consists of myelinated fibers, aiding faster communication of signals among nerve cells in the brain and to/from the periphery. Structures below these layers cover functions from memory to emotion.

Figure 1 provides a conceptual depiction of the human brain, indicating the brain's different lobes. The lobes of the brain include the frontal lobe, temporal lobe, parietal lobe and occipital lobe. The frontal lobe largely deals with processing related to action and planning; the temporal lobe handles processing for hearing; the parietal lobe involves processing for touch; and the occipital lobe involves visual processing. Folds in the lobe increase the density of nerve cells over the surface area of the brain. The folds form ridges (gyri) and fissures (sulci).



*Figure 1.* Conceptual drawing of one hemisphere of the human brain depicting the different lobes including frontal, temporal, parietal and occipital. Adapted from the Wikimedia Commons: Gray728.svg, <http://en.wikipedia.org/wiki/File:Gray728.svg>.

The brain consists of two cell types: nerve cells, or *neurons*, and glial cells. Glial cells are involved in maintenance, house-keeping tasks and provide structure; the former cell type, the neuron, is the brain's most basic processing unit. The complexity of the brain is quickly realized by considering the staggering number of neurons, estimated at 100,000,000,000 (or 100 billion) neurons (Kandel, Schwartz, & Jessell, 2000)—with many, many more interconnections between neurons. The neuron has a morphology specialized for communication using electrical signals. A neuron consists of a cell body (soma), axon and dendrites. Dendrites receive electrical signals from other neurons. Based on the type of neuron and number of dendrites, a nerve cell may receive input from anywhere between one to 100,000 neurons (Purves, et al., 2008). If the incoming signal, once integrated, is enough to alter the voltage difference across the cell's membrane, the neuron “fires”, generating an action potential, beginning at the start of the axon and

traveling its length, which can range from 0.1 mm to 3 m (Kandel, Schwartz, & Jessell, 2000) depending on the type and function of the neuron; and although a neuron at “rest” continues to fire, its firing rate is significantly reduced. The conduction velocity of the action potential along the axon ranges from 1 to 100 m/s, lasting around 1 ms at 100 mV (Kandel, Schwartz, & Jessell, 2000). If an axon is myelinated, that is, wrapped in layers of insulating tissue produced by glial cells, conduction speed increases; this technique is useful for communication across long distances between neurons. The end of the axon divides into presynaptic terminals, which come in close proximity, but do not touch, the dendrites of other neurons. Interneuronal communication occurs at the synapse where a presynaptic terminal transmits an electrical or chemical signal to a postsynaptic terminal over a gap called the synaptic cleft.

A neuron may be classified as a sensory neuron, motor neuron or interneuron. Sensory neurons facilitate our basic senses. They convey external stimuli, as sensed by receptors in the periphery, to the brain and spinal cord. The sense of touch is mediated through pseudo-unipolar neurons located in the dorsal root ganglia near the vertebrae of the spinal cord. These sensory neurons convey information from the periphery (skin, muscles, tendons and joints) to the spinal cord via a single axon, which begins at touch receptors, and ends at motor neurons and/or interneurons within the spinal cord. In turn, the spinal cord relays information to the brain—in particular, the somatosensory cortex of the parietal lobe—via interneurons for further processing. Based on the results of this processing, the environment may be acted upon through motor neurons, which send activation signals to muscles and glands.

**The human skin.** The human skin is a remarkable sensory organ that covers our entire body with functions critical to survival and perception of surrounding environments. The sense of touch is known as the “mother” of all senses in that the skin

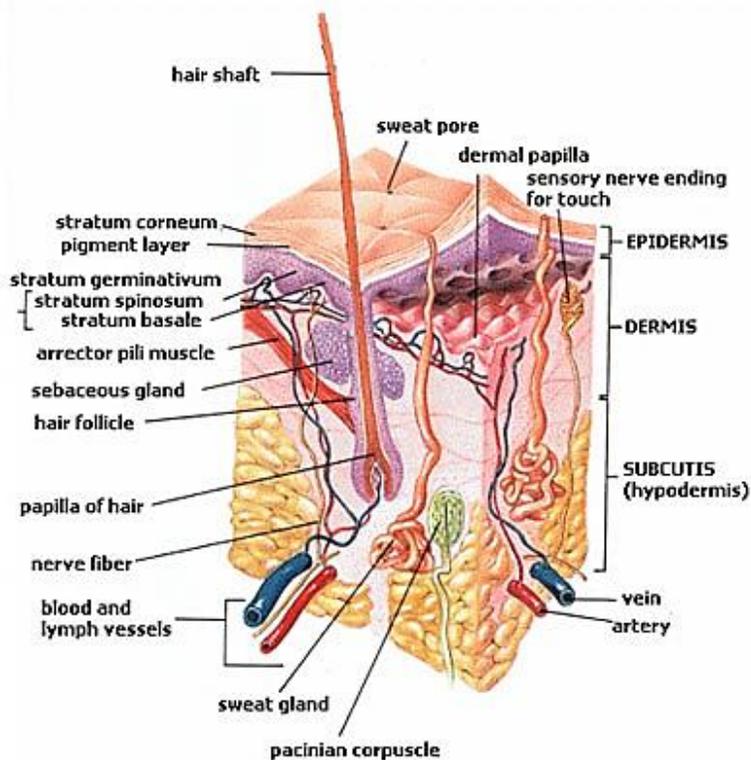
is the first sensory organ to develop, from which all other sensory organs form (Montagu, 1986). It is also the first sense to become functional during the early embryonic stage of development (Montagu, 1986). Our skin provides critical protection for our underlying soft tissue, preventing damage from harmful environmental stimuli, and preventing exposure to bacteria and heat. Therefore, it should be no surprise that without the protection provided by skin, survival would not be possible. The skin provides many more critical functions, such as temperature regulation and control—but the focus here will be the function of skin as a receptive channel for environmental information.

Sensation of environmental stimuli in contact with the skin is mediated through the skin's dense array of receptors. The skin's receptors are tuned to specific environmental stimuli such as temperature, vibrations, deformations and pain. The skin is an impressive receptive surface at 19 sq. ft. and 8 lbs. (12% of the total body weight) for the average adult male (Montagu, 1986)—the largest sensory organ in the human body. The thickness of the skin varies across the body from a 10<sup>th</sup> of a millimeter to 3-4 mm with the skin being thinnest on the eyelids and thickest on the palms and soles (Montagu, 1986).

The anatomy of the skin consists of two layers: epidermis and dermis. The epidermis is the outermost layer of the skin, and provides a protective boundary between body and environment. It is nonvascular, but contains nerves, specifically, bare nerve endings, and if the skin is hairy, mechanoreceptors. Cells in the top most layer of the skin eventually die and are shed to make room for new cells that push their way up from the bottom layer of cells of the epidermis. The dermis is a thicker layer below the epidermis, composed of connective tissue, and containing a variety of nerves and glands. The boundary between the epidermis and dermis is known as the epidermal-dermal junction. The epidermis and dermis together form what is called the cutis. Although not part of the

skin, between the dermis and muscles is subcutaneous tissue (also known as subcutis or hypodermis). This layer is composed of fat and connective tissue, and connects the skin to the muscles. It provides additional protection from harmful bacteria as well as nourishment for the dermis. Touch receptors are found throughout the epidermal-dermal junction, dermis and subcutaneous tissue. Our two skin types include glabrous (non-hairy) skin and hairy skin. Glabrous skin, found on the fingertips, lips and soles, has greater tactile sensitivity than hairy skin due to ridges formed by folds in the skin—a trick used to increase the density of receptors.

Figure 2 provides a conceptual drawing of the anatomy of the skin including its layers and some of its receptors. The following section presents a description of the mechanoreceptors of the skin. This focus was chosen, as opposed to nerve endings that mediate temperature and pain sensations, given that the applications of Chapter 5 and 6 utilize vibrotactile stimulation for somatic information delivery, which is mediated by mechanoreceptors.



*Figure 2. Anatomy of the skin including the epidermis, dermis and hypodermis, and the location of cells and structures relative to these layers. Adapted from the Wikimedia Commons: Skin.jpg, <http://en.wikipedia.org/wiki/File:Skin.jpg>.*

**Mechanoreceptors.** Mechanoreceptors are a type of touch receptor sensitive to mechanical deformations including light to hard pressure applied to the skin, movement across the skin, skin stretch, vibrations, muscle contractions, muscle tension, among other stimuli. Proprioceptors are mechanoreceptors within the muscles, tendons and joints that aid proprioception—the sense of limb position and movement. Mechanoreceptors, as well as other receptor types in the periphery, communicate with the central nervous system via the peripheral nervous system through nerve fibers and sensory neurons. Sensory neurons innervate receptors, providing a pathway to the brain and spinal cord. Touch receptors of the trunk and limbs are innervated by pseudo-unipolar cells within the dorsal root ganglia near the vertebrae of the spinal cord. There are four attributes common across all sensory

systems (Kandel, Schwartz, & Jessell, 2000): modality, location, intensity and timing, where modality is determined by the type of receptor activated; location is determined by where the activated receptors are within the skin or tissue; intensity is determined by a receptor's firing rate as well as the total number of receptors activated, dependent on the amplitude of the stimulus; and lastly, timing is determined by the duration a receptor fires, dependent upon when the stimulus is introduced and then removed, and the adaptation properties of the receptor.

Mechanoreceptors transduce mechanical energy into a receptor potential—a depolarization of membrane potential, creating a voltage difference between the inside and outside of a cell—thereby creating an action potential, or nerve impulse, that travels to the sensory neuron. The amplitude and duration of the receptor potential is dependent on the intensity of the stimuli. If a mechanoreceptor is sensitive to an applied stimulus, dependent upon its structure, stretch-sensitive sodium ion channels open, causing an influx of ions (current), which in turn generates the voltage difference, causing a receptor to send nerve impulses, or fire. The following discussion provides an overview of mechanoreceptors in terms of density variations, adaptation, and receptive field characteristics—for reviews see (Johansson & Vallbo, 1983) (Vallbo & Johansson, 1984).

The basic senses are mediated by millions of nerve endings. For touch alone, millions of nerve endings mediate the different modalities of touch. The glabrous skin of the human hand is estimated to be innervated by 17,000 sensory neurons (Johansson & Vallbo, 1979), which each may innervate anywhere from one to many receptors depending on its type. Mechanoreceptors are distributed throughout both hairy and glabrous skin. Glabrous skin contains four types of mechanoreceptors (Johansson & Vallbo, 1979): Meissner's corpuscles, Merkel disk receptors, Pacinian corpuscles and

Ruffini endings—the names of which follow their respective discoverers. Hairy skin contains Pacinian corpuscles, Ruffini endings and Merkel disk receptors, but exclusive to hairy skin are hair follicle receptors and field receptors (Vallbo, Olausson, Wessberg, & Kakuda, 1995). The density of mechanoreceptors varies across the body with denser receptor populations found in distal bodily regions compared to proximal body regions (Johansson & Vallbo, 1979); significant proximo-distal variations are present for Type I receptors with more subtle variations for Type II receptors with roughly even distributions in glabrous skin.

Mechanoreceptors will eventually adapt to stimuli, but the rate of adaptation varies depending on receptor type (Vallbo & Johansson, 1984): slowly adapting (SA) mechanoreceptors (Merkel disk receptors and Ruffini endings) fire consistently during constant pressure applied within their receptive field where firing rate increases with stimulus intensity; whereas rapidly adapting (RA) mechanoreceptors (Meissner's corpuscles and Pacinian corpuscles) fire during changes in pressure, indicating the velocity and acceleration of skin indentation.

The receptive field of a mechanoreceptor is the area of skin above the receptor that when stimulated, deforms the structure of the receptor, compressing its nerve terminal and causing an action potential. Receptive field size may be small with sharp borders (Type I) or large with obscure borders (Type II) (Johansson, 1978). Given their structure, size and position in the superficial layer of the skin, Merkel disk receptors (SA I) and Meissner's corpuscles (RA I) have small receptive fields. More specifically, since a single sensory neuron directly innervates multiple SA I or RA I receptors, it is more useful to refer to the receptive field size of the sensory neuron itself. These sensory neurons innervate the same number of receptors across the skin, and so receptor density variations create receptive field size variations, resulting in spatial resolution changes

across the body—in particular, reduced resolution from distal body parts to proximal body parts. Pacinian corpuscles (RA II) and Ruffini endings (SA II) have larger receptive fields with obscure borders given their structure and deeper position in the skin, and are less numerous compared to Type I cells. Sensory neurons innervating RA II or SA II receptors each innervate one receptor, and therefore, have one “hot spot” directly above the cell, in contrast to SA I or RA I, which have multiple hot spots since they combine many receptive fields. Type II receptors are distributed more uniformly compared to Type I receptors. Figure 3 summarizes adaptation properties and receptive field sizes across mechanoreceptors. Figure 4 and 5 provide a detailed look at receptive field size and structure for Type I and Type II receptors, respectively.

The following explores the physiology of mechanoreceptors in more detail, beginning with those in glabrous skin, followed by hairy skin, and the finally, mechanoreceptors in muscles, tendons and joints (proprioceptors).

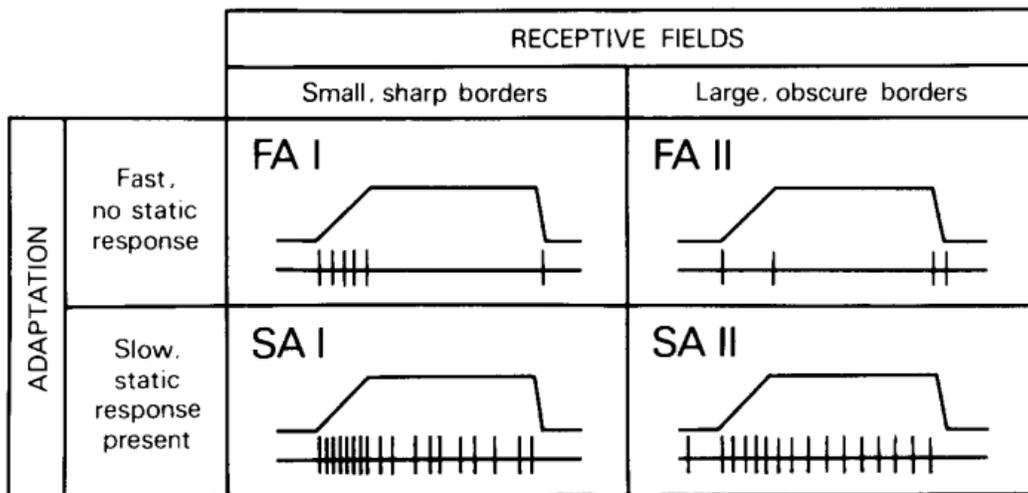
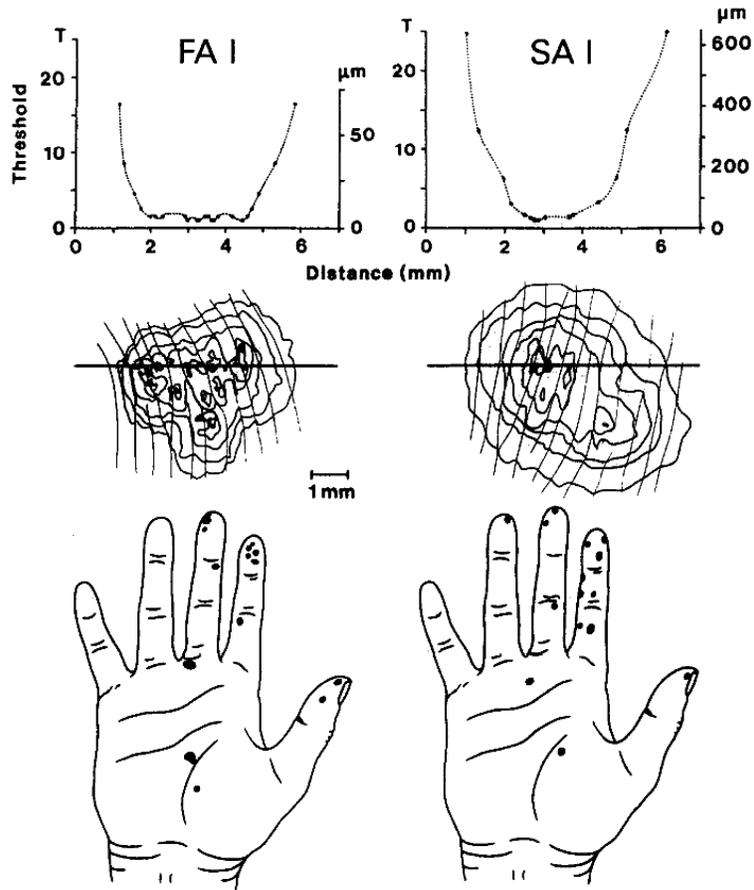
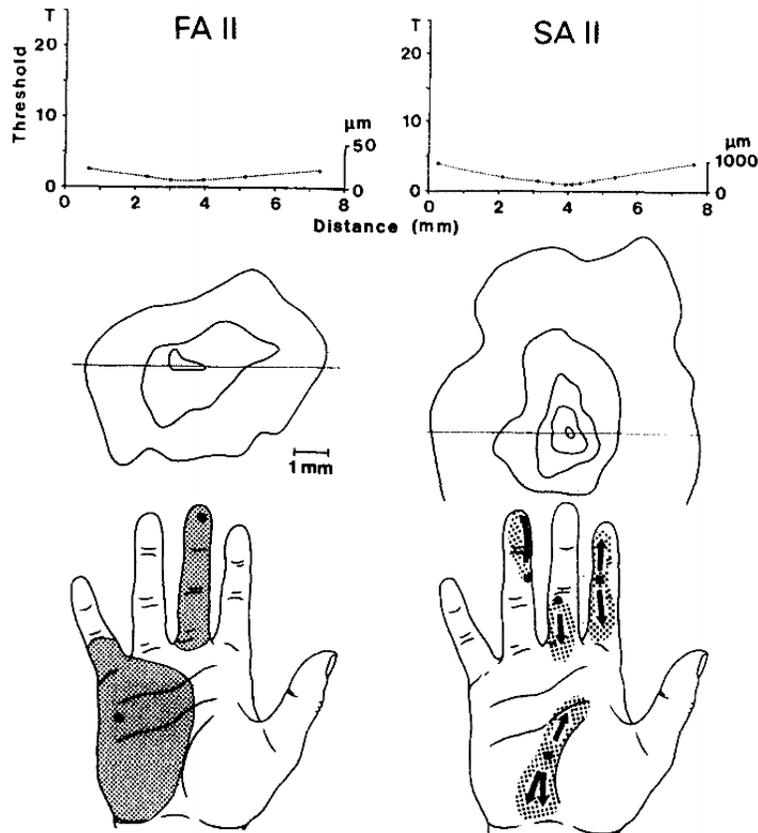


Figure 3. Adaptation characteristics for slowly adapting and rapidly adapting (or fast adapting) Type I and II receptors: SA and RA (or FA) respectively. Reprinted from “Tactile sensory coding in the glabrous skin of the human hand,” by Johansson, R. S., & Vallbo, A. B., Jan. 1983, *Trends in Neuroscience*, 6, p. 27. Copyright © 1983 by Elsevier. Reprinted with permission.



*Figure 4.* Receptive field characteristics for Type I receptors: Meissner's corpuscles (left) and Merkel disk receptors (right)—FA I and SA I, respectively. The black dots indicate receptor clusters (15 individual receptors) innervated by a single sensory neuron. As depicted, Type I receptive fields are distinct and sharp with diameters ranging between 2-8 mm. The sensitivity threshold plots show the multiple hot spots of these fields due to the innervation of multiple receptors. *Reprinted from "Tactile sensory coding in the glabrous skin of the human hand," by Johansson, R. S., & Vallbo, A. B., Jan. 1983, Trends in Neuroscience, 6, p. 28. Copyright © 1983 by Elsevier. Reprinted with permission.*

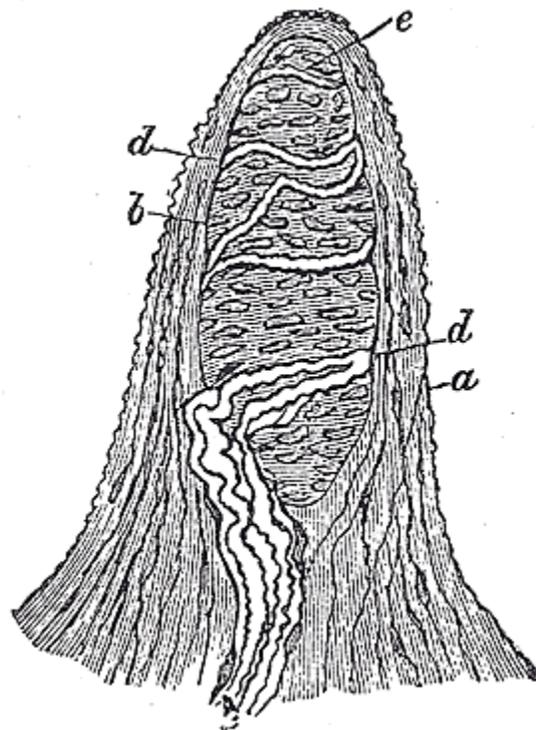


*Figure 5.* Receptive field characteristics for Type II receptors: Pacinian corpuscles (left) and Ruffini endings (right)—FA II and SA II, respectively. The black regions indicate innervation by a sensory neuron via a single receptor. As depicted, Type II receptive fields are less distinctive and much larger compared to those of Type I receptors. The arrows within the receptive fields of the Ruffini endings indicate the direction of maximum sensitivity to skin stretch. The sensitivity threshold plots show the single hot spot of these fields. *Reprinted from “Tactile sensory coding in the glabrous skin of the human hand,” by Johansson, R. S., & Vallbo, A. B., Jan. 1983, Trends in Neuroscience, 6, p. 30. Copyright © 1983 by Elsevier. Reprinted with permission.*

***Mechanoreceptors in glabrous skin.*** The Merkel cell (SA I) is a nearly rigid structure of epithelium surrounding a nerve terminal (Kandel, Schwartz, & Jessell, 2000). Merkel cells are found in clusters beneath the ridges of the glabrous skin in the epidermal-dermal junction—that is, between the epidermis (protective top layer of skin) and dermis. The structure of Merkel cells and their small receptive fields enable them to sense fine points of constant pressure, conveyed by a steady firing pattern where firing

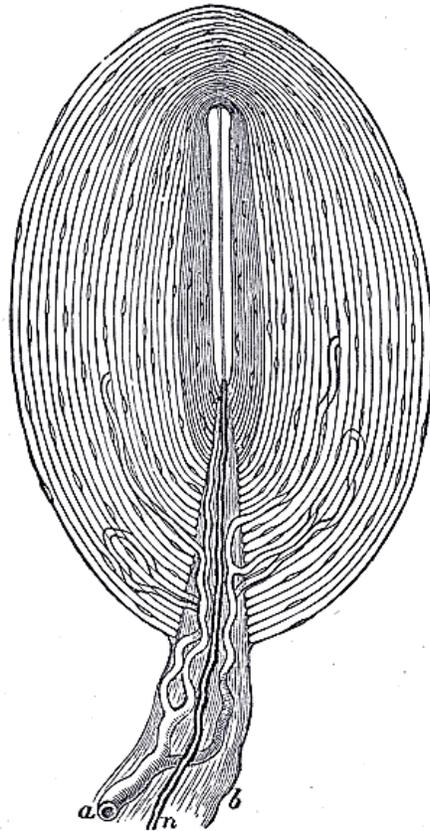
rate, along with the number of receptors activated, indicates the intensity of the applied pressure. Merkel cells are also present in hairy skin, but are found closer to the epidermis.

The Meissner's corpuscle (RA I), depicted in figure 6, is a structure of flattened cells arranged in a column within fluid where a nerve terminal wraps around the cells (Kandel, Schwartz, & Jessell, 2000). Similar to the Merkel cell, it is found in the epidermal-dermal junction in the dermal papillae. The structure of Meissner's corpuscles combined with their small receptive fields enable them to sense fine changes in pressure, conveyed by a firing rate indicative of the rate of pressure variation. Meissner's corpuscles are not present in hairy skin.



*Figure 6.* Anatomical sketch of Meissner's corpuscle as faithfully reproduced from Gray's Anatomy: (a) dermal papilla; (b) Meissner's corpuscle; (d) nerve terminal; and (e) end of nerve terminal. Adapted from the Wikimedia Commons: Gray936.png, <http://en.wikipedia.org/wiki/File:Gray936.png>.

Pacinian corpuscles (RA II), shown in figure 7, are found in the deep layers of the dermis and subcutaneous tissue in both glabrous and hairy skin. The receptor has concentric layers of thin tissue, or lamellae, with a nerve terminal contained in the fluid filled center of the structure (Kandel, Schwartz, & Jessell, 2000). Like the Meissner's corpuscle, the Pacinian corpuscle's structure allows detection of changes in pressure as opposed to constant, steady pressure. Deformation of its structure compresses the nerve terminal, generating an action potential, but the structure can quickly reshape, reducing effects of compression, and ceasing activation for constant pressure. Given their structure, Pacinian corpuscle's can detect low-amplitude high-frequency vibrations applied to the skin—even centimeters away given their large receptive fields (Kandel, Schwartz, & Jessell, 2000). Regarding vibrotactile stimulation, the frequency of a mechanoreceptor's firing rate will increase with increases in vibration frequency. The intensity of a vibration is conveyed by the number of activated mechanoreceptors given that high intensity vibrations propagate farther across the skin. The vibrotactile sensitivity of mechanoreceptors varies with Merkel cells being sensitive to low frequencies within the 5-10 Hz range; Meissner's corpuscles being sensitive to moderate frequencies within the range 20-50 Hz; and Pacinian corpuscle having the highest sensitivity to vibrations around 250 Hz, and the largest range of detectable frequencies: 60-400 Hz (Kandel, Schwartz, & Jessell, 2000).



*Figure 7. Anatomical sketch of Pacinian corpuscle as faithfully reproduced from Gray's Anatomy. The nerve terminal is denoted by n. Adapted from the Wikimedia Commons: Gray935.png, <http://en.wikipedia.org/wiki/File:Gray935.png>.*

Ruffini endings (SA II), depicted in figure 8, have a spindle-like structure in which stretching of the skin compresses the nerve terminal, causing the receptor to fire while slowly adapting to constant stimuli (Kandel, Schwartz, & Jessell, 2000). Like Pacinian corpuscles, they are located in the deep layers of the dermis and subcutaneous tissue in both glabrous and hairy skin. Their receptive fields are large with stimuli evoking larger responses when the direction of skin stretch aligns with the receptive field's direction of maximum sensitivity.

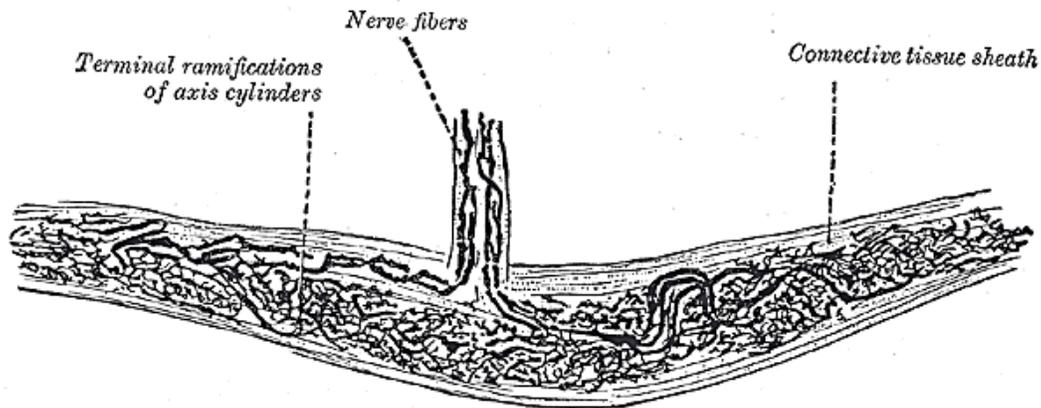


Figure 8. Anatomical sketch of a Ruffini ending as faithfully reproduced from Gray's Anatomy. Adapted from the Wikimedia Commons: Gray937.png, <http://en.wikipedia.org/wiki/File:Gray937.png>.

The combined activations of the aforementioned mechanoreceptors contribute to our haptic perception of an object (Kandel, Schwartz, & Jessell, 2000): Merkel cells respond more rapidly to higher curvature (e.g., a point) compared to flat surfaces; the activation patterns of both Merkel cells and Ruffini endings relate to the shape of an object; Meissner's corpuscles detect surface irregularities and edges; Pacinian corpuscles respond to vibrations and rapid movements; and Ruffini endings indicate when our grasp needs to be tightened to prevent slippage. All of these sensory inputs are combined to create rich, tactual experiences.

***Mechanoreceptors in hairy skin.*** The mechanoreceptors of hairy skin include Pacinian corpuscles, Ruffini endings, Merkel cells, hair follicle receptors and field receptors (Kandel, Schwartz, & Jessell, 2000). Similar to glabrous skin, receptive fields in hairy skin vary over the surface of the body with receptive field size increasing from distal to proximal bodily regions. A hair follicle receptor (RA), found only in the hairy skin of the body, is a sensory nerve that wraps around the hair follicle. This structure allows the sensory nerve to detect changes in hair position. There are three types of hair

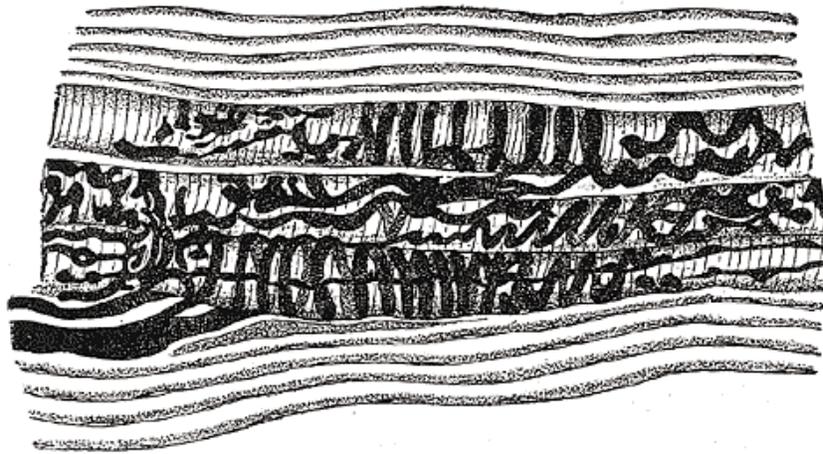
follicle receptors: down, guard and tylotrich—each of which differ in sensitivity. Lastly, field receptors (RA) detect skin stretch over the joints of the body.

*Mechanoreceptors in muscles, tendons and joints.* Proprioceptors are mechanoreceptors in the muscles, tendons and joints that sense and convey information for proprioception—our sense of limb position and movement. Proprioceptors include muscle spindle receptors, Golgi tendon organs and joint receptors. In addition, Ruffini endings, Merkel cells (hairy skin) and field receptors provide cutaneous proprioception needed for facial and lip movements (Kandel, Schwartz, & Jessell, 2000).

Muscle spindle receptors (Kandel, Schwartz, & Jessell, 2000), depicted in figure 9, are widely distributed deep within muscles. A muscle spindle receptor has a spindle-like form with sensory nerves (primary and secondary) that wrap around intrafusal muscle fibers (static nuclear bag fibers, dynamic nuclear bag fibers and nuclear chain fibers) contained within and arranged in parallel to extrafusal muscle fibers. A muscle spindle has two to three nuclear bag fibers, and around five nuclear chain fibers. A single primary muscle spindle ending wraps around the central, non-contractile regions of the static and dynamic nuclear bag fibers, and the nuclear chain fibers. A maximum of eight secondary muscle spindle endings wrap around the central, non-contractile regions of the static nuclear bag fibers and nuclear chain fibers. The distal regions of the intrafusal fibers are contractile.

When a muscle is lengthened (stretched), muscle spindles are stretched, causing their sensory nerves to stretch, and consequently, fire. More specifically, primary muscle spindle endings convey information pertaining to muscle length and the rate of change of muscle length via their firing rate. Secondary muscle spindle endings are slightly sensitive to variations in muscle length, but mostly convey information about static muscle length. Muscle shortening decreases stretch, and hence, decreases firing. This

enables muscle spindle receptors to convey information about the positions of limbs, their movements and their relative angles. Motor neurons, namely gamma motor neurons (Kandel, Schwartz, & Jessell, 2000), innervate the distal contractile regions of intrafusal muscle fibers, providing a means for varying the sensitivity of muscle spindles; in particular, gamma motor neurons are activated during muscle contraction to lengthen muscle spindles so that they may maintain sensory input—otherwise, muscle spindles would not be useful during contraction, and muscle length and rate of change could not be accurately assessed.



*Figure 9.* Anatomical sketch of a portion of a muscle spindle receptor from an adult cat as faithfully reproduced from Gray's Anatomy. Adapted from the Wikimedia Commons: Gray939.png, <http://en.wikipedia.org/wiki/File:Gray939.png>.

The Golgi tendon organ (Kandel, Schwartz, & Jessell, 2000), depicted in figure 10, connects a tendon to muscle fibers. These receptors are thin structures innervated by a single nerve fiber that splits and weaves through collagen fibers within the receptor's structure. An increase in muscle tension causes these receptors to stretch, and in turn, stretch their inner collagen fibers, which compress the sensory nerve endings. Therefore, these receptors sense muscle tension and changes in muscle tension.

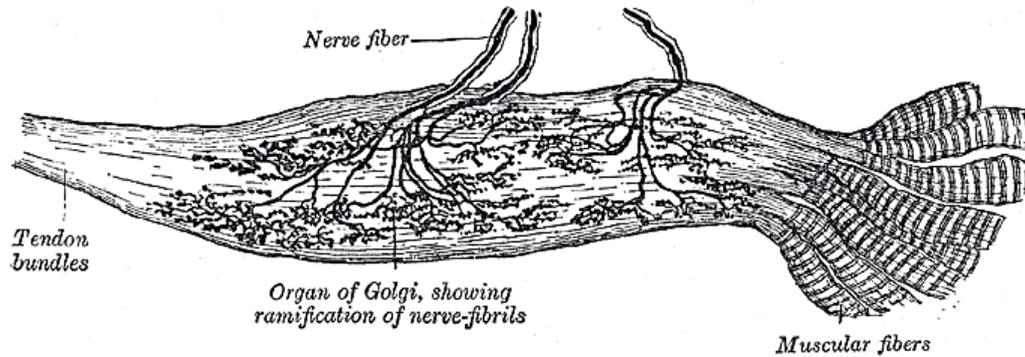


Figure 10. Anatomical sketch of a Golgi tendon organ as faithfully reproduced from Gray's Anatomy. Adapted from the Wikimedia Commons: Gray938.png, <http://en.wikipedia.org/wiki/File:Gray938.png>.

Joint receptors (Purves, et al., 2008) include the Type II mechanoreceptors of the skin, namely Pacinian corpuscles and Ruffini endings. Joint receptors are located in and near the joints of the human body. Although joint receptors don't sense accurate positional information—except for in the fingers (Purves, et al., 2008)—they signal movements of flexion or extension, and warn of joint angles beyond safe ranges of motion.

In summary, the rich input received from mechanoreceptors of the skin and proprioceptors of the muscles, tendons and joints during object grasping and exploration guides grip adjustment and fine motor control. Proprioceptors also communicate information about object shape based on the positions of fingers and limbs. In the next section, the central mechanisms are described in terms of the brain areas that receive the aforementioned peripheral signals, and how they integrate these signals to form a percept.

**Somatosensory cortex.** The parietal lobe of the brain receives and processes peripheral input from touch receptors of the skin, muscles, tendons and joints. It contains several processing areas for touch input (Kandel, Schwartz, & Jessell, 2000): the primary somatosensory cortex (S-I), the secondary somatosensory cortex (S-II) and the posterior parietal cortex. S-I is located on the postcentral gyrus (figure 11)—the ridge of a fold on

the cerebral cortex near the frontal lobe’s motor cortex on the precentral gyrus. These gyri are separated by a fissure called the central sulcus, which is where S-I begins. S-II is posterior to S-I in the parietal operculum, which lies on the lateral sulcus—a fissure dividing the temporal lobe from the frontal and parietal lobes. Lastly, the posterior parietal cortex is immediately posterior to S-I.

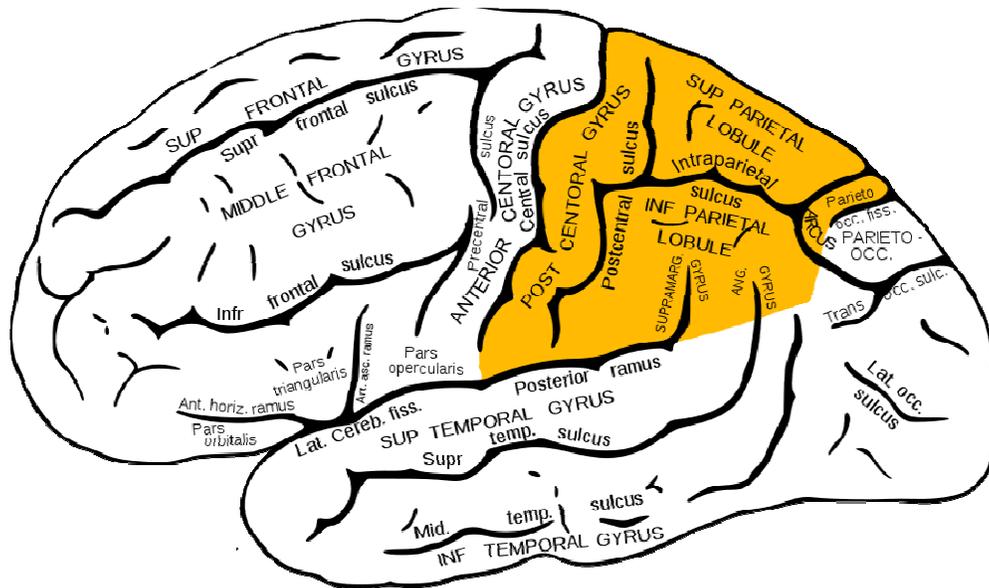


Figure 11. Surface sketch of the left cerebral hemisphere of the brain with the parietal lobe highlighted. Adapted from the Wikimedia Commons: [Gray726\\_parietal\\_lobe.png](http://en.wikipedia.org/wiki/File:Gray726_parietal_lobe.png), [http://en.wikipedia.org/wiki/File:Gray726\\_parietal\\_lobe.png](http://en.wikipedia.org/wiki/File:Gray726_parietal_lobe.png).

S-I consists of four areas (Kandel, Schwartz, & Jessell, 2000): Brodmann’s areas 3a, 3b, 1 and 2, which are arranged respectively on the postcentral gyrus. Area 3a receives proprioceptive input from the proprioceptors of the periphery, whereas area 3b receives tactile input from the cutaneous receptors of the periphery. The output of these areas is sent to the adjacent areas 1 and 2, which is where processing becomes more convergent. These areas also receive input from cutaneous and proprioceptive inputs, but areas 3a and 3b receive more afferent nerves. Areas of S-I project to S-II and the posterior parietal cortex. S-II subsequently projects to areas related to emotion and

memory. Within the posterior parietal cortex, area 5, which lies immediately posterior to S-I, receives input from S-I and associational cortices; area 7 lies immediately posterior to area 5, and receives input from S-I, associational cortices, and visual input. These areas project to the motor cortex (area 4), and are interconnected with area 5 and 7 of the contralateral hemisphere of the brain via the corpus callosum.

Similar to mechanoreceptors, the cortical neurons of S-I, S-II and the posterior parietal cortex have receptive fields. However, these fields are much larger given that cortical neurons in the somatosensory cortex receive input from many sensory neurons via interneurons. Size continues to increase as higher levels are reached from areas 3a and 3b, to areas 1 and 2, and then on to areas 5 and 7 (Kandel, Schwartz, & Jessell, 2000): for example, area 2 has much larger receptive fields (e.g., a finger or multiple fingers) compared to area 3a, 3b and 1 (e.g., fingertips)—and area 5 and 7 have even larger fields compared to area 2 (e.g., bilateral fields covering both hands via interconnections through the corpus callosum).

The topographic organization of cutaneous and proprioceptive inputs from the periphery is preserved throughout each area of S-I. This internal body map or topographic representation, called a sensory homunculus (Penfield & Rasmussen, 1950), enables accurate localization of skin and proprioceptive inputs. The amount of cortical space in S-I for a particular body part depends on innervation density—for those regions of the body that are highly sensitive and densely innervated, more cortical space is provided. For example, even though the hand is not as physically big as the abdomen, it is more densely innervated, and therefore, has a larger representation in S-I. Our internal representations within the homunculus are “plastic” in that cortical space (and receptive fields) may vary with experience.

A cortical neuron receives input of a particular modality (Merkel cells, Meissner's corpuscles, etc.) and adaptation (SA or RA) (Kandel, Schwartz, & Jessell, 2000). Within the cortex, cortical neurons are arranged in columns. In each column, cortical neurons respond to the same modality, are of the same adaptation type and correspond to the same location on the skin. Area 1 is sensitive to touch input (specifically, rapidly adapting cutaneous input), whereas area 2 integrates input from both proprioceptors and cutaneous receptors (RA and SA), and multiple modalities. Area 1 and 2, but mostly area 2, are tuned to more complex features of touch stimuli such as edge orientation, directionality of strokes, spacing of ridges, curvature, etc. In combination with proprioceptive inputs, the feature detectors of area 2 aid in three-dimensional object perception.

Topographic organization is present in only S-I of the somatosensory cortex—S-II and the posterior parietal cortex feature a functional organization given their high level representations (Kandel, Schwartz, & Jessell, 2000). Area 5 associates cutaneous and proprioceptive input to derive postures (such as the posture of the hand while grasping)—and with cutaneous, proprioceptive and visual input, the associative cortical neurons of area 7 assist with visuo-motor coordination during object grasping and manipulation.

## **Psychophysics and Perception of Vibrotactile Stimulation**

This section explores the dimensions of a vibratory signal for encoding information. For each dimension, relevant human psychophysical and perceptual results are presented, and design guidelines are described. This work forms the basis of designing vibrotactile communication systems, the results of which have been taken into account during the design and development of the applications described in Chapter 5 and 6.

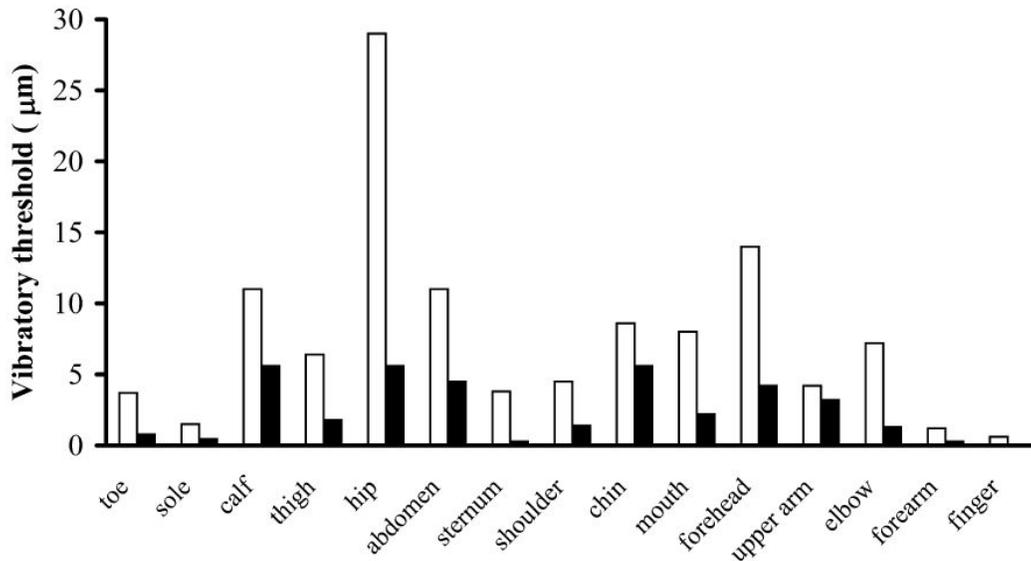
**Vibrotactile sensitivity and vibration frequency.** A rotating mass vibration motor consists of a DC motor with an off-center weight, which as it rotates, causes the unit to vibrate with a sinusoidal oscillation. This is the most common type of vibration motor for handheld portables and wearable cyberphysical systems as it is inexpensive, easy to use, small and lightweight. It is often found in cell phones and other products in the form of coin vibrating motors (also known as a pancake motors) or cylindrical motors. The number of cycles per second, measured in Hertz, is the *frequency* of the vibration signal. Upon actuation, receptors in our skin may or may not sense the vibration depending upon characteristics of the vibration and our vibrotactile sensitivity.

**Vibrotactile sensitivity.** Perceptible vibration frequencies fall within the approximate range of 20 Hz to 1,000 Hz (Gunther, 2001), where Gunther found that below 20 Hz, the vibration signal is no longer perceived as a vibration, but rather, motion. Up to and above 1,000 Hz, our sensitivity to vibrations rapidly lessens (Verrillo & Gescheider, 1992); i.e., larger amplitude thresholds (the amplitude value of which the vibration is just perceptible) are encountered. Additionally, hardware limitations come into effect: as vibration frequency increases to levels above 1,000 Hz, it may be difficult to generate amplitudes above the amplitude thresholds found at these high frequencies (Wilska, 1954). Across our body, we are most sensitive to vibrations falling in the frequency range of 150 Hz to 300 Hz (Jones & Sarter, 2008). This range of frequencies

requires the smallest vibration amplitude for perception compared to frequencies outside this range. The vibration frequency we are most sensitive to is 250 Hz (Verrillo R. T., 1963) in that, at this frequency, the lowest amplitude threshold is found. If we compare the skin's frequency detection range with that of the ear—20 Hz to 20,000 Hz—we see a substantial difference between touch and hearing in terms of frequency resolution, making hearing much more apt for frequency discrimination.

Depending on the location of vibrotactile stimulation across the surface of the skin, amplitude thresholds vary (Jones & Sarter, 2008) due to changes in the density of receptors in the skin and variations in the underlying tissue including muscle and bone structures. Wilska (1954) investigated vibration amplitude thresholds across the body for different frequency values, specifically, 50 Hz, 100 Hz, 200 Hz, 400 Hz and 800 Hz. Overall, 200 Hz achieved the lowest amplitude thresholds, which ranged from 0.07 micrometers for the fingertip (and even lower at 0.02 micrometers for 270 Hz) to larger values for the abdominal and gluteal regions, where the latter values were the largest among the smallest amplitude thresholds. The results found by Wilska suggest that vibrotactile sensitivity lessens from the distal anatomical structures (hands, feet, etc.) to the proximal structures (abdomen, hip, thighs, etc.). Figure 12 depicts a bar chart created by Jones et al. (2008), but based on original data collected by Wilska, consisting of amplitude thresholds for different bodily regions for frequencies 100 Hz and 200 Hz. Although Wilska's results are limited in that only a single contactor area size of 1 cm<sup>2</sup> was used as well as involving a limited number of subjects, the results clearly show that vibrotactile sensitivity varies with respect to both body site and frequency. Similarly, Verrillo and Chamberlain (1972) found vibrotactile sensitivity to decrease from regions of higher density to regions of lower density, specifically, from the fingerpad, to the palm, and finally, to the forearm; however, this result was found only when a surround

was used to prevent the spread of vibrations across the skin. Without a surround, the sensitivity of the palm became greater than the fingerpad, perhaps due to the activation of more mechanoreceptors by the propagating vibration, as speculated by Verrillo and Chamberlain. This suggests that the total number of activated mechanoreceptors, rather than the innervation density of the region, largely determines sensitivity.



*Figure 12.* Amplitude thresholds, measured in micrometers, for different regions of the body at 100 Hz, shown by white bars, and 200 Hz, shown by black bars. Note the difference in sensitivity between 100 Hz and 200 Hz, and how sensitivity varies across the body. Reprinted from “*Tactile displays: Guidance for their design and application,*” by Jones, L. A., & Sarter, N. B., 2008, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(1), p. 92. Copyright © 2008 by Sage Publications, Inc. Reprinted with permission.

Verrillo (1963) explored the relationship between amplitude threshold, frequency and contactor size on the palm of the right hand. Figure 13 and 14 depict amplitude threshold as a function of contactor area and frequency, respectively. In both figures, we see that for small contactor areas—specifically 0.005 cm<sup>2</sup> and 0.02 cm<sup>2</sup>—vibrotactile sensitivity is independent of frequency. This is more obvious in figure 14, but it is also shown in figure 13 by the cluster of data points at 0.005 cm<sup>2</sup> and 0.02 cm<sup>2</sup> on the

horizontal axis. Also, in both figures, we see that for small frequency values—25 Hz and 40 Hz—sensitivity is independent of contactor area. Verrillo speculated that amplitude threshold's independence of contactor area for small frequencies, and independence of frequency for small contactor areas, might be due to these stimulations activating pressure-sensitive receptors that are not responsive to changes in larger vibration frequencies or contactor areas, rather than those more sensitive to vibrations. Another important observation from both figures is that as contactor area increases (at least 0.08 cm<sup>2</sup> and above), our sensitivity to vibration increases, i.e., amplitude thresholds lessen. For these larger contactor sizes (0.08 cm<sup>2</sup> and greater), as frequency varies from low to high, the slope of the curve exhibits a U-shape—see figure 14. The dip in the curve represents maximum sensitivity within the frequency range of 200-300 Hz with a peak of approximately 250 Hz (Verrillo & Gescheider, 1992).

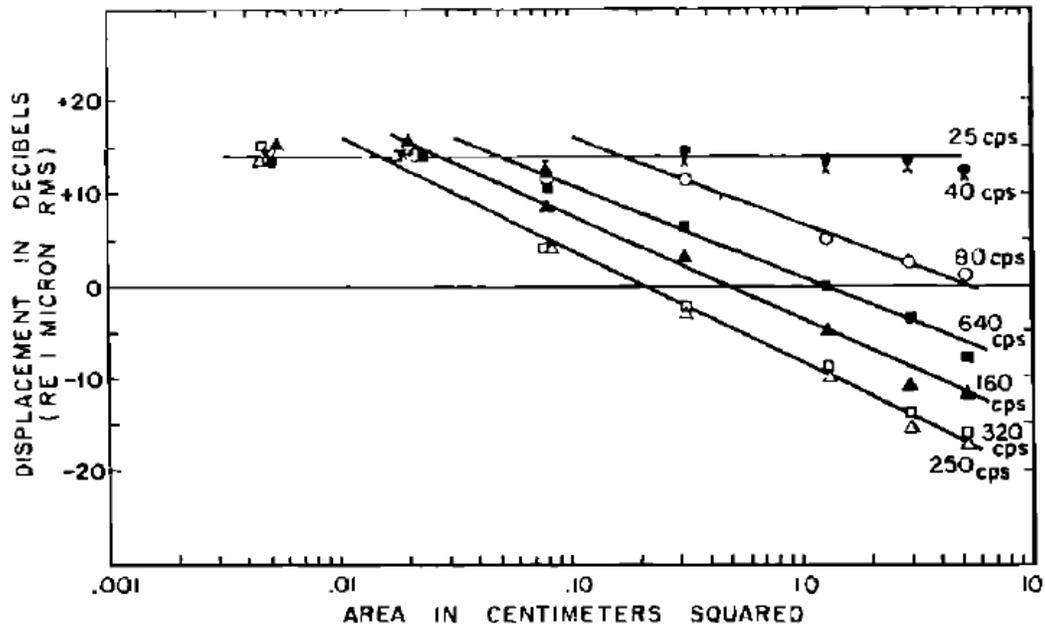


Figure 13. Plot of amplitude threshold as a function of contactor area for various frequencies. Reprinted from "Effect of contactor area on the vibrotactile threshold," by Verrillo, R. T., 1963, *The Journal of the Acoustical Society of America*, 35(12), p. 1964. Copyright © 1963 by Acoustical Society of America. Reprinted with permission.

*Dependencies between frequency, amplitude and pitch.* Both physical and perceptual dependencies exist between these dimensions. The first dependence is a physical interaction between frequency and amplitude, caused by the design of rotating mass vibration motors. For these motors, the frequency of a vibration is increased by increasing voltage, which in turn, increases the speed of the motor's rotating mass; this, subsequently, decreases the vibration's amplitude. Similarly, decreases in frequency slow the speed of the rotating mass, increasing amplitude. Therefore, for this type of vibration motor, frequency and amplitude cannot both be used to convey separate information within the context of vibrotactile communication. If these dimensions must be independent, other options exist including solenoid vibrating motors.

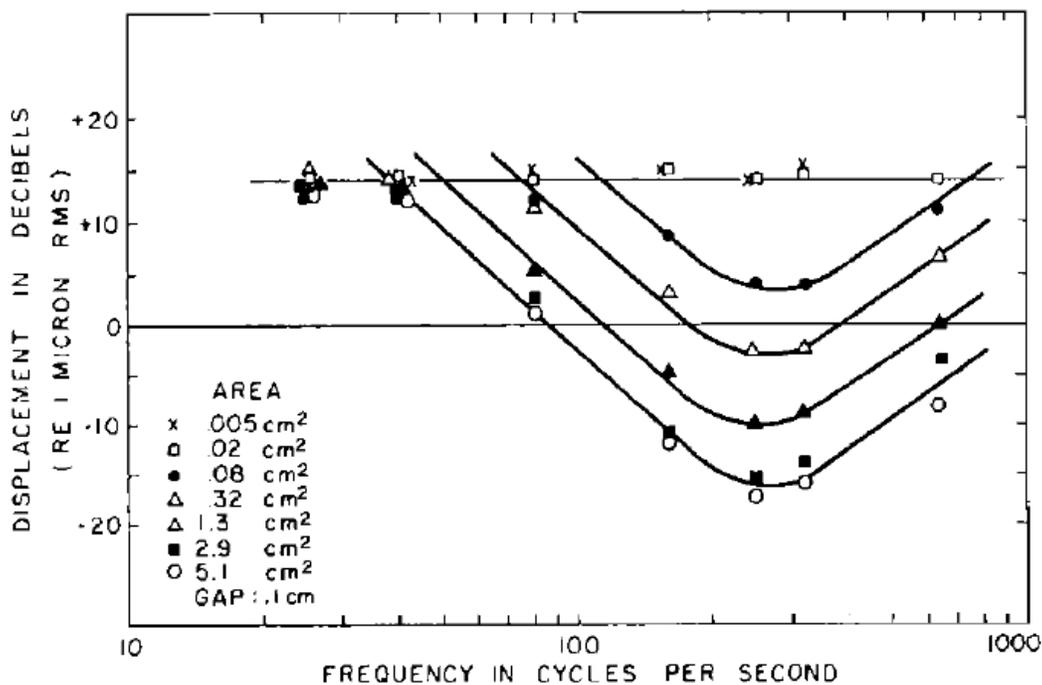


Figure 14. Plot of amplitude threshold as a function of frequency for various contactor areas. Reprinted from "Effect of contactor area on the vibrotactile threshold," by Verrillo, R. T., 1963, *The Journal of the Acoustical Society of America*, 35(12), p. 1965. Copyright © 1963 by Acoustical Society of America. Reprinted with permission.

The second dependence is a perceptual interaction between frequency, amplitude and pitch. *Vibrotactile pitch*—or perceived frequency—varies with changes in amplitude and frequency as provided by the stimulation. (In subsequent discussions, pitch refers to perceived frequency, whereas frequency and amplitude refer to the operating values of the vibration motor.) In general, as the suprathreshold amplitude increases, so does pitch. In a study conducted by Morley and Rowe (1990) where vibrations (30 Hz and 150 Hz) were delivered to the index fingertip of the left hand, most subjects perceived pitch increases with amplitude increases, even though frequency remained unchanged. However, large inter-subject variability was found with two out of the eight subjects experiencing opposite effects, and one subject experiencing no changes in pitch as amplitude changed. Moreover, conflicting results were found in a previous study by von Békésy (1962) in which for larger frequencies (at least 100 Hz and higher), pitch decreased with amplitude increases while frequency remained unchanged. Morley and Rowe present convincing claims pertaining to flaws in von Békésy’s study such as an experimental design that allowed for adaptation of vibrations, thereby possibly causing the observed decrease in pitch with increases in amplitude.

Vibrotactile pitch has also been found to change with frequency. Specifically, the density of mechanoreceptors in the skin, and underlying tissue structures, affect how vibrotactile pitch changes with frequency (Jones & Sarter, 2008): at areas of higher density, more rapid increases of pitch are perceived with increases in frequency. Hence, even when frequency remains unchanged, how we perceive it varies with body site.

***Relative and absolute frequency discrimination.*** Given our skin’s limited frequency resolution and discrimination capabilities, in addition to interactions between vibration frequency and motor design (such as contactor size), among interactions with

other vibration dimensions, such as amplitude and body site, it is a challenging parameter to use for vibrotactile information delivery. For frequency discrimination, humans excel at relative (comparative) frequency discrimination as opposed to absolute frequency discrimination (Brewster & Brown, 2004); but for vibrotactile communication, the latter may be more useful. Smaller frequencies—below 70 Hz—are more discernible than larger frequencies, and discrimination difficulty increases rapidly with frequency increases (Geldard F. A., 1960). Alternatively, frequency may be utilized in another form: amplitude modulated vibration signals—discussed in the following section. By modulating a vibration signal of a certain frequency with another signal of a different frequency, different perceptual “roughness” levels may be created. Hence, roughness might be used as another dimension wherein vibrations feel rougher (or smoother) than others. This dimension was proposed by Brown, Brewster and Purchase (2005) for communication via tactile icons, who revealed its potential as a useful dimension for vibrotactile communication; although much more useful and reliable parameters exist such as rhythm and body site (described in later sections).

**Vibration intensity.** *Amplitude*, or *intensity*, is the magnitude of a vibration, and is measured in terms of either the orthogonal displacement of the vibrating element, or skin indentation. The amplitude of a sinusoid may be defined in a number of ways, but for vibrations, the most common definition is either peak amplitude or root-mean-square (RMS) amplitude, where the latter is the standard deviation of the oscillating signal. Given the extremely small displacements of vibration motors and the large range of these values, displacements are usually visualized as a logarithmic scale using decibels with a typical reference level of one micrometer, or micron. In such a plot, one micron is 0 dB, and each +/-20 dB represents a displacement difference by a factor of 10; for example, compare 1 micron (0 dB) to 10 microns (20 dB), or to 0.1 microns (-20 dB).

Vibration amplitudes are first perceptible at their detection threshold—the smallest amplitude value that can be detected—which depends on vibration frequency and body site. The upper limit of useful amplitudes for vibrotactile communication is about 55 dB above the detection threshold, above which amplitudes may cause pain (Verrillo & Gescheider, 1992). As with frequency, the range of intensities perceptible through touch is relatively small compared to those intensities perceptible through hearing—up to about 130 dB above the detection threshold (Verrillo & Gescheider, 1992). However, regarding 55 dB as an upper bound for intensity, the useful range of vibration amplitudes for vibrotactile displays and communication devices is much less: Gunther (2001) recommends a value of around 15 dB above the detection threshold, which he describes as a “comfort zone.”

*Dependencies between amplitude, frequency and sensation magnitude.* The term *sensation magnitude*, or *loudness*, refers to the perceived vibration amplitude in contrast to the operating amplitude of the vibration motor. The effect of amplitude variations on human perception of frequency was previously discussed, but changes in frequency also affect human perception of intensity, even when intensity is kept constant. This is demonstrated in figure 15, which depicts curves of *equal sensation magnitude* (Verrillo, Fraioli, & Smith, 1969) (Verrillo & Gescheider, 1992) for the palm of the hand; that is, each curve represents the vibration amplitude needed, with respect to a particular vibration frequency, to achieve the sensation magnitude indicated by the curve. These curves are useful for adjusting vibration amplitude when a particular sensation magnitude is desired. The plot also shows that when amplitude is kept constant while increasing frequency, this will cause an increase in sensation magnitude, but only for frequencies up to approximately 250 Hz; and if frequency is kept constant while increasing amplitude, this will cause an increase in sensation magnitude.

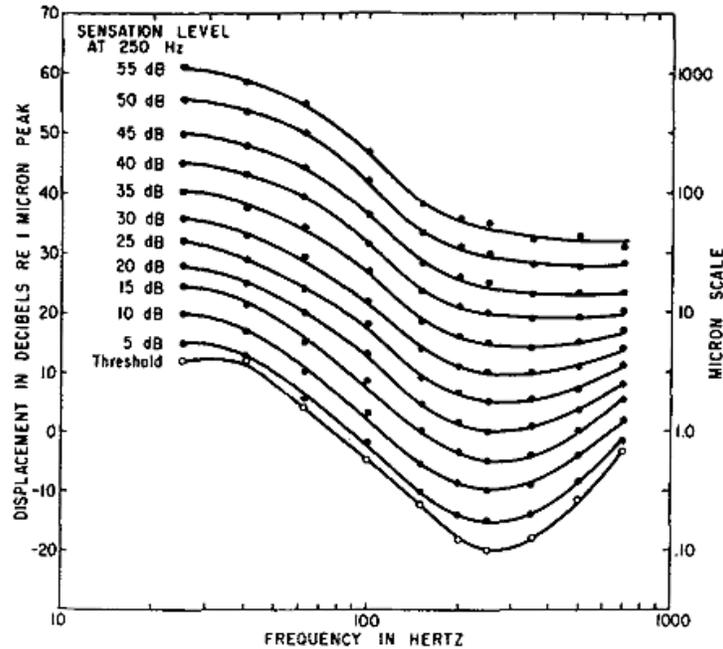


Figure 15. Plot of equal sensation magnitude curves across frequencies and vibration amplitudes with sensation magnitudes specified at the start of each curve. Reprinted from "Sensation magnitude of vibrotactile stimuli," by Verrillo, R. T., Fraioli, A. J., & Smith, R. L., 1969, *Perception & Psychophysics*, 6(6A), p. 371. Copyright © 1969 by Psychonomic Society, Inc. Reprinted with permission.

As shown in figure 15, sensation magnitude increases with increases in amplitude, but the rate of increase depends on body site. With a surround to contain propagating vibrations, Verrillo and Chamberlain (1972) found that sensation magnitude increases faster, with respect to increases in vibration amplitude, at body sites of lower innervation density compared to those sites of higher density. However, results also showed that when the surround was removed, and vibrations were allowed to spread across the skin, the sensation magnitude increased slower for the palm than it did for the fingerpad; clearly, though, the fingerpad has a higher density of mechanoreceptors compared to the palm. As with vibrotactile sensitivity, Verrillo and Chamberlain speculated that as the number of activated mechanoreceptors increases, irrespective of innervation density, so too does sensation magnitude and the rate of increase of sensation

magnitude. Supporting this claim is evidence that magnitude sensation depends on the density of actuated vibration motors (Cholewiak R. W., 1979), where increases in vibration motor density translates to activating more mechanoreceptors. Cholewiak found that as the number of vibration motors—arranged in a closely spaced 2D array—increases, so too does the sensation magnitude of the vibration, regardless of the intensity of the motors (assuming each motor is actuated with the same intensity). This effect is also seen with more sparsely spaced vibration motors.

*Relative and absolute amplitude discrimination.* As with frequency, humans are also better at relative amplitude discrimination rather than absolute discrimination. Geldard and his colleagues found 15 just noticeable intensity differences on the chest, starting at an indentation of 50 micrometers, which is the lowest vibration amplitude that can be detected 100% of the time for the chest region (1957). These just noticeable differences are depicted in figure 16, which range anywhere from 10 to 60 micrometers. In terms of absolute intensity discrimination, Geldard recommended three values spaced generously along the range of just noticeable differences; specifically, intensity values that translate well to the concepts of “soft,” “medium” and “loud” sensations (1957). As with frequency, amplitude is a challenging dimension to use for communication given its interaction with frequency as well as variations in sensation magnitude across the body due to changes in innervation density and underlying tissue structures (Geldard F. A., 1960). Moreover, intensities that are too high may be uncomfortable or painful, whereas intensities that are too low may be difficult to perceive, and increase the difficulty of perceiving other vibration dimensions (Brown, Brewster, & Purchase, 2005).

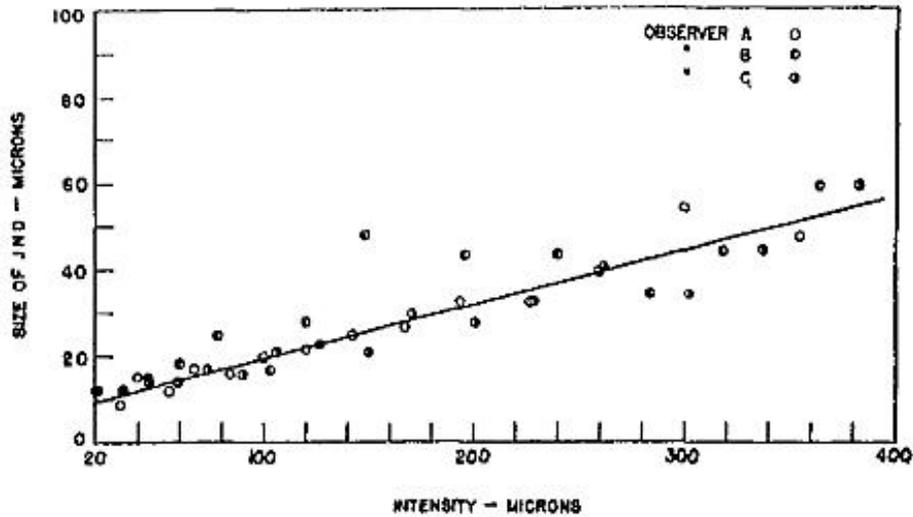
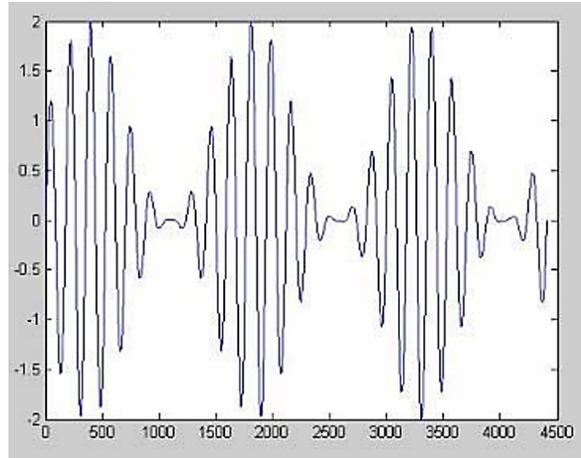


Figure 16. Plot of intensity values versus size of just noticeable differences for the chest region of three subjects. Reprinted from "Adventures in tactile literacy," by Geldard, F. A., 1957, *American Psychologist*, 12(3), p. 119. Copyright © 1957 by American Psychological Association (APA). Reprinted with permission.

As described earlier, frequency and amplitude cannot be controlled separately in rotating mass vibration motors. Intensity is altered through voltage changes affecting the speed of the rotating, off-center mass; but these changes, subsequently, alter frequency as well. In a more pragmatic study by Brown and Kaaresoja (2006), absolute recognition of vibration intensity was explored using a standard vibration motor common among cell phones (realizing, of course, frequency variations). Participants achieved 75% overall recognition accuracy on three intensity values (produced by voltages 0.93 V, 1.16 V and 1.38 V) as part of a multi-dimensional tactile icon; the other dimension was tactile rhythm. Below 0.93 V, the vibration motor was not reliable, and 1.38 V was chosen as the maximum voltage; 1.16 V was chosen to be between the aforementioned voltages. For rotating mass vibration motors, intensity offered an improvement over roughness in terms of being a parameter for vibrotactile communication.

***Varying amplitude over time: Complex waveforms as a parameter.*** It is known that humans can differentiate between simple, distinct waveforms, specifically, a sine

wave and a square wave (Gunther, 2001); but more complex waveforms may be utilized for vibrotactile communication. One possible waveform variation is *roughness* (Brown, Brewster, & Purchase, 2005). The roughness of a vibration is varied through sinusoidal amplitude modulation in which a vibration signal of a base frequency is multiplied by another vibration signal of a different frequency; see figure 17 for an example.



*Figure 17. Example of a “rough” waveform generated by multiplying a 250 Hz sinusoid with a 30 Hz sinusoid. Reprinted from “A first investigation into the effectiveness of tactons,” by Brown, L. M., Brewster, S. A., & Purchase, H. C., 2005, in Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 169. Copyright © 2005 by IEEE. Reprinted with permission.*

In the first investigation of roughness as a parameter by Brown, Brewster and Purchase (2005), TACTAID and C2 tactors were explored. A frequency of 250 Hz was chosen as the base signal, and 20 Hz and 50 Hz were chosen to be multiplied with the base as Brown et al. found that at 20 Hz, the waveform began to feel rough, and above 50 Hz, the waveform began to feel smooth. Experimentation revealed that an un-modulated sinusoid at 250 Hz feels smooth, and is distinct from rough waveforms; and as the frequency of the second waveform increases from 20 Hz to 50 Hz, the perceived roughness decreases. Using the TACTAID actuators, three roughness values were

recommended for absolute identification: an un-modulated 250 Hz sinusoid, 40 Hz modulated sinusoid and 50 Hz modulated sinusoid. Using the C2 actuators, four roughness values were recommended for absolute identification: an un-modulated 250 Hz sinusoid, 20 Hz modulated sinusoid, 40 Hz modulated sinusoid and 50 Hz modulated sinusoid. Overall, the C2 tactor produced sensations wherein roughness was easier to perceive and more intuitive for participants compared to TACTAID tactors.

In a second experiment using only the C2 tactor, roughness was used to communicate information combined with tactile rhythm. In this experiment, Brown et al. used three values—an un-modulated 250 Hz sinusoid, 30 Hz modulated sinusoid, and a 50 Hz modulated sinusoid—where 30 Hz and 50 Hz modulations indicate very rough and rough, respectively. Although roughness did not perform as well as rhythm in terms of recognition (80% recognition accuracy on average versus 93% recognition accuracy on average), it may still be a useful parameter when near perfect recognition performance is not required.

In a follow-up study, Brown, Brewster and Purchase (2006a) evaluated roughness combined with both rhythm and body site. Unfortunately, three values for roughness degraded performance (59.47% roughness recognition accuracy on average, and 48.8% complete tacton recognition accuracy on average). In a second experiment, Brown et al. found that reducing the number of roughness levels from three to two improved recognition accuracy (82.4% roughness recognition accuracy on average, and 80.56% complete tacton recognition accuracy on average), but performance was still lacking compared to tactile rhythm and body site.

It is important to note, however, that human performance using C2 tactors (solenoid vibration motors) may differ from human performance using rotating mass vibration motors (pancake or cylindrical vibration motors), where the latter is more

commonly found in cell phones. To explore these differences with respect to roughness and other parameters, Brown and Kaaresoja (2006) conducted a comparative study using a standard mobile phone vibration motor. Due to the limitations of the hardware, roughness could not be generated using the method described; it was, however, simulated by varying the speed of on-off pulses. Very short on-off pulses of equal duration were used to create “rough” (10 ms) and “very rough” (30 ms) sensations. Absolute identification of roughness using a mobile phone vibration motor was significantly worse compared to using a C2 tactor (55% versus 80% roughness recognition accuracy on average). In this regard, roughness may not be a useful parameter when using rotating mass vibration motors.

Another possible temporal variation of vibration amplitude is an *envelope*, which is the gradual increase and/or decrease of amplitude with respect to time (Gunther, 2001). In a study by Brown, Brewster and Purchase (2006b), envelopes were investigated using a TACTAID device placed on the index finger. Brown et al. found that participants could discriminate between gradual linear or exponential amplitude increases (*tactile crescendos*); gradual linear or exponential amplitude decreases; and level stimuli, i.e., no amplitude changes with time (100%, 92% and 95% recognition accuracy overall, respectively). Gradual logarithmic amplitude increases and decreases were also explored, but their performance in terms of recognition accuracy was less compared to linear and exponential variations. These results show the possible use of envelopes as a new dimension for vibrotactile communication. Moreover, *attacks* (sudden changes) and *decays* (gradual changes) could also be added before or after the envelope (Gunther, 2001). Brown et al. explored attacks prior to the envelope using *tactile sforzando-crescendos*, i.e., a short pulse with large amplitude prior to the start of a gradual increase in amplitude. A sforzando-piano is used to direct attention, and make musical envelopes

more obvious; but in Brown et al.'s study, participants found tactile sforzandos-crescendos to be confusing, and they failed to improve recognition of tactile crescendos, although they did not hurt performance. It should be noted, however, that prior to the experiment, participants were not informed about the meaning and/or purpose of the tactile sforzandos-crescendos, which may have caused the observed confusion. In applications that have utilized vibrotactile bursts as attention grabbing cues before pattern presentation, these alerts have worked well in terms of directing attention (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011).

**Vibration timing and rhythm.** The *burst duration* of a vibration is the amount of time from start to end of motor actuation. The length of a pause between vibrations delivered by the same motor is known as the *interstimulus interval*; and the time between the start of vibrations of two different motors is known as *stimulus onset asynchrony*, or SOA (van Erp, 2005). Vibrotactile pulses or bursts, of same or different burst durations, may be temporally linked and separated by pauses, to form *tactile rhythms*. Geldard and his colleagues (1957) found that below 100 ms, vibrotactile pulses (at least those of 60 Hz) are perceived as pokes or nudges. They also found that burst durations of two seconds or greater might be too slow for vibrotactile communication applications. Typical values of burst durations, as used for vibrotactile communication, fall in the range of 80 ms to 500 ms (Jones & Sarter, 2008). Vibrotactile pulses that have too short of a duration, i.e., below 50 ms, may be perceived as having too weak of intensity, even to the point of not being detected (Kaaresoja & Linjama, 2005). This may be due to how rotating mass vibration motors operate: shorter duration times may prevent vibration motors from reaching target operating values, which may explain the weak intensity perceived by participants in Kaaresoja and Linjama's study. Moreover, with shorter

durations, we may not have enough time to perceive particular dimensional values such as frequency, amplitude, body site or rhythm.

*Relative and absolute temporal discrimination.* With respect to relative discrimination of burst durations, Geldard (1957) reported that, within his recommended range of 100 ms to 2000 ms, there are 25 just noticeable differences (JNDs), the smallest of which was found to be 50 ms—see figure 18. The plot shows that for smaller temporal differences, there's a linear relationship; but for large values, the relationship is curvilinear. Hence, as burst duration increases, so must the temporal difference if two burst durations are to be distinguished. This result follows Weber's Law in that the just noticeable difference depends on the value of the burst duration. In other words, the larger the burst duration, the larger the difference must be between burst durations for accurate discrimination.

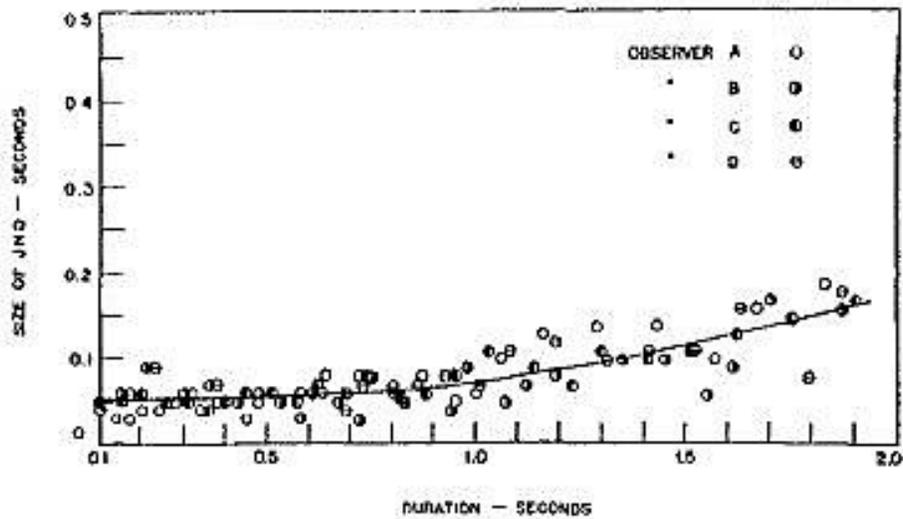


Figure 18. Plot of burst durations versus size of just noticeable differences for four subjects. Reprinted from "Adventures in tactile literacy," by Geldard, F. A., 1957, *American Psychologist*, 12(3), p. 119. Copyright © 1957 by American Psychological Association (APA). Reprinted with permission.

In terms of absolute discrimination, for accurate identification, Geldard recommended three burst durations spread across the shorter durations of figure 18 where the slope is linear—specifically, 100 ms, 300 ms and 500 ms. In a more recent study (McDaniel T. , Krishna, Balasubramanian, Colbry, & Panchanathan, 2008), participants were asked to recognize burst durations of 200 ms, 400 ms, 600 ms, 800 ms and 1000 ms of a 170 Hz vibration applied to their waist. Shorter burst durations of 200 ms and 400 ms were more easily identified compared to larger durations of 600 ms and greater, where participants often confused durations of 600 ms, 800 ms and 1000 ms; indeed, these results support Weber’s Law. Confusion might be overcome by using fewer burst durations, and a wider separation between these durations.

***Tactile rhythm.*** One popular parameter used in many applications of vibrotactile communication is *tactile rhythm*. Tactile rhythms have been successfully used in systems requiring absolute identification of vibrotactile patterns, such as tactons wherein each pattern is assigned an arbitrary meaning. With the proper design, tactile rhythms are generally easy to recognize, which is in contrast to absolute discrimination of frequency and intensity. Many other tactile rhythm designs have been evaluated through a variety of applications including navigation (van Erp & van Veen, 2001) (Lin & Cheng, 2008); tactile music (Gunther, 2001); and assistive technology for individuals who are blind (McDaniel T. L., Krishna, Colbry, & Panchanathan, 2009) (McDaniel T. L., Villanueva, Krishna, Colbry, & Panchanathan, 2010).

Absolute identification of tactile rhythm has been investigated in a number of studies. As part of their research on tactons, Brown, Brewster and Purchase (2005) explored the recognition accuracy of three tactile rhythms applied to the index finger via a C2 tactor; two parameters were explored in total with the second being roughness. The tactile rhythms consisted of a rhythm of seven short pulses, a rhythm of four long pulses,

and a rhythm of one short pulse then one long pulse. Participants achieved an impressive average recognition accuracy of 93% (see earlier discussion for roughness recognition accuracies). In a follow-up study, Brown, Brewster and Purchase (2006a) added a third parameter in addition to rhythm and roughness: body site; specifically, three equidistantly spaced vibration motors on the volar forearm with endpoints at the wrist and elbow joint. An average recognition accuracy of 96.7% was found for rhythm with 95.5% for body site. In both experiments, the rhythms remained constant, but in the latter study, vibrations were applied to the volar side of the forearm rather than the index finger. Using standard vibration motors found in pagers and cell phones, rather than C2 tactors, Brown and Kaaresoja (2006) explored the same rhythm designs that were successfully used in the aforementioned studies. During the experiment, each participant held a phone in his or her non-dominant hand as tactons were delivered wherein subjects had to recognize roughness/intensity and rhythm. An average recognition accuracy of 93% was found for rhythm, showing that similar performance is achievable for standard vibration motors compared to C2 tactors.

If more than three tactile rhythms are required, careful attention should be paid to the intuitiveness of the rhythms so that they better represent their assigned meanings in an effort to reduce cognitive load and improve recognition; moreover, base rhythms should be explored as another method to improve recognition accuracy. As an example, the analogy of a heartbeat inspired the design of tactile rhythms for use in an application to communicate interpersonal distance to individuals who are blind during social interactions (McDaniel T. L., Villanueva, Krishna, Colbry, & Panchanathan, 2010). In this system, faster heartbeats indicated a closer proximity of people in front of the user, whereas slower heartbeats indicated larger distances. Participants found the rhythms to be intuitive, and achieved an average recognition accuracy of 94.3%. A normal heartbeat

rate was used as the base rhythm in which participants were trained to compare this rhythm with all other rhythms to aid recognition; feedback from participants revealed the base rhythm to be very helpful.

Lastly, it is important to mention differences between monotonic (gradual change over time) and distinct (or discrete) tactile rhythms. Both rhythm types have been explored in a number of applications including navigation (van Erp J. B., van Veen, Jansen, & Dobbins, 2005) in which rhythm was used to inform the user of his or her distance to a target destination. For absolute identification, distinct tactile rhythms are recommended given that changes between monotonic rhythms vary too smoothly for exact rhythms to be recognized. Monotonic rhythms might be more useful for signaling a specific event, such as when a user is beginning to close in on their destination in the context of navigation.

**Body site.** The *location, locus, or body site* of a vibrotactile stimulation on the body's surface is a powerful dimension for communication given the impressive expanse of the skin. Vibrating a specific area of the skin may be used to convey information through a variety of methods; for example, tactile icons (Brewster & Brown, 2004), may arbitrarily assign concepts (meanings) to different body sites, such that when a vibration is localized, the user recalls from memory what the system is attempting to communicate based on the respective stimulated body site. Localized vibrations around the waist have been successfully used for navigation and orientation applications. For example, regarding the former application, the location of stimulation around the waist informs the user of which direction to travel (van Erp J. B., van Veen, Jansen, & Dobbins, 2005): a user simply follows the vibrations around his or her waist to travel from his or her current location to a target destination. In the latter application, specific directions, such as magnetic north, are communicated to the user through stimulating the area of the skin

nearest this direction. Pilots (Rupert A. H., 2000) (van Veen & van Erp, 2001) and astronauts (van Erp & van Veen, 2003) may utilize such a system to better orient themselves through awareness of the gravity vector. Many other applications have utilized the location of vibration on the surface of the body to convey information including virtual reality (Lindeman, Page, Yanagida, & Sibert, 2004) and assistive technology for individuals who are blind (McDaniel T. , Krishna, Balasubramanian, Colbry, & Panchanathan, 2008).

Before discussing vibrotactile spatial acuity, we will provide a brief introduction to spatial acuity based on the *two-point limen*, or *two-point threshold*. E.H. Weber (1834/1996), a German physiologist and anatomist, devised the two-point threshold task in the early 19<sup>th</sup> century to explore how sensitivity varies with respect to body region. In the two point threshold task, when two points of pressure are applied simultaneously to the skin, a subject responds with whether he or she feels one or two points. If the points are applied sequentially, then the task becomes point localization in which the subject must respond with whether the two sites of stimulation were the same or different. Weber explored two-point thresholds across the body, and verified that sensitivity depends on the body part; in fact, his results suggested that spatial acuity improves from our proximal (trunk) to distal (face, hands, etc.) body parts, with skin near joints having higher acuity compared to skin near the middle of limbs. Weber suggested that sensitivity variations across the body are related to how receptor density varies across the skin, which might be influenced by how often a body part is used to explore the environment, and the level of movement control over this limb.

Well over a century later, Weinstein (1968) conducted a study to verify Weber's two point thresholds; an updated study was needed given the lack of details with respect to Weber's experimental procedure and subject population. Weinstein explored pressure

sensitivity, and spatial acuity measures including the two point threshold and point localization, across the body for both males and females, and for the left and right sides of the body. Weinstein verified that spatial acuity improves from proximal to distal body parts, and that the amount of cortical space, and subsequently the density of receptors, is related to spatial acuity.

Within the context of vibrotactile communication, the above results are of little use, but they do provide a starting point as focus now shifts to *vibrotactile spatial acuity*; but first it is worth noting why two-point thresholds should not be used to guide the design of vibrotactile communication systems that utilize body site as a parameter: (1) pressure and vibrotactile stimuli each engage different sensory systems, each with their own characteristics; and (2) pressure and vibrotactile stimuli each affect the skin in different ways with the latter causing vibrations to propagate across the skin and through deep tissue (Jones, Held, & Hunter, 2010). The distance the vibration induced surface wave may travel depends on the characteristics of the stimulation as well as the body site, but even for small vibration motors, the surface wave may travel for many centimeters from the source (Cholewiak & Collins, 2003). Given that vibrations spread as opposed to the contained effects of point stimuli, localizing vibrations is more difficult compared to localizing points of pressure; cf. (Weinstein, 1968) (Cholewiak & Collins, 2003) (van Erp, 2005).

It's also important to differentiate between vibrotactile sensitivity and vibrotactile spatial acuity. The latter, rather than the former, should be used to guide the design of vibrotactile communication systems that utilize body site. In any case, however, they do share similarities: recall Wilska's (1954) exploration of vibrotactile sensitivity across the body discussed earlier; amplitude thresholds for different frequencies were found to lessen from proximal to distal body parts—a shared result with that of spatial

acuity using the two point threshold task; cf. (Weinstein, 1968). Although studies exploring vibrotactile spatial acuity are limited compared to vibrotactile sensitivity, of the few studies that have been conducted, much information can be gleaned. This section begins with a discussion covering four key concepts for absolute localization within a linear array of vibration motors; these are *anatomical reference points*, *endpoints*, *odd sites*, and *spacing versus numbers*. These design concepts are based on two important positioning concepts (Cholewiak & Collins, 2003) (Cholewiak, Brill, & Schwab, 2004): *place* and *space*, both of which affect the ability to localize vibrations on the skin. Following this discussion, two-dimensional arrays and relative localization will be covered.

*Anatomical reference points.* In two seminal studies, Cholewiak and his colleagues showed the usefulness of anatomical reference points for vibrotactile localization on the forearm (Cholewiak & Collins, 2003) and around the torso (Cholewiak, Brill, & Schwab, 2004) when using vibration motors that have a static surround. In the 19<sup>th</sup> century, E.H. Weber also discovered the usefulness of anatomical reference points in which spatial acuity, as measured by the two point threshold task, improves from the middle of a limb to its endpoint (joint).

In the former study (Cholewiak & Collins, 2003), vibration frequency, tactor spacing and the number of tactors were varied across the volar side of the left forearm for two subject populations: students (18 to 33 years old) and seniors (60+ years old). Here, results are summarized for the former population group, and only those experiments within the study related to anatomical reference points are described. In their first experiment, seven custom piezoceramic tactors were spaced at least 2.5 cm apart, from center-to-center, from the wrist joint to the elbow joint. Participants were asked to localize the vibrations delivered through the seven tactors at one of two possible

frequencies: 100 Hz or 250 Hz. Overall localization accuracy at wrist and elbow endpoints were better (above 65%) compared to localization accuracies at other sites (30-40%)—see figure 19. No significant difference was found for frequency. Moreover, in subsequent experiments within this study, specifically, their second and fourth experiment where frequency was evaluated in a similar way, no significant difference was found. These limited frequency variations produced only a small effect with respect to localization performance, with lower frequencies generally showing improved, albeit minimal, localization accuracy. This might suggest that suprathreshold frequency changes (and, perhaps, suprathreshold amplitude changes, as both affect sensation magnitude) have little effect on our ability to localize vibrations, given the redundancy and quantity of receptors in our skin. In their second experiment, Cholewiak and Collins centered the same array of tactors on the elbow joint. Localization accuracy at the elbow joint was still comparable to the results found from the first experiment, and superior to other sites in terms of localization accuracy—see figure 19. Also, note the impressive localization performance at the shoulder joint endpoint, but not the other endpoint, which falls on the middle of the forearm.

One might speculate that the superior localization performance at the wrist and elbow joint in the first experiment could be simply due to these points being co-located with the endpoints of the array. While endpoints may be helpful, when these results are compared to the second experiment, it is clear that the elbow joint, acting as an anatomical reference point, is assisting with localization as it now falls in the middle of the array. Moreover, the endpoint, now falling between the wrist joint and the elbow joint in the second experiment, is more difficult to localize compared to when it was co-located with the wrist joint during the first experiment. Lastly, as clearly shown by figure 19, the closer a tactor is to an anatomical reference point, the easier it is to localize.

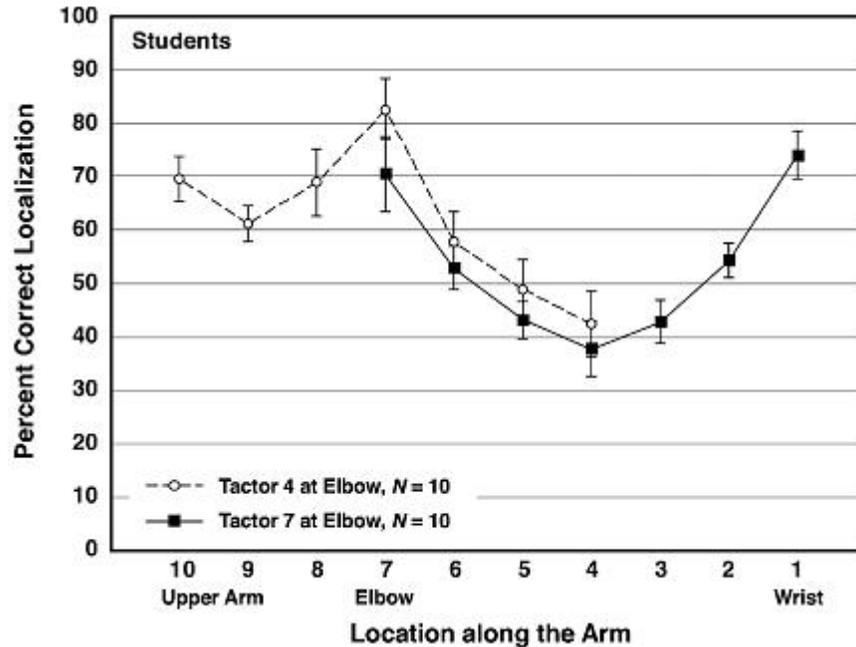


Figure 19. Localization accuracies for seven piezoceramic factors on the volar side of the left forearm either centered between the wrist and elbow joint, or centered at the elbow joint. (Note that these results pertain only to the student group.) When the factors are centered between the wrist and elbow joint, overall accuracies form a U-shaped curve, showing superior performance at the endpoints and nearby points. When the same factor array is centered at the elbow joint, the shoulder joint and elbow joint have superior performance, whereas the endpoint, opposite the shoulder, drops in overall accuracy. These results clearly show the positive effect of anatomical reference points on localizing vibrations. Reprinted from "Vibrotactile localization on the arm: Effects of place, space, and age," by Cholewiak, R. W., & Collins, A. A., 2003, *Perception & Psychophysics*, 65(7), p. 1068. Copyright © 2003 by Psychonomic Society, Inc. Reprinted with permission.

Whereas the previous study showed the usefulness of joints, specifically the wrist, elbow and shoulder joint, as anatomical reference points for localizing vibrations, another study by Cholewiak, Brill and Schwab (2004) showed the same for the navel and spine of the torso. The number of factors, separation of factors, and the orientation of factors, in addition to the waveform of vibration, were explored to assess participants' abilities at localizing vibrations around the lower abdomen and back. These results will be discussed throughout this section, but for now, discussion is limited to those results relevant to anatomical reference points. In the first experiment of the study, twelve C2

factors were spaced equidistantly around the waist such that one factor was centered on the navel, one factor was centered on the spine, one factor was centered on each side, and two factors were placed between each of the aforementioned pairs. Localization accuracy at the navel and spine were near perfect, and neighboring factors were easier to localize compared to factors at the sides and factors neighboring the sides, as depicted in figure 20. This is another example of how anatomical reference points can assist with localizing vibrations. Similar results for the torso were achieved by van Erp (2005) but for relative localization.

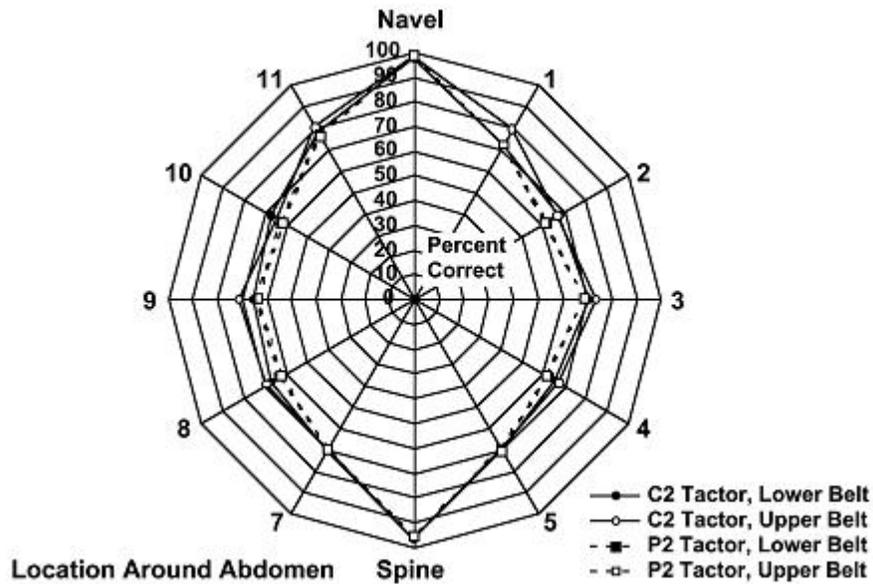
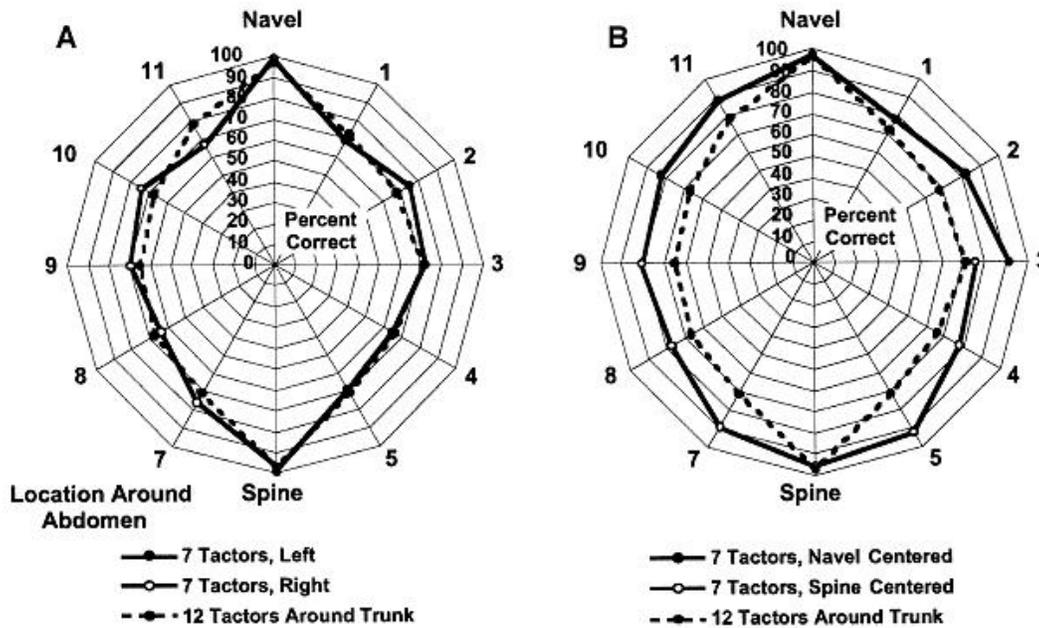


Figure 20. Localization accuracies for twelve factors around the lower or upper part of the abdomen and back. No significant difference was found for localization performance between the upper and lower torso. With respect to C2 factors, and the lower abdomen and back, the following results can be summarized: localization accuracy at the navel and spine were near perfect; nearby points were next best between 70-75%; points at sides followed with accuracies between 67-72%; and finally, points nearby sides were 64% or below. Overall accuracy was 74%. Reprinted from "Vibrotactile localization on the abdomen: Effects of place and space," by Cholewiak, R. W., Brill, J. C., & Schwab, A., 2004, *Perception & Psychophysics*, 66(6), p. 976. Copyright © 2004 by Psychonomic Society, Inc. Reprinted with permission.

**Endpoints.** An endpoint of a tactor array is the first or last tactor of a linear array of vibration motors. As these motors have one less neighboring tactor, localizing them is, in general, easier compared to tactors falling in the middle of an array. Revisiting Cholewiak, Brill and Schwab's study (2004) investigating absolute localization of vibrations around the torso, their third experiment, when compared to the first experiment, shows a significant improvement in performance when endpoints are utilized at the sides of the torso—even when tactor spacing has not changed. In their third experiment, a semi-circle of seven C2 tactors, with equidistant spacing similar to the twelve tactor array used in the first experiment, was evaluated under four different placements: centered at the navel, spine, left and right side. Localization performance was better when the semi-circle of tactors was centered at the navel or spine, compared to the sides; in both cases, anatomical reference points—the navel and spine—were exploited, but the difference in performance, as pointed out by Cholewiak, Brill and Schwab, are the artificial reference points created by the endpoints at the sides of the torso, which were easier to recognize given that only one tactor neighbors them. However, in the case where the tactor array is centered at the left or right side of the torso, the anatomical reference points overlap with the endpoints, and therefore, vibrations at the sides are not as easy to localize; results are depicted in figure 21. Moreover, these results seem to suggest that, as with anatomical reference points, tactors nearby artificial references are easier to localize.

Results may seem to conflict with figure 19, which depicts a low (between 40-50%) localization accuracy for an endpoint tactor (as well as its neighboring tactors) falling between the wrist joint and elbow joint. Even though this is an endpoint, localization performance depends on a number of factors, one of which is tactor spacing. In this case, the spacing of tactors, at 2.5 cm, may have been detrimental to performance;

compare this spacing to that of the factor arrangement in figure 21 where spacing was approximately 6.4-8.2 cm depending on the waist size of a participant. In general, however, tactor array endpoints may be used as artificial reference points to potentially improve localization accuracy.



*Figure 21.* Localization accuracies for seven C2 tactors around the lower part of the abdomen and back, centered at either the left or right side, A, or the navel or spine, B. The left and right sides showed similar performance; as did the front and back of the torso. Localization performance improved when the tactor array was centered at the navel or spine, compared to the sides. Moreover, a significant difference in performance was found between the 12 tactors, and the 7 tactors centered at the navel or spine (B). *Reprinted from "Vibrotactile localization on the abdomen: Effects of place and space," by Cholewiak, R. W., Brill, J. C., & Schwab, A., 2004, Perception & Psychophysics, 66(6), p. 980. Copyright © 2004 by Psychonomic Society, Inc. Reprinted with permission.*

**Odd sites.** Another type of artificial reference point is an *odd site* (Cholewiak & Collins, 2003). An odd site is a vibration that is intentionally different, in terms of a specific dimension such as frequency or intensity. Depending on their design, odd sites may be easier to localize given their different *feel* compared to surrounding sites of vibration.

In their investigation of localizing vibrations on the volar side of the left forearm, Cholewiak and Collins varied frequency in an attempt to create an odd site that would act as an artificial reference point (third experiment in their study). Within an array of 7 tactors, the middle tactor's frequency was varied to create an odd site: 250 Hz while all other tactors were vibrated at 100 Hz. This change in frequency provided a significant increase in localization accuracy at the odd site, but localization performance at neighboring tactors did not see the kind of improvement that might be expected. When frequencies were switched, that is, the middle tactor vibrated at 100 Hz, and all other tactors vibrated at 250 Hz, localization performance at the odd site saw less improvement; Cholewiak and Collins suspected that at 100 Hz, the vibration is much more “quiet”—and hence more difficult to localize—compared to the stronger sensation felt at the odd site when it vibrated with a frequency of 250 Hz. This result, of course, seems to conflict with their first experiment of the same study where no significant difference was found in terms of localization accuracy for vibrations of 100 or 250 Hz. It may be that in the case of the odd site, a vibration of a larger frequency may be more discernible among vibrations of smaller frequencies. It is important to note, however, that as odd sites require distinct vibration signals, overusing odd sites may reduce the desired distinction, causing confusion and reducing localization accuracy; therefore, they should be used sparingly.

*Spacing versus numbers.* Does the *number* of tactors in an array affect localization performance, or is it the *spacing* of tactors? It turns out that a fewer number of tactors may not always provide increases in localization performance. The ability to localize vibrations is a complex function of spacing, reference points and proximity to reference points. To shed some light on this question, the second experiment within Cholewiak, Brill and Schwab's study (2004) is revisited—see also the fourth experiment

of Cholewiak and Collins' (2003) study in which increases in spacing improved localization performance.

In Cholewiak, Brill and Schwab's study, the number of tactors around the entire torso were reduced from twelve to eight, and then to six, while still maintaining equidistant spacing. Upon reducing the number of tactors, and therefore, increasing spacing, localization accuracy significantly improved: first to 92% for eight tactors, and then to 97% for six tactors, with localization of vibrations at the spine and navel still exhibiting superior performance. These results will now be compared to those of the third experiment in this study in which the 12 tactor array was simplified to a 7 tactor semi-circle array while maintaining the same spacing. Although an 8 tactor array was simplified to a 7 tactor array, based on user performance, localization accuracy dropped between these experiments even though the number of tactors decreased. One difference between these tactor arrangements is the spacing; the 8 tactor array has a larger spacing between tactors compared to the 7 tactor array. Localization performance, therefore, depends on tactor separation, among other parameters, rather than the number of tactors in an array. Finally, it's important to question the role of the number of tactors in the following comparison: a significant improvement in localization performance between the 12 tactor array and the 7 tactor array (when centered at the navel or spine) was found, even though spacing remains constant. It turns out that this improvement is not due solely to the reduction in the number of tactors. As Cholewiak, Brill and Schwab suggest, localization performance increased due to the effective use of anatomical reference points, endpoints and proximity to reference points. Therefore, tactor spacing should be guided by not only the desired resolution, but also suitable localization performance. Increasing spacing may improve localization accuracy, but for regions on the body where

space is limited, taking advantage of useful factor placements to improve localization accuracy is critical.

*Vibrotactile sensitivity, age and other factors.* While Cholewiak and Collins (2003) found localization performance to change across skin with consistent vibrotactile sensitivity, in the same study, vibrotactile sensitivity does appear to have some effect on localization performance. In addition to conducting each experiment with students, the results of which were previously described, Cholewiak and Collins ran the same experiments on senior participants. They found that overall, students' localization performance was significantly better compared to that of seniors; this difference, although relatively small, shows that vibrotactile sensitivity, which has higher thresholds for seniors compared to students due to fewer touch receptors, has a small effect on the ability to localize vibrations. In any case, the localization performance of seniors, when compared to the student group, is impressive given their reduced vibrotactile sensitivity. Therefore, when choosing a site on the body to deliver vibrations that need to be accurately localized, the decision should be based less on vibrotactile sensitivity, and more on how the body site will accommodate enough space and reference points to achieve the desired accuracy and resolution. While identifying which body sites have the highest vibrotactile spatial acuity is useful, it is difficult to determine given the varying size and reference points among body parts. The chosen body site will largely depend on the resolution of the vibrotactile display: if more factors are required while maintaining high accuracy, then a larger surface of skin is needed; on the other hand, if less factors are required or localization accuracy isn't of much concern, small skin surfaces may be used. And, of course, criteria of unobtrusiveness and comfort need to be considered when choosing a body site for such stimulation. Lastly, if localization issues continue to arise, surrounds should be utilized to prevent the spread of vibrations; and avoid securing

vibration motors too tightly to the skin as this will cause vibrations to travel through bone structures (Brewster & Brown, 2004).

*Two-dimensional localization.* Our discussion regarding absolute and relative localization of vibrations on the surface of the skin has been limited to one-dimensional, linear arrays of tactors. Given that the skin is a surface, an obvious and common form factor for vibrotactile displays is a two-dimensional arrangement where tactors are placed in rows and columns. Several studies have explored absolute localization of tactors in 2D arrays on both the back and forearm. Studies in which the back has been the site of stimulation are described first.

Lindeman and Yanagida (2003) found an overall localization accuracy of 84% when encased pancake motors were arranged into a 3x3 array on the back of a chair, with a center-to-center motor spacing of 6 cm, and a vibration frequency of 91 Hz. Similarly, Jones and Ray (2008) arranged a 4x4 array of encased pancake motors on the back, with a horizontal spacing of 6 cm and a vertical spacing of 4 cm, and a vibration frequency of 115 Hz; they found an overall localization accuracy of 59% with individual localization accuracies for tactors ranging from 40-82%. Certain tactors were significantly more difficult to localize than other tactors. In contrast to the results of Lindeman and Yanagida, who found the uppermost row of tactors to be more difficult to localize, Jones and Ray found the uppermost row to be the easiest to localize, when compared to other rows, with the corners having high accuracy. Within the 4x4 array, tactors within the middle of the array in the second and third row had the lowest localization performance. However, if near perfect localization accuracy is not required, and localization can be off by at most one tactor, then overall localization accuracy improves to 95% for the 4x4 array. This, of course, will depend on the intended application. Lastly, Jones and Ray found columns to be localized more accurately than rows: 87% to 68%, respectively; they

speculated that this difference in accuracy might be due to the difference in spacing: 6 cm provides a wider spacing between columns, compared to the smaller spacing of 4 cm between rows. Similar to linear tactor arrays, inter-tactor spacing plays a prominent role in how difficult it is to localize vibrations within a 2D array. From a 3x3 to a 4x4 array, the drop in overall accuracy is most likely due to a decrease in inter-tactor spacing caused by an attempt to place more tactors within the same space.

Focusing now on the forearm, Oakley, Kim, Lee and Ryu (2006) arranged pancake motors into a 3x3 array, spaced 2.5 cm apart (center-to-center) on the dorsal side of the left forearm near the wrist; overall localization performance was 46% with individual tactor accuracies ranging from 22% to 76%. Rows of tactors (across the arm) were significantly easier to localize compared to columns of tactors (along the arm). Chen, Santos, Graves, Kim and Tan (2008) evaluated a 3x3 array of tactors spaced 2.5 cm (center-to-center) on the dorsal (first experiment), volar (second experiment) and both dorsal and volar (third experiment) sides of the left forearm near the wrist using a vibration frequency of 150 Hz. In the first and second experiment, accuracies ranged from 25-72% and 34-70%, respectively. Localization performance was slightly better for the volar side of the wrist than the dorsal side, and localization performance for tactors near the wrist were better compared to other tactors. In both experiments, tactor columns were accurately localized more often than tactor rows, which conflicts with previous results; cf. (Oakley, Kim, Lee, & Ryu, 2006).

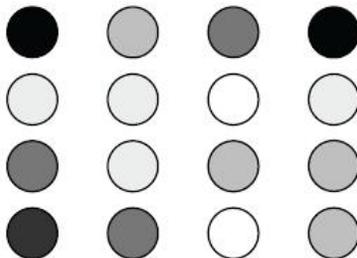
Although both of the aforementioned studies found a similar range of localization accuracies for the dorsal side of the forearm, the conflicting result of what is easier to localize, a row or a column, might be for a number of reasons, such as differences in form factor, differences in location of tactors on the forearm, or differences in the vibration signal itself (frequency and/or amplitude). Furthermore, recall that Jones and Ray (2008)

found columns to be easier to localize compared to rows, but this might have been due to spacing differences, as noted by the authors. The difference in localization accuracies for the forearm, when compared to the back, is, again, most likely due to the smaller inter-tactor spacing given the smaller surface of the forearm compared to the back.

Upon observation, localizing vibrations within a two-dimensional array seems more difficult compared to when using a linear array. Inter-tactor spacing is, of course, important, but revisiting the usefulness of reference points may provide insight into the difference between these accuracies. The usefulness of anatomical reference points for linear arrays was previously discussed, but for two-dimensional arrays, these points may be limited. In Lindeman and Yanagida's study (2003), although a column of tactors in the middle of the 3x3 array fell on the spine, localization was significantly worse for this column compared to the rightmost column. Further, for Jones and Ray's work (2008), tactors in the middle of the 4x4 array were the most difficult to localize, even though they were close to the spine. The problem might be that tactors are sharing anatomical reference points. For example, in Lindeman and Yanagida's experiment, rather than one tactor resting on the spine, three tactors rest on the spine, which increases the number of neighboring tactors for those resting at the anatomical reference point. In Chen et al.'s study, tactors near the wrist had slightly better localization performance compared to other tactors. They speculated that this could be because the wrist is an anatomical reference point, which is valid, but in addition, it could also be that the tactors at the wrist are endpoints. For both Lindeman and Yanagida's study and Jones and Ray's study, tactors on the spine did not benefit from being endpoints in addition to anatomical reference points.

As discussed, endpoints are useful for linear tactor arrays as each endpoint has only one neighboring tactor, making localization easier. In the case of two-dimensional

tactor arrays, factors at the edges could be considered endpoints. However, given that they still have a large number of neighbors, they might not be as useful as those in the linear case. Corner endpoints have three neighbors when considering diagonal factors. Endpoints between corners have five neighbors; compare this to factors falling within the two-dimensional array, which have 8 neighbors. It seems intuitive that less neighboring factors will lessen the difficulty of localizing a vibration: Jones and Ray (2008) found corner factors to have higher accuracies compared to other factors (see figure 22), and factors falling within rows or columns at the edges of the two-dimensional array generally had higher accuracies compared to those in the middle of the array. Lastly, Oakley, Kim, Lee and Ryu (2006) speculated that factor rows (across the arm) were easier to localize than factor columns (along the arm) as the sides of the arm provided an anatomical reference point.



*Figure 22. Conceptual drawing of the 4x4 factor arrangement used by Jones and Ray where the darkness of the grayscale is proportional to the localization accuracy at that factor. Reprinted from “Localization and pattern recognition with tactile displays,” by Jones, L. A., & Ray, K., 2004, in Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 36. Copyright © 2008 by IEEE. Reprinted with permission.*

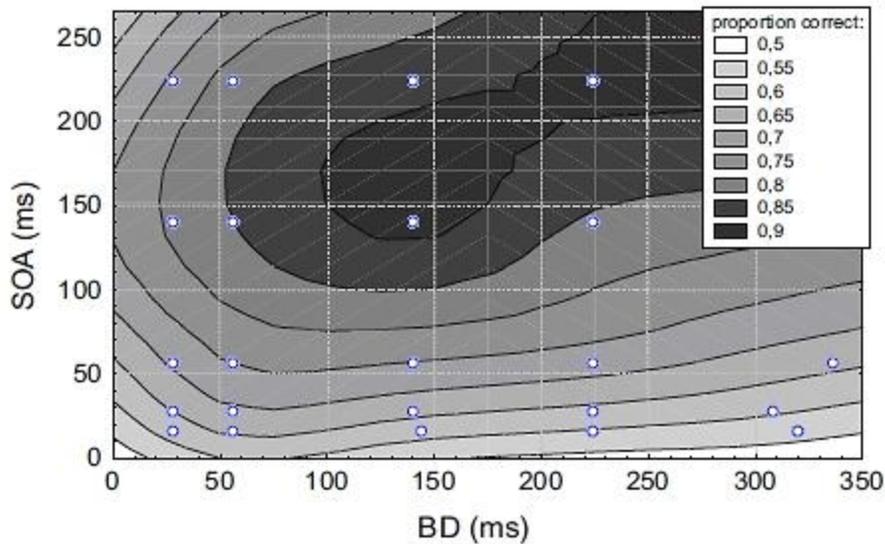
The difficulty of localizing vibrations in a two-dimensional factor array might stem from the inability to exploit reference points as well as in linear factor arrays. Nonetheless, sufficient inter-tactor spacing is critical for accurate absolute localization. Spacing will ultimately be determined by the desired resolution and where the display is

placed on the skin. If near perfect localization accuracy is not required, e.g., if localization may be off by at most one tactor, then more flexibility is possible.

***Relative Localization.*** There has been little work exploring relative localization of vibrations across the body. In a study by van Erp (2005), he explored vibrotactile spatial acuity on the abdomen and back, and its interaction with timing. In the first experiment of the study, three tactor arrangements were explored: a horizontal, linear array of 14 tactors on the back; a horizontal, linear array of 11 tactors on the abdomen; and two vertical, linear arrays of five tactors each on the abdomen—one at the midline above the navel, and another off-center at the left side of the abdomen. All tactors were miniVib-4 tactors, and operated at a frequency of 250 Hz. The procedure for relative localization was as follows: after feeling a brief vibrotactile pulse from one tactor, another tactor was vibrated with a brief pulse after a short pause of some interstimulus interval. After this presentation, participants were asked if the second pulse was to the left or right of the first pulse. This procedure was repeated. No significant difference was found between localization performance on the horizontal and vertical tactor arrays. However, a significant difference in localization performance was found between tactors located at the anatomical reference points (navel and spine), and those that were not, providing a vibrotactile spatial acuity of approximately 1 cm and 2-3 cm, respectively; cf. (Weinstein, 1968). These results show that anatomical reference points, discussed earlier for absolute localization, are also useful for relative localization. The advantages of large versus small spacing for relative localization is obvious, but further experimentation is required to learn if endpoints and odd sites are useful for relative localization.

The interaction between timing and relative localization was explored in the second experiment of van Erp's study. The apparatus and procedure is similar to the previous experiment, but rather than use a linear array of tactors, four pairs of tactors

were placed on the back, and the burst duration (BD) and stimulus onset asynchrony (SOA) were varied as participants had to relatively localize vibrations. Figure 23 shows that relative localization accuracy improves with increases in BD and/or SOA, although for small SOA values, BD has little effect on performance. While timing seems to be important for accurate localization, we may not always have the luxury of excessive pulse durations. In these situations where communication must be fast, van Erp suggests a larger spacing if larger BD and/or SOA values are not practical for an application. And although no experimental results have been gathered, it is obvious that longer burst durations—up to some extent—may help improve absolute localization performance given (1) the time required for motors to reach full intensity; and (2) the time required to direct one’s attention to the site of stimulation in order to localize the stimulation.



*Figure 23. Contour plot showing the interaction between timing parameters—burst duration (BD) and stimulus onset asynchrony (SOA)—and localization performance, where the grayscale variations denote the percentage correct; increases in darkness translate to improved accuracy. Reprinted from “Vibrotactile spatial acuity on the torso: Effects of location and timing parameters,” by van Erp, J. B. F., 2005, in Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 83. Copyright © 2005 by IEEE. Reprinted with permission.*

**Spatio-temporal patterns.** A *spatio-temporal vibration* is a vibrotactile stimulation that varies in terms of both time and space; that is, timing and body site, respectively. A simple example would be a vibration that travels across the skin over time. Given that spatio-temporal vibration patterns utilize both body site and timing/rhythm, it is often easier to create a relatively large set of perceptually distinct vibration patterns, as opposed to when using only a single dimension. Another advantage of spatio-temporal patterns is that they may be used to elicit various vibrotactile perceptual illusions to enhance the intuitiveness of a stimulation and improve recognition accuracy. One illusion is of particular interest here: *saltation*. This perceptual illusion is described in Chapter 6 where it is used in vibrotactile motor instructions to elicit apparent motion for intuitive movement cues. Spatio-temporal patterns have been used in a variety of applications including navigation (Jones, Lockyer, & Piatetski, 2006); military (Jones, Kunkel, & Torres, 2007) (Jones, Kunkel, & Piatetski, 2009); motor learning (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011) (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011); and assistive technology for individuals who are blind or visually impaired (Krishna, Bala, McDaniel, McGuire, & Panchanathan, 2010). Although studies have explored spatio-temporal pattern recognition across many different body parts from the hand (Krishna, Bala, McDaniel, McGuire, & Panchanathan, 2010) (Krishna, Bala, & Panchanathan, 2010), to the foot (Magana & Velazquez, 2008), and across the whole body (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) for specific applications, the focus here will be more commonly used bodily sites for vibrotactile communication, namely the forearm, torso and waist, and more general spatio-temporal patterns applicable to a variety of applications. In particular, Jones and her colleagues have conducted numerous

studies exploring vibrotactile pattern perception on these body parts; it is these studies that are discussed here to gain insight into how to best design spatio-temporal patterns.

Piateski and Jones (2005) mounted a 3x3 array of tactors, with a center-to-center spacing of 24 mm, on the volar side of the forearm. Two types of motors, pancake and cylindrical, were evaluated at 115 Hz and 180 Hz, respectively. Directional cues, intended for use in navigation applications, were designed, see figure 24, and presented through the tactile display. Overall recognition accuracy using cylindrical motors (93.5%) was significantly higher compared to pancake motors (85%). Pattern H was found to be the most distinct, and patterns that travelled the width of the forearm (C, D and E) as opposed to its length (A, B, and F), were easier to recognize. Piateski and Jones speculated that the improved recognition accuracy for the former patterns might be due to our utilization of the sides of the forearm as anatomical reference points; indeed, these results compare well to those found in Oakley, Kim, Lee and Ryu's study (2006) which explored two-dimensional localization on the wrist.

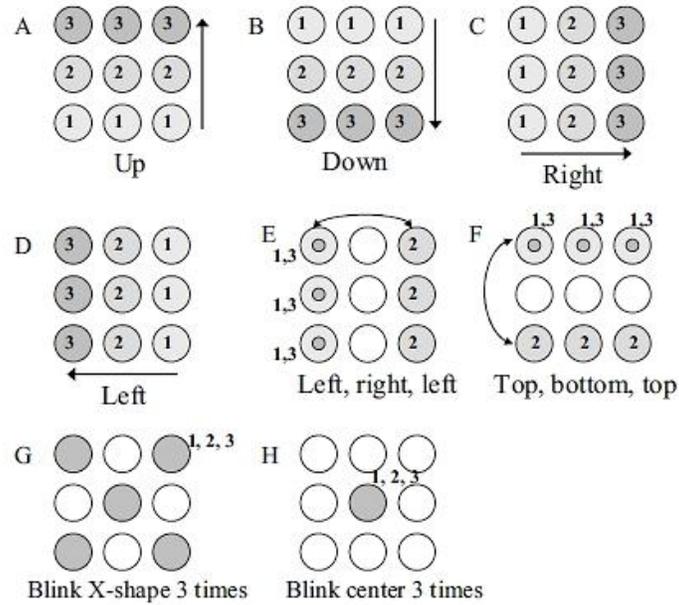


Figure 24. Spatio-temporal patterns, A through H, for Piatetski and Jones' pattern recognition experiment using the volar side of the forearm. Each circle represents a single motor within a 3x3 factor array. The arrows, numbers, grayscale variations represent activation order. For each pattern, each pulse had a burst duration of 500 ms, and an interstimulus interval of 500 ms. *Reprinted from "Vibrotactile pattern recognition on the arm and torso," by Piatetski, E., & Jones, L., 2005, in Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 92. Copyright © 2005 by IEEE. Reprinted with permission.*

In a follow-up study using pancake motors, Jones, Kunkel and Piatetski (2009) introduced several new patterns for the forearm: figure 25 depicts two sets of patterns, each consisting of eight patterns, for experiments 1A and 1B, respectively. The overall recognition for each of these sets was 62% and 85%, respectively, with the only difference between experiments being two patterns (each experiment shared six of the same patterns). A small difference in stimulus set resulted in a large difference between overall recognition accuracy as the two diagonal directions, i.e., patterns E and F, caused confusion among many of the patterns of experiment 1A given their similarities; specifically, E was often confused with C, F was often confused with D, and A was often confused with F. Patterns C, D and H had the highest recognition accuracies across the

experiments. Notice that these patterns, once again, travel across the width of the forearm.

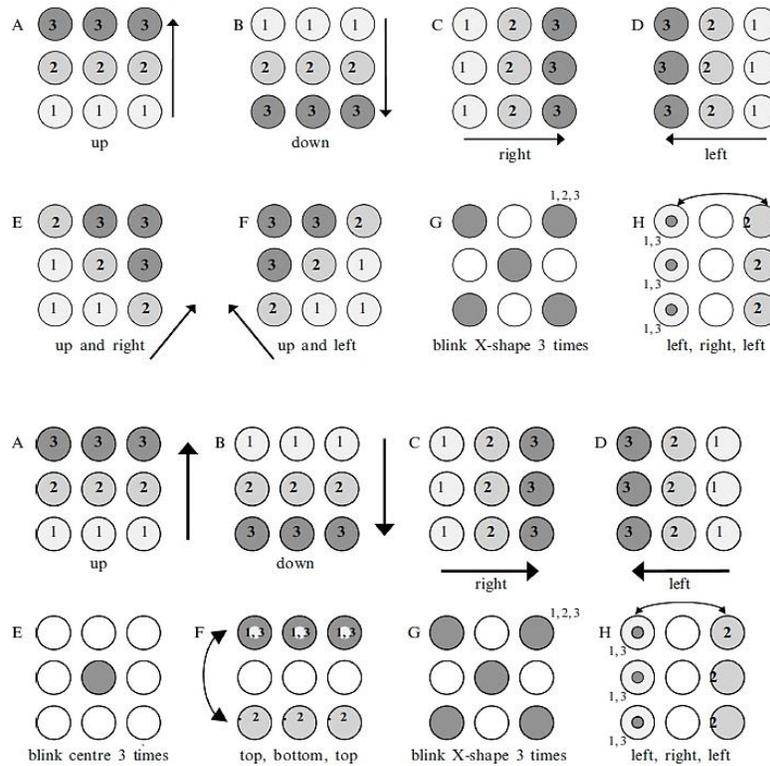


Figure 25. Two sets of spatio-temporal patterns for Jones, Kunkel and Piatetski’s pattern recognition experiment using the volar side of the forearm, where the top set, A through H, was used for Experiment 1A, and the bottom set, of the same lettering, was used for Experiment 1B. Each circle represents a single motor within a 3x3 factor array. The arrows, numbers, grayscale variations represent activation order. For each pattern, each pulse had a burst duration of 500 ms, and an interstimulus interval of 500 ms. Reprinted from “Vibrotactile pattern recognition on the arm and back,” by Jones, L. A., Kunkel, J., & Piatetski, E., 2009, *Perception*, p. 56. Copyright © 2009 by Pion Ltd, London, <http://www.enplan.com>. Reprinted with permission from publisher and first author.

Piatetski and Jones (2005) mounted a 4x4 array of pancake motors on the back with a vertical spacing of 40 mm, and a horizontal spacing of 60 mm. The patterns, depicted in figure 26, provided impressive recognition accuracy with being recognized nearly 100% of the time. In a follow-up study (Jones & Ray, 2008), the number of patterns was extended to twelve; see figure 27. The overall recognition accuracy was

95%, which is impressive given the number of patterns used. Patterns G and K were often confused with patterns L and E, respectively. Jones and Ray speculated that spatial overlap might have created the confusion. Lastly, spatio-temporal patterns have been explored for the waist. Using a linear waist-based tactile display, Jones and Ray (2008) found an impressive recognition accuracy of 99% using six patterns; see figure 28.

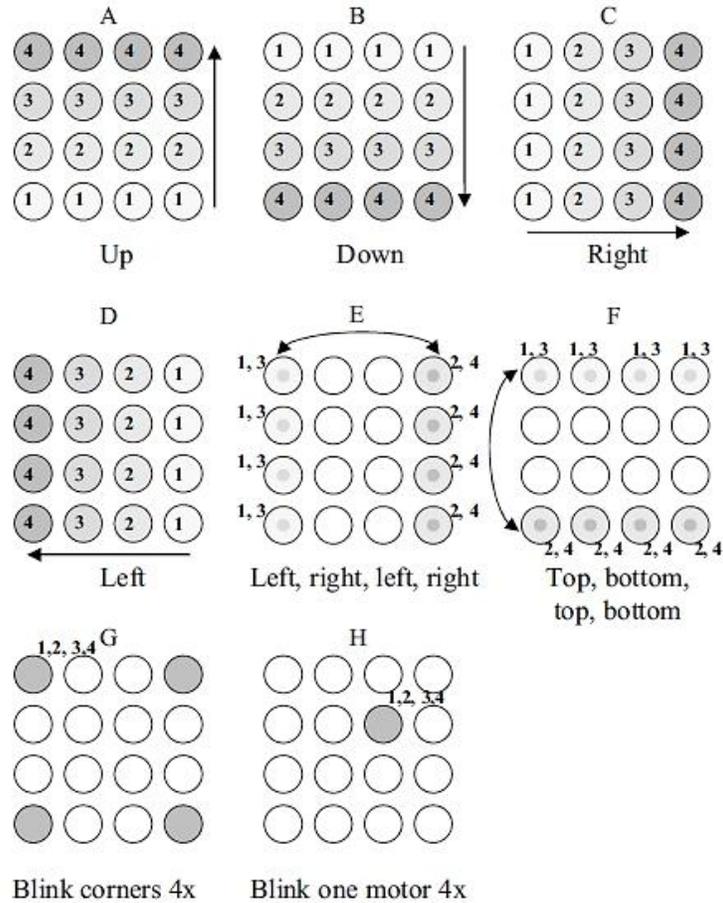


Figure 26. Spatio-temporal patterns, A through H, for Piatetski and Jones' pattern recognition experiment using the torso (back). Each circle represents a single motor within a 4x4 tactor array. The arrows, numbers, grayscale variations represent activation order. For each pattern, each pulse had a burst duration of 500 ms, and an interstimulus interval of 500 ms. *Reprinted from "Vibrotactile pattern recognition on the arm and torso," by Piatetski, E., & Jones, L., 2005, in Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 94. Copyright © 2005 by IEEE. Reprinted with permission.*

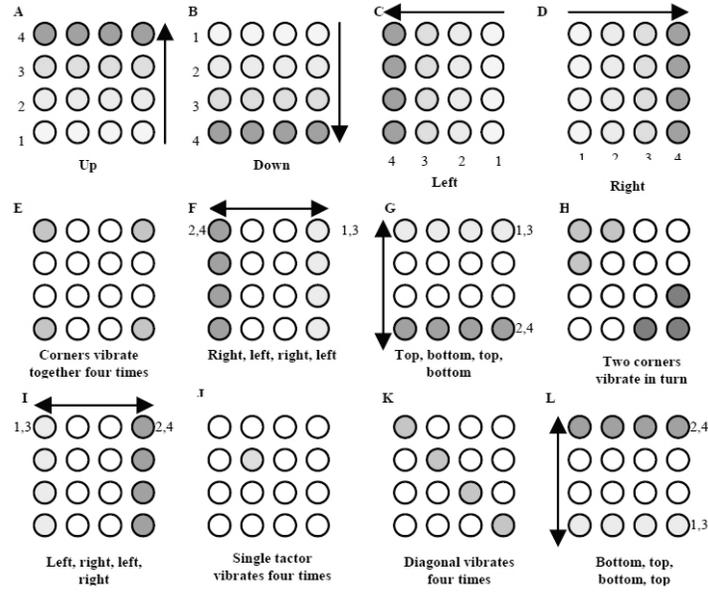


Figure 27. Additional patterns for recognition, similar in timing and activation to those of figure 26; namely F, H, K and L. Reprinted from “Localization and pattern recognition with tactile displays,” by Jones, L. A., & Ray, K., 2004, in *Proceedings of the Symposium on Haptic Interfaces for Haptic interfaces for Virtual Environment and Teleoperator Systems*, p. 37. Copyright © 2008 by IEEE. Reprinted with permission.

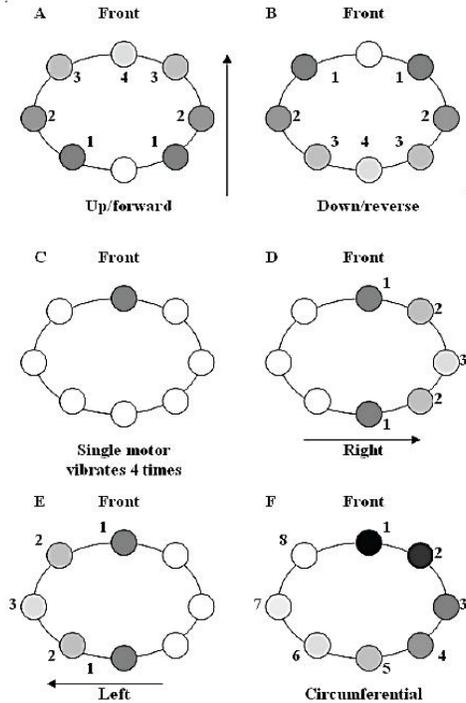


Figure 28. Spatio-temporal patterns, A through F, for Jones and Ray’s pattern recognition experiment using a belt consisting of eight, equidistantly spaced pancake motors, depicted as circles, with spacing between 80-100 mm. Reprinted from “Localization and pattern recognition with tactile displays,” by Jones, L. A., & Ray, K., 2004, in *Proceedings of the Symposium on Haptic Interfaces for Haptic interfaces for Virtual Environment and Teleoperator Systems*, p. 37. Copyright © 2008 by IEEE. Reprinted with permission.

From the aforementioned studies, it is clear that a key factor influencing spatio-temporal pattern recognition performance is the distinctness between patterns. If too much spatial overlap is present, as in patterns E & C, F & D, and A & F, in figure 25 (top set), then confusion arises; see also patterns G & L in figure 27. However, certain patterns exhibiting spatial overlap have shown to work very well, such as simple directional cues; see, for example, patterns A through D in figure 26. Overall, a distinct set of simple patterns seems to be the best choice. To achieve distinctness, the display area need not be limited to the same body part and/or display; in fact, the entire body may be used, as in Spelmezan, Jacobs, Hilgers and Borchers’ study (2009) in which vibration patterns were used to cue different movements for the purpose of motor learning. Posture

must also be taken into account (Jones, Kunkel, & Piatetski, 2009). It may be the case the vibration patterns will be delivered while the stimulated body part is in different postures. Careful attention must be paid to how the body part will change orientation during application use, and if the patterns will be invariant to these pose changes. Another key factor, described next, is naturalness.

The tactile torso-based display developed by Jones and her colleagues has also been used in a military inspired application where a variety of vibrotactile patterns, depicted in figure 29, convey different hand signals (Jones, Kunkel, & Piatetski, 2009). Overall pattern recognition accuracies of 91%, 91% and 93% were found while participants performed different tasks, namely walking, jogging and a cognitive task. Jones, Kunkel and Piatetski found that with minimal training that does not focus on teaching the mapping between stimulation and what the stimulation refers to, in this case, hand signals, performance will significantly drop when a visual reference guide, showing activation patterns, is taken away (specifically, from 98% to 75%). Proper training is therefore critical for users to sufficiently learn the mapping between stimulations and their meaning. To ease this process, patterns should intuitively and naturally represent their assigned meanings, such as directional cues conveying navigation directions. If patterns are not carefully designed or arbitrary meanings are assigned, this may only lengthen training time and increase cognitive load.

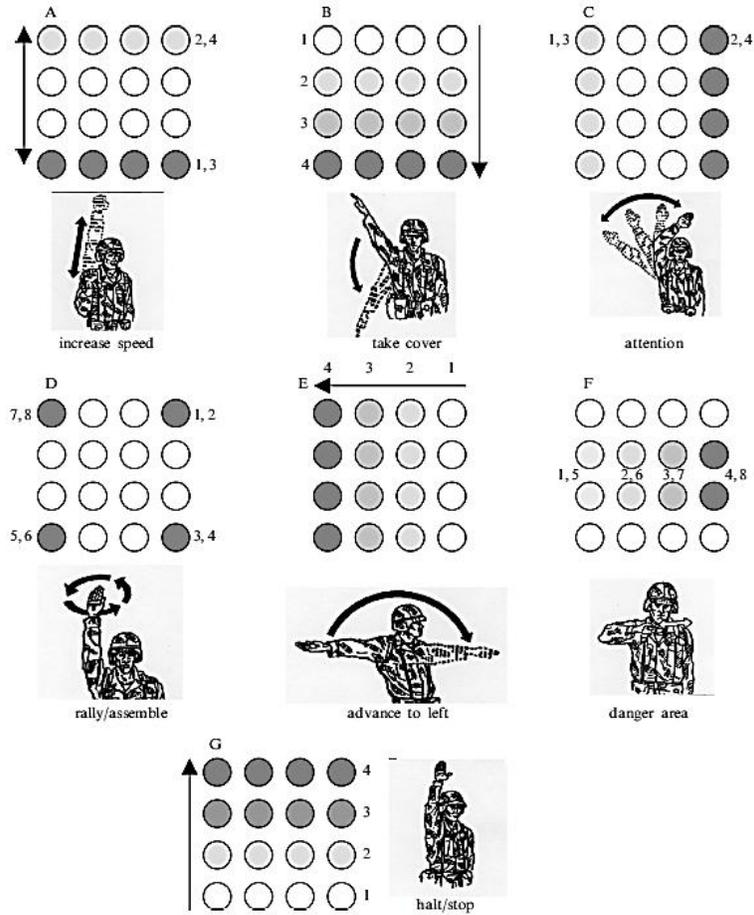


Figure 29. Intuitive spatio-temporal patterns, A through G, for Jones, Kunkel and Piatieski's pattern recognition experiment using a torso-based tactile display for military applications. Reprinted from "Vibrotactile pattern recognition on the arm and back," by Jones, L. A., Kunkel, J., & Piatieski, E., 2009, *Perception*, p. 60. Copyright © 2009 by Pion Ltd, London, <http://www.envplan.com>. Reprinted with permission from publisher and first author.

**RELATED WORK**

A variety of approaches have been proposed toward using the sense of touch as a channel for information delivery and communication. These approaches may be categorized in several ways including the type of interaction: *human-to-human*, *human-computer* or *mediated interpersonal interaction* in which two or more people interact indirectly through a computerized device. Other categorizations may be made in terms of specific design parameters including the *level of abstraction* or *signal parameter association*.

The level of abstraction of the information to be conveyed may range from *literal* (low level) to *symbolic* (high level). In a *literal translation*, information is directly presented to the sense of touch from a different modality—typically vision or hearing. To accommodate the new modality, cross-modal transformations are applied wherein the original message remains largely unchanged. This includes computerized systems for *sensory substitution* such as tactile-vision and tactile-audio assistive aids as well as human-to-human interaction approaches including tadoma and tactile sign language. Theoretically, no learning should be required, but given inter-modal differences including sensory and perceptual differences, it often takes significant training and practice to become acclimated with the new sensory input. Toward the other end of the spectrum, a *symbolic* mapping provides a high level of abstraction in that it represents information in a conceptualized, often metaphorical form. Here, the user does not have access to a high resolution channel as with literal translations; instead, the computer communicates a high-level representation of the data in the form of discrete messages (or patterns, cues, etc.) whose associations must first be learned.

*Signal parameter associations* range from *arbitrary* to *intuitive* mappings of meanings to the different dimensions of a stimulation (i.e., a meaning is assigned

arbitrarily or intuitively to each dimension of, e.g., a vibration signal, which might include frequency, amplitude, duration, etc.), where intuitive, or *natural*, associations are subjective but decided based on heuristics and empirical results including participant feedback. With an arbitrary mapping, there is no correlation between the stimulation and its associated meaning. Such a design strategy may present an extraordinarily high learning curve when large sets of stimuli are used (Geldard F. A., 1957); but even small sets of stimuli that have arbitrary signal parameter associations (Enriquez, MacLean, & Chita, 2006) tend to have much greater learning curves compared to intuitive associations (Chan, MacLean, & McGrenere, 2005). Intuitive associations provide a clear relationship between stimulation and its meaning, supporting faster learning given that stimuli naturally relate to their intended meaning. Ideally, stimulations should naturally elicit their intended meaning, i.e., without additional training. In general, however, outside of simple sensory substitutions and therapeutic mediated haptic interpersonal communication technologies that simulate touching (DiSalvo, Gemperle, Forlizzi, & Montgomery, 2003) (Bonanni, Vaucelle, Lieberman, & Zuckerman, 2006), *literal stimulations* that elicit their intended meaning are difficult to provide. As an example of this challenge, consider Spelmezan, Jacobs, Hilgers and Borchers' investigation (2009) of vibrotactile stimulation for eliciting motor movements for snowboarding. They found that participants' natural interpretation of vibrotactile stimulations without prior training varied considerably and were often vague; therefore, a consensus was difficult to obtain for most patterns, most likely due to inter-subject sensory, perceptual and experiential differences.

In the context of a theoretical design framework for somatic information delivery, both a high level of abstraction and intuitive signal parameter association are critical to ensure a reasonable amount of training and practice for mastery; and although

eventual mastery may be possible with extensive training and practice regardless of how practical or intuitive a language design might be, there is a tradeoff between *learning effort* and *perceived value of learning the language*. If a high learning curve isn't worth the user's time nor energy, it may not be practical for the user to pursue, and therefore, he or she may lose motivation. A high level of abstraction and intuitive signal parameter association are not the only attributes of a theoretical framework for designing a *functional* (expressive, configurable and expandable) and *practical* (easy to learn and easy to use) somatic communication language; such a framework must be *versatile* and support the design of *expandable*, *efficient*, *rich* and *robust* languages, defined as follows:

- **Versatile:** Framework can be used to design languages relevant and applicable to diverse application domains.
- **Expandable:** It is simple and straightforward to add novel *communication units* as well as combine existing units.
- **Efficient:** Fast communication speeds are possible to ensure usefulness within a variety of application domains including those in which high level concepts must be conveyed in real time.
- **Rich:** Expressive communication possibilities, even from a small set of communication units, through the use of *context* (environmental settings including *body site*) and *stimulation variations* (similar to *tonal variations* found in natural language).
- **Robust:** Stimulations are hard to miss and/or redundant. This includes the use of attention-grabbing cues; redundant spatial, temporal or spatio-temporal signals (without repeating the entire signal); and/or adapting the signal to the conditions of the environment to ensure successful delivery.

For the remainder of this section, approaches are categorized in terms of the *type of information* delivered to the skin: *literal translations* (e.g., *sensory substitution*), *alphanumeric information* (letters, numbers and common words) and *conceptual information* (metaphorical representations). Note that overlap may be present between categories. Table 1 presents a summary of proposed approaches, including Somatic ABC's, with those attributes that have been met, as indicated by a checkmark, and those which have not been met, as indicated by a blank entry. In the following sections, approaches are described and related to table 1. For presentation purposes, literal translation approaches are summarized under the heading *Literal Translations* in table 1; this is similarly done for alphanumeric approaches. As shown by the table, the proposed approach, Somatic ABC's, meets all of the aforementioned attributes, which is later verified by both design considerations and experimental results. The following three sections cover each type of information delivery, and provide justification for the criteria given in table 1.

Table 1

*Summarization of Proposed Approaches in terms of Desired Design and Performance Criteria for Achieving Functional and Practical Somatic Languages*

	High Level of Abstraction	Intuitive Mapping	Versatile	Expandable	Efficient	Rich	Robust
<b>Literal Translation<sup>1</sup></b>	-	✓	-	✓	✓	✓	✓
<b>Alphanumeric<sup>2</sup></b>	✓	-	✓	✓	-	✓	✓
<b>Mediated Channel<sup>3</sup></b>	✓	✓	-	✓	✓	-	-
<b>Tactile Icons<sup>4</sup></b>	✓	-	✓	✓	✓	-	-
<b>Haptic Glyphs<sup>5</sup></b>	✓	✓	-	✓	✓	-	✓
<b>Haptic Icons<sup>6</sup></b>	✓	✓	✓	✓	✓	-	-
<b>Somatic ABC's<sup>7</sup></b>	✓	✓	✓	✓	✓	✓	✓

*Note.* A checkmark indicates that the proposed approach has met the respective criterion; an entry that has been left blank indicates that the proposed approach has not met the respective criterion. Proposed approaches include literal translation approaches<sup>1</sup>: tadoma, tactile sign language, haptics (Lahtinen, 2008), Optacon (Linvill & Bliss, 1966), Optohapt (Geldard F. A., 1966), TVSS (Bach-y-Rita, 1972) and tactile-audio substitution systems; alphanumeric approaches<sup>2</sup>: Braille, Morse code and Vibratase (Geldard F. A., 1957); mediated haptic interpersonal communication technology<sup>3</sup>: HandJive (Fogg, Cutler, Arnold, & Eisbach, 1998), InTouch (Brave & Dahley, 1997), ComTouch (Chang, O'Modhrain, Jacob, Gunther, & Ishii, 2002), Shake2Talk (Brown & Williamson, 2007), HIM (Rovers & van Essen, 2004) and Contact IM (Oakley & O'Modhrain, 2002); tactile icons<sup>4</sup> (Brewster & Brown, 2004) (Brown, Brewster, & Purchase, 2005) (Brown, Brewster, & Purchase, 2006a); haptic glyphs<sup>5</sup> (Roberts & Franklin, 2005) (Borst & Baiyya, 2007); haptic icons<sup>6</sup> (Enriquez & MacLean, 2003) (MacLean & Enriques, 2003) (Chan, MacLean, & McGrenere, 2005) (Enriquez, MacLean, & Chita, 2006); and Somatic ABC's<sup>7</sup>.

### **Literal Translations via Touch**

*Literal translation* is the direct presentation of sensory input to an alternative modality (here, the focus is tactile-vision and tactile-audio) after application of a cross-modal transformation in which the content of the original message is largely unchanged. Literal translations are found in both human-to-human communication such as tadoma, tactile sign language and tactile fingerspelling; and computerized systems for sensory substitution, all of which are described below. In human-to-human interactions, cross-modal transformations to convert from one modality to another, involve literally feeling the stimuli originally intended for (a) sight, as in the case of tactile sign language; or (b) both sight and hearing, as in the case of tadoma; whereas sensory substitution systems apply an algorithmic transformation (e.g., discretization, resolution reduction, bandpass filtering, etc.) to the original input data.

Although these approaches provide many benefits including intuitive associations and rich communication possibilities that are robust (redundant and interactive), expandable and efficient (table 1), their inherent literal conversions between modalities using cross-modal transformations create challenges; in particular, their low level of abstraction usually produces high learning curves given the introduction of a new sensory input to a potentially suboptimal modality. Significant training and practice are often needed to overcome perceptual limitations created by mismatched spatial and/or temporal acuity between modalities. Moreover, the versatility of these methods is limited in that they are not practical for a wide array of applications: tadoma, tactile sign language, tactile fingerspelling, and social-haptic communication require use of the hands; and sensory substitution systems tend to work best in controlled conditions that lack environmental noise and interference. The following sections describe the

aforementioned approaches for human-to-human communication and sensory substitution, respectively.

**Human-to-human literal translation.** For individuals with severe visual and auditory impairments, touch offers a useful communication channel over the remaining basic senses of taste and smell. For individuals with hearing impairments who have learned sign language, if vision begins to deteriorate, tactile sign language, tactile fingerspelling and social-haptic communication are obvious extensions (described later). Another technique, less in use today, is Tadoma—a method for tactile speechreading (reading of the lips and other features of speech production including throat vibrations and mouth/nasal airflow). Ultimately, the communication method chosen will depend on many factors related to an individual's condition including preference, education and age of onset (Reed, Durlach, & Delhorne, 1992). The communication rate (words/s) of tactile sign language (ASL) is higher compared to Tadoma, which both have a higher rate than tactile fingerspelling (Jones & Lederman, 2006).

**Tadoma.** Tadoma is a form of human-to-human communication in which a Tadoma user feels actions of the speech production process by placing a hand on the face of a talker. The precise position of the hand in contact with the face varies between users, but roughly the thumb is placed across the lips, the middle three fingers are placed along the jaw, and the little finger is placed on the throat (figure 30). As the partner speaks, tactile and kinesthetic sensations of lip and mouth motions, throat vibrations, and airflow are indicative of articulations (Jones & Lederman, 2006). Tadoma was first used in the United States in the 1920s by American school teacher Sophie Alcorn to teach students who were deaf-blind. Since its introduction in the United States, until the 1960s, Tadoma was extensively used for the education of individuals who were deaf-blind for speech reading and production; but after this period, its use has steadily declined to where only a

handful of Tadoma users are in the United States (Reed, Durlach, & Delhorne, 1992). Tadoma is named after its first students: Winthrop ‘Tad’ Chapman and Oma Simpson.

Tadoma has several limitations. Given its low level of abstraction, Tadoma requires considerable training and practice to sense and perceive features of speech production through touch. In particular, training typically occurs through an extensive education program over many years. Less accessible features of speech production, such as tongue position, can create interpretation problems, which has prompted researchers to explore kinesthetic and tactile displays, in the form of mechanical skull models, to emulate the speech production actions of a speaker (Reed, Rabinowitz, Durlach, & Braida, 1985). Lastly, Tadoma requires close physical contact, and is limited to human-to-human interaction, which reduces versatility.

However, Tadoma allows individuals who are deaf-blind to experience rich communication through haptic perception of speech, showcasing the potential of touch as a communication channel. While not as efficient as listening to speech, it still provides communication speeds of roughly half that of the normal conversational speaking rate (Reed, Durlach, & Delhorne, 1992).



*Figure 30. Visual demonstration of Tadoma method for haptic human-to-human communication by individuals who are deaf-blind. Reprinted from "Natural methods of tactual communication," by Reed, C. M., Durlach, N. I., & Delhorne, L. A., 1992, In I. R. Summers (Ed.), Tactile Aids for the Hearing Impaired (p. 220), London, England: Whurr Publishers Ltd. Copyright © 1992 by John Wiley and Sons. Reprinted with Permission.*

***Tactile sign language and fingerspelling.*** Users of tactile sign language typically adapt their fluency in sign language, learned early in life with the onset of deafness, to haptic reception during the onset of blindness (Reed, Durlach, & Delhorne, 1992). In tactile sign language, the receiver's hand(s) is placed in contact with the sender's hand(s) as signs are produced (figure 31). Through the many degrees of freedom of the hands, rich and expressive signs have been designed through use of handshape, location, orientation and movement (Reed, Durlach, & Delhorne, 1992). Many sign languages exist, such as American Sign Language (ASL) and Pidgin Sign English (PSE); the language taught to an individual who is deaf-blind will depend on his or her onset of impairments, education and environment (Reed, Durlach, & Delhorne, 1992).

In terms of perceptibility, isolated signs are more easily recognized compared to signs within sentences (Reed, Delhorne, Durlach, & Fischer, 1995); Reed et al.

speculated that this might be due to isolated signs carrying more meaning in addition to more processing time as handshapes are held for longer durations when isolated. Tactile sign language is slower and less accurate compared to the visual perception of sign language: 1.5 signs/s compared to 2.5 signs/s, respectively (Reed, Delhorne, Durlach, & Fischer, 1995). And although it requires considerable training and practice, it's a relatively fast communication method compared to Tadoma and fingerspelling, even approaching communication rates of spoken language, making it a useful form of haptic human-to-human communication. It is therefore more widely taught than Tadoma and fingerspelling.



*Figure 31. Visual demonstration of tactile sign language for haptic human-to-human communication by individuals who are deaf-blind. Reprinted from "Natural methods of tactual communication," by Reed, C. M., Durlach, N. I., & Delhorne, L. A., 1992, In I. R. Summers (Ed.), Tactile Aids for the Hearing Impaired (p. 225), London, England: Whurr Publishers Ltd. Copyright © 1992 by John Wiley and Sons. Reprinted with Permission.*

In tactile fingerspelling, the receiver places his or her hand in contact with the sender's hand, who produces static or dynamic handshapes (figure 32), each of which are uniquely associated with a letter, and produced sequentially to build a word. The

handshapes depend on the chosen manual alphabet—in the United States, the American One-Handed Manual Alphabet (AOHMA) is the most commonly used manual alphabet (Reed, Durlach, & Delhorne, 1992).

Accurate perception of handshapes, either tactually or visually (the latter in the case of individuals who are deaf), is possible by experienced users at natural, manual production rates of 2-6 letters/s (Reed, Delhorne, Durlach, & Fischer, 1990). Fingerspelling is considerably slower compared to normal speaking rates at about a quarter the speed, but has higher communication rates compared to other alphanumeric approaches such as Morse code and Vibratese (Reed, Delhorne, Durlach, & Fischer, 1990).



*Figure 32. Visual demonstration of tactile fingerspelling for haptic human-to-human communication by individuals who are deaf-blind. Reprinted from “Natural methods of tactual communication,” by Reed, C. M., Durlach, N. I., & Delhorne, L. A., 1992, In I. R. Summers (Ed.), Tactile Aids for the Hearing Impaired (p. 223), London, England: Whurr Publishers Ltd. Copyright © 1992 by John Wiley and Sons. Reprinted with Permission.*

***Social-haptic communication.*** For over 25 years (1980—2007), Riitta M. Lahtinen, with help from Russ C. Palmer and colleagues, developed, expanded, and evaluated a theoretical framework for social-haptic communication (Lahtinen, 2008) to facilitate human-to-human communication for individuals with severe visual and hearing impairments. The approach was developed and evaluated around the communication between an individual who was deaf-blind, and an individual with normal vision and hearing. Although the basis of theory is haptic communication, it may readily be combined with spoken/written language and/or sign language—ultimately, the forms of communication used will depend on a user’s preference and degree of visual and hearing impairment.

The basis of the framework is a *haptice*—a social-haptic message conveyed through touch. Each haptice is composed of *haptemes*—building blocks, or lower order dimensions, used to construct a haptice. Examples of haptices include confirmation (yes/no); social quick messages (identifying oneself, turn-taking, feedback, etc.); guidance and orientation (directions, pointing, etc.); drawing/signing on the body (block characters, layout of a room, etc.); emotional expressions; expressions of art, music and games; among many other social-haptic messages. The building blocks (haptemes) of haptices include duration, intensity, repetitions, rhythm, movements, direction of movements, body site, orientation between sender and receiver, distance between sender and receiver, social body space and the modalities involved.

Haptices may be considered the “words” of a social-haptic language, which may subsequently be combined to create rich sentences. Other features of natural, spoken language, such as context and intonation, have been explored within Lahtinen’s social-haptic communication framework. Similar to spoken language, intonation may be used to change the meaning of haptices. Possible intonations include pressure variations,

direction variations, speed variations, frequency variations, and duration variations, among others. Somatic ABC's also borrows intonation from spoken language, and it is utilized for both changing the meaning of words and to ensure the successful delivery of words.

Context also applies to social-haptic communication in that it influences the meaning of haptices based on the social situation in which it is delivered. Body site, i.e., where the stimulation is delivered to the skin, is dependent on the social setting, relative location between the sender and receiver, and the message being delivered. In contrast, Somatic ABC's utilizes body site as a channel to change the meaning of words for enriching vocabularies while limiting training time.

Empirical results thus far have shown haptices within social-haptic communication to be an effective, rich method for haptic human-to-human communication for individuals who are deaf-blind. Depending on the severity of a user's impairment, it may be more efficient than tactile fingerspelling and tactile signing. It can easily be expanded to accommodate more messages, which naturally occurs over time as users become familiar with the communication paradigm, and the need for more haptices arises. It eventually allows for discreet communication as movements become smaller over time as the sensory and perceptual capabilities of touch become accustomed to the new input. However, given its low level of abstraction, it does require extensive training and practice to build large vocabularies. And like the aforementioned human-to-human literal translation approaches, its versatility is limited to human-to-human social-haptic communication.

**Computerized systems for sensory substitution.** Sensory substitution systems may be divided between those for tactile-vision, tactile-audio, tactile-tactile and tactile-vestibular. Tactile-tactile devices are applicable to remote touch applications, or assistive aids for individuals with tactile sensory impairments in which tactile information, detected through pressure sensors at the affected site (e.g., the hand), is presented to an unaffected body site (Kaczmarek, Webster, Bach-y-Rita, & Tompkins, 1991). Tactile-vestibular systems assist those with impaired balance by providing orientation information, detected through motion sensors, and presented via vibrotactile stimulation around the waist (Wall III & Weinberg, 2003) or electrotactile stimulation on the tongue (Bach-y-Rita & Kercel, 2003).

Given the extensive research efforts toward tactile-vision and tactile-audio devices, focus will be given to these two sensory substitution areas. The cross-modal transformations involved in tactile-vision and tactile-audio are pictorial-to-tactile and frequency-to-location translation, respectively. Tactile-vision systems are first presented, followed by tactile-audio devices.

***Tactile-vision translation.*** Tactile-vision sensory aids may be divided between devices aimed at converting printed material to touch, described first; and those for converting general visual images to touch, such as the Tactile-Vision Substitution System (TVSS).

***Tactile perception of print.*** Improving the accessibility of printed material for individuals who are blind is a problem that has been investigated for decades. Today, pages are captured via a mounted camera, such as with the iCare Reader (Hedgpeth, Black, & Panchanathan, 2006); or with a point-and-shoot style camera, such as with the Intel Reader<sup>3</sup>, which also has a mounted camera option. After visual capture, optical

---

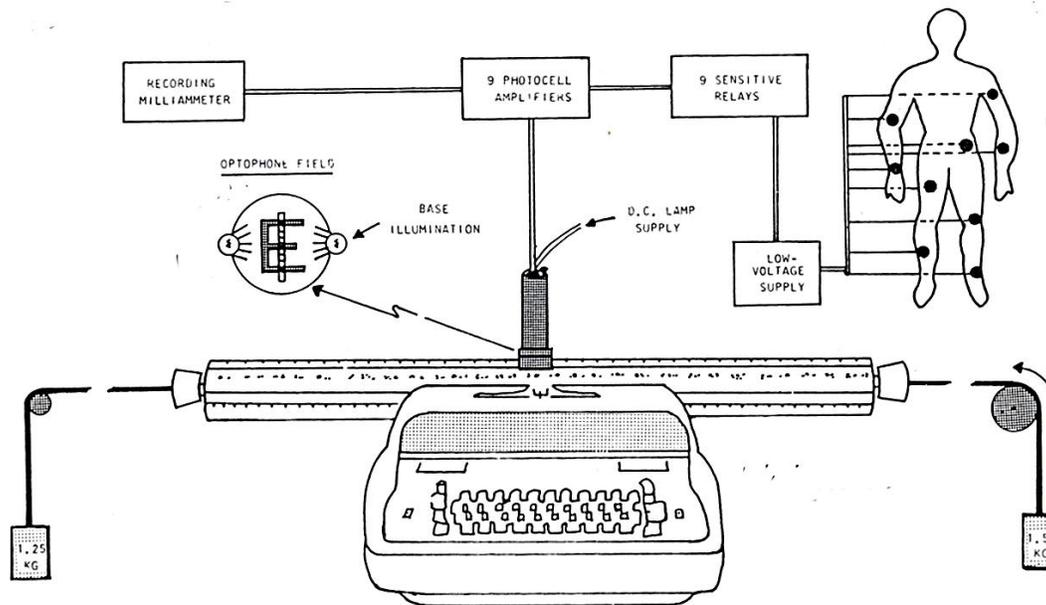
<sup>3</sup> <http://www.careinnovations.com/Products/Reader/Default.aspx>

character recognition (OCR) is applied to the text of the page, and results are outputted through audio. Before these solutions, however, tactile-vision sensory substitution was explored where the visual image of a character was directly presented to the skin of a user after a visual-to-tactile transformation. Two such devices, each utilizing different presentation approaches, are described next.

The Optacon (OPTical-to-TActile CONversion), proposed by Linvill and Bliss (1966), is a handheld optomechanical device in which the visual image of a printed symbol (letter, number, etc.), captured using an 8x12 array of photosensitive cells, drives an 8x12 array of vibrating pins. (These dimensions were eventually increased to 6x24 to accommodate more lenient camera placement for visual capture of characters.) Each photosensitive cell is coupled to a vibrating pin in which the white of a page or black of a character turns the corresponding motor off or on, respectively, as the sensor is moved across the text of a page. Through a controlled computer simulation in which text was scrolled across the tactile display of the Optacon, Linvill and Bliss found reading rates of 20 words per minute after 17 hours of training, and 30 words per minute after 50 hours of training. In another study through a training program by the manufacturer, Telesensory Systems Inc., who manufactured the device for over two decades since 1971, 10-12 words per minute was achieved by participants after nine days, eventually reaching 30-50 words per minute with continued training (Craig & Sherrick, 1982). Although reading rates are considerable slower than visual reading rates and intensive training is required due to Optacon's low-level, literal translation, the target user population found it very useful for reading printed text.

Another computerized approach for reading printed material through touch, developed around the same time as the initial version of the Optacon, is the Optohapt (Geldard F. A., 1966). The setup of this system is depicted in figure 33. Printed text

scrolls across a vertical array of photosensitive cells through use of a typewriter augmented with a motor and weights to ensure slow and smooth text scrolls across the sensors. Each photosensitive cell is coupled to a vibration motor, each attached to a different body site: one on the abdomen, and two on each arm and leg. Vibration motors are driven much the same way as with the Optacon, thereby creating unique spatially and temporally varying vibration patterns as characters are moved across the vertical sensors.



*Figure 33. System setup of the Optohapt. Nine vibration motors, spread across the body, are driven by a vertical array of photocells which detect passing characters. Optohapt enables users to literally feel a character through unique spatio-temporal vibration patterns. Reprinted from "Cutaneous coding of optical signals: The optohapt," by Geldard, F. A., 1966, Perception & Psychophysics, 1, p. 379. Copyright © 1966 by Psychonomic Society, Inc. Reprinted with permission.*

Given the spatial and/or temporal similarities between characters when translated to touch, punctuation marks were much easier to discern compared to alphanumeric characters. Therefore, to ease recognition, coding was found necessary in which alphanumeric characters were associated with unique symbols to achieve perceptual separation in terms of spatial and temporal characteristics. This latter approach was not

evaluated, but in any case, the low level of abstraction and arbitrary signal parameter association of Optohapt would likely involve high learning curves.

*Tactile-vision sensory aids.* In 1968, Bach-y-Rita and colleagues developed the Tactile-Vision Substitution System (TVSS)—an optomechanical device that converts captured visual images to touch stimulations felt on the back via a 20x20 matrix of 400 vibrating solenoids (Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969). The original setup is shown in figure 34 where a dental chair hosts the back display, driven by input from a television camera mounted on tripod with controls for manually adjusting pan, tilt, zoom, aperture and focus. Each captured image of video input is divided into blocks of pixels, each coupled to an actuator, vibrating only if the corresponding intensity is above a threshold.

Early studies (Bach-y-Rita, Collins, Saunders, White, & Scadden, 1969) (White, Saunders, Scadden, Bach-y-Rita, & Collins, 1970) (Bach-y-Rita, 1972) with TVSS provided much insight related to perception and learning of visual input mediated by the skin, and the general usability of substitution systems for tactile-vision. Visual lines in horizontal, vertical and diagonal orientations could be immediately discerned; as could simple motions such as back-and-forth movement. After preliminary training, subjects could accurately discriminate between basic shapes (circles, squares, etc.) when allowed to move the camera for active exploration. Passive exploration in which no camera movement was allowed resulted in poor recognition performance (50-60% recognition accuracy). Subjects could also distinguish between common objects such as a telephone, chair, etc. Considerable time, between 5 – 15 minutes, was needed to identify new objects, but this latency decreased with repeated presentations—as did the time to recognize novel objects. Through these initial experiments, subjects encountered visual

concepts through touch including perspective, shadow, occlusion and the relationship between size and distance.

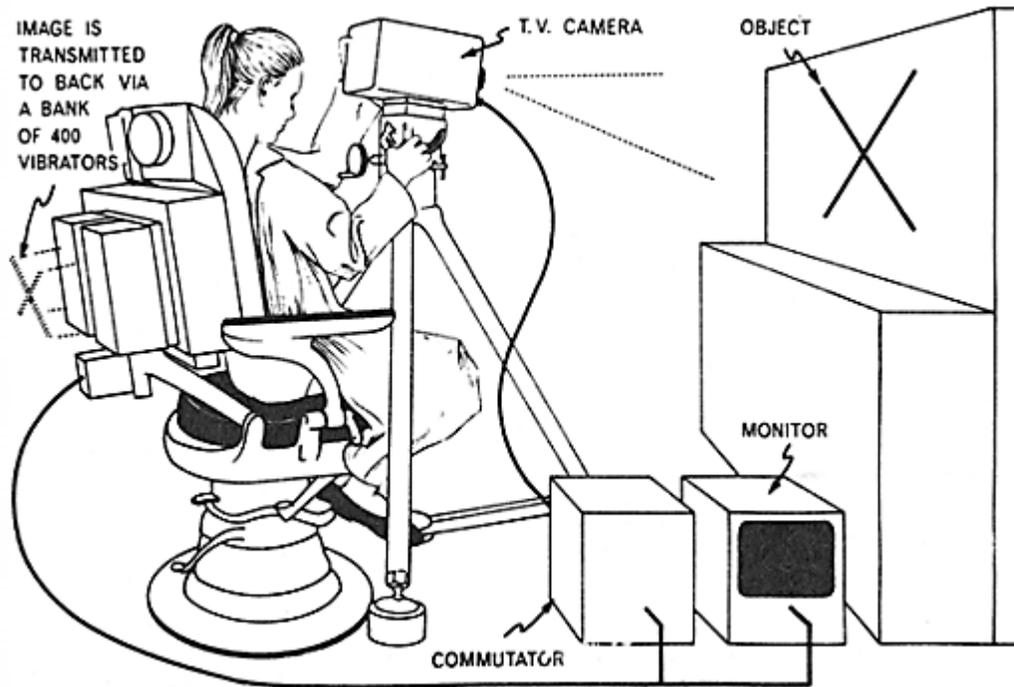


Figure 34. Original apparatus of TVSS with vibrotactile display on dental chair, moveable camera and monitor for visual observation. Reprinted from “Seeing with the skin,” by White, B. W., Saunders, F. A., Scadden, L., Bach-y-Rita, P., & Collins, C. C., 1970, *Perception & Psychophysics*, 7(1), p. 23. Copyright © 1970 by Psychonomic Society, Inc. Reprinted with permission.

To overcome the limitations of the first prototype—in particular, its bulkiness and high power demands—Bach-y-Rita and colleagues developed a portable TVSS (1972) utilizing an electrotactile display (8x8 matrix of electrodes). The new prototype enhanced wearability and portability, providing users with improved interaction with their environment, leading to discoveries of hand-“eye” coordination. Bach-y-Rita subsequently explored smaller electrotactile displays (7x7) for sensory substitution for the fingertip (Frissen-Gibson, Bach-y-Rita, Tompkins, & Webster, 1987); and then finally, the tongue (Bach-y-Rita, Kaczmarek, Tyler, & Garcia-Lara, 1998). Compared to

the fingertip, the tongue offered ideal conditions for electrical stimulation, requiring less voltage given that touch receptors are close to the tongue's surface, and good electrical contact is afforded by the mouth's saliva. Since the late 1990's, the aforementioned tongue-based display for sensory substitution, eventually called the Tongue Display Unit (TDU), a conceptual drawing of which is depicted in figure 35, has evolved into the BrainPort vision device<sup>4</sup>: a wearable, portable, rechargeable tongue-display driven by a small head-mounted camera system.

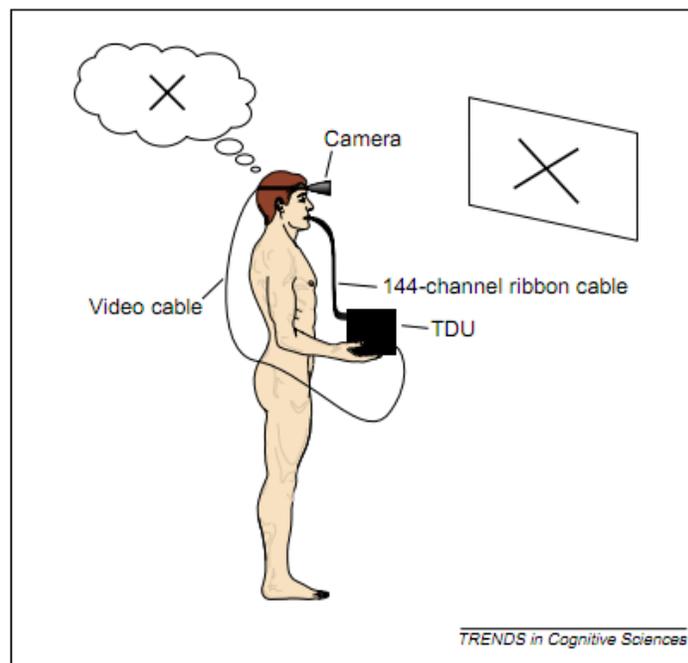


Figure 35. Tongue Display Unit (TDU) version of TVSS. An electro-tactile display placed on tongue conveys a “blurred” image captured by the head-mounted video camera. Reprinted from “Sensory substitution and the human-machine interface,” by Bach-y-Rita, P., & Kercel, S. W., Dec. 2003, *Trends in Cognitive Sciences*, 7(12), p. 543. Copyright © 2003 by Elsevier. Reprinted with permission.

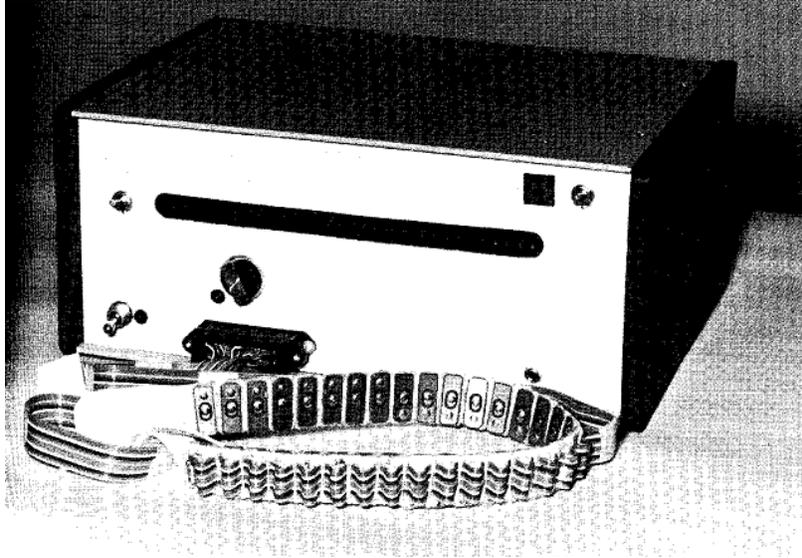
The basis of TVSS is neuroplasticity in which the brain's touch centers interact with the visual cortex to eventually reorganize vision centers to visually interpret

<sup>4</sup> <http://vision.wicab.com>

incoming somatosensory signals (Bach-y-Rita & Kercel, 2003) (Bach-y-Rita, 2004). As the brain interprets incoming sensory signals as patterns of impulses, any type of input can be delivered to a modality for interpretation as long as the receptors mediating that modality are sensitive to the stimuli.

The main shortcoming of TVSS is the intensive training, both in terms of camera control and tactile perception, for proficient use—many of the subjects involved in the aforementioned studies underwent 40+ hours of training. Moreover, the internal detail of objects and cluttered backgrounds are difficult to perceive. As of 2011 (over forty years since its inception), clinical trials are underway, preparing the device for eventual commercialization.

***Tactile-audio translation.*** Tactile-audio sensory substitution systems convert sound (environmental sounds and/or speech sounds) into vibrotactile or electrotactile stimulation through bandpass filtering and noise suppression algorithms. These devices are usually geared toward speech, and are utilized by individuals who are deaf to improve speech production, and/or lip and speech reading. The teletactor (Saunders, Hill, & Franklin, 1981), figure 36, was an early tactile-audio device that used bandpass filters to convert acoustic information, gathered from a microphone, into 32 electrotactile stimulations delivered using a wearable waistbelt. Electrotactile stimulations carried timing information related to the speech signal, and the intensity at each location corresponded to the intensity within the respective frequency band.



*Figure 36. Teletactor device for audio-to-tactile conversion using 32 bandpass filters and a waistbelt of 32 electrodes for electrotactile display. Reprinted from “A wearable tactile sensory aid for profoundly deaf children,” by Saunders, F. A., Hill, W. A., & Franklin, B., 1981, *Journal of Medical Systems*, 5(4), p. 267. Copyright © 1981 by Springer Netherlands. Reprinted with permission.*

More recently, the Tactaid line of devices<sup>5</sup>, by Audiological Engineering Corporation (AEC), were commercially available for 25 years since the early 1980's. These devices were available in different models including Tactaid I, Tactaid II and Tactaid VII. Each device was portable and wearable, with a flexible, attachable vibrotactile display. The Tactaid I provided only one channel of communication, but presented rhythmic and temporal information related to acoustic input. An additional channel was later added with Tactaid II, enabling users to differentiate between different types of sounds based on signal frequency. More channels were subsequently added with Tactaid VII in which seven channels enabled rich sound differentiation that was ideal for speech training. Numerous studies have since explored the effectiveness of the Tactaid devices (Weisenberger & Percy, 1994) (Reed & Delhorne, 1995). As with tactile-vision, intensive training and practice are needed for proficient use of tactile-audio sensory aids.

---

<sup>5</sup> <http://www.tactaid.com/>

### **Alphanumeric Information Delivery via Touch**

Approaches whose communication units are alphanumeric provide rich communication languages that are robust, versatile and expandable, but ultimately suffer from their inherent arbitrary associations and lack of efficiency in terms of fast conceptual communication (table 1). Examples of approaches that fall under this category include Braille, Morse code and vibratase, all of which represent alphanumeric units (letters, numbers and/or common words) with some form of arbitrarily assigned stimulation applied to the skin, either passively or actively. Although these approaches provide a high level of abstraction in which characters are converted into a coded form for perception, it is the arbitrary association between form and meaning that creates extraordinary high learning curves. These high learning curves are further exacerbated by the large character sets that must be encoded and learned. Moreover, since communication occurs at the rate of character transmission (or common words at best), high level concepts (objects, places, ideas, etc.) may take significant amounts of time to convey when using these verbose methods; even when impressive word rates are achievable, these approaches are not as useful for applications requiring real-time communication at a conceptual level rather than verbal level.

Although alphanumeric approaches utilize a high level representation, each communication unit represents a low-level concept—in particular, letters, numbers or common words. Because these languages themselves represent language, they are highly versatile, rich and expandable. Moreover, this representation provides for robustness in that if a character is missed during passive or active interaction, the respective word may still be perceived through use of the context provided by other characters of the word, and surrounding words.

**Braille.** Braille was developed in 1825 by Louis Braille as a method for individuals who are blind to read and write. Each character of the alphabet, in addition to punctuation marks, numbers and commonly grouped letters such as AND, TH and CH, is represented by a Braille cell—a simple, structured pattern (3x2) of raised dots (figure 37). Braille cells are written horizontally and read sequentially, similar to written characters. But unlike written characters, which are perceived visually with a large field of view and high acuity, Braille is read with the fingertips, and is therefore perceived within a small field of view and with much lower spatial acuity—necessitating the need for simple patterns (Foulke, 1982). Braille topics ranging from reading behavior and ability; haptic perception of the tactile patterns; and display variations, have been extensively explored through user studies (Foulke, 1982).

Braille reading rates are much lower compared to visual reading rates: on average, visual reading rates are two to three times faster than the reading rate of experienced Braille users (Foulke, 1982); but exceptions do exist where extraordinary Braille readers have reading speeds that compete with visual readers. As described, the arbitrary associations of meaning to dot patterns result in high learning curves. Unlike Morse code and vibratase (described next), both of which are passive interaction techniques, Braille is perceived through active exploration as fingers glide across Braille cells. This interactivity further strengthens Braille’s robustness as characters and words may be revisited.

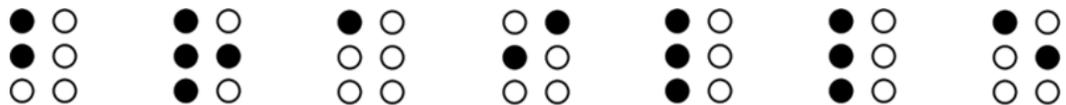


Figure 37. Example of Braille cells spelling the word ‘braille’. Adapted from Wikimedia Commons File: “File:800px-Braille.png”.

**Morse code.** Morse code was invented and refined from 1832 to the mid 1800's by American inventor Samuel F. B. Morse, and his assistant, Alfred Vail. Developed as a method to write from a distance (telegraphy), opening and closing a switch on the operator side generates patterns of tick marks, recorded by a mechanical pen on the receiver side, whose associations to letters and numbers could be looked up using a codebook to decipher messages. These tick marks, or codes, are patterns of *dots* (short tick) and *dashes* (long tick). In 1850, "writing from a distance" was replaced by the more efficient auditory presentation of Morse code using beeps; after which, auditory Morse code became the form most commonly used. Morse code users with hearing impairments, however, opted for an alternative communication channel through touch by placing their hands on the speakers generating the Morse code beeps (Tan, Durlach, Rabinowitz, Reed, & Santos, 1997). Since 1832, Morse code underwent several refinements, and after satisfying international requests, International Morse Code (figure 38) became the standard representation of the code.

Morse code is still in use today in aviation (station identification), navy (communication during radio silence), amateur radio and assistive technology. Regarding the latter, Morse code offers a promising alternative to both human-to-human communication and access to computer applications (e.g., word processors) for individuals with severe physical impairments in that it enables communication through simple binary muscle movements (King, 2000). Based on the user's physical impairments and preferences, binary muscle movements may involve the movement of a limb up/down or left/right; the blink of an eye; or the puff/sip of a straw.

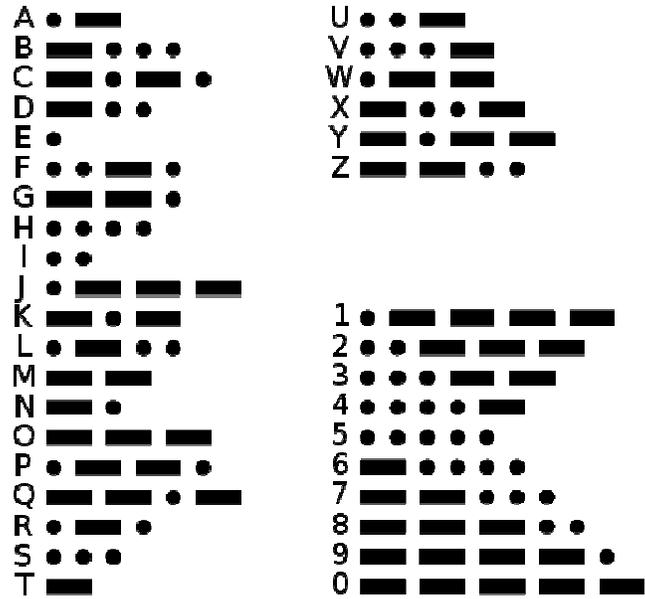


Figure 38. International Morse Code. Adapted from the Wikimedia Commons file "File:500px-International\_Morse\_Code.svg.png".

Since the late 1800's, researchers have studied the learning rates associated with Morse code. More recently, Tan, Durlach, Rabinowitz, Reed and Santos (1997) compared Morse code reception at different words rates (from 12 to 24), after 70-80 hours of training, between experienced (+20 years) and inexperienced Morse code users across different modalities: kinesthetic (up and down movements of a finger), vibrotactile (vibratory pulses) and auditory (beeps). Tan et al. found auditory word rates to be twice those of vibrotactile presentation, the latter of which was 1.3 times that of kinesthetic presentation. Tan et al. argued that this might be due to the auditory modality's superior response time and accuracy to dynamic signals. Experienced Morse code users outperformed novices as they utilized chunking to perceive messages at higher levels compared to beginners who concentrated on low-level signal parameters to build up letters and words. And thus, the expert Morse code users' abilities in auditory Morse code

perception transferred to other modalities given that they could focus on high level meanings rather than low level signal parameters.

Like other alphanumeric approaches, Morse code is limited by its high learning curves and slow communication speeds. The sequential pattern of Morse code messages through a 1-bit display, combined with pauses between individual pulses, results in an inefficient presentation: roughly 480 milliseconds per letter (Tan, Durlach, Rabinowitz, Reed, & Santos, 1997). This is in contrast to vibratense, described below, which utilizes multiple points of contact and a more efficient design to improve word rates.

**Vibratense.** In the 1920's and 1930's, research in vibrotactile communication began in the context of developing tactile-audio sensory substitution systems for individuals with hearing impairments. In the 1940's and 1950's, Geldard and his colleagues began to question previous research in vibrotactile communication (1957). In particular, they speculated that the learning challenges faced by users of tactile-audio devices, was due to sensory limitations encountered when directly presenting an acoustic signal to the skin. They argued that previous work failed to ask one fundamental question: what are the communication capabilities and limitations of the skin? Their research efforts were the first attempt at finding a language of the skin. They began by first identifying dimensions of a vibration signal that might be used to convey information: dimensions including frequency, amplitude, duration, locus, waveform, as well as spatio-temporal patterns such as perceptual illusions. Moreover, their aim was to develop a fast, vibrotactile communication channel that could replace Morse code; motivated by the fact that Morse code is inherently limited by pauses between pulses, whereas a new vibrotactile communication method could avoid this, potentially achieving faster communication rates.

To explore the first-order dimensions of vibrations, namely frequency, amplitude and duration, Geldard and his colleagues conducted a number of psychophysical studies (1957) (1960)—the results of which are still used today—to uncover *just noticeable differences* (JND's) and the limits of human perceptual capabilities. Their findings were subsequently used as the basis for the *vibratese* language (1957): an alphanumeric, vibrotactile encoding of letters and numbers (figure 39). Three intensities, three durations and five loci were chosen to encode meaning into vibration signals. Note that vowels, which occur most frequently in written and spoken language, have been assigned the shortest duration, whereas numbers, which occur less frequently, have been assigned the longest duration. This design choice was made to ensure fast communication speeds for alphanumeric information delivery. The vibratese language was evaluated through a user study involving three participants wherein after 12 hours of training (spread across a couple of days), participants had sufficiently associated the vibration signals to their respective meaning, and progressed to learning words and sentences. Given vibratese's arbitrary mapping combined with the unavoidable large set of stimulants that must be learned due to the alphanumeric representation, specifically 40 patterns for letters, numbers and common words, the learning curve is high. In theory, vibratese may allow for communication rates of up to 67 words/min (much lower in practice, however—one participant who received extended training hit a plateau at 38 words/min); this is in contrast to communication speeds of experts in Morse code (24 words/min, or a little higher for some).

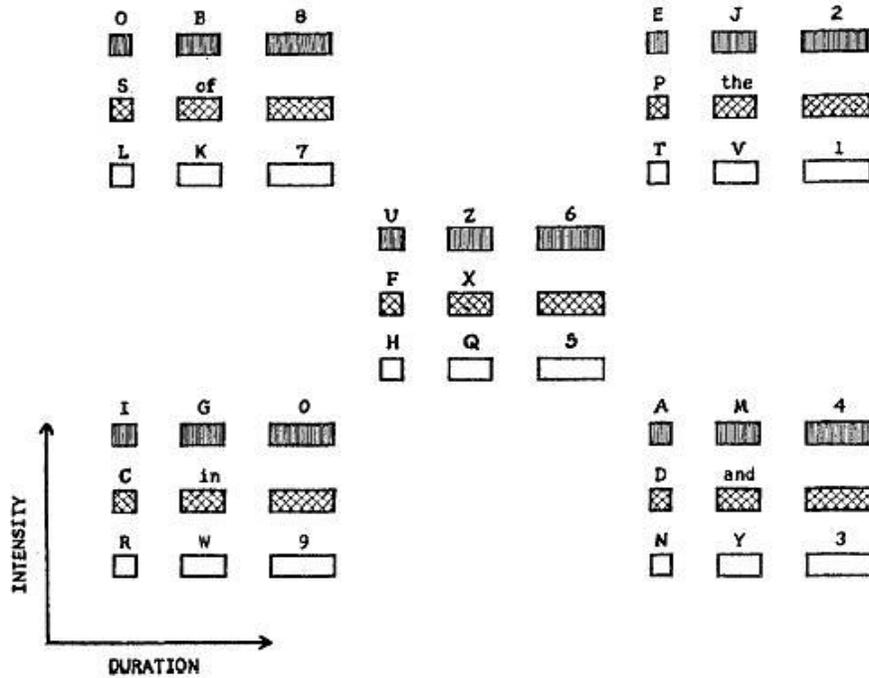


Figure 39. Vibratese language for alphanumeric communication through vibrations. Encoding utilizes three intensity values (vertical axis), three duration values (horizontal axis) and five locations spread across the skin of the chest, creating 45 different vibrotactile patterns. Only 40 of these patterns were used: 26 for letters A-Z, 10 for numbers 0-9 and four for common words including *the*, *and*, *of* and *in*. Reprinted from "Adventures in tactile literacy," by Geldard, F. A., 1957, *American Psychologist*, 12(3), p. 120. Copyright © 1957 by American Psychological Association (APA). Reprinted with permission.

## **Automated Conceptual Information Delivery via Touch**

A variety of approaches have been proposed for concise, automated delivery of conceptual information to the skin. Most of these approaches incorporate a high level of abstraction and intuitive signal parameter association in their design, and are often versatile, expandable and efficient. However, existing theoretical frameworks are lacking in terms of enabling the design of rich and robust somatic languages (table 1). Existing approaches do not take into account variations in contextual usage of communication units and how these variations influence meaning, as in natural language, to enrich communication. Here, *context* refers to environmental variations including the body site of the stimulation; that is, the conditions under which a somatic signal is presented. Regarding body site, we consider the surface of the human body being influential regarding the meaning of communication units based on the site of stimulation (i.e., *body context*). This concept provides expressive communication possibilities, as we'll later see, for enriching languages with limited vocabularies. Moreover, existing approaches have not yet explored enriching somatic communication languages through dynamically changing the meaning of communication units by varying parameters of stimuli (i.e., *stimulation variations*). Stimulation variations have a counterpart in natural language called tonal variations in which tonal changes dynamically vary the meaning of our spoken words.

Lastly, existing approaches have paid little attention to improving the robustness of somatic communication for ensuring reliable communication (with the exception of haptic glyphs which utilize active exploration). Repeated presentations of stimuli are commonly performed (Brown, Brewster, & Purchase, 2005), and although this aids robustness through redundancy, such a presentation scheme may not work for applications requiring real-time communication or interaction. A more efficient solution

is the use of an attention grabbing signal, *somatic alert* or *somalert*, at the start of a stimulation as well as encoding redundancy in the signal itself such as four, rather than three, brief pulses in saltation signals (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011). Attention grabbing signals have been explored to some extent (Brown, Brewster, & Purchase, 2006b), but with adverse results due to insufficient training: participants were not told about these priming signals, and therefore, perceived them as part of the stimulation. The robustness of a somatic language must also be considered in adverse environmental conditions such as those with noise or high cognitive load (workload). Approaches have proposed automatic cross-modal transitions to audio from touch when vibrotactile perception is no longer reliable (Brewster, Chohan, & Brown, 2007); i.e., in the presence of ambient vibrations such as those experienced while moving or while riding in a vehicle. An alternative might be to alter signal dimensions (such as intensity or the number of active actuators) to ensure successful delivery of a message, particularly in situations where other modalities might be overloaded. Workload conditions of high cognitive load create reliability issues (Chan, MacLean, & McGrenere, 2005), and therefore stimuli must be carefully designed to ensure successful delivery; in this scenario, somatic alerts, redundancy and stimulation variations based on the workload must all be employed to ensure successful delivery.

The proposed theoretical framework, Somatic ABC's, has been designed to meet these criteria (table 1) to overcome the previously mentioned limitations, achieving a framework for the design, development and evaluation of functional and practical languages for somatic information delivery. Over the following sections, related computerized approaches for conceptual information delivery are reviewed including approaches for mediated haptic interpersonal communication, tactile icons, haptic glyphs and haptic icons.

**Mediated haptic interpersonal communication technology.** *Haptic interpersonal communication*, between two or more people, refers to any type of information exchanged through the modality of touch. If the communication channel is *mediated*, this information exchange happens through a *mediator* (computerized system) that transfers the information from one person to the other. Technology to facilitate haptic interpersonal communication can be divided into four broad categories: *therapy*, *gaming*, *general communication* and *computer supported collaborative work* (CSCW). Note that these categories are not distinct; often a technology will be applicable to two or more categories, whether or not this was intended by the designer. Many proposed technologies for mediated haptic interpersonal communication provide a medium to develop novel communication possibilities. Some of the most significant contributions in this respect are described next.

An early entertainment device, HandJive (Fogg, Cutler, Arnold, & Eisbach, 1998), consisted of two interconnected spheres where one is held in each hand. The spheres can be shifted from their upright position either forward or backward, together or separately, allowing nine possible combinations. Shifting a sphere causes the corresponding sphere to shift on the interaction partner's HandJive, but from side to side rather than forward or backward. This protocol prevents users from fighting for control, i.e., users can create and send HandJive haptic signals while still receiving cues from their partner without being interfered with or interrupted. HandJive has been proposed as a general haptic communication tool using the *tactilese* language. Essentially, the smallest units (position of the spheres) are used to create simple movements (patterns), and in turn, combined to create complex movements (routines).

More general devices for mediated haptic interpersonal communication whose intended purpose are for implicit or nonverbal communication beyond specific

applications such as entertainment or therapy may be gathered under the broad category of *general communication*. As communication tools, these devices provide a medium for exchanging concrete to abstract information; and often, new forms of expression develop through use of this new somatic communication channel, augmenting existing communication systems to provide redundant and/or complementary information.

One of the earliest mediated haptic interpersonal communication devices for general communication was InTouch (Brave & Dahley, 1997). The system consisted of two three-pin rollers, each controlled by one interaction partner. To communicate, a user moves the rollers with his or her hand, and the other roller changes accordingly in real-time. The communication channel is two-way (bi-directional), and its input and output signals are mapped to the same channel (symmetric I/O mapping), operating on the principle of a *shared object*; i.e., it is as if there is only one object being manipulated. InTouch enabled two types of interaction: *passive*, where the user's hand is placed on the device to feel what his or her partner is communicating; or *active*, where both users manipulate the object, and perceive its output, simultaneously. A pilot test revealed the usefulness of the device to convey abstract, subtle communication cues, such as those found in intimate communication, as opposed to communication in general.

Another example is ComTouch (Chang, O'Modhrain, Jacob, Gunther, & Ishii, 2002): a vibrotactile glove for complementing verbal information exchanged during a phone conversation. When a user, e.g., *user A*, applies pressure through use of the glove, a vibrotactile signal is sent to his or her partner, e.g., *user B*, where the intensity is proportional to the amount of applied pressure. *User B* feels the vibration at his or her index finger's metacarpophalangeal joint. *User A* also feels a vibration, but on his or her index finger's proximal interphalangeal joint in the form of a feedback signal, enabling a way to assess the intensity of the signal being sent, and readjust the pressure accordingly.

During an audio conversation, experimenters observed that participants created their own novel *tactile gestures* usually to (1) *emphasize* what they were saying by applying pressure while saying certain words or phrases; (2) indicate *turn-taking* by sending a vibratory signal before speaking; and/or (3) *mimic* the other user by exchanging the same vibrotactile pattern, which could be used to indicate presence or acknowledge each other. Although simple in its conceptual design, ComTouch's addition of a tactile channel provided a powerful form of expression, complementing the auditory channel with nonverbal information.

Shake2Talk, a cell-phone based system designed and developed by Brown and Williamson (2007), used gesture-based inputs, such as strokes, taps, etc., to create *audio-tactile messages* (figure 40). For example, a tapping gesture may generate the sound of gentle tapping and the sensation of someone tapping; such a message may be interpreted as the caller asking the recipient to call back soon. In a user study involving six couples (Brown, Sellen, Krishna, & Harper, 2009), some couples developed a vocabulary, assigning meanings to certain messages; the majority of couples used the multimodal messages for coordination, e.g., "I'm on my way over", but the messages were also used for awareness/reassurance, play and social touch.

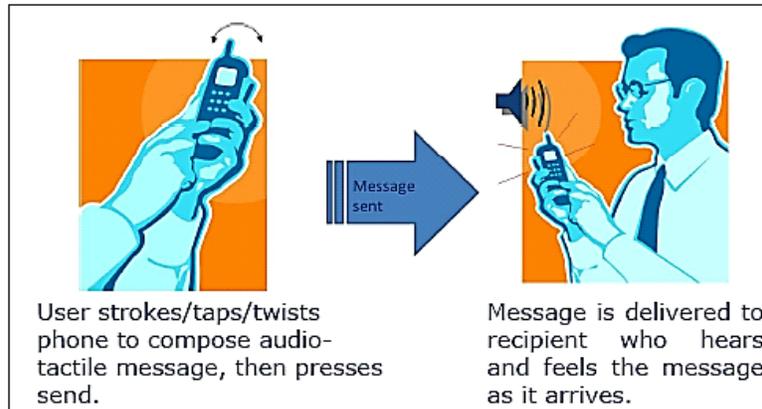


Figure 40. Conceptual design of Shake2Talk where speech and haptic interactions are used to create multimodal messages. Reprinted from “Shake2Talk: Multimodal messaging for interpersonal communication,” by Brown, L., & Williamson, J., 2007, In I. Oakley and S. Brewster (Eds.), *Haptic and Audio Interaction Design* (p. 44), LNCS 4813. Copyright © 2007 by Springer-Verlag Berlin Heidelberg. Reprinted with permission.

Some work has explored augmenting instant messages (IM) with haptic signals to communicate nonverbal cues. For the Haptic Instant Messenger (HIM), Rovers and van Essen (2004) augmented simple emoticons (happy, sad, etc.) with haptic icons in the form of vibration signals to enrich instant messages. Contact IM (Oakley & O'Modhrain, 2002) used instant messages in addition to force-feedback, provided by a Phantom haptic device, to enable users to chat while throwing a virtual ball to each other, providing a familiar, yet subtle and abstract form of expression.

Although many of the aforementioned technologies provide interesting and unique haptic communication channels within their respective applications, designs tend to be application-oriented and functionally confined, lacking the versatility and richness needed for a generic framework for somatic information delivery.

**Tactile icons.** Tactile icons, or *tactons*, are a more general, abstract and versatile methodology for vibrotactile communication proposed by Brewster and Brown (2004), who defined them as “structured, abstract messages that can be used to communicate messages non-visually” (p. 15). Here, *abstract* refers to an arbitrary association between the vibrotactile stimulation and its conceptual meaning. Tactons are the tactile counterpart to icons: visual symbols or representations that convey abstract messages. Brewster and Brown proposed three types of tactons:

- A *compound tacton* consists of two or more concatenated, *simple* tactons, where simple refers to the use of a single dimension to convey a message, such as rhythm or intensity. By concatenating simple tactons that each represent a basic action, object or concept, more detailed and specific messages may be created.
- A *hierarchical tacton* begins with, and adds to, inherited properties from *base tactons*. For example, a base tacton representing an incoming call may signal this with a particular tactile rhythm; a tacton inheriting from this base tacton may slow the tempo to represent a loved one calling, whereas a faster tempo may represent the incoming call of a boss.
- A *transformational tacton* arbitrarily associates meaning to different dimensions of the vibration signal. For example, the type of call (phone, text, etc.) could be associated with rhythm, and the ID of the caller could be associated with intensity. The transformational tacton design is the most widely used given its simplicity and ease of use.

The vibrotactile patterns found in vibratese (Geldard F. A., 1957) may be considered an early form of transformational tactons. Since the time of vibratese and its early psychophysical studies exploring just noticeable differences, scientists and researchers have continued to explore human haptic perception of the dimensional values

of vibration signals including frequency, amplitude, duration, body site and spatio-temporal patterns. This is particularly important for tactons since they rely on our ability to learn and recognize individual dimensional values. In this regard, three types of actuators are most commonly explored: Engineering Acoustics' C2 tactors<sup>6</sup>; Audiological Engineering Corporation's TACTAID actuators<sup>7</sup>; and pancake vibration motors.

Tactile rhythm and body site have both been successfully used to design reliably recognizable tactons. Brown, Brewster and Purchase (2005) proposed three tactile rhythms: one rhythm of seven short pulses, another rhythm of four long pulses, and a rhythm of one short pulse then one long pulse. These rhythms have been evaluated on the fingertip of the index finger using a C2 tactor with 93% overall recognition accuracy (Brown, Brewster, & Purchase, 2005); the volar side of the forearm using a C2 tactor with 96.7% overall recognition accuracy (Brown, Brewster, & Purchase, 2006a); and in the palm of the non-dominant hand using a standard vibration motor (within a mobile phone) with 95% overall recognition accuracy (Brown & Kaaresoja, 2006). These rhythms have inspired similar designs in many applications, and have themselves been successfully applied: Lin and Cheng (2008) used the aforementioned rhythms for creating tactons for use in pedestrian navigation to convey the direction of travel (*turn right, turn left* and *stop*) where tempo was used to convey the distance to the next change in direction. Barralon, Ng, Dumont, Schwarz and Ansermino (2007) designed three tactile rhythms to convey alert levels in physiological monitoring of anesthetized patients; these rhythms included a single long pulse, two short pulses and three very short pulses. Overall recognition accuracy of rhythms was 96.3% when delivered around the waist using C2 tactors. The location of the vibration around the waist corresponded to another physiological cue: one of six different physiological events based on which of the six

---

<sup>6</sup> <http://www.eaiinfo.com>

<sup>7</sup> <http://www.tactaid.com>

vibration motors were actuated (95.1% overall recognition accuracy). Brown, Brewster and Purchase (2006a) also explored body site as a tacton parameter: localization accuracy was investigated for three equidistantly spaced C2 tactors on the volar forearm with endpoints at the wrist and elbow joint (95.5% overall recognition accuracy). Both of the aforementioned explorations of body site as a tacton dimension utilized the results of vibrotactile localization studies by Cholewiak and his colleagues (Cholewiak & Collins, 2003) (Cholewiak, Brill, & Schwab, 2004).

Intensity and more complex waveform variations, including *roughness* and *envelopes*, have also been explored in the context of tactons; but these parameters have not been as successful compared to tactile rhythm and body site in the context of tacton learning and recognition. In particular, intensity, as well as frequency, are not recommended given the limited human perceptual resolution of these parameters in addition to their unwanted interaction in standard vibration motors. Shieh and Wu (2008) explored human perception of four intensity values (low, high, increasing or decreasing intensity) combined with four two-pulse (short-short, short-long, long-short or long-long) tactile rhythms. The intensity variations are *envelopes*, or gradual changes in intensity over time (Gunther, 2001). They found a higher overall recognition accuracy for rhythm (90.97%) compared to intensity (74.7% for envelopes, 86.11% for the two static levels, or 80.90% overall). With respect to envelopes, more promising results have been found by Brown, Brewster and Purchase (2006b). Using a TACTAID actuator placed on the index finger, they found overall recognition accuracies of 100% and 92% for gradual linear/exponential increases and gradual linear/exponential decreases, respectively. Differences between actuators and/or stimuli presentation durations might have attributed to these differences in accuracy; in particular, Brown, Brewster and Purchase used a

longer presentation duration of two seconds compared to the shorter duration of 550 ms used by Shieh and Wu.

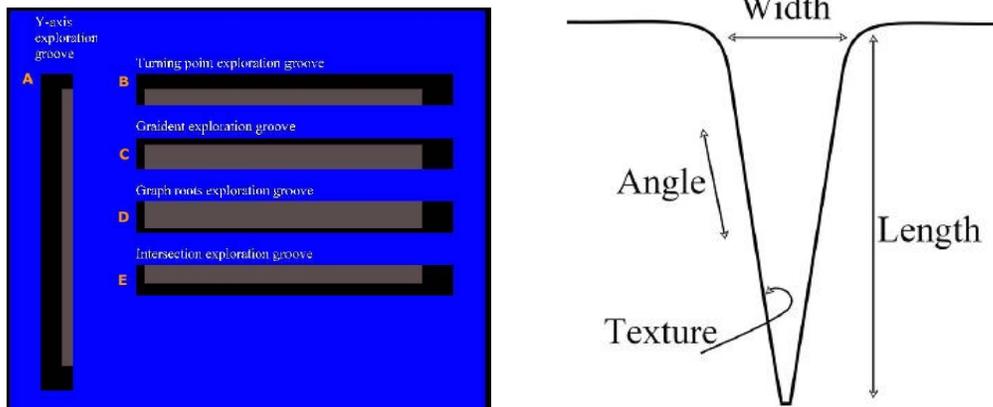
Sinusoidal amplitude modulation (multiplication of a signal with a base frequency with another signal of a different frequency) can be used to create vibration signals by varying perceived “*roughness*”. Roughness as a tacton parameter has received much exploration, but studies have shown that it is not as useful compared to rhythm or body site, especially when standard vibration motors are used. In particular, Brown and her colleagues have extensively explored roughness for tactons using both C2 tactors (Brown, Brewster, & Purchase, 2005) (Brown, Brewster, & Purchase, 2006a) and standard vibration motors available in mobile phones (Brown & Kaaresoja, 2006); in the latter experiment, roughness was simulated through speed variations of on-off pulses due to hardware limitations. Brown and Kaaresoja found a decrease in overall roughness recognition accuracy: from 80% (for C2 tactors) to 55% (for standard vibration motors). In the aforementioned physiological monitoring application (Barralon, Ng, Dumont, Schwarz, & Ansermino, 2007), roughness was used to communicate a change in the direction of the level of alert: “roughness” indicated an increasing alert level, and “smoothness” indicated a decreasing alert level; 88.7% overall recognition accuracy was achieved.

Given the arbitrary associations between stimuli and meaning, high learning curves may be encountered particularly when large sets of stimuli are used; further, designs are limited in terms of the number of tacton parameters and dimensional values that can be used without sacrificing recognition accuracy. Tactons are general enough to be applied to a variety of application domains, and therefore, it is a versatile framework; however, as satisfactory recognition accuracy is achievable with at most two or three parameters with a few dimensional values each, it is difficult to create rich, expressive

languages from tactons (refer back to table 1). Furthermore, in terms of robustness, repeated presentations are not an option during real-time communication; somatic alerts and built-in signal redundancy needs to be further explored.

**Haptic glyphs.** Glyphs are visual symbols or shapes that convey information where typically multiple parts of the glyph encode related and complementary information which may be constant or dynamic based on incoming input data (Roberts & Franklin, 2005). This is in contrast to another form of visual communication through symbols, i.e., icons, which are static and convey a single meaning. Inspired by glyphs, Roberts and Franklin (2005) proposed *haptic glyphs*, or *hlyphs*, in which meaning is associated with force feedback parameters such as attraction/repulsion, friction or vibration, depending on the location of the device's interaction point in 2D or 3D space. Haptic glyphs may be explored actively, passively, or as a combination of both where users are guided to different sections of the hlyph for active exploration within a limited space. Roberts and Franklin presented several design principles for hlyphs in that they should be (1) well-structured such that the act of exploration is intuitive and easy to perform including straightforward transitions between parts of the hlyph; (2) compound/multifaceted in that meaning is associated with multiple hlyph parts to ensure rich communication possibilities; (3) self-contained such that hlyphs stand alone, cover a small area (to simplify exploration in addition to memorization) and have no "holes" to ensure a user's interaction point will not "escape" during active exploration, which could create confusion and frustration; (4) endogenous to ensure ease of exploration and reduce frustration. Active exploration using a force feedback device is easiest while exploring concave surfaces whereas convex surfaces create opportunities for the interaction point to lose contact with the object; this may be prevented by exploring the inside of an object, or by using a boundary to surround the outside of an object; (5) pre-attentive in that hlyph

parts haptically “pop out” to grab one’s attention; (6) conceptual such that the associations of meaning to force feedback parameters and hlyph components are intuitive; and (7) intuitive such that users implicitly understand how to explore and navigate the components of the hlyph. Based on their proposed theoretical framework and design principles, Roberts and Franklin proposed two example hlyphs: synoptic hlyph and cavern hlyph (figure 41). In the synoptic hlyph, different parts of a graph (maximum and minimum values, roots, etc.) are each associated with a groove where the positions of raised and lowered surfaces within the groove communicate quantitative data. In the cavern hlyph, a virtual valley with adjustable width, length, texture and angle, is used to convey information through the assignment of meaning to the aforementioned parameters.



*Figure 41. Two examples of haptic glyphs (hlyphs). Synoptic hlyph (left) where grooves with raised and lower surfaces communicate attributes of a graph including minimum and maximum values, roots, turning points, gradients and intersections; and the cavern glyph (right) where the metaphor of a cavern or valley is used to communicate information by associating it with those attributes shown. Adapted from “Haptic Glyphs (Hlyphs) - Structured haptic objects for haptic visualization,” by Roberts, J. C., & Franklin, K., 2005, In Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 373. Copyright © 2005 by IEEE. Adapted with permission.*

Other variants of glyphs exist that are not defined in terms of force feedback, but rather, vibrotactile stimulation. Moreover, as the most general definition of a glyph is a

visual symbol or shape that conveys information, other interpretations have been proposed. Osawa (2006) proposed tactile glyphs in which spatial variations of vibrotactile stimulation across both hands determined the meaning of the tactile glyph; the meaning of a tactile glyph is defined only by the comparison of spatial vibration patterns between the left and right hands. This design was proposed for a multimodal immersive learning environment for teaching programming concepts.

Borst and Baiyya (2007) (2009) proposed haptic glyphs for collaborative virtual reality through vibrotactile stimulation using a two-dimensional array of vibration motors. Rather than use a visual heads-up display to communicate the position and orientation of remote users, a palm-sized vibrotactile array was used where the following parameters were adjusted to create haptic glyphs: shape (parametric curve or line segment) in which vibrotactile stimulation temporally varied from one end point to the other; position of the haptic glyph on the display; orientation; scale; count of times the shape was traced; duration of trace; and the intensity profile of the vibrotactile stimulation. Moreover, the type/priority of the haptic glyph may be conveyed by its presentation. A haptic glyph may take precedence over and interrupt a haptic glyph that is currently playing; or a haptic glyph may be superimposed onto a haptic glyph that is currently playing. Any of the aforementioned parameters may be used to assign meaning to the haptic glyph; for example, in the application of collaborative virtual reality, Borst and Baiyya used the position and orientation of a haptic glyph in the shape of a line segment to communicate the position and orientation of a remote user; and the intensity profile of the haptic glyph was used to communicate the identity of the remote user. User localization and recognition of the proposed haptic glyphs were evaluated in terms of position, orientation and intensity recognition performance. Intensity profiles that rose then dipped, dipped then rose, or remained constant, were used. Each haptic glyph traced

10 times at one second per trace. Participants excelled in terms of localization, and orientation estimation was satisfactory at 21.1 degrees. Intensity profile seemed to be challenging with an average accuracy of 25% (3 out of 12). The authors also explored individual parameters, i.e., rendering of position, orientation or intensity irrespective of the other parameters; for position and intensity, the shape of the vibrotactile stimulation was rendered as a point. They found a noticeable difference between these conditions with individual presentations, at least for orientation and intensity, providing better performance with significant differences. The authors speculated that the challenges associated with perceiving the proposed haptic glyphs could be due to multiple vibration motors being simultaneously activated as part of the line segment. This is likely given that vibrations propagate and multiple vibration motor actuations can influence the perceived magnitude (Cholewiak R. W., 1979). As with other forms of computerized communication where multiple parameters are used to code meaning, careful attention must be paid to the interaction between parameters and how they influence human perception.

Although haptic glyphs are useful for specific applications, their versatility and richness is limited given their structure in that the full extent of the surface area of the skin is not exploited for communication and contextual cues (refer back to table 1). Their robustness, however, is strong, particularly for force feedback-based haptic glyphs (Roberts & Franklin, 2005) in which an active exploration environment enables users to explore the components of hlyphs at their own pace. However, haptic glyphs that use a “tracing” function to apply repeated presentations of passive, vibrotactile stimulations (Borst & Baiyya, 2007) (Borst & Baiyya, 2009) are not as useful for real-time communication.

**Haptic icons.** Haptic icons, or *hapticons* (Enriquez & MacLean, 2003) (MacLean & Enriques, 2003), in their most general definition, are haptic signals, tactile or kinesthetic, in which meanings have been intuitively or arbitrarily assigned to individual dimensions of parameters, to be communicated through any type of haptic display. Most work related to hapticons has explored programmable forces, defined by waveform, amplitude, frequency, and/or duration, delivered passively through a 1 degree-of-freedom force-feedback knob. In this regard, Enriquez and MacLean (2003) proposed a development environment for hapticons in which these signals could be visually designed or recorded in real-time. Simple waveforms could be superimposed, concatenated, and locally/globally adjusted. Generated waveforms could be played back through a knob at a set speed or actively explored at the user's own pace. Waveforms were recorded for playback as users manipulated the knob along a single axis. Subsequently, MacLean and Enriquez (2003) explored human haptic perception of haptic icons to provide insight into what parameters and dimensional values might be most useful for communication. They created a set of haptic icons that varied along amplitude (12.3, 19.6, 29.4 millinewton meters), frequency (0.5, 5, 20, 100 Hz) and waveform (sine, square, sawtooth) for periodic waveforms (duration was kept constant). Participants perceived and rated the similarity between stimulations by sorting stimuli into groups (different trials varied the number of groups) based on their own notion of similarity. Multidimensional scaling (MDS), a dimensionality reduction technique, was applied to the perceived similarities (or dissimilarities) between stimulations, mapping them into a new Euclidean space in which axes represent the most salient features. Within a two dimensional space, MacLean and Enriquez found frequency to be the most salient. Using lower frequencies and a smaller range of frequencies (specifically, 3, 7, 10, 16, 25 Hz), a MDS of participants' dissimilarity ratings revealed some saliency with respect to a

smooth (sine) waveform and a “jerky” (square/sawtooth) waveform. MacLean and Enriquez speculated that larger and more extreme values of frequency masked waveform, and therefore, smaller, more contained values of frequency should be used to ensure accurate perception of other parameters.

Chan, MacLean and McGrenere (2005) explored intuitive, vibrotactile haptic icons in the context of application sharing among remote users. They proposed a novel turn-taking protocol that used haptic icons, in the form of vibrotactile stimulations delivered through an augmented Logitech iFeel mouse, to convey information about *changes in control* of an application, *being in control* of an application and *waiting for control* of an application. Meaning was intuitively mapped to the dimensions of vibrotactile signals using a metaphorical interpretation—see table 2. The design of vibrotactile stimuli for conveying cues for *in control*, as shown in table 2, were decided using MDS. Frequency (20, 60, 100 Hz), amplitude (500, 2000, 5000, 8000; values given by Immersion Studio development environment and dependent on frequency) and duration (a single 1000 ms presentation, or a 700 ms burst, followed by a 100 ms delay, and then another 700 ms burst) were varied to create 24 stimuli that participants sorted based on similarity (the set also included both *changes in control* cues to ensure differentiability from *in control* cues). *Waiting for control* cues were not included as the authors were confident about their distinctness and intuitiveness. The recognition accuracy of the proposed cues was explored under various workload conditions: each participant had to identify the aforementioned cues while performing, in a random order, a visual task (puzzle), audiovisual task (puzzle plus listen for a specific word) and control condition (no task). An average of three minutes was required during the learning phase, which yielded an overall accuracy of 95% regardless of condition. As expected, workload had a significant effect on detection time: on average, participants took longer to respond

during the visual task (as compared to the control condition) as well as during the audiovisual task (as compared to the visual task). Involving more modalities and complicating the task seemed to further stretch attentional and cognitive resources. Cues may be designed to be more *intrusive* to “break through” distractions and improve detection time. The authors found one *in control* cue, IN++ (see table), to have consistent detection times regardless of condition (control versus high workload); this was expected since the cue was designed to be more intrusive since another user needs to urgently acquire application control. Overall, the impressive learning time can be directly linked to the intuitive metaphorical mapping between stimulation and meaning. This is in contrast to lengthy learning times for haptic icons utilizing arbitrary mappings (Enriquez, MacLean, & Chita, 2006), described next. Although the proposed methodology is useful for a limited set of cues, building a rich, versatile language would be challenging.

Table 2

*Design of Haptic Icons for Application Sharing*

Family	Icon ID	State	Haptic Sensation	Study 2 Label
Change of Control	CH+	User has gained control of the shared application	A short, weak buzz followed by a short, strong buzz	Awake
	CH-	User has lost control of the shared application	A short, strong buzz followed by a short, weak buzz	Asleep
In Control	IN	User is in control of the shared application	A just-noticeable, periodic vibration	Low Stress
	IN+	User is in control, but someone has gently requested control	A noticeable, but not unpleasant, periodic vibration	Medium Stress
	IN++	User is in control, but someone has urgently requested control	A very noticeable, somewhat unpleasant, periodic vibration	High Stress
Waiting for Control	WAIT	User has gently requested control	A periodic, quick, light tap	Bored
	WAIT+	User has urgently requested control	Two quick, light taps, delivered periodically	Really Bored

*Note.* The design consists of the haptic stimulation, *haptic sensation* in table, and assigned meaning and metaphor, *state* and *label* in table, respectively. Reprinted from “Learning and identifying haptic icons under workload,” by Chan, A., MacLean, K., & McGrenere, J., 2005, In *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, p. 435. Copyright © 2005 by IEEE. Reprinted with permission.

To achieve a rich communication language, there must be an underlying framework for combining stimulations to create more complex, but intuitive, stimulations

that map to higher-level concepts. Enriquez, MacLean and Chita (2006) proposed *haptic phonemes*, the smallest communication unit of a haptic signal (specifically, a simple waveform with a fixed frequency and amplitude). Haptic phonemes are assigned meaning, and then combined to create *haptic words*. Enriquez, MacLean and Chita's haptic phoneme development follows three guiding principles: (1) Differentiable: haptic phonemes should be distinct when used separate or together with other phonemes (i.e., as part of a haptic word); (2) Identifiable: the learned meaning of a haptic phoneme should be easy to recall; and (3) Learnable: the mapping between haptic phoneme and meaning should be natural and easy to learn. Haptic words may be created by (1) concatenating haptic phonemes, or (2) superimposing haptic phonemes.

In contrast to Chan, MacLean and McGrenere's study (2005), Enriquez, MacLean and Chita explored arbitrary assignments of meaning to haptic phonemes delivered through a haptic knob. Five waveform variations (triangle, square, three intermediates) and five frequencies (3, 7, 13, 18, 21 Hz) were used to create 25 stimuli with amplitude adjusted across all stimuli for equal sensation magnitude. These stimuli were subsequently sorted, based on similarity, by participants, and then dimensionally reduced through MDS. To ensure perceptual distinctness, nine stimuli were decided upon by selecting those with large separations along salient axes within the new dimensionality reduced space: waveforms included triangle, square and one intermediate with frequencies of 7, 10 and 18 Hz. The arbitrary association of meaning to signal dimensions included concepts *grass*, *flower* and *tree* assigned to frequencies, and *blueberry*, *strawberry* and *orange* assigned to waveforms. Next, the learnability and identification performance of the nine stimuli were assessed. Participants were asked to learn the mapping between sensation and meaning of the phonemes, and were tested on their ability to identify these associations through a sorting task. The association accuracy of

waveform and frequency were 73% and 81%, respectively, with large inter and intra-subject variation.

Most of the incorrect responses were due to learning associations incorrectly; indeed, four subjects struggled with incorrectly learning associations. Rather than more require extensive training to master arbitrary mappings, intuitive associations would most likely help in addition to reducing training time; indeed, Enriquez, MacLean and Chita's study required an average of 25 minutes of training, whereas Chan, MacLean and McGrenere's study required only three minutes of training and demonstrated higher recognition accuracy. Although the studies are not identical in terms of display device and information being communicated, this loose comparison clearly shows the benefits of natural, intuitive signal parameter associations. Moreover, Enriquez, MacLean and Chita found that, for a specific parameter such as frequency or waveform, intermediate dimensional values were more difficult to recognize than endpoints. This observation reflects the discussion earlier within the context of tactons: rhythm, body site and spatio-temporal patterns are more useful communication parameters compared to frequency, amplitude, duration or waveform, given that large sets of *distinct* haptic signals are easier to create as the latter parameters span a linear range, requiring users to memorize often closely spaced values and overcome just noticeable differences.

Haptic icons and phonemes lack two features important to a theoretical framework for somatic information delivery: richness and robustness (refer back to table 1). Haptic phonemes do not take into account contextual cues, such as the environment and/or body site, nor how meaning at the phoneme level influences word creation. Moreover, robustness is not included in the framework; there is some discussion by Chan, MacLean and McGrenere (2005) who demonstrated the challenge of creating intrusive, attention grabbing haptic icons, but this challenge still remains.

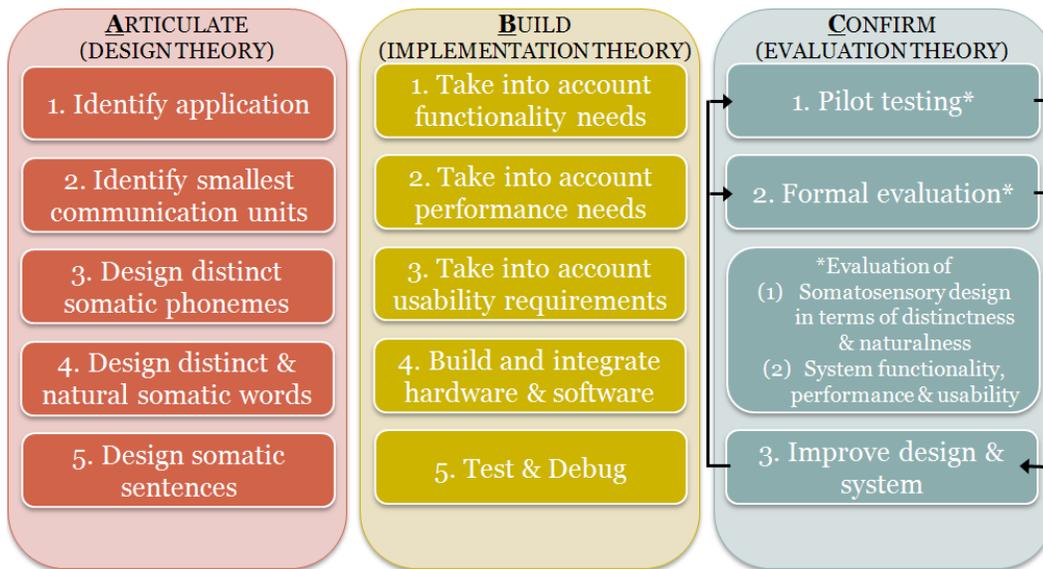
## Chapter 4

### SOMATIC ABC'S

The proposed theoretical framework, *Somatic ABC's* (figure 42), supports the design, implementation and evaluation of high level, intuitive, versatile, expandable, efficient, rich and robust somatic languages. There are three components to the process of creating a somatic language: *articulate*, *build* and *confirm*, or the ABC's of somatic language construction. Each is associated with an underlying design, implementation or evaluation theory, respectively. The proposed design theory guides the construction of the building blocks of somatic languages, and how to combine them into higher level constructs, eventual forming a somatic language that is both functional and practical. Implementation theory covers a more practical perspective of building and integrating actuators into an overall system, and system-level design considerations for functionality, performance and usability. Lastly, the proposed evaluation theory discusses how to effectively evaluate a somatic language for distinctness and naturalness: two attributes that are critical for somatic languages.

The proposed terminology (figure 43) is general to accommodate any modality of touch, from tactile to kinesthetic. *Somatic phonemes* (*somatemes*) are combined to create *somatic words* (*somatocepts*), which are combined to create *somatic sentences* (*somatences*). For the abbreviation of somatic word, the suffix '-cept' was inspired from a natural word's ability to evoke a general *concept*, that may be made more specific with context. These components make up our *somatic language* (*somatuage*). Terminology that targets specific touch modalities may also be used. For tactile stimulation (deformations or movement across the skin), the language building blocks become *tactile phonemes* (*tactemes*), *tactile words* (*tactacepts*), *tactile sentences* (*tactences*); all of which make up our *tactile language* (*tactuage*). For kinesthesia, we have *haptic*

*phonemes (haptemes), haptic words (haptocepts), haptic sentences (haptences) and haptic language (haptuage).* Both may be further narrowed if needed. For example, for vibrotactile stimulation, we have *vibrotactile phonemes (vibrotemes), vibrotactile words (vibrocepts), vibrotactile sentences (vibrotences), and vibrotactile language (vibrotuage).* Similar naming conventions may be used for other submodalities of touch including electro tactile, temperature, etc. Although the phrase *haptemes* has been coined before (Lahtinen, 2008), it was proposed for human-to-human interaction. And although the phrase *haptic phonemes* was first introduced for haptic icons (Enriquez, MacLean, & Chita, 2006), the proposed definition of a haptic phoneme, within the Somatic ABC's framework, is different from that defined by Enriquez, MacLean and Chita. In the following sections, each theoretical component is described.



*Figure 42.* Somatic ABC's theoretical framework to support the creation and evaluation of functional and practical somatic languages, and their integration into larger systems. The theoretical framework has been defined in general, rather than for a specific modality of touch, so that it may be applied to any type of touch-based stimulation, from tactile to kinesthetic. The framework consists of three theoretical components: articulate (design theory), build (implementation theory) and confirm (evaluation theory). Each theoretical components involves multiple steps that support and guide the creation of a somatic language.

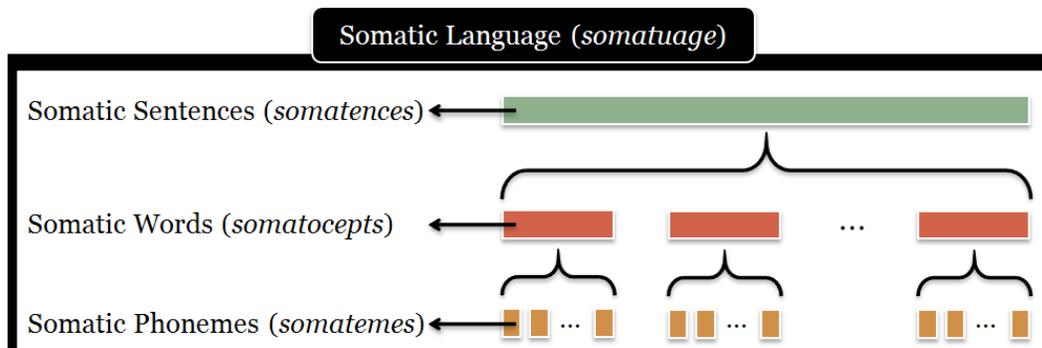


Figure 43. General terminology, defined in terms of somatic (body related) stimulation, to accommodate any modality of touch, from tactile to kinesthetic. *Somatic phonemes (somatemes)* are combined to create *somatic words (somatocepts)*, which are combined to create *somatic sentences (somatences)*. These components make up a somatic language (somatuage).

### Somatic Language Articulation

The first step of incorporating touch-based information delivery into a computerized system is describing and designing, herein referred to as *articulating*, a somatic language. To aid articulation, Somatic ABC's provides a theory of design involving five steps (figure 42): (1) identify application; (2) identify smallest communication units of application; (3) design distinct somatic phonemes; (4) design distinct and natural somatic words; and (5) design somatic sentences. The proposed design theory is inspired by natural, spoken language. Natural language is an integral part of our well-being providing an expressive communication tool that we utilize in just about every part of our lives from social interactions to acquiring or disseminating knowledge. It provides a means to communicate with others either directly through social interactions or indirectly through reading/listening to what others have wrote or recorded. Natural language is a promising candidate to use as a basis for a framework for somatic language creation given its versatility, richness and well-structured communication constructs (i.e., words and sentences). However, these attributes come with a price: the complex phonological, syntactical and grammatical rules that govern language use, combined with an arbitrary

association of meaning to words (with the exception of onomatopias), make learning a new language difficult.

Therefore, how might natural language inspire somatic language design, contributing its desirable properties while avoiding high learning curves? To achieve this, Somatic ABC's commonalities to natural language are metaphorical. Somatic ABC's design theory does not attempt to approach the complexity of natural language in terms of its phonological, syntactical and grammatical rules; nor does it attempt to approach the versatility and richness of natural language. It does, however, borrow metaphorical interpretations of natural, spoken language concepts; these similarities and contrasts are outlined in figure 44 where language, phonemes, words and sentences are compared. The creation of a somatic language begins at the highest level where the scope and needs of the somatic language are identified; next, a bottom-up approach is taken in which somatic phonemes, words and sentences are designed.

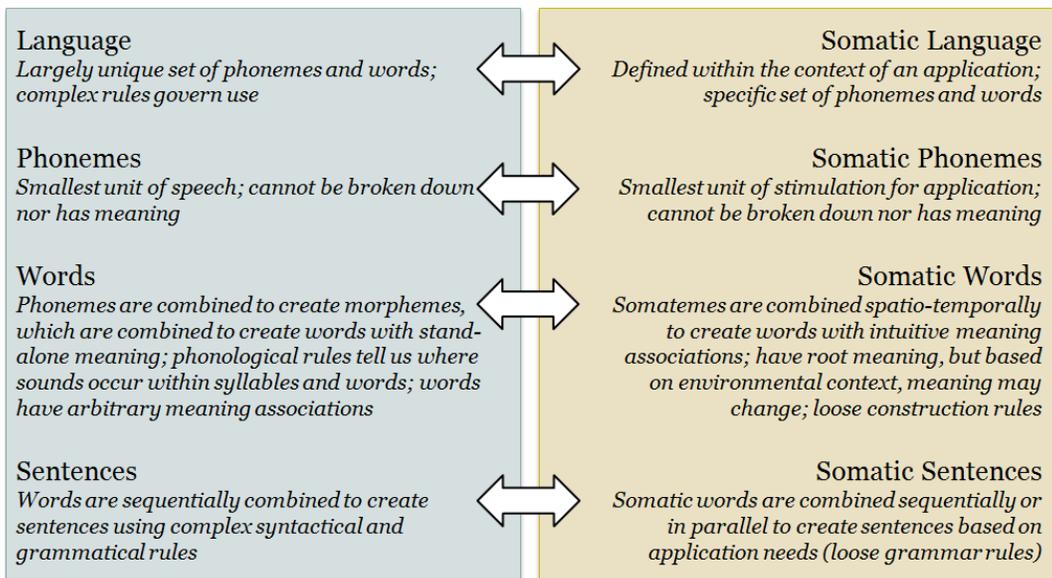


Figure 44. Somatic language defined in terms of a metaphorical interpretation of natural language. Similarities and contrasts are shown between natural language and somatic language for different communication components including phonemes, words and sentences.

**Somatic language.** A universal somatic language with no limitations regarding conceptual communication would be ideal, but such a language does not yet exist; and if it did, an arbitrary signal parameter association would most likely be needed, which as previously shown, requires extensive training and practice to learn. Instead, learning may be simplified through use of intuitive signal parameter associations, but at the cost that *each* application now has its own somatic language for touch-based delivery of information; this, however, is not very different from natural language in that most countries across the globe have their own unique language with words and phrases whose meaning stem from societal and cultural norms. Often within the same country, multiple dialects are be found. And just as languages borrow words and phrases from other languages, so, too, may somatic languages whose applications share similarities. Therefore, the first step in articulation is to *identify* the application for which the somatic language is intended for. Although this step seems trivial, it is nonetheless important as it defines purpose and scope.

Within an application, information to be delivered to the user via touch should be summarized in terms of *discrete communication units* that may be associated with a touch-based signal parameter; further, the *smallest* units of communication within an application are recommended to achieve small word vocabularies with expressive communication possibilities. This, however, is not a strict guideline, and higher level communication units may be selected.

Any application where conceptual information will be conveyed may be applicable to Somatic ABC's. Moreover, any application in which a continuous range of data may be discretized becomes applicable to the framework. This information delivery requirement limits applications to those with levels of abstractions above literal translations. For example, sensory substitution approaches are not applicable to Somatic

ABC's given that a continuous stream of input is directly presented to the user. On the other hand, a variety of applications do meet this requirement, enhancing the *versatility* of Somatic ABC's in that it can accommodate the creation of languages for a diverse range of applications. Somatic languages themselves can be versatile; communication units will largely determine versatility where use of the smallest communication units may provide the greatest applicability of the language.

**Somatic phonemes.** Metaphorically, in natural, spoken languages, phonemes are the *building blocks* of communication in that they are the smallest speech units used to form words. Natural languages across the globe have different sets of phonemes, and therefore, there exists no universal phonemic language. Similarly, when considering the extraordinarily large and diverse range of application domains in which somatic communication technology could be applied, achieving a universal set of somatic phonemes is not likely to be feasible. Rather, each application of somatic information delivery may utilize its own set of somatemes for word and sentence construction; and just as phoneme sets of certain natural languages may overlap and share similarities, so, too, may somatemes of similar somatic communication applications.

**Phonemes versus somatemes.** Phonemes and somatemes are similar in that they are metaphorical building blocks for spoken and somatic language, respectively, and either cannot be broken down into smaller unit (figure 44). Somatemes, like phonemes, do not have meaning until they are combined to create words. This is in contrast to Enriquez, MacLean and Chita's approach (2006) in which meaning is associated at the level of *haptic phoneme*, which they define as the smallest communication unit of a haptic stimulation. The issue of meaning association at the phonemic level (i.e., *meaningful phonemes*) is that it limits how phonemes may be combined to create useful, expressive word vocabularies. This is clearly demonstrated when attempting to create

more complex spatio-temporal words from intuitive phonemes: the overall spatio-temporal pattern should deliver a single meaning, rather than its individual components, which may be spatially and/or temporally complex themselves. The proposed somatemes lack an associated meaning to prevent any restrictions to word creation. As an example, consider vibrotactile stimulation in which a vibroteme is a localized vibration with a simple waveform and a fixed intensity, frequency and duration. In theory, a rich somatic language could be created from a very small set of distinct vibrotemes (or somatemes) when combined spatially and/or temporally to create a vocabulary.

*Designing somatemes.* During the initial stages of articulating a somatic language, a critical step is the design of distinct somatemes (figure 42); i.e., the smallest physical stimuli that will eventually form somatic words. During the previous steps of Somatic ABC's, an application has been selected, and its smallest conceptual communication units identified. These conceptual communication units represent the somatic words of the language. If the proposed somatic words utilize parameters such as spatial variations (body site), temporal variations (e.g., rhythm) or spatio-temporal variations, then only a small set (e.g., one or two) of somatemes may be needed. This is because parameters such as body site or pauses between stimuli (in the case of rhythm), don't affect the low level somateme parameters, which, depending on the modality, might include

- Speed, indentation and/or duration for tactile stimulation (pressure or movement across the skin) via tactemes.
- Frequency, amplitude, duration and/or waveform for vibrotactile stimulation via vibrotemes.
- Force, degrees of freedom, speed, duration, frequency, waveform and/or refresh rate for kinesthetic stimulation via haptemes. Here, body site is

defined in terms of not only *which* body part is moved, but also *how* it is moved. Therefore, a hapteme applied to the hand which causes the hand to move in different directions may all be considered the same hapteme but applied under different contexts.

If words are defined in terms of parameters that do affect signal components, e.g., intensity variations or burst duration variations, then more somatemes will be needed. Such word definitions, however, may be unlikely given that intuitive signal parameter associations are easier for higher order stimuli.

**Somatic words.** In natural, spoken language, phonemes are combined to create words where phonological rules guide their placement. Similarly, somatic words are created from temporal concatenation of somatemes (in addition to spatial presentation and/or spatio-temporal presentation of somatemes), but without complex rules such as those that enforce natural language. Rules that govern somatic words are largely dependent on the needs of the designer and application. Within Somatic ABC's, somatic words enable rich, efficient and robust delivery of information through touch.

**Vocabulary.** A set of somatic words is the vocabulary of a somatic language, which may be expanded with novel application-specific words, or those borrowed from other languages if they both are intended for applications that share similarities. Just as natural languages borrow words and phrases from each other, so, too, may somatic languages. Borrowed somatic words are referred to as *general words*. Somatic words may also be strictly intended for a specific-application, or limited to the aforementioned application by its design; these are therefore referred to as *application-specific words*.

For learnability and usability, vocabulary size should be small; but a small vocabulary size should not deter designers desiring rich, expressive somatic languages. Within the framework of Somatic ABC's, small vocabularies facilitate rich

communication possibilities through context and stimulation variations (described below). Moreover, the ease at which application-specific or general words may be added provides designers with an expandable vocabulary, which is further enriched through sentence construction.

In many instances, vocabularies may need to be augmented with stimuli that does not follow the proposed somatic language construction nor relate to the intended application. This is similar to natural language in which we have words, or rather *sounds* that are not made up of phonemes, but are still used to convey meaning nonetheless. In somatic languages, these “sounds” are referred to as somatic alerts given their usefulness for directing or “grabbing” the attention of a user; or indicating the beginning and/or end of a transmission.

***Using context to alter meaning.*** In natural language, a word typically conveys a general concept (person, object, event, etc.); it isn't until it is perceived in a specific context (e.g., a social setting or topic of conversation), and delivered with a specific intonation, that it conveys an exact meaning. This is also true for somatic words in that they convey general concepts until *felt* within a context. Here, context refers to the environment including the *body site* in which the stimulation is applied. That is, a somatic word has a general, conceptual meaning, but once applied to the body within a particular environment, a specific meaning is given.

***Stimulation variations.*** In spoken language, tonal variations are commonly used to ensure a word or sentence is successfully delivered (e.g., increasing loudness at a cocktail party), or to alter the meaning of a word or sentence (e.g., make its delivery sarcastic). Common tonal variations include changes in pitch (low versus high), loudness (volume or intensity changes), speed (duration) and rhythm (Crystal, 2007). In somatic languages, transformations that are the equivalent to tonal variations in natural language

are referred to as *stimulation variations*. It is important to note that stimulation variations are *not* affecting the original somatemes per se—but rather they conceptually alter the signal at the word level. Recall that somatemes are the smallest physical building blocks of communication within a somatic language, and therefore, they should not be *individually* altered to avoid confusion. Stimulation variations operate on whole words or sentences rather than at a phonemic level, and are designed to facilitate intuitive, relative recognition to simplify learning.

*Using stimulation variations to ensure successful delivery of a message.* In spoken language, given the environment or setting in which communication is occurring, we may need to introduce tonal variations to ensure the successful *delivery* of our message. Spoken words are communicated through the auditory channel, which may be noisy, in which case the loudness of a spoken word may be increased to ensure delivery. For somatic languages, the communication channel is the body, so we must be aware of:

- Sensitivity differences across the skin; if one body part is less sensitive compared to another, a more intense signal may be needed.
- Underlying tissue and bone beneath the skin; bone structures may inadvertently conduct stimulations.
- Surface area differences across the skin; in the case of vibrotactile stimulation applied to different body parts, the spacing and number of motors may need to change to accommodate variations in body part size and skin area.
- Range of motion and degrees of freedom of a limb; kinesthetic stimulation applied to one body part may not be applicable to another body part that varies with respect to range of motion, degrees of freedom, structure and joints.

Given these attributes of the communication channel, we may need to apply stimulation variations to accommodate spatial, structural and sensory variations. Moreover, these variations should be expected to occur across users: older users will have lower sensitivity compared to younger users, whereas some users may have smaller or larger body sizes compared to others. Designers should also expect perceptual differences across users that must be accommodated individually.

The environment may also interfere with the delivery of the signal. A common problem, particularly in portable systems, is ambient noise. Most cell phone users can attest to missing an incoming call when the vibrating ring tone is not felt—which commonly occurs while walking or riding in a vehicle. In these scenarios, ambient vibrations experienced while moving might be circumvented through an increase in vibration intensity to ensure successful delivery. This is akin to increasing the loudness of voice during a noisy cocktail party.

*Using stimulation variations to alter the meaning of a message.* In spoken language, tonal variations are often used to alter the meaning of a word or sentence; for example, the same word or sentence can be made to sound serious or sarcastic with subtle intonations. Likewise, *stimulation variations* can be used to change the meaning of somatic words or sentences. Stimulation variations should be applied to signal parameters whose variations are indicative of their respective meaning. For example, changing the tempo of a vibrotactile rhythm should intuitively convey the intended meaning carried with temporal variations.

Stimulation variations should not introduce significant demands in terms of learning and training—these may be achieved through natural, relative comparisons. In natural language, intonation is usually clearly perceived through relative comparison of prosodic variations. Likewise, in somatic languages, stimulation variations that alter

meaning should utilize the simplicity of relative recognition through comparisons with a *base signal*. For example, a base rhythm, followed by a rhythm with a noticeable tempo change, would simplify recognition.

***Designing somatic words.*** Referring to figure 42, the next step in articulation is the design of *distinct* and *natural* somatic words. As discussed, somatic words are not governed by complex phonological rules like the words of natural language. Although design rules are largely left to the designer based on the needs of the application, there are guidelines that should be followed in terms of *level of abstraction*, *signal parameter association* and *parameter value selection*.

*Level of abstraction.* Abstraction levels, which were previously discussed, are reiterated here in the context of articulation. Recall that the level of abstraction varies from *literal translation* (low level) to *symbolic* (high level). At the lowest level, literal translations are without abstraction; that is, information is conveyed directly to the same or alternative modality often after a cross-modal transformation that largely retains the original content. Symbolic representations utilize a higher level of abstraction to encode information in an often metaphorical, conceptualized form. It is paramount that somatic words utilize a high level of abstraction. Although literal translations provide a rich channel of information delivery, extraordinarily high learning curves are encountered due to sensory and perceptual differences between modalities. By abstracting the data stream, only a discrete set of communication units need to be learned. If these are distinct and natural, learning is improved. Moreover, for Somatic ABC's to be useful, applications must have identifiable, discrete communication units to enable word-level encoding; since literal translations communicate a raw, continuous input data stream, they do not meet this criterion. Therefore, *some* level of abstraction is needed to at least identify discrete words.

*Signal parameter association.* Recall that signal parameter associations range from *arbitrary mappings* (no relation between stimulation and its assigned meaning) to *intuitive mappings* (stimulation is representative of meaning, or even elicits intended meaning when felt). Signal parameter association has an important influence on the learning curve of somatic words: as part of the discussion on related work, learning curves for arbitrary mappings (Geldard F. A., 1957) (Enriquez, MacLean, & Chita, 2006) were noticeably higher compared to intuitive mappings (Chan, MacLean, & McGrenere, 2005). The aforementioned literature suggests that stimulations whose parameters intuitively encode meaning seem to support faster learning through their inherent naturalness. Therefore, intuitive signal parameter associations are critical if somatic words are to be easy to learn and use.

Which association type, arbitrary or intuitive, supports larger word vocabularies? Although large word vocabularies have been achieved with arbitrary signal parameter associations (Geldard F. A., 1957), extensive is often required training. Intuitive signal parameter associations may help build large vocabularies while reducing training time. Although no user studies have explored particularly large word vocabularies built from intuitive signal parameter associations, some relative large sets have been explored (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) exhibiting promising training times and accuracies. In practice, however, we must assume that there is a limit to the vocabulary size at which point impractical training times, extensive practice, and reduced distinctness between communication units will be encountered. Somatic ABC's circumvents this issue with promoting small word vocabularies that are just as rich and expressive as large vocabularies through use of contextual extensions and stimulation variations, as previously described.

*Parameter value selection.* Ultimately, whichever parameters of a signal are selected for encoding meaning should facilitate distinctness and naturalness. Concerning the latter, the selected parameters will largely depend on both the modality and application, but in general, spatial (body site), temporal (rhythm) and spatio-temporal variations (particularly, saltation) have proven useful for creating natural signal parameter associations in applications ranging from navigation to motor learning. Lower order parameters, such as frequency, intensity, duration and waveform for vibrotactile stimulation, may not be as useful for instilling natural meaning—with the exception of intensity for stimulation variations.

Concerning distinctness, careful attention must be paid to both somateme and somatic word articulation (as well as when designing stimulation variations) to ensure phonemes and words are distinct and recognizable from each other. Although somatic languages enforce no complex language rules, human psychophysical and perceptual data should be used as a guideline during articulation. For those parameters that will remain constant (ignoring, for a moment, stimulation variations), a good rule of thumb is to select a value that humans are most sensitive to. For example, our sensitivity to vibrations is maximum at 250 Hz (Verrillo R. T., 1963). For parameters that will be varied to convey meaning, unique values should be chosen such that they are perceptually separable. For example, vibrotactile stimulation's first order dimensions such as frequency, amplitude and duration, at first glance, seem to benefit from a wide selection of possible parameter values, but humans struggle to learn and identify more than a handful of frequency, amplitude or duration values (Geldard F. A., 1957).

Spatial, temporal (rhythm) and spatio-temporal parameters tend to provide more opportunities for achieving separable stimuli that are perceptually distinct. But given their diverse and extensive parameter value possibilities (configurations, ordering, timing,

etc.), spatio-temporal patterns may provide the most distinct and separable patterns. Care must still be exercised as spatio-temporal patterns that seem distinct during articulation, may be perceptually similar once delivered to users. Also, overlap and actuator sharing should be minimized as much as possible to further enhance separability (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011). If spatio-temporal patterns are to elicit certain perceptual illusions, such as apparent motion through saltation (Geldard & Sherrick, 1972), then existing design guidelines should be closely followed.

Lastly, robustness must be an integral part of articulation at the word-level. Previously, somatic alerts were discussed as a way to build in robustness. Higher order dimensions, such as spatio-temporal patterns, afford greater robustness than simpler dimensions. Regarding the latter, repeating a stimulus or long durations are obvious design approaches to ensure a message is noticed and perceived accurately. However, such methods are time consuming, and not practical for most real-time applications. Spatio-temporal patterns, on the other hand, may be expanded in ways that improve perceptibility while not excessively increasing time.

**Somatic sentences.** In natural language, words are spoken sequentially to create rich, expressive sentences governed by complex syntactical and grammatical rules. Somatic sentences, on the other hand, combine words sequentially or in parallel, and their construction largely depends on the needs of the application. How words are combined should be intuitive based on the application, and the spacing (pauses) between words should be sufficient (but not excessive) for accurate, timely perception of individual somatic words. Somatic sentences, combined with context and stimulation variations, have the potential to convey a rich content through limitless variations. The benefit of somatic sentences is that users need not learn each unique sentence; once words are learned, understanding their spatial arrangements and temporal concatenations should

quickly follow assuming individual words can be recognized easily and timely. The same is true for natural language in that sentences never heard or seen before can be easily understood assuming the receiver understands the individual words of the sentence and its context. Moreover, once a somatic word is learned irrespective of context, the specificity of its meaning will generally be straightforward to derive based on context and stimulation variations. This, too, is similar to natural language in which known words used in novel contexts or when delivered with familiar intonations (but unique to the word), are generally easy to understand.

Somatic words and sentences provide the efficiency needed for real-time use in a variety of applications. Compared to alphanumeric communication, conceptualized information delivery generally provides a faster means of presentation. Obviously, faster communication methods might exist once different modalities are considered (vision, hearing, etc.), but in terms of somatic communication speeds, presenting information at a conceptual level provides reasonable and practical transmission speeds. Even when considering touch alone, modality, of course, matters: particular modalities, such as vibrations, provide a more efficient communication channel compared to, say, temperature variations or chemical reactions due to presentation times and delays between stimuli presentation (Geldard F. A., 1957).

### **Building a System for Somatic Language Communication**

The first three building steps of Somatic ABC's (refer back to figure 42) should be performed simultaneously, taking into account *functionality*, *performance* and *usability* needs during the selection of actuators, and their integration into a larger system. Inspired by design requirements for vibrotactile wearable systems (Lindeman R. , Yanagida, Noma, & Hosaka, 2006) (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011), table 3 generalize these design requirements to any type of somatic information delivery system regardless of portability or wearability. The following sections describe each of the three aforementioned requirements for building somatic language communication systems. Once these requirements are taken into account during component (actuator, form factor, etc.) selection and integration planning, the remaining steps of the implementation theory of Somatic ABC's may be executed: hardware and software development and integration, followed by testing and debugging to ensure the system is operating as intended.

Table 3

*Functionality, Performance and Usability Requirements during the Construction of a Somatic Language Communication System.*

Functionality	Performance	Usability
<input type="checkbox"/> Expressive	<input type="checkbox"/> Durable	<input type="checkbox"/> Easy to learn
<input type="checkbox"/> Scalable	<input type="checkbox"/> Reliable	<input type="checkbox"/> Easy to use
<input type="checkbox"/> Reconfigurable	<input type="checkbox"/> Efficient	<input type="checkbox"/> Comfortable
<input type="checkbox"/> Portable*	<input type="checkbox"/> Long battery life*	<input type="checkbox"/> Discreet
	<input type="checkbox"/> Fast wireless communication*	<input type="checkbox"/> Easy to don/doff*
		<input type="checkbox"/> Doesn't hinder movement*

\*Applicable to portable systems (wearable or handheld)

*Note.* Criteria for functionality include *expressiveness, scalability, reconfigurability, and portability*. Criteria for performance include *durability, reliability, efficiency, long battery life* and *fast wireless communication*. Lastly, criteria for usability include *easy to learn, easy to use, comfort, discreetness, easy to don/doff, and doesn't hinder movement*.

**Functionality requirements.** After articulation, the first step is to choose a relevant actuator whose functionality meets the needs of the application and language design. An actuator should be selected that supports the modality and parameter values of the somatic language. *Expressiveness*, originally proposed by Lindeman, Yanagida, Noma and Hosaka (2006) within the context of vibrotactile displays, is generalized here to refer to an actuator that supports access to application required dimensions and values. Moreover, the interaction between dimensions must be taken into account to avoid unintentional parameter variations. For example, in standard DC vibration motor, frequency and amplitude cannot be varied independently due to hardware limitations—that is, changing one alters the other. If these parameters must be varied independently, more advanced vibration motors may be sought.

Scalability and reconfigurability have been recommended for vibrotactile displays (Lindeman R. , Yanagida, Noma, & Hosaka, 2006), and elaborated on for vibrotactile belts (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011), but may be generalized across any modality of touch. Scalability is the capability of actuators to be added or removed from a system without performance degradation. The importance of scalability depends on the application; specifically, how many body sites will receive tactile or kinesthetic stimulation, and how will this number vary. If multiple actuators are required with separate or simultaneous activation, the system must support increases in actuators. Reconfigurability refers to the ease of system modification including altering parameters values (via an Application Programming Interface) and changing the location of actuators on the body. Reconfigurability may be more useful for designers and developers, but its relevance to users should not be ignored, particularly when considering sensory, perceptual and body proportion differences across users where reconfigurability and adaptability may enhance usability.

Lastly, portability is another important criterion of functionality, but this attribute's relevance depends on the application. Desktop applications obviously do not apply here; portability is reserved for wearable or handheld systems that are intended to be used "on-the-go". With portability comes stringent performance and usability design requirements not found for desktop or stationary systems, described below.

**Performance requirements.** The performance attributes of table 3 are borrowed from vibrotactile belt design (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011), but generalized here for somatic information delivery systems of any modality. Real world use necessitates a rigid, durable device to prevent breakage. To ensure consistent, repeatable system operation, reliability is critical. And actuators must allow for efficient presentation speeds otherwise users will not have the patience to wait

for information to be delivered. Without these performance criteria, usability will decline. Lastly, for portable systems, a long battery life and fast wireless communication speed are advantageous and aid usability.

**Usability requirements.** In terms of usability (refer to table 3), somatic communication systems should facilitate the initial stages of familiarity and learning, and be user friendly to support and welcome continued use while minimizing frustrations. These attributes are attainable through Somatic ABC's high level of abstraction and intuitive signal parameter associations, as previously described. Wearable systems should be comfortable; and regardless of portability, if a system is used in public, it should be discreet in terms of physical appear and noise to avoid distracting people nearby (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011). For portable systems, Lindeman, Yanagida, Noma and Hosaka (2006) described a wearable vibrotactile system of limited cumber as one that is easy to don and doff, and doesn't hinder movement. These attributes apply to portable somatic information delivery systems in general.

### **Confirm Somatic Language through Evaluation**

After articulation and implementation, the somatic language must be evaluated for two important traits: *distinctness* and *naturalness*:

- **Distinctness:** Users should easily perceive the differences between somatic words to improve recognition time and accuracy, and reduce confusion and frustration. Somatic words that are distinct, even after context changes and stimulation variations, aid usability in terms of both ease of learning and ease of use.
- **Naturalness:** To further enhance learning, and potentially improve recall and reduce cognitive load, somatic words should be perceptually intuitive. Somatic words need not be natural to the extent of literal stimulations, but learning should be minimized in that it is quick and straightforward.

In addition to distinctness and naturalness, the aforementioned functionality, performance and usability design requirements must be confirmed:

- **Functionality:** Does the expressiveness, scalability and reconfigurability of the implementation satisfy the needs of the application? If the application requires a system that must be used “on-the-go”, is the implementation portable in either a handheld or wearable form factor?
- **Performance:** Does the durability, reliability and efficiency of the implementation satisfy the needs of the application? For portable systems, many applications require sufficient battery life and fast wireless communication—are these features present?
- **Usability:** Is the system easy to learn, and subsequently, easy to use? Is the system comfortable and discreet? If portable, is the system easy to don and doff, and by how much is movement hindered?

**Pilot and formal evaluations.** The aforementioned criteria and requirements should be evaluated in two stages: pilot testing, and then a formal evaluation. A pilot test is a quick, informal user study involving two to three participants. As a first step, the purpose of the pilot test is to obtain initial insight into the distinctness and naturalness of the somatic language. In this regard, participants should be representative of the target user population, unless difficult to procure. Moreover, experimental settings need not be completely realistic for initial test runs. On the other hand, for formal evaluations, enough participants of the target user population should be obtained to ensure representative data for analysis including significance testing. The experimental setting should be as close to those in which the system will be used, but it is often useful to begin in a controlled laboratory setting, and then progress to more complex test environments in subsequent formal evaluations. Both pilot and formal testing should involve three stages, described below, when assessing the psychophysical response of a somatic language design:

- Familiarization: Participants are introduced to the somatic language, and invited to feel communication units in an interactive, casual session.
- Training: Participants are randomly presented somatic words to recognize in a timely manner. The experimenter confirms correct guesses, and corrects misclassifications. Each participant progresses to the testing phase only after a certain level of performance is reached during a training trial, which is typically accuracy anywhere between 70-90% depending on the application. If performance is not reached, another training trial is repeated.
- Testing: the testing phase is similar to training with the exception that no feedback is given by the experimenter, and it is more extensive in terms

of the number of trials to confirm that participants have indeed learned and mastered the proposed somatic language.

To avoid excessive user study durations when evaluating both words and sentences, the training phase for somatic sentences may be skipped. Participants may be familiarized, trained and tested on somatic words, and then briefly familiarized with combinations of words (somatic sentences) before being tested since sentences should theoretically require no additional learning (beyond a brief familiarization).

**Objective evaluation.** Somatic word and sentence recognition accuracy, misclassifications and the number of training trials provide objective insight into both the distinctness and naturalness of somatic languages. A confusion matrix can help visualize which words were easy to recognize, which were difficult to recognize, and which were confused. For somatic words that lie on a continuous, but discretized, range of data, “off by one” misclassifications may be satisfactory depending on the application. Care must be taken to design a vocabulary where each somatic word is distinct; otherwise, stimuli will be confused resulting in reduced recognition accuracy, which can easily lead to confusion, frustration and reduced usability. Hesitation may be a likely sign that the proposed somatic words and/or sentences are not intuitive; therefore, response time during recognition should be recorded and assessed.

**Subjective evaluation.** A post-experiment questionnaire may be used to obtain (1) user feedback regarding usability criteria (table 3) via Likert scales; (2) an assessment of the naturalness of each somatic word by first creating an ordered list of those found most natural to least natural, followed by grouping those that might be described as ‘excellent’, ‘acceptable’ or ‘unacceptable’ in terms of naturalness. Somatic words rated ‘excellent’ are intuitive and further improvements would be insignificant; those rated ‘acceptable’ could be improved to enhance intuitiveness; and those rated ‘unacceptable’

need major improvement as they were not intuitive, possibly causing increased training time, hesitation and/or confusion. Note that distinctness influence naturalness: somatic words that are too similar may often be confused, reducing their intuitiveness. Another alternative to ordering somatic words in terms of naturalness is the use of mean ratings via Likert scales. Lastly, experimenter observations combined with user comments during and after the experiment may be useful for drawing connections between objective and subjective results.

Pilot tests to initially gauge functionality, performance and usability criteria are also recommended. Such quick tests will help with any preliminary adjustments regarding implementation: form factor, actuators, etc. After pilot testing is completed, design and implementation changes should be made before the formal evaluation begins. The importance of this step cannot be stressed enough: pilot testing will help reveal design and implementation flaws and shortcomings that must be changed or refined to reduce issues during formal evaluation.

**APPLICATION #1: AUDIO-HAPTIC DESCRIBED MOVIES**

As most content portrayed in a movie is visual, it is not surprising that movies are largely inaccessible to individuals who are blind or visually impaired. Although access to conversations, sound effects and musical scores enables partial comprehension, a lack of visual information (appearances, interpersonal interactions, facial expressions, etc.) prevents a viewer from completely interpreting and appreciating a film. The accessibility of visual cues may be improved through an audio description (Benecke, 2004), also known as descriptive video service (DVS), which is a narration that describes a film's visual content largely inaccessible by audio only. The descriptions of the narrator are added to the existing audio track while avoiding overlap with conversations, sound effects and, to some extent, musical scores.

Since first being developed in the 1970's by Gregory Frazier (Snyder, 2005), audio descriptions have proven useful for improving the accessibility of films, television, plays, museum tours and sporting events (Whitehead, 2005). Through a corpus-based analysis across 91 audio described films, Salway (2007) found the most frequently used words; and through contextual analysis, identified content commonly portrayed by these words. The most frequently used words fell into one of the following categories: actions, objects, scenes, characters and their body parts; and were utilized to convey information pertaining to characters' appearances, locations, interactions, emotional states and their focus of attention.

Audio descriptions have several major drawbacks given the diversity of the content they are designed to describe; for example, consider the diversity of films and television shows, which may vary in terms of genre, structure, scenes and characters. The

following list describes scenarios in film and television where audio descriptions are limited:

- As audio descriptions should avoid overlapping all dialogue within a film to avoid confusion and distraction, scenes with continuous dialogue present challenges during description given limited dialogue-free gaps. This is also true for scenes with abundant sound effects or musical scores that are important for understanding a film. For example, these might be sound effects that pertain to the actions of characters in absence of dialogue; or background music that sets the tone of a scene, or pertains to the emotional state of a character. Many films have extensive dialogue at least in particular scenes, and many television programs, such as soap operas, games shows and news programs are mostly dialogue.
- Fast paced films, such as action films or other genres that have action sequences, are difficult to narrate given the slow communication speeds of audio descriptions compared to visual depiction. Hence, during scenes with short sequences, each of multiple character actions or events, it is difficult to verbally describe all relevant information in the allotted time. Even when time is available during silent sections of a film (between scenes, between dialogue, etc.), audio descriptions are still abridged to fit within silent gaps.
- Although audio descriptions aim to convey only the most pertinent of visual cues that are critical to understanding and enjoying a film, some films require extensive audio descriptions due to their complexity and wealth of relevant visual content. In such scenarios, the viewer is aurally overloaded with verbal descriptions, making the movie viewing experience tiring (Benecke, 2004).

- Lastly, given the limited time in which an audio description is presented, verbal descriptions are abridged; but descriptions usually provide enough information to acquire a vague idea of what is happening in a scene. Often, however, more details are useful for aiding interpretation and visualization of scenes within a film. For example, movements and positions of people and objects are commonly found in audio descriptions, but this information is presented relatively, losing accuracy—for example, “John enters the room” or “Mike stands next to Susan”. As another example, the rich, communicative expressions of the face are summarized such as “Doug smiles” or “Julia frowns”. In these scenarios, the richness of the visual content is lost, which could attribute to poor understanding and visualization as well as reduced enjoyment.

To communicate the aforementioned visual cues missed by audio descriptions, an alternative modality may be employed. Given that vision is unavailable, and hearing perhaps overloaded, touch offers a promising channel for presenting information during movie viewer. Several approaches have been explored, described in the following section, for enhancing the realism of movies with veridical touch sensations to convey character experiences and emotions. Although these approaches may improve the experience and entertainment of movie watching, movie comprehension may still be challenging without access to a film’s content such as visual, non-verbal cues.

The following presents an overview of opportunities where haptics can augment audio described films to overcome the limitations of narrated video media. In particular, the aforementioned scenarios where audio descriptions are limited are revisited:

- During situations in which the auditory modality is not available for receiving audio descriptions, such as during continuous dialogue, sounds effects or music,

pertinent information may be offloaded to touch. Care must be taken when haptic descriptions overlap with audio (discussed later).

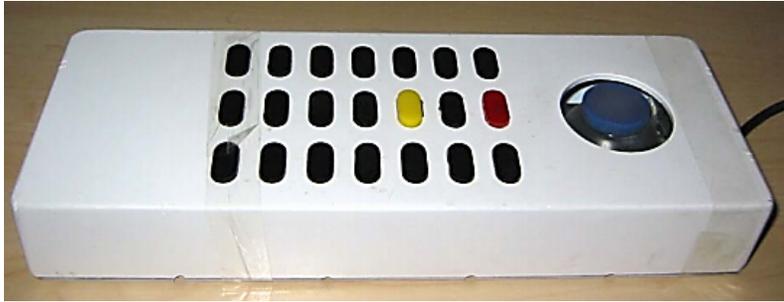
- Although the bandwidth of touch is not as high as vision, utilizing both haptic and audio descriptions during fast paced scenes may enable the presentation of more relevant visual content compared to using only one modality. For example, during an action sequence, audio descriptions might communicate the actions of characters (e.g., “John begins to run away”) while haptic descriptions convey movements (e.g., a vibration patterns indicating the direction in which John is running).
- For films that overwhelm users with seemingly continuous verbal description of visual content, information may be offloaded to the sense of touch to ease the burden on the viewer’s auditory modality.
- Lastly, haptic descriptions can complement audio descriptions by conveying additional, but relevant information. For example, audio descriptions tend to convey movements and positions relatively. However, to accurately visualize a scene, and appreciate the rich social interaction dynamics that occur, knowledge of detailed character positions and movements are useful. In this regard, haptic descriptions may convey more precise positions and movements using the surface of the skin. Other visual, non-verbal cues, such as facial expressions, may be made more accessible by providing further detail through touch.

As a first step toward these goals, the position of onscreen characters were targeted in terms of (1) their location across the screen; (2) their distance from the camera; and (3) their movement across the screen. The positions and movements of characters were associated with vibrations delivered around the waist using a belt of vibration motors. Vibrations felt around the waist are known to be intuitive for conveying

directional information (van Erp J. B., van Veen, Jansen, & Dobbins, 2005). Moreover, a useful method for conveying distance-based information via touch is through tactile rhythms (McDaniel T. L., Villanueva, Krishna, Colbry, & Panchanathan, 2010) where temporal patterns of vibrotactile pulses represent different conceptual distances such as close, middle and far.

### **Related Work**

Haptic feedback has been extensively explored toward enhancing the realism and immersiveness of virtual reality, simulations and gaming, through kinesthetic and tactile stimulations that mimic or relate to those found in reality. A well-known example is force-feedback joysticks and vibrating controllers for gaming used to enhance enjoyment through realism and immersion. Within the last decade, researchers and designers have begun to explore haptically augmenting video media, in particular, films and television shows. O'Modhain and Oakley (2003) proposed interactive television, or *Touch TV*, and presented criteria for such systems including ease of use and integration into existing television use; rich haptic feedback for versatility; and affordability. They developed a haptic remote control with a two degree-of-freedom knob (figure 45), which was used to enhance cartoons by enabling viewers to feel onscreen activity (e.g., the buzz of a bee) as well as interact with the visual content (2004). O'Modhain and Oakley termed this *presentation interaction* in which viewers can alter the presentation of content, but not its structure. For example, in one cartoon, viewers watched as a character rode a bee, and felt the movements of the bee across the screen via a haptic knob. By interacting with the knob, viewers could influence the bee's movements.



*Figure 45. Haptic remote control for adding haptic feedback and interactivity to video media. Reprinted from “Adding interactivity: Active touch in broadcast media,” by O’Modhrain, S., & Oakley, I., 2004, In Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 293. Copyright © 2004 by IEEE. Reprinted with permission.*

Gaw, Morris and Salisbury (2006) developed an authoring tool for adding haptic feedback to video media. The authoring environment allowed playback of video content at variable speeds while recording the movements of a haptic (force-feedback) device with the option of adding and editing force vectors to create sharp impulses where needed (e.g., to simulate an impact). While watching the authored movie, a viewer holds the joystick of the haptic device, and experiences movements and other actions related to onscreen characters.

Lemmens, Cromptvoets, Brokken, van den Eerenbeemd and de Vries (2009) were the first to explore the use of wearable haptic technology for eliciting the emotions of onscreen characters to enhance immersion. They proposed a tactile jacket consisting of 64 vibration motors evenly distributed across the torso and arms (figure 46). The basis of their approach is that the bodily reactions that accompany emotions (e.g., those experiencing fear will often feel chills down their spine) may be simulated through spatio-temporal vibration patterns, and used to elicit their respective emotions. A user study was conducted to explore if simulated bodily reactions enhance a movie viewer’s emotional immersion. Participants watched seven different movie clips (each targeting a specific emotion such as love or fear); their presentation was randomized, but clips

without actuation were always presented before those with augmented vibrotactile patterns. The patterns were drawn from a set of 40, each inspired by idioms or interactions found in social touching. Questionnaire results revealed that participants experienced greater immersion and emotional response when viewing clips with haptics, although potential order effects need to be investigated.



*Figure 46. Movie enhancing tactile jacket with 62 vibrotactile actuators distributed across the torso and arms. Reprinted from “A body-conforming tactile jacket to enrich movie viewing,” by Lemmens, P., Crompvoets, F., Brokken, D., van den Eerenbeemd, J., & de Vries, G. J., 2009, In Proceedings of the Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, p. 7. Copyright © 2009 by IEEE. Reprinted with permission.*

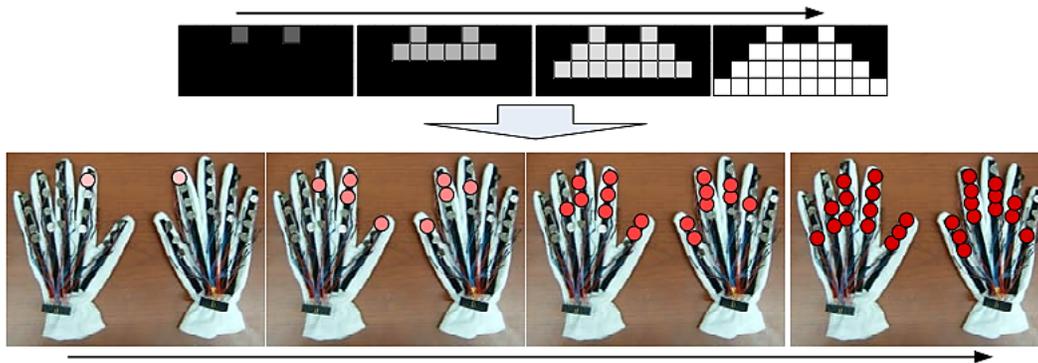
Cha, Oakley, Ho, Kim and Ryu (2009) proposed a framework for encoding, decoding and broadcasting haptic media in MPEG-4 videos. Haptic sensations as part of the movie viewing experience included passive spatio-temporal tactile stimulations (what they called *linear haptic media*); and active haptic exploration of 3D objects and surfaces (what they called *nonlinear haptic media*). The framework is composed of three components: content authoring during which haptic media is created; transmission of media, such as streaming over a network; and user viewing and haptic interaction during which a viewer feels linear haptic media through a glove-based system (figure 47); and

haptic interactivity is mediated through a hybrid haptic device (figure 48). The glove-based system consisted of two wireless gloves each with vibration motors attached to the back of the fingers and hand. Tactile content is authored on top of the existing video using an authoring/editing tool (Kim, Cha, Oakley, & Ryu, 2009) in which lines may be drawn onscreen using tactile brushes of different sizes (where size is related to vibration intensity). This input is used to create a low resolution tactile video where each “pixel” is mapped to a vibration motor on the glove—that is, 10x4 “pixels” of the tactile video are mapped to 10x4 vibrators on the glove. A force-feedback device may be used to mediate haptic interactivity; Cha et al. utilized a hybrid device combining force-feedback with tactile stimulation delivered to the tip of a finger.

Rahman, Alkhaldi, Cha and Saddik (2010) proposed authoring YouTube videos with tactile content to be displayed by a vibrotactile jacket embedded with vibration motors. They created an authoring/editing environment where spatio-temporal vibration patterns are specified to create a tactile video (a low resolution grid of tactile pixels where intensity is depicted by brightness). The haptic media is then converted into XML, and embedded into any YouTube video. Anyone with a tactile display can feel the haptic stimulations embedded in the video while viewing its synchronized audiovisual content. The authored haptic content could represent experiences of those onscreen; for example, the impact of a boxer being punched.

Among the aforementioned approaches that have been proposed, those using force-feedback devices might have too high a cost for the average consumer. In this case, tactile displays present a promising alternative. The vibrotactile displays and authoring/editing tools of Lemmens et al. (2009) and Cha et al. (2009) could be used to display both emotional content (what characters are experiencing) as well as non-verbal cues such as position and distance of onscreen characters (the focus of this work).

Lemmens et al. focused on emotional content, and did not explore character position and distance. Cha et al. explored movement to some extent, but thorough psychophysical testing was not conducted to learn how well users can localize vibrations as they are associated with onscreen positions. Moreover, communication of distance of objects in a scene was not explored. Lastly, these systems were not geared toward individuals who are blind or visually impaired—a user population who experiences movies very differently compared to sighted movie goers; in particular, the integration of haptic media with both the movie *and* audio descriptions must be investigated.



*Figure 47. Tactile stimulation felt over time using a glove-based system for experiencing authored tactile video. Red dots represent actuation, and their intensity correlates with vibration intensity. Reprinted from “A framework for haptic broadcasting,” by Cha, J., Oakley, I., Ho, Y.-S., Kim, Y., & Ryu, J., July-Sept. 2009, IEEE Multimedia, 16(3), p. 21. Copyright © 2009 by IEEE. Reprinted with permission.*

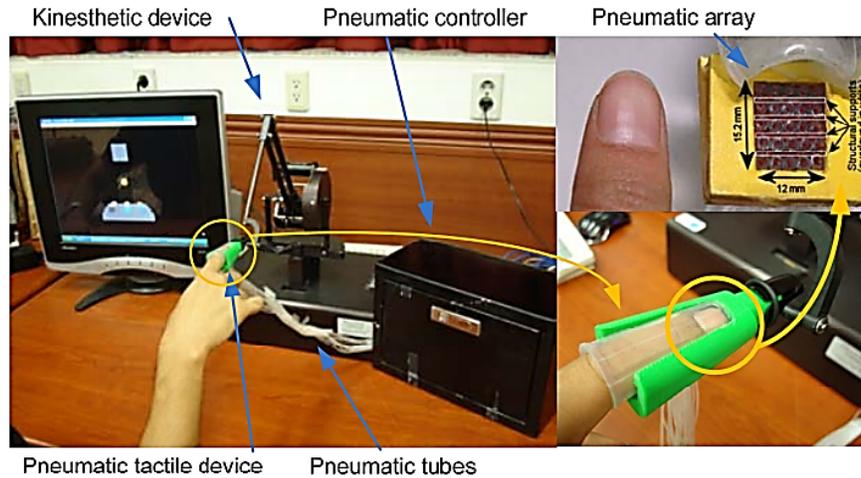


Figure 48. Hybrid haptic device for experiencing haptic movies using a force-feedback device combined with a tactile display for the finger. Reprinted from “A framework for haptic broadcasting,” by Cha, J., Oakley, I., Ho, Y.-S., Kim, Y., & Ryu, J., July-Sept. 2009, *IEEE Multimedia*, 16(3), p. 24. Copyright © 2009 by IEEE. Reprinted with permission.

### Proposed Approach

This section presents the proposed somatic language for communicating the non-verbal cues of position, distance, and movement of characters within movie scenes. The details of applying the theories of Somatic ABC’s to design, develop and evaluate the proposed somatic language are covered including discussions related to design and performance criteria.

**Articulate.** The application was identified as augmenting audio described films with haptics; in particular, complementing audio descriptions with vibrations to convey the position, distance and movements of characters across the screen. The scope of the proposed language will be limited to communicating the position and movements of one character at any given moment, although scenes may involve multiple characters. Moreover, as a first step, scenes will be limited to those with dialogue with limited movements involving two or three characters. More complex scenes with fast paced movements (such as action scenes) will be explored as part of future work.

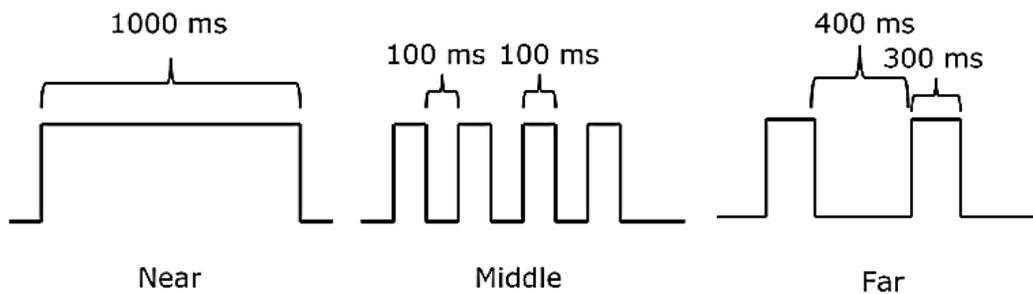
For any given scene, positions and movements from one camera perspective will be conveyed, rather than changing the perspective of the somatic communication system each time camera perspective is altered (which could be multiple times within a single scene). This is similar to how audio descriptions setup a scene. Lastly, the communication units of the proposed somatic language, which will be used to form a haptic description for films, may overlap with audio descriptions and/or audio content from the film itself. A film with both an audio description and complementary haptic description (*audio-haptic description*) will be referred to as an *audio-haptic described film*.

The qualitative distance of a character from the camera provides a useful choice of communication units in that it ensures a small word vocabulary for simplified learning. Three distances are proposed: *close*, *middle* and *far*. While a finer discretization is possible, more distances may not enhance the visualization of a scene and complicate learning with a large vocabulary. Recall that high level constructs, although not as rich as low level representations, simplify learning. Intuitive signal parameter associations must also be achieved. To meet this requirement, tactile rhythms (repetitive temporal variations of vibrotactile stimulation) were used given their success at communicating interpersonal distances (McDaniel T. , Krishna, Colbry, & Panchanathan, 2009) (McDaniel T. L., Villanueva, Krishna, Colbry, & Panchanathan, 2010).

The next step in Somatic ABC's is the design of distinct vibrotemes. To create three distinct rhythms (representing the words of the language), three distinct vibrotemes were selected: short vibrotactile pulses of duration 1000 ms, 300 ms and 100 ms. Because a higher order dimension, namely rhythm, was utilized in word creation, alteration of low level vibroteme parameters was avoided, and so only a few vibrotemes were needed.

Each vibrotome varies with respect to duration—vibration frequency, amplitude and waveform remain constant across these stimulations.

Using the aforementioned vibrotomes, three distinct and intuitive vibrotactile words are proposed, depicted in figure 49; these words represent the three previously described communication units of conceptual distance. As shown from the structure of these rhythms, distinct representations were sought through varying both burst and gap duration between the rhythms. Indeed, extensive pilot testing revealed perceptual distinctness. Natural vibrotactile words were sought through a design influenced by common radar systems in which the length of pauses between audible beeps indicate the distance of an approaching threat. If the threat is far, beeps are spaced far apart; as the threat approaches, the pauses shorten until the stimulus becomes a continuous beep.



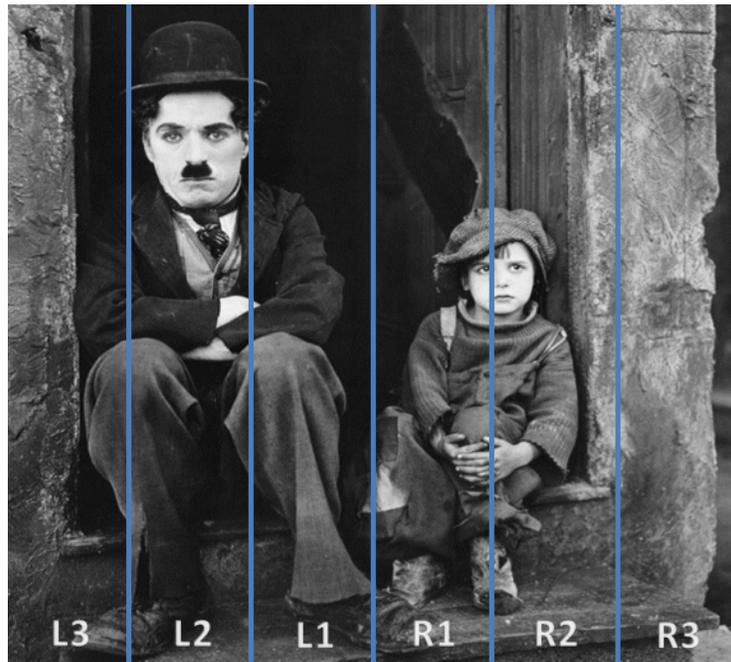
*Figure 49.* Proposed vibrotactile words for communicating the distance of a character from the camera for use in audio-haptic descriptions. Each rhythm is one second in duration. The rhythm representing a distance of near is a steady vibration; the rhythm representing a distance of far consists of well separated bursts of short duration; and the rhythm representing a distance of middle falls between these extremes: it consists of very short bursts presented in rapid succession. These rhythms are modeled after radar systems where as a threat becomes closer to a target, the tempo of audible beeps increases until steady. *Reprinted from “Audio-haptic description in movies,” by Viswanathan, L. N., McDaniel, T., & Panchanathan, S., 2011, In C. Stephanidis (Ed.), HCI International 2011 – Posters' Extended Abstracts (p. 417), LNCS 173, Berlin, Heidelberg: Springer. Copyright © 2011 by Springer Berlin Heidelberg. Reprinted with permission.*

To enrich the proposed vocabulary, Somatic ABC's use of context was employed; in particular, vibrotactile words felt at different body sites represented

different locations across the screen. Therefore, by combining stimulation at a specific body site with tactile rhythm, a more accurate position of characters within a scene may be presented. By varying these stimulations both spatially and temporally, complex character movements may be conveyed. To mediate the presentation of vibrotactile words through body context, a vibrotactile belt offers a promising communication modality. Vibrations around the waist have been shown to intuitively convey directional information—e.g., where to move next for navigation applications (van Erp J. B., van Veen, Jansen, & Dobbins, 2005), or where people are standing in social interactions for aids for the visually impaired (McDaniel T. , Krishna, Balasubramanian, Colbry, & Panchanathan, 2008). To ensure accurate localization, a limited number of vibration motors were used; therefore, a discretization of positions across the screen was needed. Six regions (figure 50) were chosen as they provided a good tradeoff between resolution and ease of recognition. Each region maps to a vibration motor around the waist.

Another design decision relates to the placement and spacing of vibration motors around the waist. Accurate localization of vibrations is needed to ensure ease of use and low cognitive load. Cholewiak, Brill and Schwab (2004) explored vibrotactile localization around the waist using vibrotactile belts varying in terms of the number of motors, and their placement and spacing. They found that vibrotactile stimulation near anatomical reference points were more easily localized compared to other sites. Moreover, they found that end points also simplified localization. Therefore, the placement of motors within the proposed belt design incorporated two vibration motors near the navel (one slightly to the left, *L1*, and the other slightly to the right, *R1*); one at each side (*L3* and *R3*) with only one neighboring motor; and to further enhance resolution while maintaining satisfactory localization accuracy, a motor between *L1* and *L3*, and between *R1* and *R3*, were added. Since the motors are associated with a linear display

(characters across a screen), *L3* and *R3* were placed slightly before the sides to lessen the “curved” feel of the display. Essentially, vibrations felt on the left side of the viewer’s waist correspond to a character on the left side of the screen, and vice versa.



*Figure 50. Division of screen into six regions of equal width for audio-haptic described movies. Regions are labeled for reference. Adapted from the Wikimedia Commons: Chaplin\_The\_Kid.jpg, [http://en.wikipedia.org/wiki/File:Chaplin\\_The\\_Kid.jpg](http://en.wikipedia.org/wiki/File:Chaplin_The_Kid.jpg).*

Individual vibrotactile words were temporally combined to create sentences representing the movements of characters across the screen, and/or away/toward the camera. Within a vibrotactile sentence, words presented to the same motor—that is, variations in distance only—were separated with a 100 ms gap. For vibrotactile sentences where words occur across different motors, no gap was necessary between subsequent words. A gap of at least one second was introduced between sentences to separate movements. These design choices were found to work well during pilot tests. Although a small vocabulary was utilized, through sentence creation and context, a rich and

expressive somatic language was formed; and since communication is at a high level of abstraction—in particular, communication of concepts related to position, distance and movement—communication of this information is efficient and capable of keeping up with regular playback speeds of films as verified through extensive testing.

Regarding versatility, the proposed design applies to any video media involving movements belonging to characters or objects. The design may also be applicable to social interaction assistants for the blind or visually impaired (McDaniel T. , Krishna, Balasubramanian, Colbry, & Panchanathan, 2008). Lastly, given the inherent redundancy in movements—e.g., with simple linear movements, the most important characteristics are the start and end points—the proposed somatic language was found to be robust through pilot testing. Even with more complex movements, not all vibrotactile words need to be accurately perceived to understand movements of characters.

**Build.** After articulating a somatic language to design audio-haptic descriptions, the stimulations were implemented in a custom vibrotactile belt for information delivery through a custom audio-haptic movie viewer (figure 51 depicts the system setup). The audio-haptic description system was built under the guidance of Somatic ABC's implementation theory. Design and performance requirements were identified and closely monitored during construction. First, the design and implementation of the vibrotactile belt and its software are described, followed by a description of the audio-haptic movie viewer software.



*Figure 51.* System setup for viewing audio-haptic described movies. Setup involved a personal computer, headphones and custom vibrotactile belt. Audio and haptic media are delivered through headphones and vibrotactile belt, respectively.

***Vibrotactile belt.*** This section presents the design and implementation of a custom vibrotactile belt for displaying haptic media. Additional hardware and software details can be found elsewhere (Edwards, et al., 2009) (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011).

***Hardware description.*** The system architecture of the proposed vibrotactile belt is depicted in figure 52. The system consists of three main components (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011): control module, tactor module and the belt itself. The belt is made of flat nylon webbing (1.5 in. by 8 in.) worn by adjusting its length through a buckle, which simplified donning and doffing. A buckle-based implementation allowed for “one-size-fits-all” wearability and comfort, which also helped tactor modules maintain close contact with the waist; this is in contrast to Velcro-based implementations which are geared toward specific waist sizes and often loosen during the duration of individual uses. The belt form factor provides a naturally discreet device in that it integrates well with existing wardrobes.

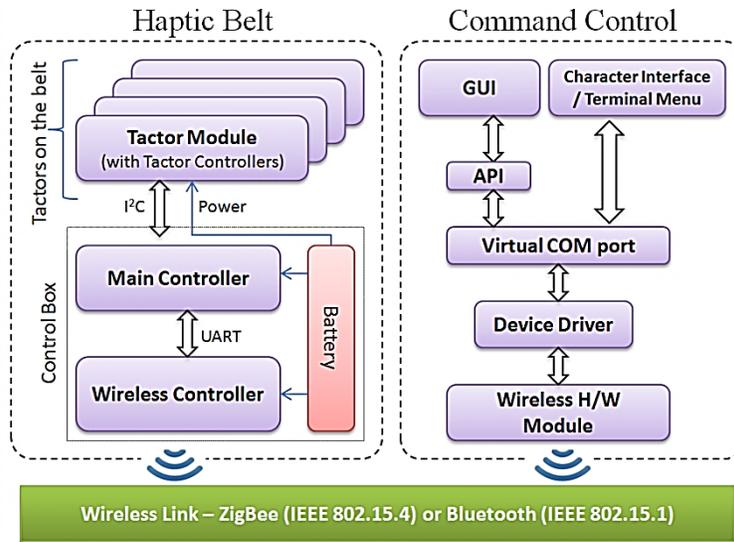
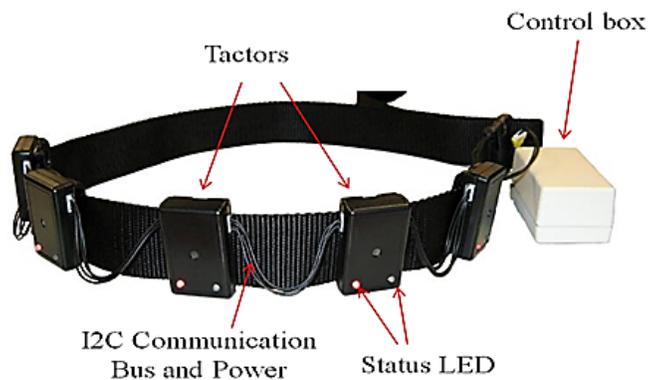


Figure 52. System architecture of vibrotactile belt and command control. Reprinted from “Design, implementation, and case study of a pragmatic vibrotactile belt,” by Rosenthal, J., Edwards, N., Villanueva, D., Krishna, S., McDaniel, T., & Panchanathan, S., 2011, *IEEE Transactions on Instrumentation and Measurement*, 60(1), p. 117. Copyright © 2011 by IEEE. Reprinted with permission.

The control module of the vibrotactile belt consists of a microcontroller (Arduino Funnel IO with ATmega168); wireless module (Bluetooth, IEEE 802.15.1); power supply; and enclosure (3.15 in. by 1.58 in. by .79 in.) with a pocket clip to easily attach onto belt and slide into position. It provides fast, reliable, long range wireless connectivity between the vibrotactile belt and a personal computer. The power supply is a small, rechargeable Polymer Lithium Battery (3.7 V, 800 mAh) with a long battery life—specifically, up to six hours of continuous use when fully charged.

The tactor modules of the vibrotactile belt consist of a microcontroller (Atmel ATtiny88); vibration motor; and enclosure (2.125 in. by 1.375 in. by 0.58 in.) also with a pocket clip. The vibration motor is a coin vibrating motor with a diameter of 12 mm; when the system is powered, vibration motors run at a frequency of 150 Hz. The enclosures of the control and tactor modules assist with system durability and rigidity.

Functionality design criteria of expressiveness, scalability and reconfigurability motivated two key design choices: on-board management of vibrotactile stimulation by individual tactor modules, and a “plug-and-play” style I<sup>2</sup>C communication bus. Communication between the control module and tactor modules is mediated by an I<sup>2</sup>C bus of four wires: two for power, one for data and one for clock. Tactor modules may be plugged into (or removed at) any place along the I<sup>2</sup>C bus with up to 16 tactors supported for useful scalability. At startup, bus addresses are dynamically assigned, enhancing reconfigurability as novel arrangements may be created depending on the requirements of applications. The tactor modules themselves manage storage and processing of activation commands sent by the control module, allowing efficient use of the control module’s processing time. Expressiveness is achieved through the versatility of the vibrations: different body sites may be stimulated based on which tactors are actuated; timing variations may be used to create unique tactile rhythms; pulse-width modulation may be used to vary vibration intensity; and lastly, these individual dimensions may be combined to create rich spatio-temporal vibrations. The final version of belt is shown in figure 53.



*Figure 53.* Vibrotactile belt implementation depicting tactor modules, control module, communication bus and belt. The status LEDs on each tactor module were used for debugging efficiency. Reprinted from “Design, implementation, and case study of a pragmatic vibrotactile belt,” by Rosenthal, J., Edwards, N., Villanueva, D., Krishna, S., McDaniel, T., & Panchanathan, S., 2011, *IEEE Transactions on Instrumentation and Measurement*, 60(1), p. 117. Copyright © 2011 by IEEE. Reprinted with permission.

*Software description.* The firmware of each factor module receives and processes commands from the control module. The control module's firmware was designed to allow for maximum reconfigurability: it provides functions for creating new belt configurations, and storing and using existing user-defined spatio-temporal vibration patterns. These patterns are created through a graphical user interface (GUI) designed with learnability and usability in mind. The GUI was implemented on both a desktop computer and portable platform (PDA), and provides basic functionality in terms of connecting/disconnecting to the belt, creating patterns, and storing/activating patterns on the belt. Figure 54 depicts the GUI on the portable platform (left) where patterns are created using dropdown selections; in the right of the figure, a tactile rhythm authoring tool was created to simplify authoring of haptic patterns.

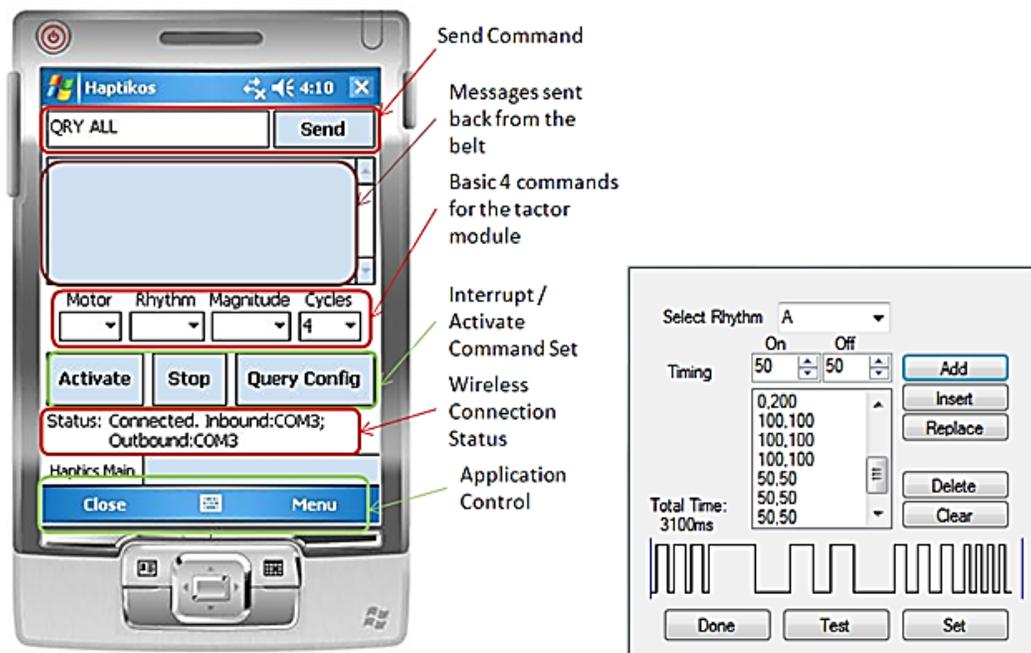


Figure 54. Graphical user interface of vibrotactile belt command console on PDA (left), and tactile rhythm authoring tool (right). Reprinted from “Design, implementation, and case study of a pragmatic vibrotactile belt,” by Rosenthal, J., Edwards, N., Villanueva, D., Krishna, S., McDaniel, T., & Panchanathan, S., 2011, *IEEE Transactions on Instrumentation and Measurement*, 60(1), p. 119. Copyright © 2011 by IEEE. Reprinted with permission.

**Audio-haptic movie player.** For loading and playing audio-haptic described films, a custom viewer was implemented (figure 55). The viewer was developed in C#, and uses a DLL to connect to and send commands to the belt via an Application Programming Interface. As shown in the figure, the GUI provides options for connecting/disconnecting to the belt; loading different movies or movie clips; pausing or stopping playback; and toggling haptic description on/off.

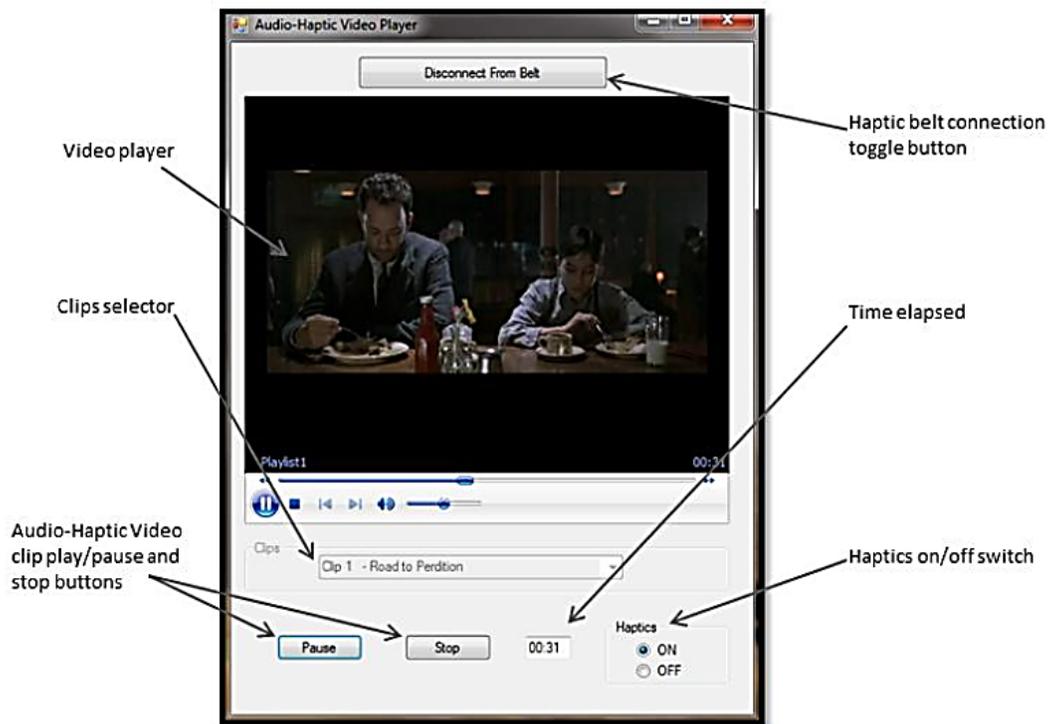


Figure 55. Graphical user interface of audio-haptic described movie player. Reprinted from "Enhancing movie comprehension for individuals who are visually impaired or blind," by Viswanathan, L. N., 2011, Thesis (M.S.), Arizona State University, p. 88. Copyright © 2011 by Viswanathan. Reprinted with permission from author.

**Authoring.** Films are manually authored with haptic descriptions by creating a haptic track on top of the existing video and audio track. The haptic track consists of timing information and actuation commands. When designer-specified points along a film's timeline are reached during playback, stored actuation commands are sent to the belt. Haptic descriptions comprise the aforementioned vibrotactile words to convey character locations across the screen, the relative distance of characters from the camera, and movements of characters within a scene. In particular, haptic descriptions communicate which vibrotactile words (stored in the control module's firmware) to present along the timeline of the described film. Sequentially presenting words form vibrotactile sentences without incurring additional storage space in the control module's memory.

**Confirm.** The proposed somatic language was evaluated through a user study conducted in collaboration with Lakshmie Narayan Viswanathan as part of his Master's Thesis (2011). The study was approved by ASU's Institutional Review Board. The aim of the study was to assess the proposed haptic descriptions for complementing audio described movies. This experiment constitutes the formal evaluation described as part of Somatic ABC's evaluation theory. Extensive pilot testing was conducted during design and implementation, the results of which influenced the design of the final system evaluated here.

**Subjects.** Ten participants (five males and five females) were recruited for this study. Each participant was awarded a monetary compensation of \$25 for participating. Of the ten, four participants were totally blind, four were legally blind with low vision, and two were visually impaired with low vision. Ages ranged between 20 and 65 with the following breakdown: four were between the ages of 20 and 29, one was between the ages of 30 and 39, four were between the ages of 40 and 49, and one was between the age

of 50 and 59. Nine of the ten participants stated that they watched movies with audio descriptions. All participants stated that they watched movies; average number of movie viewings per year was estimated at a little over one hundred films.

*Materials and apparatus.* Experimental equipment consisted of the custom vibrotactile belt, custom audio-haptic movie viewer, and stereophonic headphones. Seventeen audio described movies were selected, and from each film, a single clip was chosen for haptic authoring. Clips were selected to satisfy the following: character movements within a conversational scene involving a maximum of three characters. Clips had an average duration of 2 minutes. The majority of films fell under the genre of drama although action and comedy were also present. Clips did not contain camera movements. Table 4 provides a summary of the clips selected. For each clip, a haptic track was created; initial positions of characters were encoded, and any subsequent movements were encoded. Initial positions were presented during character introductions by the audio description. For example, “John enters the room, and walks across” would be accompanied by a tactile rhythm for John’s initial position, followed by spatio-temporal variations across the waist as John walked across the room. Participants indicated the number of times each of the seventeen movies had been viewed, and how well they remembered them using a 5-point Likert scale. For each participant, twelve clips were used for the study. These twelve were selected from those films participants had not seen—if this was not possible, the least remembered films were selected.

Table 4

*List of Clips from Audio Described Movies for User Study*

<b>Film Title</b>	<b>Start Time</b>	<b>Duration</b>
<b>Road to Perdition (2002)</b>	<b>01:15:47</b>	<b>1 min, 10 s</b>
<b>Munich (2005)</b>	<b>01:31:27</b>	<b>1 min, 56 s</b>
<b>(500) Days of Summer (2006)</b>	<b>00:09:52</b>	<b>0 min, 53 s</b>
<b>The Ultimate Gift (2006)</b>	<b>00:24:04</b>	<b>1 min, 7 s</b>
<b>Cinderella Man (2006)</b>	<b>00:09:50</b>	<b>1 min, 40 s</b>
<b>Blind Dating (2006)</b>	<b>00:12:02</b>	<b>1 min, 54 s</b>
<b>Evan Almighty (2007)</b>	<b>00:36:25</b>	<b>1 min, 33 s</b>
<b>Wanted (2008)</b>	<b>00:23:17</b>	<b>1 min, 42 s</b>
<b>The Incredible Hulk (2008)</b>	<b>00:25:21</b>	<b>1 min, 46 s</b>
<b>Public Enemies (2009)</b>	<b>02:08:54</b>	<b>2 min, 24 s</b>
<b>The Bounty Hunter (2010)</b>	<b>00:12:58</b>	<b>1 min, 22 s</b>
<b>Inside Man (2010)</b>	<b>00:52:29</b>	<b>1 min, 16 s</b>
<b>Iron Man 2 (2010)</b>	<b>00:36:10</b>	<b>2 min, 12 s</b>
<b>Eat Pray Love (2010)</b>	<b>00:24:45</b>	<b>1 min, 33 s</b>
<b>Salt Director's Cut (2010)</b>	<b>01:07:56</b>	<b>1 min, 8 s</b>
<b>The Karate Kid (2010)</b>	<b>01:11:51</b>	<b>2 min, 11 s</b>
<b>The Social Network (2010)</b>	<b>00:23:17</b>	<b>2 min, 2 s</b>

*Note.* For each movie title, the start time and duration of the selected clip is listed. Adapted from "Enhancing movie comprehension for individuals who are visually impaired or blind," by Viswanathan, L. N., 2011, Thesis (M.S.), Arizona State University, p. 85. Copyright © 2011 by Viswanathan. Adapted with permission from author.

***Procedure.*** A within-subject design was used for this study—complete details of which may be found in (Viswanathan L. N., 2011). Two conditions were explored: *audio-only* and *audio-haptic*. The audio-only condition is the control condition where participants perceived only the audio of the clips including their audio descriptions. In the audio-haptic condition, audio-haptic described clips were experienced with both audio (film audio track and audio description) and haptic description. The control condition was used to assess whether haptic descriptions complemented audio descriptions by adding relevant, useful information in addition to enhancing enjoyment. Each participant completed both conditions, but the order was counterbalanced across participants to eliminate order effects. Half of the participants first completed the audio-only condition, and then the other half completed the audio-haptic condition.

*Audio-only condition.* The audio-only condition began with a familiarization phase in which participants listened to an audio described clip for acclimation. Of the twelve clips selected for each participant, if some had been seen, the most remembered clip of these was selected for familiarization. After the initial presentation, participants could request the clip to be repeated a maximum of two times. After familiarization, participants began the testing phase where five audio described clips were sequentially presented in a random order. After listening to each clip, participants were asked to describe what happened during the clip in terms of:

- *Context: location of scene, ambience and topic of conversation*
- *Number of characters in the scene*
- *Locations (position and distance) and movements of characters in the scene*

After each clip, participants were asked questions related to:

- *Perceived understanding of the clip*

- *Perceived concentration to understand the clip*
- *Perceived complexity of the clip*

For these questions, ratings were recorded using a 5-point Likert scale in which a rating of '1' represented low and a rating of '5' represented high.

*Audio-haptic condition.* The audio-haptic condition consisted of two parts completed in the following order: psychophysical analysis of the proposed somatic language, and perceptual analysis of audio-haptic described films. In the first part, participants' recognition accuracy of the proposed vibrotactile words and body context were assessed; recall that vibrotactile words were rhythms indicative of a character's relative distance from the screen, and body context employed vibrotactile stimulation at different body sites around the waist to convey a character's position across the screen. Participants were first familiarized with the vibrotactile belt and the location of vibration motors around the waist. Each vibration motor, from *L3* to *R3*, was vibrated in sequence with a rhythm not used in the study. Presentations were repeated if requested by the participant. During this time, the experimenter explained how the site of stimulation relates to a character's position across the screen. Next, participants were familiarized with the proposed tactile rhythms and how they relate to a character's relative distance to the screen. Each rhythm was presented at *L1*, and repeated when requested.

During training, twelve patterns were randomly presented (three rhythms each presented four times where each body site was covered twice). Participants were asked to recognize the dimensions of the pattern, and respond with the location of the vibration around their waist (*L3* through *R3*), and the distance the rhythm represented (*close*, *middle* or *far*). The experimenter provided feedback to confirm correct guesses, and correct those guesses that were incorrect. To move on to testing, 80% recognition

accuracy along each dimension needed to be achieved; otherwise, training was repeated for a maximum of one time.

The testing phase was similar to training with a few differences. First, more patterns were presented: 24 patterns were randomly presented (three rhythms were presented eight times each where each body site was covered four times). During testing, the experimenter refrained from providing feedback. Finally, each participant was asked questions related to the aforementioned psychophysical analysis:

- *Ease of learning vibration patterns*
- *Intuitiveness of vibrotactile stimulation for representing a character's location across the screen*
- *Intuitiveness of vibrotactile stimulation for representing a character's distance from the camera*

The ratings of questions were recorded using a 5-point Likert scale in which a rating of '1' represented low and a rating of '5' represented high.

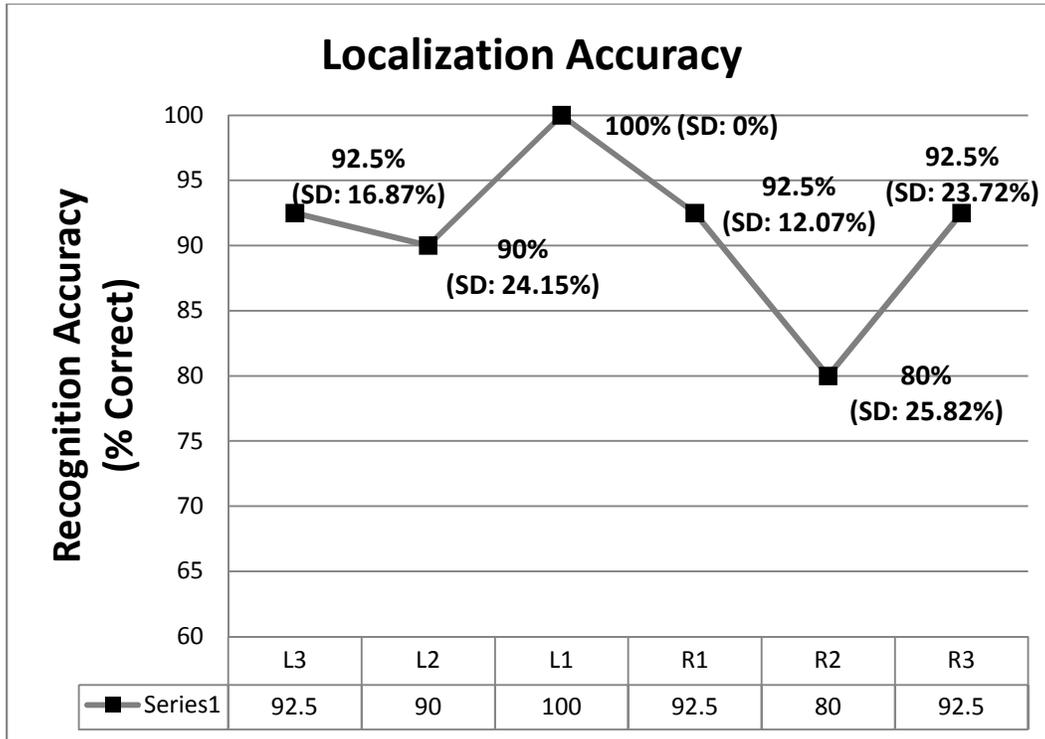
In the second part of the audio-haptic condition, participants were assessed for their understanding of both the observed audio-haptic described clips and the details of the presented haptic descriptions. Participants were first familiarized with spatio-temporal vibration variations associated with character movements across a scene. Each participant was presented with two sample movements (not part of the study), which could each be repeated twice when requested. This was followed by familiarization with an audio-haptic clip. As before, of the twelve clips selected for each participant, if some had been seen, the most remembered clip of these was selected for familiarization. Five audio-haptic described clips were selected for testing (clips not seen before, or those least remembered). As in the audio-only condition, after each clip, participants were asked to describe what happened in terms of context, number of characters, and their positions and

movements. They were also asked questions related to perceived understanding, concentration and complexity for each clip (identical to audio-only condition). After the audio-haptic condition, participants were asked questions related to their perception of the usability and effectiveness of the proposed system in terms of

- *Ease of wearing the belt*
- *Comfort of the belt*
- *Ease of associating the vibrotactile patterns with characters on the screen*
- *Ease of finding the location of a character across the screen*
- *Ease of finding the distance of a character within a scene*
- *Ease of combining haptic descriptions with audio information*
- *Degree to which haptic descriptions obstructed audio*
- *Information added to clip by haptic description with the goal of enhancing understanding of clip*

The ratings of questions were recorded using a 5-point Likert scale in which a rating of '1' represented low and a rating of '5' represented high.

**Results.** Figure 56 and 57 summarize results from the first part of the audio-haptic condition—that is, localization (body context) and rhythm (vibrotactile word) recognition accuracy, respectively.



*Figure 56.* Average localization accuracy with standard deviation (SD) across participants and rhythms for L3 through R3 (first part of audio-haptic condition). The average localization accuracy across body sites was 91.25% (SD: 19.43%). *Adapted from “Enhancing movie comprehension for individuals who are visually impaired or blind,” by Viswanathan, L. N., 2011, Thesis (M.S.), Arizona State University, p. 97. Copyright © 2011 by Viswanathan. Adapted with permission from author.*

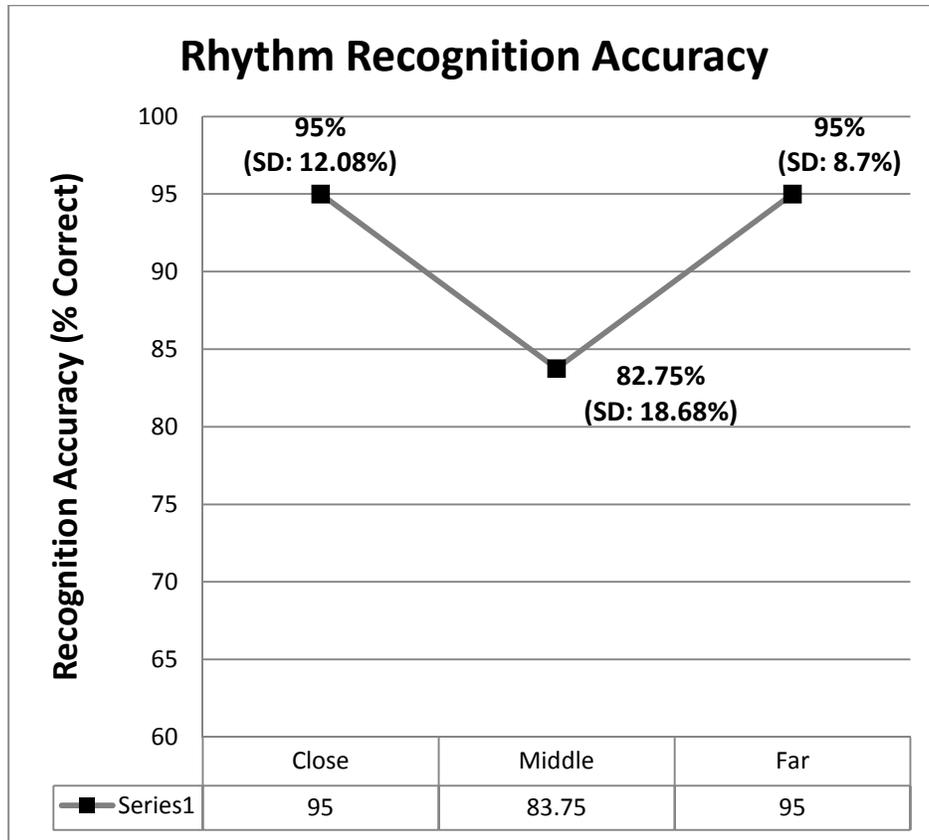


Figure 57. Average rhythm recognition accuracy with standard deviation (SD) across participants and body sites for close, middle and far rhythm (first part of audio-haptic condition). The average rhythm recognition accuracy across rhythms was 91.25% (SD: 14.37%). Adapted from “Enhancing movie comprehension for individuals who are visually impaired or blind,” by Viswanathan, L. N., 2011, Thesis (M.S.), Arizona State University, p. 98. Copyright © 2011 by Viswanathan. Adapted with permission from author.

During the second part of the audio-haptic condition, participants were asked to recognize (1) the locations of characters across the screen as indicated by stimulated body site (body context); (2) the distance of characters from the camera as indicated by tactile rhythm (vibrotactile words); and (3) the movements of characters as indicated by spatio-temporal variations of vibrations (vibrotactile sentences). Participants were asked to provide a high-level description of each movement including where it began, its direction, and where it ended. Moreover, each of (1)-(3) must be correctly associated with the character to which it belongs. For the audio-haptic condition, recognition accuracies

of (1) location, (2) distance and (3) movement, were 66.73% (SD: 13.61%), 71.73% (SD: 9.23%) and 85.9% (SD: 10.54%), respectively. This is in contrast to the audio-only condition in which (1) location and (2) distance of characters could not be accurately estimated from audio alone. Although this result was expected, participants could somewhat detect and describe movements in the audio-only condition using footsteps and other sound cues provided by the stereophonic headset. For the audio-only condition, movement recognition accuracy was 48.69% (SD: 18.01%).

Participants were also asked to rate their perceived understanding, concentration, and the complexity of each clip. These results are summarized below in figure 58.

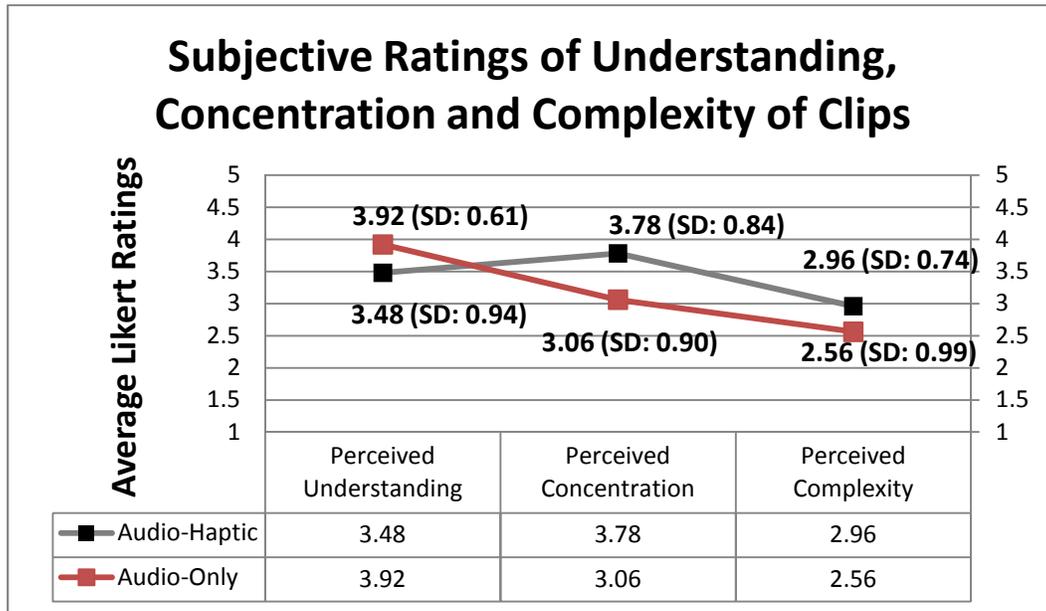


Figure 58. Likert ratings for participants’ perception of their understanding of an audio-haptic described clip, their needed concentration and the overall complexity of a clip. Likert ratings are averaged across participants. Adapted from “Enhancing movie comprehension for individuals who are visually impaired or blind,” by Viswanathan, L. N., 2011, Thesis (M.S.), Arizona State University, p. 101. Copyright © 2011 by Viswanathan. Adapted with permission from author.

Results of Likert scale questions pertaining to the proposed vibrotactile language:

- *How easy was it to learn the vibration patterns?* 3.7 (SD: 1)
- *How intuitive was the information about the location of a character presented?* 3.8 (SD: 0.9)
- *How intuitive was the information on the distance of a character presented?* 3.9 (SD: 0.8)

Results of Likert scale questions related to the audio-haptic condition

- *How easy was it to wear the belt?* 4.2 (SD: 0.9)
- *How comfortable was the belt?* 4 (SD: 0.9)
- *When experiencing vibration(s) with the belt, how easy was it to associate them with an actor on screen?* 2.9 (SD: 0.7)
- *While listening to the movie clips, how easy was it to find the location of an actor across the breadth of the screen with the belt?* 3.4 (SD: 0.9)
- *While listening to the movie clips, how easy was it to find the distance of an actor from the screen with the belt?* 3.6 (SD: 1.1)
- *How easy was it to combine the information received through the vibrations with that of audio?* 2.8 (SD: 0.9)
- *How much were the vibrations obstructing your attention to audio?* 3.4 (SD: 1)
- *Do you think that the information presented through the belt added to the understanding of the clip?* 3.5 (SD: 0.9)

Regarding the final question (answered at the end of the study), participants could chose to answer 'no', and refrain from giving a rating. The average rating that is reported above was computed only from those participants who answered 'yes'. Only two of the ten

participants answered 'no'. All of the aforementioned questionnaire results are from the original data of (Viswanathan L. N., 2011).

*Discussion.* Somatic ABC's was applied to design, develop and evaluate a complete system and somatic language for authoring and playing audio-haptic described films. Somatic ABC's evaluation theory guided the design of a user study to understand how well key design requirements (distinctness and naturalness) and implementation requirements (functionality, performance and usability) were met. Ultimately, these results provide insight into the usefulness of Somatic ABC's theoretical framework for articulating, building and confirming somatic languages. This section discusses how well these requirements were met in relation to the results of the experiment, beginning with implementation requirements.

*Functionality and performance assessment.* During implementation of the proposed vibrotactile belt and its firmware, performance and design criteria were accounted for including functionality, performance and usability. The expressiveness of the system enabled the creation spatio-temporal vibration patterns needed to communicate character positions and movements in scenes. Intensity variations were not used, although the system provides this functionality. Complete control of the timing of the presentation and duration of vibration patterns eased design and implementation. The scalability of the belt provided the flexibility to quickly test different designs, ultimately helping to decide on a six-tactor belt. Reconfigurability was achieved through position adjustable tactor modules, and firmware (with API) that allowed user-defined patterns and configurations to be developed and stored. Given the versatility of the vibrotactile belt, afforded by the aforementioned functionality characteristics, it has since been used in a variety of other applications including dance instruction (Rosenthal, Edwards,

Villanueva, Krishna, McDaniel, & Panchanathan, 2011) and conveying interpersonal position in social interactions.

In terms of performance, the vibrotactile belt has been durable and reliable with efficient response times during both initial testing and the current user study. Although battery life has been an issue, this problem is avoided for audio-haptic described films since users are seated at a computer, and hence, the belt can be plugged in and charging while the movie is viewed. For applications requiring portable use where long battery life is needed, a battery of 2000 mAh or greater will be used rather than 800 mAh.

*Subjective assessment.* In terms of usability, an easy-to-learn and easy-to-use system was sought by focusing on the distinctness and naturalness of vibration patterns; in addition to a comfortable, easy to don/doff belt design. System usability was assessed both objectively and subjectively. Objective results are in the form of recognition accuracies, which are discussed later. Subjective usability results were collected through a questionnaire. Participants found the vibrotactile belt comfortable (4) and easy to wear (4.2). Many users have commented on the discreetness of the belt in terms of its likeness to a waist belt, and the option of wearing it under clothes.

Regarding learnability, participants found the patterns easy to learn (3.7) due to their naturalness: both vibrotactile stimulations for the location of a character across the screen, and for the distance of a character from the camera, were perceived as being intuitive (3.8 and 3.9, respectively). Participants also found the system easy to use; in particular, participants found it easy to perceive the location and distance of characters in a scene using the proposed vibratory design (3.4 and 3.6, respectively). Associating the vibrotactile stimulations with the correct characters was met with some difficulty, but overall, satisfactory performance was achieved (2.9). Participants found that the haptic descriptions somewhat obstructed their attention to audio (3.4), and that both modalities

were slightly challenging to combine (2.8). Challenges related to attentional allocation and intermodal integration become clear when considering the perceived understanding, concentration and complexity of clips from both conditions, described below. Lastly, eight of the ten participants found that haptic descriptions added to their understanding of the clips (3.5), which was the main goal of this work. Of the other two participants that found otherwise: one participant said the amount of low vision that remained was enough to watch audio described films without haptics; and the other participant said haptics was distracting (although this participant's performance and overall results might have been influenced by distractions created from participant's cell phone which rang throughout the study).

Participants rated their perceived clip-wise understanding and concentration in addition to the perceived complexity of each clip. As shown in figure 58, the overall perceived understanding was lower for audio-haptic described films than for audio-only described films (not significant,  $S=2$ ,  $p>0.05$ , but approaching); and both perceived concentration (not significant,  $S=1$ ,  $p>0.05$ , but approaching) and complexity (not significant,  $S=2$ ,  $p>0.05$ , but approaching) were higher for audio-haptic described films than for audio-only described films. (The binomial sign test was used for significance testing.) This isn't to say that the proposed system was not effective; these results were expected given the novel communication channel added by touch, together with the short training time. Since haptic communication of position was new for participants, it increased concentration and complexity; which in turn shifted attention away from audio, increasing chances for missing pertinent contextual information, thereby decreasing perceived understanding. As previously mentioned, however, participants felt that the information provided through haptics did indeed add to their understanding of each clip in terms of movements and interactions between characters and their environment.

The aforementioned subjective results are promising when considering the short training times of participants; results clearly indicate that users perceived the system as both usable and useful, but objective results also need to be examined.

*Objective assessment.* The distinctness of the proposed communication units, along with their contextual variations, was assessed through two parts, as previously described in the procedure. In the first part, localization and recognition accuracy of location and distance were assessed; in the second part, movement recognition. Overall localization accuracy (91.25%) is impressive compared to related studies (McDaniel T. L., Krishna, Colbry, & Panchanathan, 2009) (McDaniel T. L., Villanueva, Krishna, Colbry, & Panchanathan, 2010) given both the shorter vibration durations and the older participant population as used here. No significant difference,  $F(5,54)=1.12$ ,  $p=0.3608$ , was found between localization accuracies for recognizing the body site of vibrotactile stimulations around the waist. This reveals that no particular body site, as stimulated by the proposed six-factor belt, was more difficult to localize compared to other body sites. Clearly, however, certain body sites have higher average localization accuracy—in particular, vibration modules at the midline (*L1* and *R1*) and at the endpoints (*L3* and *R3*) are greater than those found for points between (figure 56), which was expected given the insights provided by psychophysical studies that have explored localization around the waist (Cholewiak, Brill, & Schwab, 2004)—these differences, however, were not significant.

Overall rhythm recognition accuracy was impressive at 91.25%; and the design of the rhythms was well-received by many participants who found the metaphor of the stimulation intuitive and natural. No significant difference,  $F(2, 27)=2.22$ ,  $p=0.1285$ , was found between recognition accuracies of the proposed rhythms. This shows that no particular rhythm design was more difficult to recognize than the others; however, some

average accuracies are greater than others; in particular, rhythms representing close and far have higher average accuracies than middle distance (figure 57). Indeed, participants found close and far rhythms to be intuitive and indicative of the concept they represented. Participants also commented that rhythms for middle and far could be further separated.

Participants performed very well in terms of correctly associating and recognizing the movements of characters during audio-haptic described films—overall recognition accuracy was 85.9%. This is in contrast to the audio-only condition in which participants achieved an overall movement recognition accuracy of 48.6%. The mean increase from audio-only to audio-haptic was statistically significant,  $t(9)=7.2$ ,  $p<0.01$ , two-tailed, showing that recognizing movements with just audio was extremely difficult. These results correlate well with subjective results in that participants found it easy to make associations between the vibrations and characters on the screen.

Given that movements were often composed of many locations and distances as characters moved about a scene, recognizing individual locations and distances was more challenging than recognizing movements. Specifically, overall localization and rhythm recognition accuracy were 66.73% and 61.75%, respectively. It is interesting to note that subjective results revealed that participants found it easy to recognize location and distance; this is most likely due to a sufficient number of these cues being perceived during movements to estimate the overall movement. In any case, understanding the movements of characters (start position, direction of movement and end position) during scenes where many location and distance changes occurred seems to be more important. This information could not be gleaned from audio-only described films.

## **Conclusion and Future Work**

In this application, Somatic ABC's was utilized to create a somatic language and system for audio-haptic described films—in particular, haptic descriptions communicating the position and movement of characters within a scene. The versatility of Somatic ABC's enabled straightforward application of the framework's theories for creating a novel language for haptic descriptions. The framework's unique natural language inspired design methodology enabled the creation of an expandable, efficient, rich and robust somatic language:

- The proposed somatic language is expandable in that more somatic words could be added to its vocabulary—that is, the degree of discretization may vary. Pilot testing and formal evaluation have revealed that three words (rhythms) worked well, but fewer (no less than two) or more (probably no more than four as it may be harder to obtain perceptual separation without extensive training) may be selected depending on preference and the application.
- The haptic presentation of locations and distances as part of character movements were easily synchronized with video and audio media. The conceptual, high level description of positions afforded a concise representation that was efficiently conveyed and perceived by users.
- The small somatic word vocabulary of three distances supported distinctness and learnability, and was further enriched through the use of body context to convey the locations of characters across the screen. Through Somatic ABC's body context, an expressive communication channel was created to help visualize positions, movements and dynamic interactions of characters within a movie scene.

- Vibrotactile sentences were created by temporally combining vibrotactile words and context, enabling the expression of limitless movements. Moreover, since movements are composed of many positions, the language is robust in that not all individual positions need to be accurately identified to estimate a movement.

In addition to the above mentioned attributes, the somatic language's high level of abstraction and intuitive signal parameter association helped articulate distinct and natural vibrotactile words, which eased learning and enhanced usability. Participants appreciated the naturalness of the language, and found it easy to recognize its words and sentences. Lastly, the language itself is versatile, being easily extended to other applications such as interpersonal positioning (direction and proxemics) in social interaction assistants (McDaniel T. , Krishna, Balasubramanian, Colbry, & Panchanathan, 2008) for individuals who are blind; among other non-verbal socio-communicative cues in this application area such as head nodding, body language and hand gestures.

Overall, the proposed system was found usable by participants, and was well received. In terms of the application of audio-haptic description, this work represents the first step toward descriptive video services that use haptic descriptions. Possibilities for directions of future work include:

- Optimal integration of haptic descriptions with the audio track (including audio description) of a film must be explored. Similar to audio descriptions, the placement of haptic descriptions should correlate with the onscreen activity they represent. But precise placement is variable, and may be adjusted to reduce overlap with audio descriptions and/or the audio of a film. Optimal placement in terms of its effects on the perception of a scene and cognitive load must be further explored.

- Cross-modal integration and interaction between haptics and audio should be investigated. Clearly, these two modalities combine to construct a percept of a scene. Understanding how information from both channels integrate and interact will help guide the design and insertion of haptic descriptions into audio described films.
- Reducing the redundancy between haptic and audio content requires careful attention. Haptic descriptions can complement audio descriptions and the audio of a film, but redundant descriptions, while not adding information content to a film, may enhance a film in terms of enjoyment and experience. This claim must be further explored.
- Haptic descriptions for other non-verbal cues, such as facial expressions, body language, among other socio-communicative cues, need to be investigated to enrich this novel channel for descriptive video.
- In the present study, haptic descriptions were developed for conversational scenes. How these haptic descriptions may be applied to other genres and fast-paced scenes should be explored. Somatic ABC's stimulation variations might be useful in this context; in particular, tempo increases and decreases could be utilized depending on the pace of movements within a scene as the film progresses.

**APPLICATION #2: VIBROTACTILE MOTOR INSTRUCTIONS AND  
FEEDBACK**

Movement is integral to both action and perception. Seemingly simple yet coordinated and controlled complex movements enable us to act upon and perceive our environment. Efficient and effective perception of our surroundings relies not only on limb movements for grasping, holding and haptically exploring objects through exploratory procedures (Lederman & Klatzky, 1987), but also eye movements (saccades), both voluntary (changing our direction of eye gaze between fixations) and involuntary (fast jumps between pertinent visual features when looking at a scene or object), for extracting visual details.

Movement is just as important within interpersonal interactions. Speech is articulated through complex mouth and tongue movements, and vibrations of the vocal folds within the larynx (voice box). But speech is only one component of social interactions, making up less than half of the information transmitted (Knapp, 1978). The remaining information is conveyed through non-verbal cues including posture, hand gestures, eye gaze, social touching and facial expressions. The building blocks of facial expressions are called *facial action units*, and include curling the lips, wrinkling the nose, raising the cheeks, blinking, and winking, among many others facial movements.

Clearly, movement is critical to our survival. And in this same sense, we often strive to learn more complex movements in an effort to enrich our lives and health. For example, we may learn movements as part of a skillset for a new career, exercise regimen or physical activity; or we may need to relearn movements when we are out of practice, or during physical rehabilitation after a motor impairment. While learning novel movements, learning progress is influenced by the learning style of the trainee and the pedagogy of the instructor.

As learners, we tend to prefer one of three styles of motor learning: visual, auditory or kinesthetic (Kane, 2004). Visual learners prefer a visual demonstration to visually map the viewed movement onto their frame of reference for mimicry; auditory learners prefer clear and detailed descriptions of movements with discussion; and kinesthetic learners are “hands-on” in that they learn best through practice. Since most learners are visual learners, traditionally pedagogical instruction of motor skills constitutes visual demonstration and verbal description (Kennedy & Yoke, 2009) accompanied by visual, verbal and/or physical feedback. Physical feedback by a trainer is commonly provided through gentle touches guiding or correcting movements and posture, either through direct manipulation of limbs, or directing the trainee’s attention to the source of error.

Traditional motor instruction occurs within one of two settings: individualized instruction or group instruction. In general, one-on-one instruction allows close, uninterrupted interaction between a trainer and trainee. This environment helps trainers adapt their pedagogy to align with the learning preferences of the trainee—a technique that is much more difficult in group settings involving many students. Individualized instruction also supports real-time visual, verbal and/or physical feedback throughout training; whereas in group settings, feedback is only sparsely available given the divided attention of the trainer. Group instruction also suffers from the large interpersonal distance between a trainer and trainee where many students must watch and listen to the instructor over other students in the class. Therefore, it is no surprise that students tend to learn motor skills more effectively when instructors are nearby (Kennedy & Yoke, 2009), possibly due to increased accountability and motivation, feedback from the instructor, and clearer, more personal instructions. Unfortunately, since individualized instruction is

inaccessible to most students due to cost, group instruction is the most common setting for motor learning.

Context-specific limitations exist that encompass both individualized and group settings. For example, in swimming (Förster, Bächlin, & Tröster, 2009), snowboarding (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) and many other sports where extreme physical trainer-trainee separation is present, real-time feedback is unavailable. In many situations, instructor feedback will need to interrupt a motor performance when modalities are unavailable for communication, and attention is occupied, such as while playing a musical instrument (van der Linden, Johnson, Bird, Rogers, & Schoonderwaldt, 2011). Limited feedback slows the learning process as error information must be available for motor learning to occur.

There are two types of feedback (Schmidt & Wrisberg, 2000): *intrinsic feedback* is performance-relevant sensory information that occurs naturally as a result of a movement; whereas *extrinsic feedback* (or *augmented feedback*) is delivered from an outside source, such as an instructor or electronic device. There are two types of extrinsic feedback (Schmidt & Wrisberg, 2000): *knowledge of results* is feedback related to performance in terms of how the performer achieved the desired movement, or met the overarching goal; whereas *knowledge of performance* pertains to the details of the performance of the movements involved. Unless noted, feedback here will refer to extrinsic feedback of either type. While intrinsic information can provide error information, augmented feedback is critical when we do not have access to intrinsic feedback, or when it is insufficient. An example of the former is when an individual with a sensory or perceptual impairment is attempting to learn a motor skill, but his or her impairment prevents access to relevant intrinsic feedback. An example of the latter is an attempt to learn a complex motor skill, the details of which we are not familiar with. In

these cases, augmented feedback will enable gains and improvements related to motor learning and performance. Moreover, when feedback is related to a learner's progress toward his or her goal, it can provide motivation and increased effort (Schmidt & Wrisberg, 2000). Lastly, positive feedback can reinforce good performance, thereby improving learning.

To address the aforementioned limitations of traditional motor instruction techniques, computerized delivery of motor instructions and feedback offers a promising alternative. Researchers have explored various modalities including visual, auditory and kinesthetic to mediate this communication. Virtual reality and force-feedback systems have been limited by cost and portability. Virtual reality systems, as well as audio implementations, are also limited by obstructing modalities that may already be occupied or even unavailable. Considering these disadvantages, the tactile modality provides an alternative option that is unobtrusive and discreet with affordable and portable implementation possibilities.

Vibrotactile stimulation for both motor instruction and feedback is proposed. Vibrotactile instructions are pre-defined, spatio-temporal stimulations representing motor movements at a high level (e.g., which movement to perform next as part of a regimen) or low level (e.g., detailed instructions that convey how to perform a movement—that is, which limbs to use and how). These instruction-based approaches only cue a user to perform a specific movement, and are not linked to actual motor performance (Drobny, Weiss, & Borchers, 2009). We propose low-level instructions for targeting fundamental movements (Behnke, 2006), the building blocks of human motion, through natural, saltatory vibration patterns.

Vibrotactile stimulation for feedback is driven by measures of a user's motor performance represented at a high level (knowledge of results, such as whether the

movement was performed successfully) or at a low level (knowledge of performance, such as detailed error information related to position or speed). The proposed approach targets real-time knowledge of performance by presenting errors related to position (angles of limbs) and speed (angular rate of change).

Both computerized vibrotactile instruction and feedback may complement learning within the classroom and/or at home. For novices, trainer-specified vibrotactile instructions may help with following a recommended regimen within a fast-paced group class, and help beginners understand the individual movements involved within complex movements. Vibrotactile instructions and feedback can also bridge the gap between large interpersonal distances: trainees separated from trainers during physical activity (e.g., swimming, snowboarding, etc.), can continue to receive real-time instructions and feedback, either automated or delivered manually by the trainer.

Vibrotactile feedback driven by motor performance may be useful for both novices and experts—the latter of which may be more interested in further mastery of movements they have learned. Vibrotactile feedback can provide automated, real-time feedback within any type of instructional setting including practice at home. In contrast to physical feedback by a trainer, it can also provide feedback for multiple limbs by stimulating possibly many different parts of the body simultaneously (Lieberman & Breazeal, 2007). In terms of feedback frequency and amount, the following issues must be considered (Schmidt & Wrisberg, 2000):

- Feedback that is too frequent can create a dependency in which the learner relies too heavily on the feedback; and therefore, may experience performance difficulties in the absence of guidance.
- Feedback that is too frequent can also lose its reinforcing power.

- For novices, too much feedback per performance may be overwhelming, causing users to lose focus on how best to improve their motor movements.

Computerized systems for vibrotactile feedback may be designed to account for these limitations of feedback. The frequency of feedback may be reduced over time, and its decline may be coupled with motor learning. Feedback bandwidth may increase over time with improvements in motor learning; but in the beginning, the feedback given per performance should focus on the most important attribute that needs improvement (Schmidt & Wrisberg, 2000).

### **Related Work**

This section describes virtual reality (VR) and augmented reality (AR) visual, acoustic, kinesthetic and/or vibrotactile approaches for complementing traditional motor learning. This section focuses on vibrotactile approaches, but related visual, acoustic and kinesthetic approaches are briefly visited first.

**Virtual reality.** Since the 1990's, many VR and AR systems have been proposed for a variety of application-specific motor learning tasks including physical therapy, dance, exercise and calligraphy. In this section, two VR systems are described that demonstrate the basic approach used within many of these systems: mimicking the movements of a virtual instructor with real-time visual feedback. A detailed review of virtual environments for motor learning and rehabilitation can be found in (Holden, 2005).

Yang and Kim (2002) proposed a novel interaction paradigm for virtual reality-based motor learning systems called *Just Follow Me* (JFM), which utilizes a *ghost* metaphor. In JFM, the user views his or her virtual avatar through a head mounted display, superimposed with the instructor's avatar (or ghost). The ghost then moves out

of the user's virtual body, after which the user follows the ghost to mimic the movements of the instructor. Yang and Kim developed a virtual reality system for learning movements involved in calligraphy. Users wore a head mounted display for the first-person view of JFM, and Polhemus trackers were used to capture movements.

Chua et al. (2003) have proposed a virtual reality system for learning Tai Chi movements (figure 59) by mimicking the motions of a virtual instructor seen through a head mounted display. A user's movements were captured in real-time using the Vicon system where IR cameras captured motion using reflective markers placed on the body. Setup involved the placement of 41 reflective markers on a Spandex suit, followed by a calibration phase. The three-dimensional locations of markers were used to find the relative positions of a user's limbs, which were then rendered and displayed through a head mounted display for real-time visual feedback. Users always saw their virtual representations in first-person, but the virtual instructor could be superimposed or displayed outside of the user's body, but always facing away.



*Figure 59. View of Tai Chi system setup (left) and virtual scene (right) depicting the virtual avatars of the student (user) and instructor. Reprinted from “Training for physical tasks in virtual environments: Tai Chi,” by Chua, P. T., Crivella, E., Daly, B., Hu, N., Schaaf, R., Ventura, D., Camill, T., Hodgins, J., & Pausch, R., 2003, In Proceedings of IEEE Virtual Reality, p. 87. Copyright © 2003 by IEEE. Reprinted with permission.*

**Acoustic systems.** Acoustic-based instruction and feedback systems employ musical rhythms and sound feedback for enhancing motor learning. Some research has investigated the perceptual characteristics of acoustic feedback within this application, and its intermodal integration (Effenberg, 2005). Several approaches have been developed; two of which are presented below.

Takahata et al. (2004) developed a sound enhanced instruction and feedback system for learning karate. Students learned and practiced movements while musical rhythms were played to help with timing and motivation. Students wore accelerometers on their wrists and waist to capture motion, which drove the generation of sounds indicative of movement timing and intensity.

Saltate!, developed by Drobny, Weiss and Borchers (2009), is an acoustic feedback system for learning dance. A sensor module for the shoe was implemented,

which detects steps using a force sensing resistor. The system recognizes steps as correct or incorrect based on the rhythm of the music. Feedback is provided in the form of acoustic beats where incorrect/correct steps cause linear increases/decreases in volume, respectively, to direct attention to mistakes and enhance motivation.

**Kinesthetic systems.** Kinesthetic or force-feedback devices have been explored for VR and AR motor learning and feedback. The most common form factor is a graspable joystick, handle or knob, but other implementations, such as exoskeletons and robots, have been investigated. Approaches typically support motor learning and feedback through either guidance and/or resistance. One popular application of automated haptic guidance and resistance is motor rehabilitation, which began to garner interest in the 1990's—a detailed review can be found in (Hesse, Schmidt, Werner, & Bardeleben, 2003). Other applications include skill training and exercise. Two approaches are presented here: an approach for teaching dance, and another for percussion training.

Kosuge, Hayashi, Hirata and Tobiyama (2003) explored human-robot coordination for teaching dances involving a partner. They developed a robotic dance partner, Ms DancerR, whose wheeled base enabled omnidirectional movement, and whose body force sensor between the base and body enabled detection of forces by a human dance partner. Based on the forces exerted by a human dance partner, Ms DancerR could recognize steps, and move accordingly.

Gindlay (2008) developed the Haptic Guidance System (HAGUS) to record and playback wrist movements involved in playing percussion instruments. In particular, drum playing was explored and simple wrist movements—flexion and extension—were implemented. As rhythms are played, a drumstick, actuated by a servo motor, plays back

the recorded rhythm. Users learn the rhythm through haptic guidance and/or listening to the beats.

**Vibrotactile systems.** Vibrotactile instructions and feedback have been explored for a variety of applications including music—violin bowing (van der Linden, Johnson, Bird, Rogers, & Schoonderwaldt, 2011) and piano playing (Huang, et al., 2010); sports and recreation—swimming (Förster, Bächlin, & Tröster, 2009), snowboarding (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) and dancing (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011); and physical therapy (Lindeman R. W., Yanagida, Hosaka, & Abe, 2006) (Lieberman & Breazeal, 2007) (Kapur, Jensen, Buxbaum, Jax, & Kuchenbecker, 2010).

Vibrotactile feedback was investigated for correcting improper bowing during violin playing. The MusicJacket, developed by van der Linden et al. (2011) and depicted in figure 60, applies vibrotactile stimulation to the arms, wrists and torso to guide straight bowing movement and correct poor posture related to holding the instrument. The system uses a portable motion capture system by Animazoo, which computes the relative, three dimensional positions of limbs using orientation data sensed by on-body inertial measurement units (IMUs) and calibration.

Another haptic system that complements motor learning for music playing is the Mobile Music Touch (MMT) system (Huang, et al., 2010). MMT is an instruction-only system that cues which finger to use next within a piano song. Cueing is mediated through a wireless, vibrotactile glove with a vibration motor placed near each finger's metacarpophalangeal joint. The system is intended to support subconscious learning away from the piano while performing other, unrelated tasks.



*Figure 60. MusicJacket system being used by a young violin student. Reprinted from “Towards a real-time system for teaching novices correct violin bowing technique,” by van der Linden, J., Schoonderwaldt, E., & Bird, J., 2009, In Proceedings of IEEE International Workshop on Haptic Audio Visual Environments and Games, p. 82. Copyright © 2009 by IEEE. Reprinted with permission.*

Förster et al. (2009) developed a waterproof, wrist-worn, vibrotactile instruction and feedback system for cueing swim strokes and speed adjustments while swimming. The presence of a vibration indicated speed: fast when vibration was present, slow when vibration was absent. Stroke type was cued based on the duration of the vibration: short versus long. Rosenthal et al. (2011) proposed a wireless vibrotactile belt for dance instruction. Vibrations around the waist cued different dance steps (step forward, step back, step right, step left, etc.) based on spatio-temporal variations indicative of these movements.

In 2009, an extensive study was published by Spelmezan et al. (2009) that explored vibrotactile instructions for snowboarding movements. These instructions were

intended to be used by remotely located coaches for sending movement instructions and feedback in real-time. The project involved three stages: a lab-based open response paradigm to create the initial instruction set; lab-based snowboarding simulation; and an in the wild study involving actual snowboarding. To discover movements naturally elicited by different vibrotactile stimulation designs, an open response paradigm was employed involving a large set of sample vibration patterns. Participants preferred saltation patterns for movements, in contrast to single, localized vibrotactile pulses, given the former's directionality. Saltation (Geldard & Sherrick, 1972), or the *cutaneous rabbit*, is a perceptual illusion of apparent motion in which a train of quick vibrotactile bursts, fixed at only a few body sites, is perceived as a train of evenly spaced phantom vibrations. This illusion provides vivid sensations of quick, evenly spaced bursts hopping across the skin—hence the name, the cutaneous rabbit. It is a robust illusion, occurring under many configuration variations including the number of motors, spacing of motors, duration of bursts, pauses between bursts and stimulated body sites.

Although responses varied across participants, the most common responses were used to form a set of vibrotactile instructions for snowboarding movements. The experimental setup and system is depicted in figure 61, and the placement of the vibration motors is shown in figure 62. Spatio-temporal saltation patterns were used to cue snowboarding-specific movements including learning the body forward, backward, left or right; turning the upper body left or right; and stretching or flexing the legs.

Spelmezan et al. also discovered that participants perceived vibrations as pushing or pulling the limb they stimulated. They suggested that patterns can be designed under the *push* or *pull metaphor* in that they either push a limb or pull a limb, respectively. Spelmezan et al. used the push metaphor during their study.



*Figure 61. System and experimental setup of the tactile motion instruction system. Reprinted from "Tactile motion instructions for physical activities," by Spelmezan, D., Jacobs, M., Hilgers, A., & Borchers, J., 2009, In Proceedings of the 27th International Conference on Human Factors in Computing Systems, p. 2248. Copyright © 2009 by Association for Computing Machinery, Inc. Reprinted with permission.*

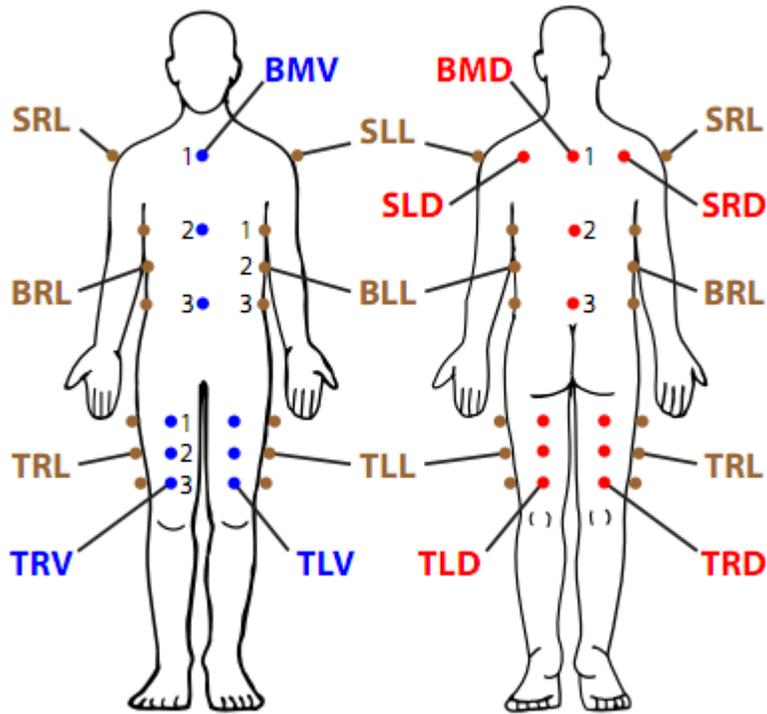
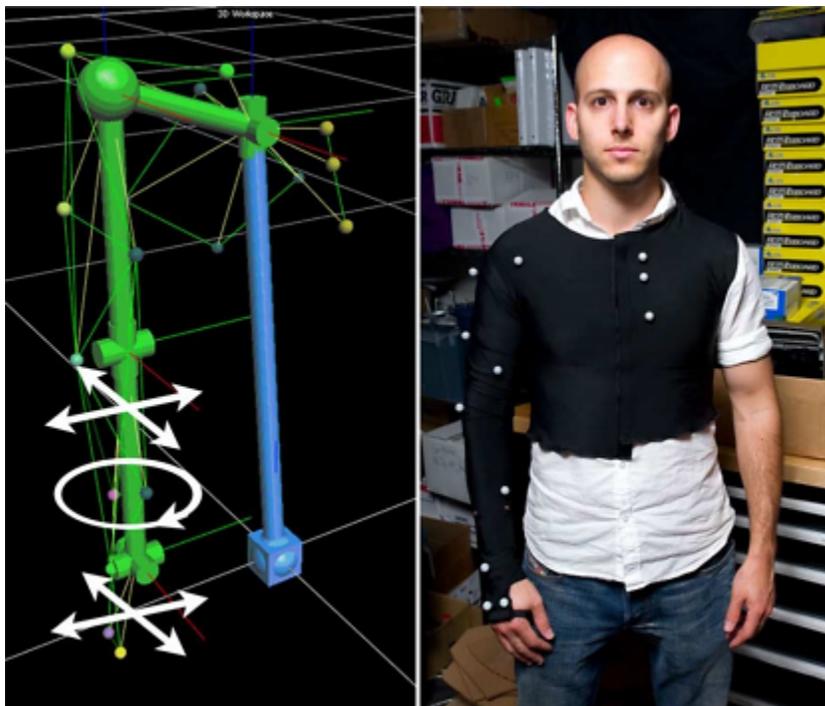


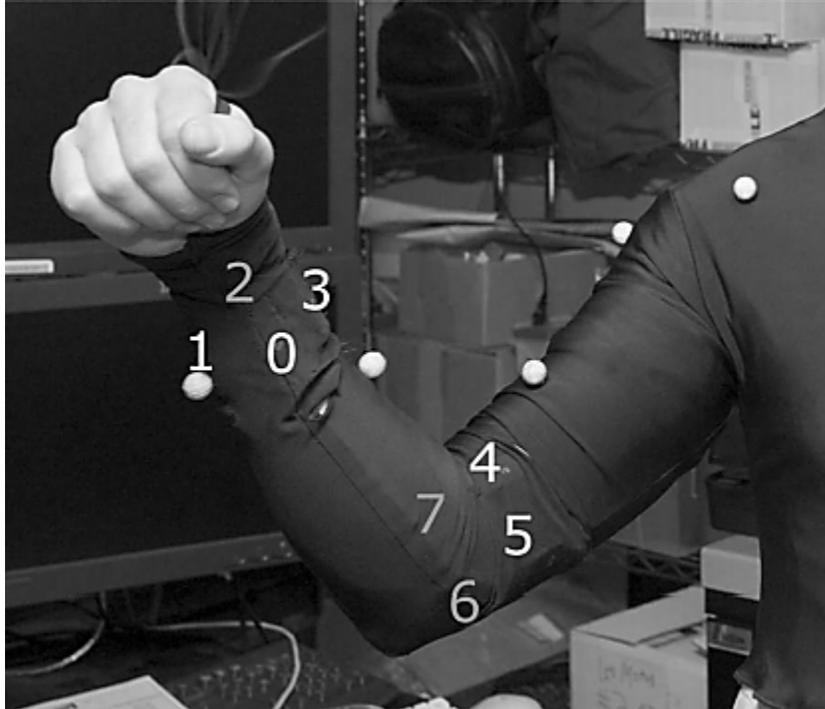
Figure 62. Tactile motion instruction system’s placement of vibration motors configured into groups based on the movements they represent. Groups are indicated by three letters: The first letter refers to the Thigh, Body or Shoulder; the second letter refers to the Left, Medial or Right; and the last letter refers to the Lateral, Ventral or Dorsal. Leg flexes or stretches use motors on the back or front of the legs, respectively; shifting weight to the left/right or front/back uses motors on the left/right leg or back/front of the body, respectively; and lastly, upper body rotations circle around the torso. *Reprinted from “Tactile motion instructions for physical activities,” by Spelmezan, D., Jacobs, M., Hilgers, A., & Borchers, J., 2009, In Proceedings of the 27th International Conference on Human Factors in Computing Systems, p. 2245. Copyright © 2009 by Association for Computing Machinery, Inc. Reprinted with permission.*

A number of vibrotactile feedback systems have been proposed for physical therapy. The TactaPack (Lindeman R. W., Yanagida, Hosaka, & Abe, 2006) is a physical therapy device consisting of wireless, wearable modules. Each module contains an accelerometer for motion sensing, a vibration motor for real-time feedback, and components for processing, power and wireless capabilities. Vibrotactile stimulation replaces the nudges of a physical therapist, warning of limbs exceeding or not reach recommended accelerations established during calibration.

Another vibrotactile feedback system for complementing traditional physical therapy is the Tactile Interaction for Kinesthetic Learning (TIKL) system (Lieberman & Breazeal, 2007): a wearable system that indicates when joints are in error (with respect to the movements of an instructor) through vibrotactile stimulation where intensity is proportional to the amount of error. Movements are captured using the vision-based Vicon motion capture system. Marker placements are shown in figure 63, and motor placements are depicted in figure 64.



*Figure 63. TIKL marker placement for use with the Vicon system for capturing arm movements. Reprinted from “TIKL: Development of a wearable vibrotactile feedback suit for improved human motor learning,” by Lieberman, J., & Breazeal, C., 2007, IEEE Transactions on Robotics, 23(5), p. 920. Copyright © 2007 by IEEE. Reprinted with permission.*



*Figure 64. Motor placement for TIKL. Numbers ‘0’ through ‘7’ indicate locations of vibration motors (placed under fabric for direct contact with skin). Placement is influenced by the design’s use of the push metaphor for correcting joint errors for fundamental movements. Actuation of motors ‘0’ and ‘2’ cue wrist extension and flexion, respectively; motors ‘1’ and ‘3’ cue wrist abduction and adduction, respectively; motors ‘4’ and ‘5’ cue elbow extension and flexion, respectively; motors ‘5’ and ‘7’ cue shoulder adduction and abduction, respectively; and finally, saltation patterns (clockwise or counterclockwise around the vibration motors for the wrist) cue clockwise or counterclockwise forearm rotation. Reprinted from “TIKL: Development of a wearable vibrotactile feedback suit for improved human motor learning,” by Lieberman, J., & Breazeal, C., 2007, *IEEE Transactions on Robotics*, 23(5), p. 921. Copyright © 2007 by IEEE. Reprinted with permission.*

Kapur et al. (2010) developed a sleeve augmented with vibrotactile actuators for movement feedback during stroke rehabilitation. Movement is captured in real-time using an Ascension electromagnetic motion capture system with three sensors, and rendered on screen along with a virtual representation of the target movement. Joint errors are communicated in real-time through vibrotactile stimulation using the push metaphor. Fundamental movements in consideration here include elbow flexion and extension; shoulder flexion and extension; shoulder abduction and adduction; and shoulder rotation.

Many of the aforementioned approaches focus either on vibrotactile instructions or feedback, and therefore, do not bridge the divide between these different modes, continuing to rely on other modalities for augmentation. Moreover, vibrotactile instructions of previous approaches are application-specific, such as instruction sets for snowboarding (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) or dancing (Rosenthal, Edwards, Villanueva, Krishna, McDaniel, & Panchanathan, 2011); and therefore, may be difficult to generalize to other motor learning applications. Although vibrotactile feedback has been shown to be helpful for motor learning (Lieberman & Breazeal, 2007), there is no clear bridge between previously proposed vibrotactile feedback designs and vibrotactile instructions—that is, how can such a system both cue users of movements to perform, and then follow up with feedback driven by motor performance. Novices may benefit most from both instructions and feedback, while experts may benefit more from detailed feedback. A system that incorporates both modes, therefore, may benefit from versatility, scalability and usefulness.

Moreover, most of these approaches rely on visual or electromagnetic motion capture systems, which are bulky, expensive and lack portability. Electronic sensing devices, such as accelerometers and inertial measurement units (IMUs), offer a more cost-effective solution while improving upon portability, discreetness and usability with sufficient sampling speeds and accuracy. One example is van der Linden et al.'s use of inertial measurement units (IMUs) for driving application-specific, vibrotactile feedback for correcting bowing and posture during violin playing. Although this feedback design worked well for the given application, a more general feedback design is needed for applicability and versatility. Lastly, vibrotactile feedback designs have largely focused on positioning errors, ignoring corrections for the speed of movement, which for many applications, may be just as important.

## **Proposed Approach**

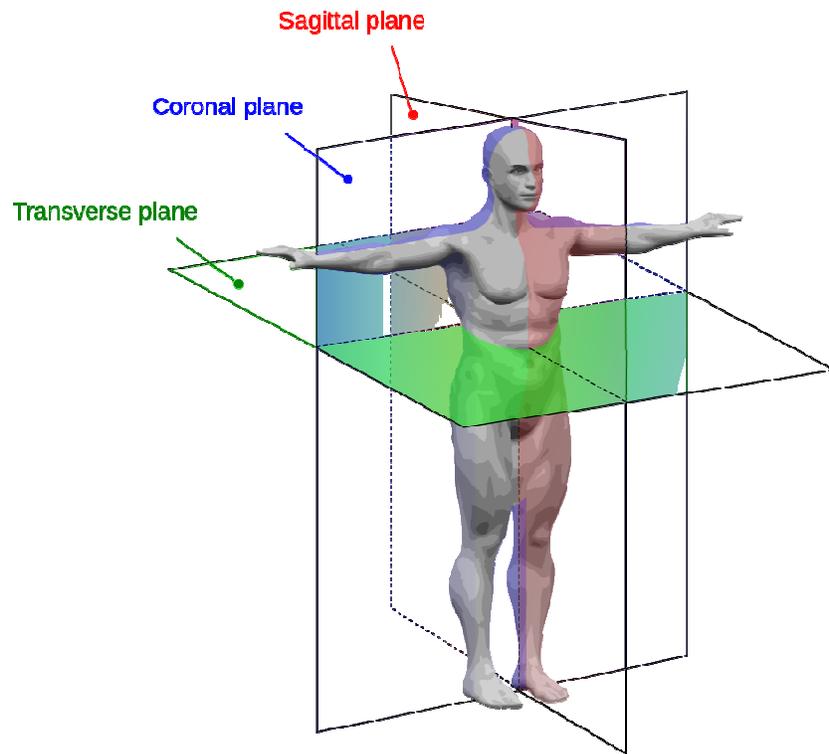
This section presents the proposed somatic language for communicating vibrotactile instructions and feedback to complement traditional pedagogy for motor learning. The following sections discuss how Somatic ABC's was applied to design, develop and evaluate vibrotactile motor instructions and feedback (in this order).

**Articulate: vibrotactile motor instructions.** The design steps of Somatic ABC's were employed to create a novel language for vibrotactile instructions. The first step was to identify an application, and define its scope. The application here is motor learning through instruction and feedback using vibrotactile stimulation. The scope is limited to low level descriptions of movement to enable novices to learn the intricate motions involved in more complex movements. We have also limited our system implementation to provide stimulation related to one movement at a time, and only movements involving the right arm (specifically, joints below the shoulder joint). These choices were made as they were within the current hardware limitations; and were a necessary first step to explore the distinctness and naturalness of vibrotactile instructions and feedback for simple movements before progressing to more complex, concurrent movements involving the whole body.

The next step of Somatic ABC's was to identify the smallest communication units of the language. As previously described, existing vibrotactile instruction sets are application-specific in that they are designed for use within a particular application—this design choice, therefore, limits their versatility and applicability. Rather than design new vibrotactile instructions for each motor learning scenario, a more efficient approach is to design a generic instruction set that can be applied across diverse application areas. Fundamental movements, the building blocks of human motion, were chosen as the proposed somatic language's communication units to ensure a small word vocabulary for

ease of learning. Although these are our most basic movements, there are only five: flexion, extension, abduction, adduction and rotation (Behnke, 2006). However, when these movements are carried out concurrently and sequentially, rich movements are formed. Consider the human body in the anatomical posture (standing straight with arms by the side and palms facing forward). Within this posture, each fundamental movement occurs within a specific cross-sectional plane (figure 65). Rotations occur within the transverse plan (or horizontal plane) either toward the sagittal plane (pronation) or away (supination). Flexion and extension occur within the sagittal plane during which the joint angle decreases (flexion) or increases (extension). Abduction and adduction occur within the coronal plane (or frontal plane) either toward the sagittal plane (adduction) or away (abduction). The fundamental movements explored as part of this work are visually depicted in figure 66.

In terms of the level of abstraction, fundamental movements provide a good balance between vocabulary size and expression. More complex movements would necessitate a larger vocabulary, complicating learning. A lower level of abstraction might be difficult to discretize, increasing vocabulary size. Achieving an intuitive signal parameter association is also important; to accomplish this, vibrotactile saltation patterns were chosen. Saltation, described previously, has been shown to work well for motor learning applications (Spelmezan, Jacobs, Hilgers, & Borchers, 2009) given their inherent directionality from apparent motion.



*Figure 65.* Anatomical planes of the human body. Three planes divide the human body into different halves: sagittal plane, coronal plane and transverse plane. Adapted from the Wikimedia Commons: *Human\_anatomy\_planes.svg*, [http://en.wikipedia.org/wiki/File:Human\\_anatomy\\_planes.svg](http://en.wikipedia.org/wiki/File:Human_anatomy_planes.svg).

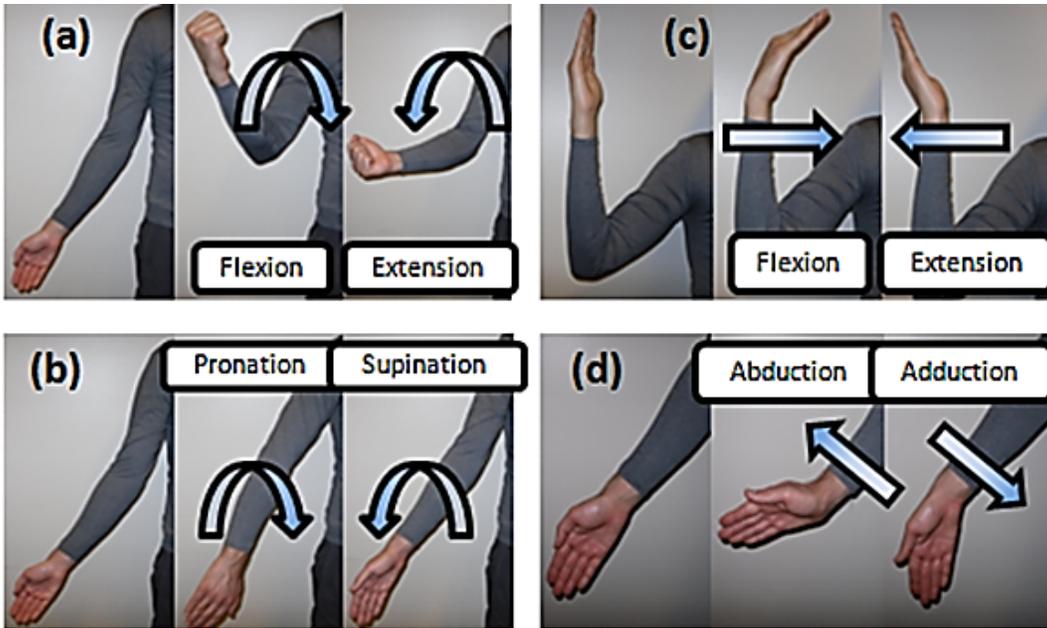
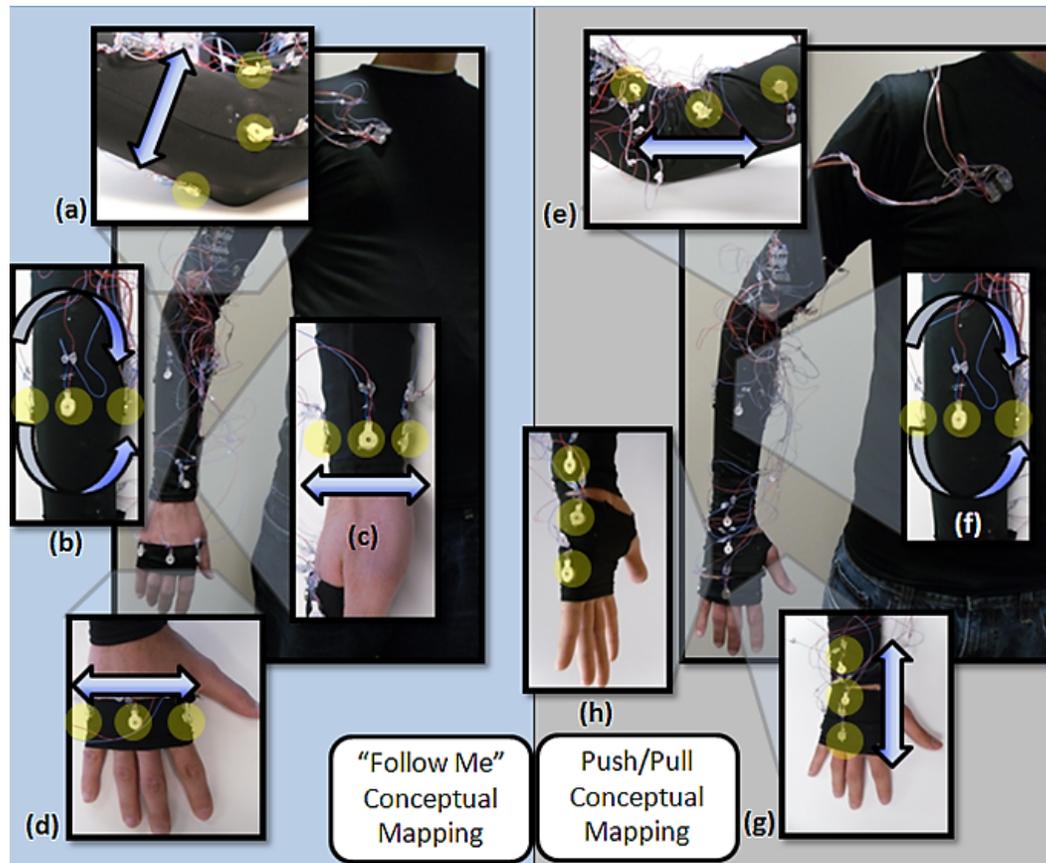


Figure 66. Visual depiction of the right arm fundamental movements involved in this work: (a) elbow flexion and extension; (b) wrist flexion and extension; (c) forearm pronation and supination; and (d) wrist abduction and adduction. *Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In Proceedings of the 19th ACM International Conference on Multimedia, p. 545. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.*

Given that saltation patterns inherently vary both spatially and temporally, only one vibrotome was needed to create the six vibrotactile words that make up the proposed instruction set for fundamental movements: flexion, extension, abduction, adduction, pronation and supination. This vibrotome was a short 100 ms vibration with fixed frequency and amplitude. This vocabulary was also augmented with a single vibrotactile alert appended to each word to direct attention to the incoming instruction; it is similar to the proposed vibrotome, but with a longer duration (500 ms). The purpose of the vibrolert is to aid language robustness by making the stimulations hard to miss.

The proposed vibrotactile words target specific joints across the body through the use of body context to enrich vocabulary. The appended alert also helps to direct attention to the body site that will soon receive instruction. Figure 67 depicts the

proposed vibrotactile words; note that two variations (*conceptual mappings*) are proposed based on how these instructions are taught to students: the “follow me” concept, and the push/pull metaphor.



*Figure 67.* Depiction of conceptual mappings “follow me” and push/pull for vibrotactile instructions: “follow me” is given by (a)-(d), and push/pull is given by (e)-(g). Highlights indicate motor locations, and arrows indicate directionality of vibrotactile stimulation during traversal of the skin. Rotations utilize four motors (motor on volar side of forearm is occluded in the figure) whereas all other movements utilize three. The rotation patterns depicted in (b) and (f) are identical—both use the “follow me” conceptual mapping. To create the saltation effect, motors are actuated four times each (three times each for rotations) with a brief pulse of 100 ms, separated by a 60 ms pause. The total duration of the proposed words is 2.56 seconds except for rotations, which are 3.04 seconds in duration since five motors are actuated rather than three; vibrotactile instructions for rotations come full circle to end on the start motor. *Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In Proceedings of the 19th ACM International Conference on Multimedia, p. 547. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.*

Under the “follow me” conceptual mapping, users were instructed to follow a vibration’s direction as it traversed across the skin. Vibration motors were arranged along cross-sections of skin such that saltation patterns ran tangential to the path of motion trajectories. This design was settled upon after pilot testing revealed its naturalness for following vibrations across the skin. Figure 67 (a)-(d) depicts the proposed vibrotactile words for “follow me” after they’ve been specified and configured for joints through body context—detailed measurements of motor spacing and placement relative to anatomical locations is given in (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011). Extensive pilot testing helped narrow the design space to settle upon these patterns in terms of distinctness and naturalness for the “follow me” conceptual mapping. An overview of the main results of pilot testing is given below (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011):

For elbow flexion/extension, saltation felt most natural when delivered to the volar aspect of the middle of the forearm or more proximal, near the elbow joint. The middle of the forearm should be avoided, however, to prevent confusion with vibrations for forearm rotations; as should more distal regions to avoid confusion with vibrations for wrist movements. For forearm rotations, saltation (conveyed by at least four motors) felt most natural anywhere on the forearm; but the middle portion is recommended to avoid vibrations for wrist and elbow movements. For wrist flexion/extension, saltation felt most natural when delivered to either side of the wrist joint (we used the medial side when the back of the hand is anterior to the palm of the hand). For any wrist movement, it is recommend to avoid placing motors across the wrist joint and onto the forearm as rotational movements will cause the forearm to move within the worn fabric, misaligning a configuration with its respective movement; in other words, if vibration patterns are to work well for any arm (or, limb, body, etc.) posture, then careful attention must be paid

to spatial variations of motors as movements are performed. Also, avoid placing motors on the palm as it may be obtrusive. Lastly, for wrist abduction/adduction, saltation felt most natural when delivered to the back of the hand on or below the knuckles, where the generous surface area provides sufficient spacing between individual motors, as well as with vibrations targeting wrist flexion/extension. In general, to improve distinctness, vibration patterns targeting different fundamental movements, e.g., rotations versus elbow flexion/extension, should not share motors, and be as far apart as possible. Lastly, within a configuration, motors must be spaced such that directionality is easily perceived. (p. 546) *Copyright © Association for Computing Machinery, Inc. Reprinted with permission.*

Under the push/pull metaphor, users were instructed to interpret vibration directionality as either pushing or pulling a limb. This is similar to Spelmezan et al.'s approach (2009), but different in that Spelmezan et al. used one or the other, whereas our approach combined them into push/pull to halve the number of motors to reduce cost and simplify design. Vibration motors were arranged across the involved joints. This design was well received during pilot testing, feeling natural for the chosen conceptual mapping. Figure 67 (e)-(h) shows how vibration motors are spaced and placed for the push/pull metaphor—again, refer to (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011) for detailed measurements. Pilot testing explored the naturalness and distinctness of configurations for the push/pull metaphor. The main results of pilot testing are summarized below (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011):

For elbow flexion/extension, saltation felt most natural when delivered to the volar aspect of the arm across the elbow joint, with the center motor on the elbow joint. Motors should be generously spaced apart so that when the arm is fully flexed, the

vibration pattern for extension may still be easily perceived. Vibration patterns for rotations were most intuitive when explained and delivered under the “follow me” concept... so no push/pull version is proposed. For wrist flexion/extension, saltation felt most natural when delivered to either the palm or back of the hand, but it is recommended to avoid the palm; and as described before, for wrist movements, motors should not be placed posterior to the wrist joint (and hence onto the forearm) to avoid complications arising from forearm rotations. For wrist abduction/adduction, saltation felt most natural when delivered to the lateral side of the hand when the back of the hand is anterior to the palm. (p. 547) *Copyright © Association for Computing Machinery, Inc. Reprinted with permission.*

The proposed somatic language exploits context to enrich the small set of words representing fundamental movement instructions. Through Somatic ABC’s use of context, a variety of joints may be actuated across the body. As described, this work focuses on the wrist and elbow joint of the right arm. Moreover, the language may be enriched through another means: sentence creation. More complex movements may be conveyed through spatial, temporal or spatio-temporal combinations of vibrotactile words to create rich sentences representing almost any type of human movement. Vibrotactile sentences for motor instruction will be explored as part of future work; in particular, the psychophysics of temporal and spatial variations of the words within vibrotactile sentences will be explored.

The high level of abstraction affords efficient communication of movements at least for novices learning the basic structure of more complex movements. Expert users whom already know the movements, but wish to perfect them in terms of coordination, control and precision, may benefit more from the proposed vibrotactile feedback,

described after the following implementation details and formal evaluation of the proposed vibrotactile motor instructions.

**Build: vibrotactile motor instructions.** A platform was built to realize and evaluate the proposed vibrotactile motor instruction set. In this section, the custom hardware, firmware and software of this platform, called the *haptic suit*, is presented.

***Haptic suit for vibrotactile motor instructions.*** This section first describes the hardware details of the haptic suit, followed by its firmware, and then finally, its software.

***Hardware description.*** The haptic suit is depicted in figure 68, and hardware components are given in figure 69. Hardware details are summarized below (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011):

The sleeve is part of a compression shirt (Men's medium; 84% polyester, 16% spandex). A LilyPad Arduino (ATmega328) microcontroller is powered using a LilyPad LiPower and a 2000 mAh Polymer Lithium Ion battery. To deliver power, stranded wires are used to reduce resistance. Thin, flexible, solid core wires are used to trigger motors. Wires are slack to provide flexibility when altering configurations, and to enable subjects to easily move while wearing the system. The microcontroller controls vibration motors (pancake motors; 150 Hz), attached with a small dab of hot glue that is easily removed when spatially altering motors. Motors are not directly connected to the microcontroller, but instead, are connected through nested 8-bit address latches (model#: 74HC259N). Within our implementation, latches are nested for two levels, enabling one microcontroller to support over 200 motors. Between a latch and a motor (each latch supports 7 motors) is a driver (Hi V & A Darlington Transistor Array; model#: ULN2004ANE4). (pp. 547-548) *Copyright © Association for Computing Machinery, Inc. Reprinted with permission.*

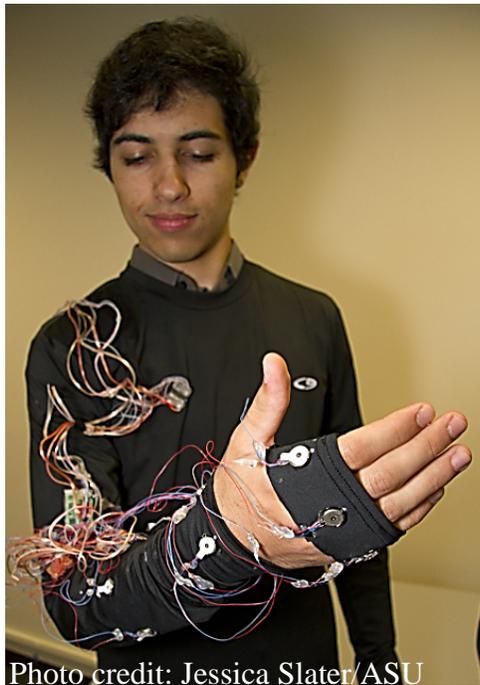


Photo credit: Jessica Slater/ASU

*Figure 68.* Haptic suit being worn and demonstrated by a student researcher.

A tight-fitting compression shirt was chosen as the form factor to ensure constant contact between the vibration motors and the wearer for reliable communication. The use of a hierarchy of 8-bit latches supported scalability—our current implementation uses 14 motors, but supports the addition of up to 200 motors, spread across the entire form factor. Motors can also be easily removed if needed. Each motor is held down with a small dab of glue to easily reconfigure its placement, particular for experimentation. Portability is supported through fast wireless connectivity and long battery life. System operation was found to be reliable and efficient for accurate presentation of spatio-temporal vibration patterns, including complex rhythms involving vibrations with durations as short as 50 ms. Although the thin wires used here support flexibility such that movement is not hindered, they can break easily if snagged on surrounding objects. To improve durability, an outer sleeve or shirt may be worn, which also improves

discreetness. Lastly, pilot testing revealed the suit to be comfortable, easy to don/doff, easy to learn and use. The latter two criteria largely depend on the distinctness and naturalness of the vibration patterns, which were supported through sufficient motor spacing and intuitive motor configurations and actuations using saltation.

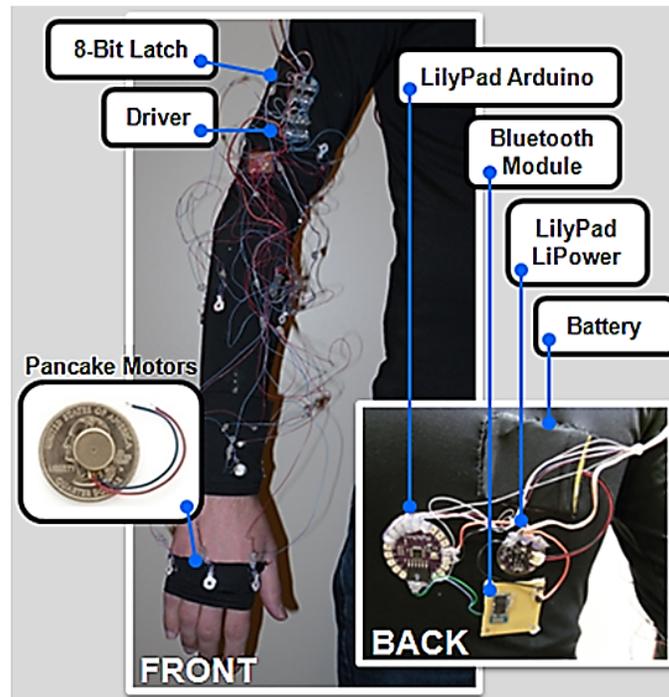
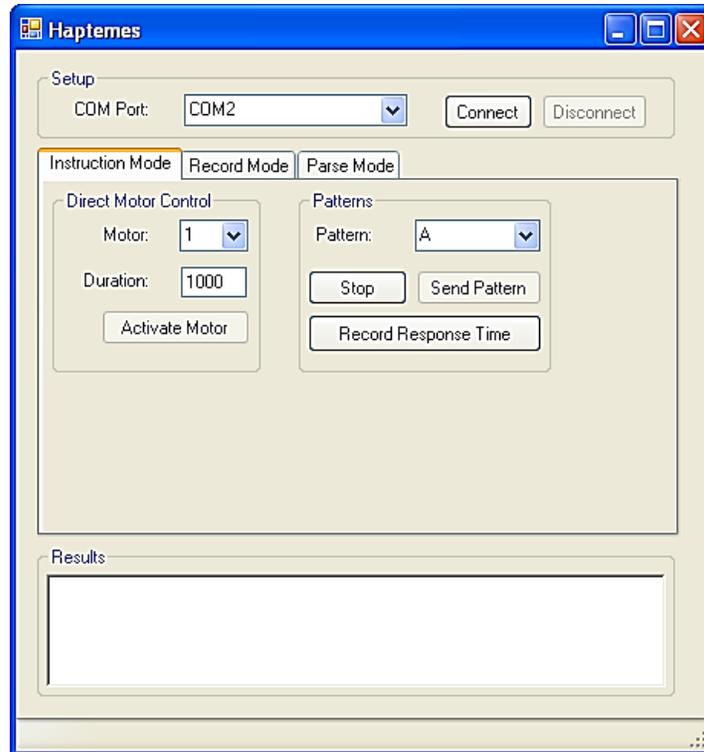


Figure 69. Hardware details of haptic suit. Back: Microcontroller, power supply and wireless connectivity; Front: Communication buses and vibration motors. *Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In Proceedings of the 19th ACM International Conference on Multimedia, p. 548. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.*

*Firmware description.* Firmware was developed using the Arduino development environment. A communications manager was implemented to receive input commands over serial via Bluetooth. These commands are parsed and used to trigger pre-defined vibrotactile motor instructions. Instructions were defined for each saltation pattern described in figure 67 where each definition was a sequence of motor actuations.

Although motor instructions must be implemented manually through the hard coding of timing variations, the firmware still supported reconfigurability and provided expressive control over defining rich spatio-temporal stimulations. As part of future work, a graphical authoring tool is being developed that will allow designs to be quickly and easily realized, and then outputted in a standard format for upload and storage onto a microcontroller.

***Vibrotactile motor instruction software.*** A graphical user interface, developed in Visual C#, was built for communicating between a computer and the haptic suit. The GUI allows users to connect to the haptic suit to trigger pre-defined vibrotactile motor instructions that were previously uploaded to the microcontroller. Pre-defined patterns are given letters, 'A', 'B', 'C', etc. for simplicity. Individual motors may be actuated at variable durations. Response times may be recorded for experimentation or training. Figure 70 shows a screenshot of the GUI.



*Figure 70.* Graphical user interface for triggering vibrotactile motor instructions and individual motors through the haptic suit. Results pertaining to system operation are displayed in real-time in the results box for debugging purposes. Other tabs, ‘Record Mode’ and ‘Parse Mode’, are used for vibrotactile feedback (described later).

**Confirm: vibrotactile motor instructions.** A formal evaluation was conducted to assess the distinctness and naturalness of the proposed vibrotactile motor instructions—and ultimately, evaluate the usefulness of Somatic ABC’s toward creating the proposed somatic language for cueing movements (for the intended use of motor learning). The evaluation performed by this study was psychophysical in that recognition and timing responses were assessed rather than exploring how the proposed system enhances motor learning compared to traditional instruction. The latter study was found to be outside of the scope of the present work, but will be conducted as part of future work. The follow excerpt is from McDaniel, Morris, Villanueva, Viswanathan and Panchanathan (2011), where the study was originally published:

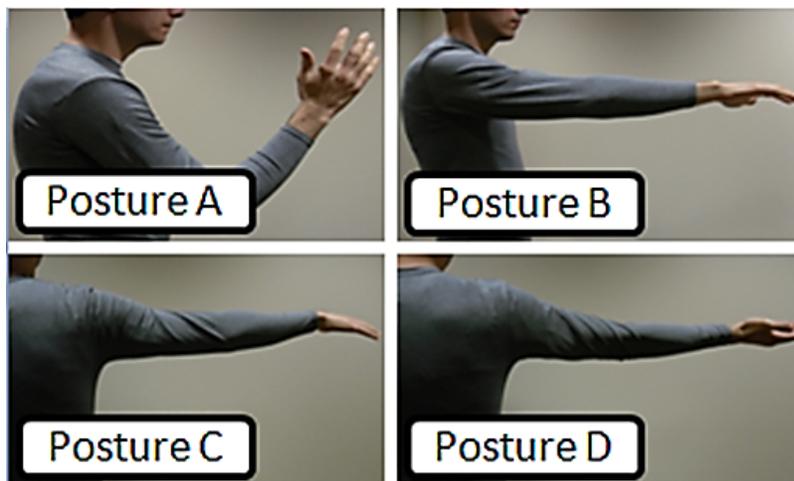
**Aim.** The purpose of this study is to explore the naturalness of the proposed kinematic-vibrotactile mapping [the proposed vibrotactile words]; in particular, we wish to explore how the “follow me” concept and push/pull metaphor affect naturalness. Naturalness is primarily investigated through subjective feedback, but learning rate, recognition accuracy, and response time may also shed light on the usefulness of the conceptual mappings. It is important to note that the intuitiveness of a conceptual mapping is closely linked to motor spacing and placement (configuration); we’ve accounted for this through extensive pilot testing to find the most useful and natural configurations for each fundamental movement of the two conceptual mappings. Moreover, we cannot assume that vibration patterns, after being learned in one posture, will generalize to different postures. Ideally, however, we’d prefer *posture-free* vibration patterns that generalize well to other postures after being mastered in one training posture. To this end, we explore how well the proposed vibration patterns generalize to novel postures (various arm postures) depicted in figure 71.

**Subjects.** The experiment involved 20 subjects, all Arizona State University students, divided between two conditions. The “follow me” condition involved 8 males and 2 females (age range: 19 to 27; mean: 24); and the push/pull condition also involved 8 males and 2 female (age range: 20 to 34; mean: 25). No subjects had motor or tactile impairments.

**Procedure.** Subject information including age, sex, height and weight was collected. The experiment was briefly explained to participants, after which they donned the wearable system [figure 67 (a) or (b) depending on assigned condition]...The experiment consisted of three phases: a familiarization, training, and two-part testing phase. The experimenter explained the randomly assigned condition, which was either the “follow me” or push/pull conceptual mapping. During the entire study, with the

exception of the second part of testing, subjects were asked to remain standing with their arms by their sides (training posture). During the familiarization phase, each vibration pattern of the assigned conceptual mapping was sequentially presented; before each presentation, the experimenter demonstrated the movement and explained the stimulation, relating it to its conceptual mapping. To avoid confusion, layman terminology...was used to specify fundamental movements: for example, 'wrist up' rather than 'wrist extension'. For simplicity, since wrist abduction/adduction is depended upon the posture of the hand with respect to the sagittal plane, they are taught in posture B [figure 71]...and assumed to remain the same across different postures, 'Wrist Left' and 'Wrist Right', respectively...Once completing the first pass through the patterns, the vibration patterns were delivered once more. During the training phase, training trials were repeated unless the subject scored a recognition accuracy of at least 80% (7 out of 8 patterns) during a trial. A single training trial involved the random presentation of all eight vibration patterns, once each. Participants were told to respond with the movement the vibration cued, as quickly, but also as accurately, as possible. The experimenter informed the subject about the correctness of each response; if the movement was incorrect, the experimenter demonstrated the correct movement, and presented the pattern once more. During each phase, the experimenter recorded learning rate (training phase only), response correctness and response time. Learning rate is the number of training trials required before the subject passes on to testing. The correctness of each response is used to derive recognition accuracy, or the percentage of correct responses. Response time is the duration between the start time of the presentation of the pattern, and the time at which the subject began performing the correct movement; if incorrect movements were performed first, but then corrected by performing the correct movement, within a time limit of 15 seconds, the response was marked as correct.

The first part of the testing phase was similar to the training phase with the exception that four trials (32 presentations total with four random presentations per pattern) were performed for each subject, and no feedback was given. During the second part of the testing phase, four new postures [figure 71]...were introduced. The experimenter demonstrated each posture...Each vibration pattern was presented once for each posture, for a total of 32 presentations. Presentation pairs (*posture, vibration pattern*) were randomized. Before each presentation, the participant was informed which posture to change to, after which, the pattern was presented. No feedback was given. Finally, subjects were asked to fill out a questionnaire. (pp. 548-549) *Copyright* © *Association for Computing Machinery, Inc. Reprinted with permission.*



*Figure 71. Novel postures used in user study to assess if recognition performance of vibrotactile motor instructions is independent of postures. Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In Proceedings of the 19th ACM International Conference on Multimedia, p. 548. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.*

Results related to learning rate, recognition accuracy and response time were collected, summarized (with descriptive statistics) and analyzed for relevant significant differences using Analysis of Variance. The following excerpt from McDaniel, Morris, Villanueva, Viswanathan and Panchanathan (2011) provides a summary of the results:

**Results.** The mean average number of learning trials was 1.9 (SD: 0.99) and 1.4 (SD: 0.7) for “follow me” and push/pull conditions, respectively. Recognition accuracies and classifications for each vibration pattern are summarized in [figure 72].... For the “follow me” and push/pull conditions, the overall recognition accuracy for the first part of testing was 97% (SD: 8.8%) and 98% (SD: 6.1%), respectively; and 98% (SD: 8.1%) and 94% (SD: 14.5%) for the second part. Mean response times for each vibration pattern are summarized in [figure 73].... For the “follow me” and push/pull conditions, the overall response time for the training phase was 3.6 s (SD: 1.59 s) and 2.8 s (SD: 0.72 s), respectively; for the first part of testing, 2.9 s (SD: 0.96 s) and 2.5 s (SD: 5.9 s); and for the second part of testing, 2.9 s (SD: 0.86 s) and 2.5 s (SD: 0.59 s)...[Table 5] summarizes results from the post-experiment questionnaire where subjects rated a series of questions using a Likert scale from 1 (low/difficult) to 5 (high/easy)...[Table 5] summarizes results pertaining to the subjective naturalness of each vibration pattern, where subjects rated each pattern’s naturalness as ‘excellent’ (perfect or near perfect), ‘acceptable’ (satisfactory) or ‘unacceptable’ (needs improvement). (p. 549) *Copyright © Association for Computing Machinery, Inc. Reprinted with permission.*

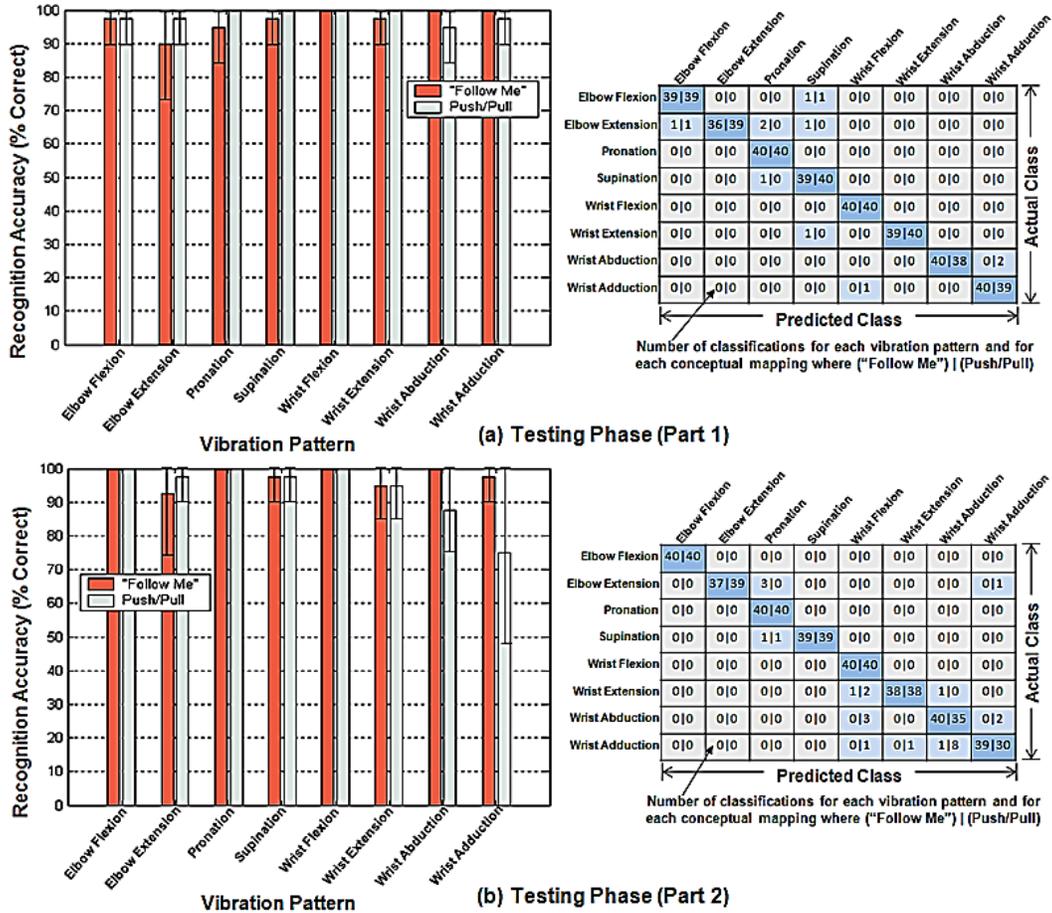


Figure 72. Summary of mean recognition accuracies averaged across participants (left) and misclassifications displayed using confusion matrices (right) for vibrotactile motor instructions under (a) the first phase of testing, and (b) the second phase of testing (novel postures). Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In *Proceedings of the 19th ACM International Conference on Multimedia*, p. 550. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.

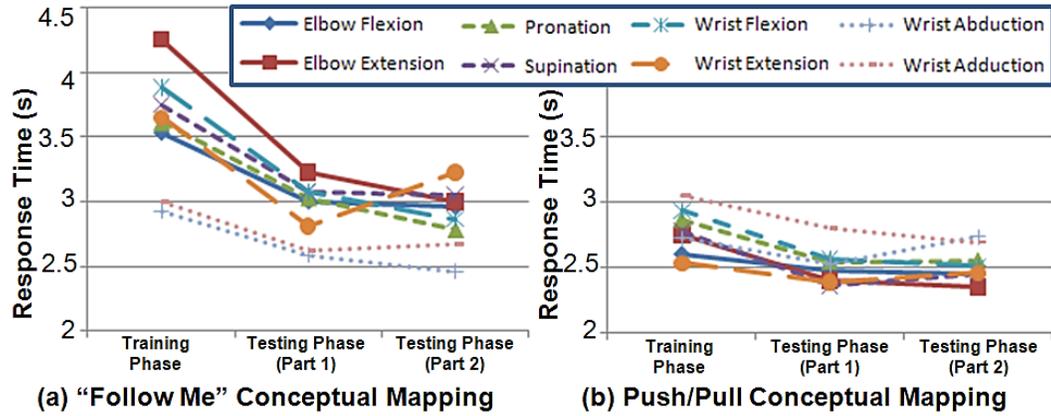


Figure 73. Mean response times averaged across participants for each vibrotactile motor instruction under training, testing (first phase) or testing (second phase—novel postures). Results for both conceptual mappings—(a) “follow me” and (b) push/pull—are shown. Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In *Proceedings of the 19th ACM International Conference on Multimedia*, p. 550. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.

Table 5

*Mean Responses to the Post-Experiment Questionnaire for Vibrotactile Motor Instructions*

Questions	Follow Me		Push/Pull	
	Mean	SD	Mean	SD
1) How easy was it to put on the suit?	3.6	0.42	4	0.67
2) How easy was it to take off the suit?	3.6	0.96	3.7	0.95
3) How easy was it to perform the movements with the suit on?	4.7	0.67	4.9	0.32
4) How comfortable was the suit?	4	0.82	3.9	0.88
5) How lightweight was the suit?	4.9	0.32	5	0
6) How silent were the suit's vibration motors?	4.1	0.57	4.1	0.57
7a) How easy was it to learn the vibration pattern for 'wrist left' (Wrist Adduction)?	4.9	0.32	4	1.25
7b) ... for 'wrist right' (Wrist Abduction)?	4.9	0.32	4	1.05
7c) ... for 'wrist up' (Wrist Extension)?	4.3	0.67	4.9	0.32
7d) ... for 'wrist down' (Wrist Flexion)?	4.4	0.70	4.7	0.48
7e) ... for 'rotate right' (Supination)?	4.6	0.69	4.4	1.07
7f) ... for 'rotate left' (Pronation)?	4.6	0.69	4.4	1.07
7g) ... for 'elbow flex' (Elbow Flexion)?	4.1	0.88	4.9	0.32
7h) ... for 'elbow extend' (Elbow Extension)?	4	0.82	4.9	0.32
8a) How easy was it to recognize & respond to vibration for 'wrist left' (Wrist Adduction)?	5	0	3.8	1.03
8b) ... for 'wrist right' (Wrist Abduction)?	5	0	4	0.82
8c) ... for 'wrist up' (Wrist Extension)?	4.4	0.94	4.7	0.67
8d) ... for 'wrist down' (Wrist Flexion)?	4.5	0.96	4.7	0.67
8e) ... for 'rotate right' (Supination)?	4.3	0.63	4.3	1.06
8f) ... for 'rotate left' (Pronation)?	4.2	0.75	4.3	1.06
8g) ... for 'elbow flex' (Elbow Flexion)?	4	0.94	4.8	0.42
8h) ... for 'elbow extend' (Elbow Extension)?	3.9	0.99	4.8	0.42

*Note.* Conceptual mappings: “follow me” (left) and push/pull (right). Layman terminology was used to describe movements. Each question was based on a Likert scale (1 through 5). Reprinted from “Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements,” by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In *Proceedings of the 19th ACM International Conference on Multimedia*, p. 551. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.

Table 6

*Subjective Evaluation of Naturalness of Vibrotactile Motor Instructions*

Vibration Patterns	Excellent	Acceptable	Unacceptable
'Wrist Left' (Wrist Adduction)	9 0	1 6	0 4
'Wrist Right' (Wrist Abduction)	9 0	1 8	0 2
'Wrist Up' (Wrist Extension)	2 5	7 5	1 0
'Wrist Down' (Wrist Flexion)	2 4	7 6	1 0
'Rotate Right' (Supination)	6 7	4 2	0 1
'Rotate Left' (Pronation)	6 7	4 2	0 1
'Elbow Flex' (Elbow Flexion)	1 9	8 1	1 0
'Elbow Extend' (Elbow Extension)	1 8	8 1	1 1

*Note.* Votes for either 'excellent', 'acceptable' or 'unacceptable' are out of ten where the following convention is used: "follow me" | "push/pull". Reprinted from "Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements," by McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S., 2011, In *Proceedings of the 19th ACM International Conference on Multimedia*, p. 551. Copyright © 2011 by Association for Computing Machinery, Inc. Reprinted with permission.

**Discussion.** The aforementioned results provide insight into the distinctness and naturalness of the proposed vibrotactile words. The following discussion, originally presented in (McDaniel, Goldberg, Villanueva, Viswanathan, & Panchanathan, 2011), provides a detailed discussion of the results:

*Learning rate.* The average number of learning trials did not differ significantly between conditions,  $t(18)=1.30$ ,  $p>0.2$ , two-tailed, showing that both conceptual mappings were easy to learn.

*Recognition accuracy.* For the first part of testing, the overall recognition accuracy (across subjects) of each vibration pattern (and for either condition) is impressive at 90% or better, with most accuracies being in the high 90's [figure 72(a)]. Moreover, a one-way repeated measure ANOVA revealed that recognition accuracies between vibration patterns did not differ significantly,  $F(7,63)=1.52$ ,  $p>0.05$ , and

$F(7,63)=0.93$ ,  $p>0.05$ , for the “follow me” and push/pull conditions, respectively. This shows that within each condition, patterns were distinct and easy to recognize. For the second part of testing in which novel postures were introduced, the overall recognition accuracy (across subjects and postures) of each vibration pattern (for either condition) is impressive given no prior training on the novel postures...most accuracies are 90% or better [figure 72(b)], showing that most patterns, for either condition, were still distinct and easy to recognize even for new postures. However, for the push/pull condition, wrist abduction and adduction were both below 90% at 88% (SD: 13.1%) and 75% (SD: 28.9%), respectively. A two-way repeated measure ANOVA revealed that the main effects for vibration pattern and posture were both significant,  $F(7,63)=5.14$ ,  $p<0.0002$ , and  $F(3,27)=4.33$ ,  $p<0.05$ , as well as their interaction,  $F(21,189)=3.1$ ,  $p<2\times 10^{-5}$ . Regarding the main effect of pattern type...[figure 72(b)] suggests lower recognition accuracy for wrist adduction compared to other patterns, regardless of posture. Although we observed slight difficulties with recognizing this pattern while in posture A, B and C, it was posture D that presented the biggest challenge. Regarding the main effect of posture, we observed posture D to have lower overall recognition accuracy, regardless of pattern type, when compared to other postures. However, we observed that the patterns of wrist abduction and adduction created the most problems for participants while in posture D (interaction effect). Overall wrist abduction and adduction accuracy, while in posture D, were both very low at 50% (SD: 52.7%) each. As shown in the confusion matrix [figure 72(b)]...all five misclassifications of wrist abduction occurred in posture D, whereas half (five out of ten) misclassifications of wrist adduction occurred in posture D; most of the confusion happened between wrist movements. Subjective feedback confirmed the difficulty of recognizing wrist abduction and adduction patterns in posture D for the push/pull condition: many subjects commented that wrist abduction and

adduction for push/pull were very difficult to recognize while in posture D due to the (rotated) hand posture. Indeed, in (8) [table 5]...we see that wrist abduction/adduction were the lowest rated among other patterns in the push/pull condition.

*Response time.* After training, overall response times for either condition and for any pattern were impressive, at roughly three seconds or less. [figure 73]...shows a general decrease in overall response time (across subjects) for vibration patterns as subjects progressed from training to the first part of testing; then seemingly stabilizing between the first and second part of testing with some small increases or decreases depending on the pattern and condition. A two-way repeated measure ANOVA revealed the main effect of phase type to be significant,  $F(2,18)=15.87$ ,  $p<1.1\times 10^{-4}$  and  $F(2,18)=15.53$ ,  $p<1.21\times 10^{-4}$ , for “follow me” and push/pull conditions, respectively...this suggests that with continued exposure to the patterns, reaction times improved, with perhaps the exception of the transition between the two parts of testing. This may be due to the introduction of the novel postures, or perhaps more time was needed before we saw further improvements in terms of response time. We hypothesize that over long term use, users will continue to become more proficient at recognizing and responding to the patterns. Only for the “follow me” condition was the main effect of pattern type significant,  $F(7,63)=4.13$ ,  $p<8.61\times 10^{-4}$ . Indeed...we see that patterns for wrist abduction and adduction were recognized faster on average compared to other patterns [figure 73]. This coincides with subjective feedback: see (8) [table 5]. As expected, this indicates that more natural patterns...will lead to faster response times [table 6]. No significant interaction effects were found for either condition.

*Posture-free vibrations.* With the exception of wrist abduction/adduction for the push/pull condition, based on the impressive recognition accuracies when novel postures were introduced, along with consistent response times, we see that the proposed

conceptual mappings and configurations generalize well to new postures that are different from the training posture. This is important as we cannot expect users to re-learn vibration patterns for every new posture they might encounter, which would be unrealistic for many applications. However, we cannot ignore that the vibration pattern for wrist abduction/adduction did not perform well for every posture. We hypothesize that the ideal solution will involve both conceptual mappings, utilizing the most natural patterns.

*Subjective feedback.* For the “follow me” condition, vibration patterns for wrist abduction/adduction were rated higher in terms of learnability and distinctness [table 5]...as well as naturalness [table 6]...where all but one subject rated the patterns as ‘excellent’ in terms of naturalness; whereas wrist abduction/adduction for the push/pull condition received no ‘excellent’ ratings—mostly ‘acceptable’ or ‘unacceptable’. As previously mentioned, subjects felt the latter vibration patterns to be too similar and close to those of wrist flexion/extension. It seems obvious, then, that wrist abductions and adductions should be cued using the “follow me” conceptual mapping with the respective configuration. This will allow for sufficient spacing between wrist flexion and extension vibrations. Wrist flexion/extension under the push/pull condition received higher ratings for learnability and distinctness...as well as naturalness...compared to the “follow me” condition. Most ratings for the naturalness of wrist flexion/extension, for the “follow me” condition, fell under ‘acceptable’; many subjects felt the vibration patterns were more appropriate for rotations, although these patterns were rarely misclassified as such [figure 72]... The ideal configuration would have motors in a straight line such that the directionality is tangential to the arc of the motion; however, due to the curvature of the skin around the arm, especially around the wrist joint, there is a tradeoff between motor spacing and the curvature of the directionality. Enough spacing is required to provide the

illusion of apparent motion, but with larger spacing, motors will cover a greater circumference around the arm. This is an inherent problem when using the “follow me” conceptual mapping to design configurations for flexion and extension, at least where there is limited flatness. Therefore, the conceptual mapping of push/pull seems to be a better option for movements of flexions and extensions. For elbow flexion/extension, there is a clear preference for the push/pull version...As shown [table 6]...most ratings were ‘excellent’ whereas most ratings for the “follow me” condition were ‘acceptable’. As mentioned, for the “follow me” conceptual mapping, these patterns share the same problem as those for wrist flexion/extension. Indeed, we see that most misclassifications were with rotations...Lastly, most subjects felt vibration patterns for rotations to be intuitive, easy to learn, and easy to recognize. It is therefore clear that a combination of patterns from the two conceptual mappings explored here is needed rather than using one concept to explain all kinematic-vibrotactile mappings. The most effective patterns from each conceptual mapping should be used: “follow me” wrist abduction/adduction, push/pull wrist flexion/extension, push/pull elbow flexion/extension, and “follow me” rotations. (pp. 549-551) *Copyright © Association for Computing Machinery, Inc. Reprinted with permission.*

**Conclusion and future work: vibrotactile motor instructions.** Somatic ABC’s was applied to design, develop and implement a somatic language for vibrotactile motor instructions. This language was evaluated through a psychophysical study that assessed distinctness and naturalness. Overall, participants found the proposed vibrotactile motor instructions easy to learn and recognize given their distinctness and naturalness. Augmenting the proposed somatic language with vibrotactile feedback is a clear extension to this work, which is described in the following section. Further research questions related to vibrotactile motor instructions include:

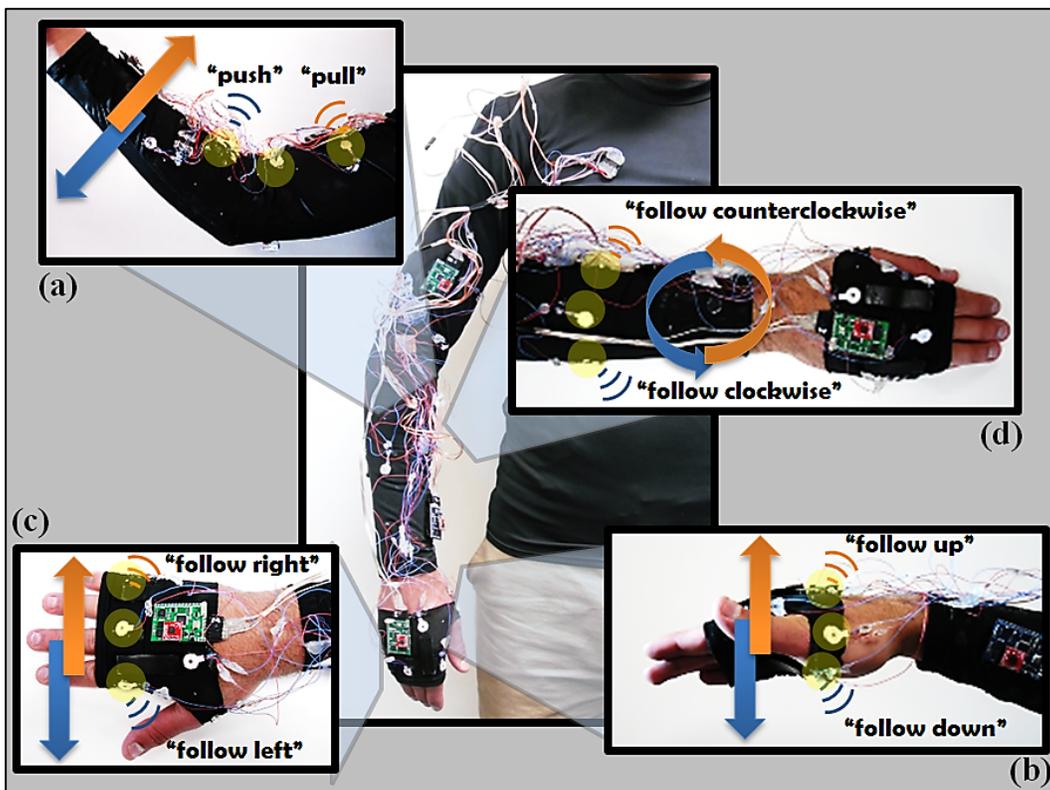
- Learning, recognition and response effects when different conceptual mappings (namely, “follow me” and push/pull) are combined into the same pedagogy. Although this study revealed that one conceptual mapping may not be optimal for all movements, future work must explore ways in which these conceptual mappings may be integrated without introducing unwanted perceptual effects.
- Spatio-temporal presentations of vibrotactile motor instructions will be explored to learn how to cue more complex movements consisting of spatially and/or temporally overlapping fundamental movements. The perceptual effects of different scenarios (concurrent and/or sequential presentation) will be assessed for different movements across the arm.
- Generalization of the results found here will be explored across the body at different joints and structures; in particular, vibrotactile motor instructions for flexion/extension, abduction/adduction and rotation will be applied to different joints capable of articulating these same fundamental movements. These results will provide insight into the potential of body context for this application.

**Articulate: vibrotactile feedback.** The proposed somatic language for vibrotactile motor instructions was extended to accommodate two types of feedback: (1) positioning errors in terms of joint angle and degree of rotation; and (2) speed errors in terms of the angular speed of joints and rotational movements. These two features of movement were selected for feedback as they represent important error information when learning and perfecting movements. In the following sections, the design of vibrotactile feedback for positioning and speed errors is described.

*Feedback for positioning errors.* The proposed design is inspired by interactions during physical therapy where a therapist applies gentle nudges to guide or direct attention to limbs that need adjustment. We extend the aforementioned somatic language (instruction set) by one vibrotactile word for positioning feedback. The word is a vibrotactile rhythm built from the sequential presentation of one vibroteme (a vibration of duration 120 ms with fixed amplitude and frequency), each separated by a gap of 120 ms. The rhythm feels like quick, gentle nudges guiding a limb to specific angle, after which the vibration ceases to indicate that the target position has been reached. Pilot testing revealed the frustrations of reaching a precise angle. These frustration were alleviated when a padding (acceptable amount of error), such as +/- 5 degrees, was introduced, improving system usability. Ultimately, this padding will be application-specific, dependent upon how much precision is required. Pilot testing confirmed the naturalness of the “tapping”; but participants also found a steady vibration to be natural and effective for positioning feedback. The latter is recommended for applications requiring precise positioning with small errors as pauses between bursts of the former rhythm may increase chances of passing over small paddings wherein the target angle lies.

Body context was employed to enrich the proposed vibrotactile word by applying it to different joints for joint-specific positioning feedback. The final spacing and

placement of motors, determined after extensive pilot testing evaluating naturalness and distinctness, is shown in figure 74. This design is not independent from the vibrotactile motor instructions described before—the proposed instruction and feedback designs are intended to be used together, particularly for novices; expert users have the option of using only feedback for perfecting movements. Moreover, similar to vibrotactile motor instructions, the pedagogical approach to teaching these feedback signals is inspired by either the “follow me” or push/pull conceptual mapping—which also influences the placement of motors. The following describes the feedback design for each fundamental movement of interest here: elbow flexion/extension, wrist flexion/extension, wrist abduction/adduction and forearm pronation/supination.



*Figure 74.* Motor configurations and conceptual mappings used for vibrotactile positioning feedback. Pulses indicate locations of vibration motors, and arrows indicate intended direction of movement based on the stimulated body site and conceptual mapping. Movements depicted include (a) elbow flexion/extension; (b) wrist flexion/extension; (c) wrist abduction/adduction; and (d) forearm supination/pronation.

Positioning feedback for elbow flexion/extension shares the motor configuration (and conceptual mapping) used for cueing elbow flexion/extension movements under the push/pull metaphor (figure 67). This configuration was chosen for augmenting with positioning feedback as participants found it more natural than “follow me” during formal evaluation. After perceiving an instruction to flex or extend the arm at the elbow joint, the user moves to what he or she believes to be the correct position, and then stops. If the angle is in error, the user feels one of two types of vibrotactile feedback: gentle nudges on the volar side of the forearm taught to be perceived as pushing the forearm for extension; or gentle nudges on the bicep taught to be perceived as pulling the forearm for flexion—these are depicted in figure 74(a). These stimulations are felt until the user reaches a pre-defined position determined by the relative angle between the forearm and upper arm.

Positioning feedback for wrist flexion/extension moved the motor arrangement of figure 67(c) onto the medial side of the hand (near the thumb) just anterior to the wrist joint. As shown in figure 74(b), the spacing was slightly widened such that endpoints fall on the palm and back of the hand. These changes were made mainly to accommodate space needed for motion sensors—but ultimately, pilot test participants found the updated configuration to provide feedback signals that were vivid and natural given their direct stimulation of the hand. The “follow me” conceptual mapping was chosen over the push/pull metaphor as study participants found either of these to work satisfactorily for wrist flexion/extension (table 5). After perceiving an instruction to flex or extend the hand at the wrist joint, the user moves to what he or she believes to be the correct position, and then stops. If in error, the user feels either gentle nudges on the palm (just anterior to the wrist joint) taught to be followed to flex; or gentle nudges on the back of the hand (just anterior to the wrist joint) taught to be followed to extend—depicted in

figure 74(b). These stimulations are felt until the user reaches a pre-defined position determined by the relative angle between the hand and forearm.

Positioning feedback for wrist abduction/adduction moved the configuration of figure 67(d) distally along the fingers to accommodate space needed for motion sensors. This configuration and conceptual mapping was decided as formal evaluation of instructions found it more natural than the push/pull metaphor for wrist abduction/adduction. After perceiving an instruction to abduct or adduct the hand, the user moves to what he or she believes to be the correct position, and then stops. If in error, the user feels either gentle nudges on the medial side of the hand (on the index finger of the right hand) taught to be followed to adduct the hand; or gentle nudges on the lateral side of the hand (on the little finger of the right hand) taught to be followed to abduct the hand—depicted in figure 74(c). These stimulations are felt until the user reaches a pre-defined position determined by the relative angle between the hand and forearm.

Positioning feedback for adjustments to forearm pronation or supination shares motors of the configuration and conceptual mapping of figure 67(b). The “follow me” conceptual mapping was shown to be more intuitive than the push/pull mapping during pilot testing conducted as part of the investigation of vibrotactile motor instructions. After perceiving an instruction to rotate the forearm clockwise or counterclockwise, the user rotates to what he or she believes to be the correct position, and then stops. If the degree of rotation is in error, the user feels either gentle nudges on the medial side of the forearm taught to be followed to rotate the arm counterclockwise (pronation); or gentle nudges on the lateral side of the forearm taught to be followed to rotate the arm clockwise (supination)—these are depicted in figure 74(d). These stimulations are felt until the user

reaches a pre-defined position determined by the orientation of the hand relative to the upper arm.

Pilot testing supported the naturalness of the aforementioned proposed vibrotactile feedback signals for correcting positioning error. Pilot tests also showed the useful interactivity of the proposed feedback signals. Typically, on first attempts at reaching a target angle, users would overshoot, but then follow the feedback signal back, eventually finding the pre-defined target angle. Participants appreciated the system's interactivity in which they could get a feel of the position they needed to move into. They would then attempt to memorize this target position, and on subsequent tries, try to reach it without activating the feedback signals. This would usually take two to three tries with feedback signals providing slight adjustments. Upon reaching the target angle, and holding its position for less than a second, a vibration signal runs up the length of the arm, indicating that the user has achieved the correct position. Although concurrent feedback for multiple joints is possible, it may overwhelm novice users; recall that too much feedback may distract students, causing them to lose focus of more important errors that need to be reduced. In any case, concurrent feedback for positioning errors will be explored as part of future work. Vibrotactile sentences in the form of sequential corrections for positioning errors across different joints will also be explored.

***Feedback for speed errors.*** A novel approach for vibrotactile feedback for correcting errors related to angular speed is presented here. Similar to vibrotactile feedback for positioning errors, feedback for correcting speed errors informs users of the *direction* to make adjustments; specifically, speed up or slow down, rather than conveying an exact speed that needs to be reached. This is similar to feedback for positioning errors in that the feedback signal simply conveys that the current position needs to be increased or decreased. But just as with positioning feedback, speed feedback

may indicate when the correct speed is reached (plus or minus an acceptable amount of error) by either ceasing feedback vibrations or displaying novel vibration patterns indicative of achieving the correct movement.

Two presentation techniques for speed feedback were considered and evaluated through pilot tests: Real-time feedback or near real-time feedback. The design of both presentation schemes uses a vibrotactile rhythm for communicating speed information. For each presentation technique, this rhythm was presented to the elbow of the right arm so that it is common across all movements, and improves distinctness by avoiding body sites where instructions or positioning feedback were presented. During pilot tests, this design concept was preferred by participants, over more localized feedback for speed corrections; many participants commented that it simplified use (while still maintaining naturalness) as they knew where to expect the incoming feedback stimulation.

For real-time feedback, the tempo of a vibrotactile rhythm was coupled to the speed of a user's movement, and displayed in real-time while the user moved. As an alternative, we also discretized this range of speeds into categories of slow, moderate and fast. Through extensive pilot testing, neither approach for real-time feedback was found to work well given that the rhythm varied too quickly throughout movements. We speculate that the variation was caused by the short range of motion involved in the movements investigated as part of this work, combined with acceleration and deceleration at the start and stop of movements. Given the short range of motions involved, and our sampling rate of 8 samples per second, time and data constraints were insufficient for real-time speed feedback.

We therefore opted for the second presentation style of near real-time feedback. After a user feels a vibrotactile instruction, makes a movement, and then stops, he or she feels a vibrotactile rhythm communicating the needed speed adjustment. After a short

pause and a vibrotactile alert (1 s in duration with fixed amplitude and frequency) applied to the hand indicating to begin moving, the user then responds with an updated speed. This process repeats until the target speed is reached, after which a vibrotactile stimulation is felt representing that the goal has been achieved (a vivid and distinct vibration running up the length of the arm). Pilot test participants appreciated this interactivity, and were able to reach target speeds quite easily.

The design of the proposed vibrotactile rhythm utilizes a single word constructed from one vibrotème: a vibration of duration 200 ms with fixed amplitude and frequency, separated by gaps of 500 ms. On its own, this “base” rhythm does not convey any information until its tempo changes through use of stimulation variations. These are applied to the proposed vibrotactile rhythm to create rhythms that inform the user of speed adjustments when relatively compared to the base rhythm. The stimulation variation applied here is a change in tempo: a decrease in tempo by half, or an increase in tempo by double; these tempo changes form a *slow down* and *speed up* rhythm, respectively, depicted in figure 75, concatenated to the base rhythm to indicate how the previous movement needs to be adjusted. Stimulation variations were chosen to enrich the proposed somatic language without increasing learning demands. No new vibrotactile patterns must be memorized; users simply compare tempo changes of a base rhythm to learn if they need to speed up or slow down their next movement. Early pilot testing revealed the ease of recognizing these relative comparisons with only brief familiarization.

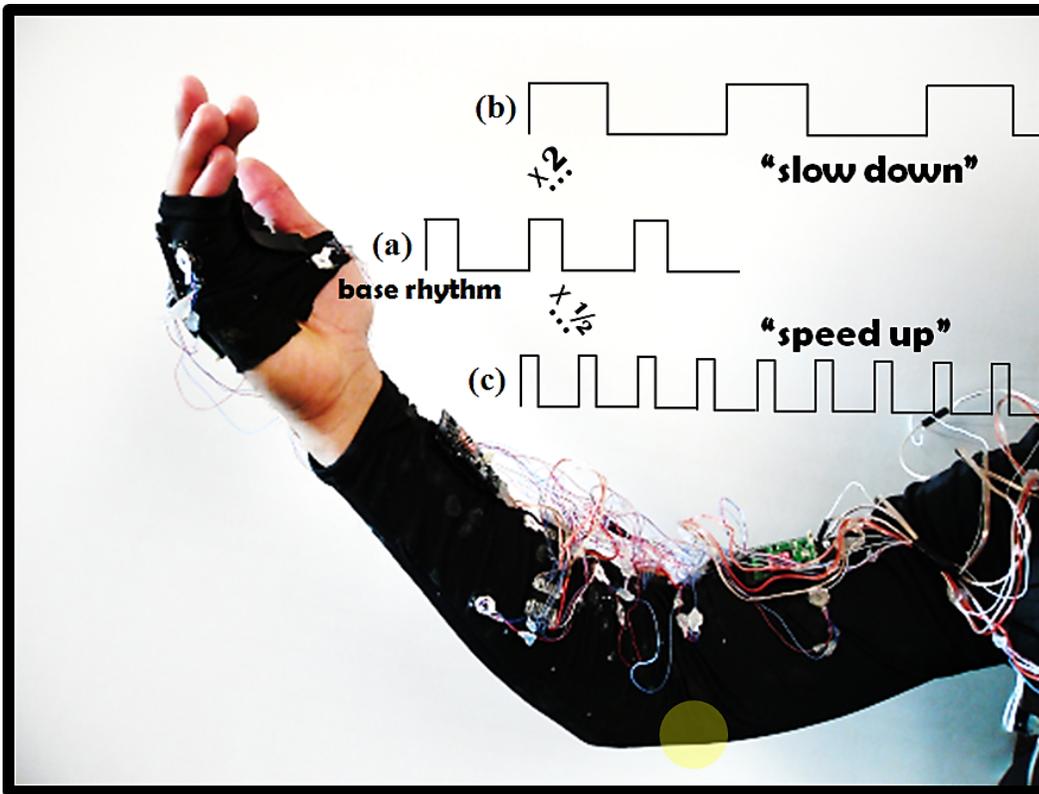


Figure 75. Visual representation of vibrotactile rhythms used in the proposed vibrotactile feedback for correcting speed errors: (a) base rhythm consisting of three pulses (200 ms each, separated by 500 ms) with a total duration of 2.1 seconds at 1.428 pulses/second; followed by (b) “slow down” rhythm consisting of three pulses (400 ms each, separated by 1000 ms) with a total duration of 4.2 seconds at 0.714 pulses/second; or followed by (c) “speed up” rhythm consisting of nine pulses (100 ms each, separated by 250 ms) with a total duration of 3.15 seconds at 2.857 pulses/second. These rhythms are presented to a vibration motor near the elbow.

**Build: vibrotactile feedback.** Somatic ABC’s implementation theory was employed to augment the proposed haptic suit for vibrotactile motor instruction with hardware and software for vibrotactile feedback. This new implementation is termed the *haptic feedback suit*. In the follow sections, the hardware and firmware details of the haptic feedback suit are described; followed by a description of the software for sending commands and recording/parsing movements.

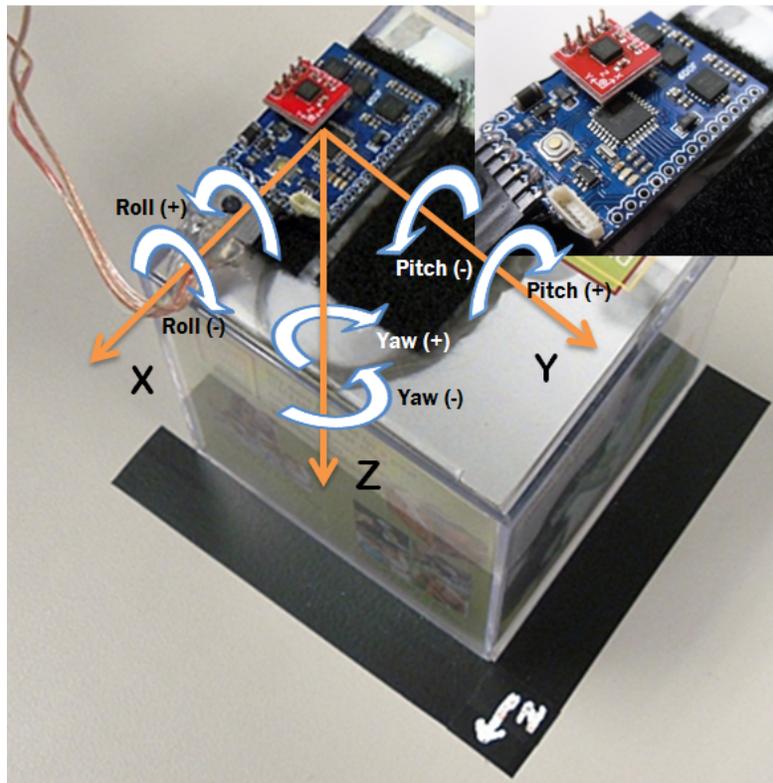
*Haptic feedback suit.* First, hardware augmentations made to the haptic suit to build the haptic feedback suit are described. Next, firmware details including calculation of the relative positions between limbs, is described.

*Hardware description.* The proposed haptic suit for vibrotactile motor instructions was augmented with motion sensing capabilities to drive vibrotactile feedback. Rather than rely on bulky and expensive visual motion capture, inertial measurement units (IMUs) were used. IMUs sense both acceleration (translation) and angular velocity (rotation) through an accelerometer and gyroscope, respectively, to measure the orientation (roll, pitch and yaw) of the sensor relative to Earth. Relative orientations between IMUs may be calculated by comparing these measurements. IMUs require a magnetometer (or GPS) to compensate for accumulated errors in yaw due to drift.

The IMU chosen for this work was the ArduIMU+ V2 (Flat), which has the following electronic components and features: triple-axis accelerometer (ADXL335), triple-axis gyroscope (LPR530AL, LY530ALH), Arduino-compatible on-board processor (Atmega328@16mhz), power regulation, protection circuitry, serial port output, and status LEDs. Triple-axis magnetometers were not included, but later purchased (HMC5843) and connected to each IMU. The on-board Atmega328 provided local, efficient processing needed for real-time motion capture. Existing firmware, the Attitude Heading Reference System (AHRS), was uploaded and used on each IMU for calculating orientation (roll, pitch and yaw), which could be efficiently and reliably communicated over serial. Figure 76 depicts the ArduIMU+ V2 (Flat) with its local coordinate system, and the directions of roll, pitch and yaw, overlaid.

Accurate calibration requires the IMU's x-axis to face magnetic north while the sensor lays flat and motionless. Calibration takes just a few seconds. Once a calibration

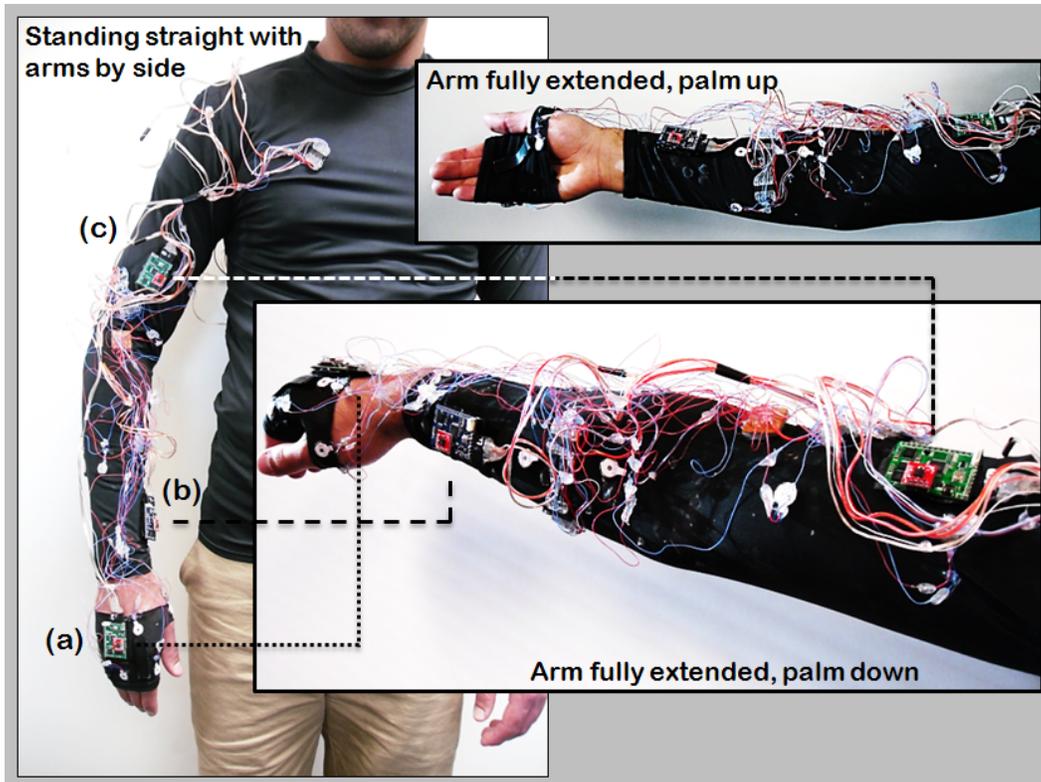
file is created, the sensor's firmware may be updated so that calibration is loaded, rather than recreated, upon each startup, which takes 10-15 seconds to stabilize sensors. We used the cube shown in figure 76 for calibration. This cube was also used to measure sensing inaccuracies by changing the IMU's orientation in 90° increments with respect to each axis. Roll and pitch are accurate up to  $\pm 2^\circ$ , whereas yaw is accurate up to  $\pm 10^\circ$  due to sensing inaccuracies created from the difficulty of mounting the magnetometers perfectly flat. A newer version of this IMU, ArduIMU+ V3 (Flat), comes with a built-in magnetometer, which will reduce these errors. In any case, these errors did not affect experimental results as the study explored position *variations* as opposed to static positions.



*Figure 76.* IMU with coordinate system overlaid and inset for detail. Rotations around the x-axis, y-axis and z-axis change position with respect to roll, pitch and yaw, respectively. A cube, shown in figure, was used for calibration and assessing sensor errors.

To capture the relative positions of the hand, forearm and upper arm, an IMU was attached to each rigid part of the arm. Three IMUs were calibrated, tested for accuracy, and mounted onto the haptic suit via Velcro at three body sites (figure 77). These sensors were integrated into the existing haptic suit through custom-built serial ports connecting each IMU to the LilyPad microcontroller via serial. Only one IMU could take advantage of the LilyPad's single hardware serial port; the other sensors were forced to communicate to the LilyPad via a software serial interface, but without loss of sampling speed. The IMU's introduced additional power requirements, adding a second battery to the power supply. Wireless capabilities were removed due to a hardware malfunction. The existing infrastructure in place for vibrotactile motor instructions did not change; and stimulation for vibrotactile feedback shared vibration motors used by the proposed somatic language for vibrotactile motor instructions.

Referring to figure 77, IMUs (a) and (b) are used to calculate the angle between the hand and forearm (relative to pitch) for detection and estimation of the degree of wrist flexion/extension; IMUs (a) and (b) are again used to calculate the angle between the hand and forearm (relative to yaw) for detection and estimation of the degree of wrist abduction/adduction; IMUs (a) and (c) are used to calculate the relative rotation (roll) between the upper arm and hand/forearm for detection and estimation of the degree of forearm pronation/supination; and lastly, IMUs (b) and (c) are used to calculate the angle between the forearm and upper arm (relative to pitch) for detection and estimation of the degree of elbow flexion/extension.



*Figure 77.* IMU placement on the haptic suit. An IMU is placed on (a) back of the hand; (b) medial side of forearm; and (c) bicep of upper arm. Regarding (b), placement remains at medial side when standing or when arm is fully extended with the palm facing down (see figure); but as the arm rotates such that the palm is facing up (see figure), the forearm rotates within the garment, causing the position of IMU (b) to change relative to the forearm—in particular, it is now on the volar aspect of the forearm. This is desired, however, to enable accurate sensing of elbow flexion/extension.

The gyroscopes of the IMUs saturate at high speeds, but quickly recover after a few milliseconds. During recovery time, sensor readings may be turned off in the IMUs firmware to avoid recording inaccurate readings. The rated value of speed before saturation is  $300^\circ/\text{s}$ ; but experimentally, we found capture of speeds at around  $120^\circ/\text{s}$  or below to be more reliable. For this reason, we limited movements involved in system usage and evaluation to this range to proactively avoid saturation from fast movements.

Using software serial ports, the system is scalable in that more IMUs can be added to potentially record full body movements; but more power may be needed as the

number of sensors increases. Portability, comfort, ease of movement and ease of donning/doffing were not affected by the addition of the IMUs (according to pilot testing). Durability could be improved by enclosing sensors within plastic cases (this will be done as part of future work). Current battery life is satisfactory at well over two hours with rechargeable batteries. In the following section, the firmware of the haptic feedback suit is described, and connections between implementation and the remaining design requirements are drawn.

*Firmware description.* The existing firmware of the haptic suit was augmented with functions to sample IMU readings, compute relative angles between sensors and actuate motors for vibrotactile feedback linked to motor performance. Upon entering feedback mode, the IMUs are sensed at roughly 8 samples/s (approximately, every 120 ms). This sampling rate was largely influenced by the demands of other processing components of the firmware, such as managing actuation of vibration motors, and sensing and computational demands for not just one but three IMUs. However, at 8 samples/s, sufficient resolution was achieved for the movement speeds that are of interest here. At each sampling instance, each IMU is sensed to capture its current roll, pitch and yaw. These values are used to rotate unit vectors to match the current orientation of each IMU. These rotation matrices include

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\text{roll}) & -\sin(\text{roll}) \\ 0 & \sin(\text{roll}) & \cos(\text{roll}) \end{bmatrix}, \quad (1)$$

$$R_y = \begin{bmatrix} \cos(\text{pitch}) & 0 & \sin(\text{pitch}) \\ 0 & 1 & 0 \\ -\sin(\text{pitch}) & 0 & \cos(\text{pitch}) \end{bmatrix}, \quad (2)$$

$$R_z = \begin{bmatrix} \cos(yaw) & -\sin(yaw) & 0 \\ \sin(yaw) & \cos(yaw) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

for rotation (roll) around the x-axis, rotation (pitch) around the y-axis, and rotation (yaw) around the z-axis, respectively, where *roll*, *pitch* and *yaw* are in radians. These are multiplied with a unit vector, represented as  $\mathbf{v}_o$ , to obtain an updated vector,  $\mathbf{v}_u$ , that aligns with the sensors current orientation (with respect to the axis of the given unit vector),

$$\mathbf{v}_u = R_z R_y R_x \mathbf{v}_o \quad (4)$$

These rotated vectors may then be compared to compute their relative angle using

$$\theta = \cos^{-1} \left( \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right) \quad (5)$$

where  $\mathbf{a}$  and  $\mathbf{b}$  are the vectors being compared. Figure 78 depicts which axes of the IMUs are used toward calculating specific angles related to the fundamental movements.

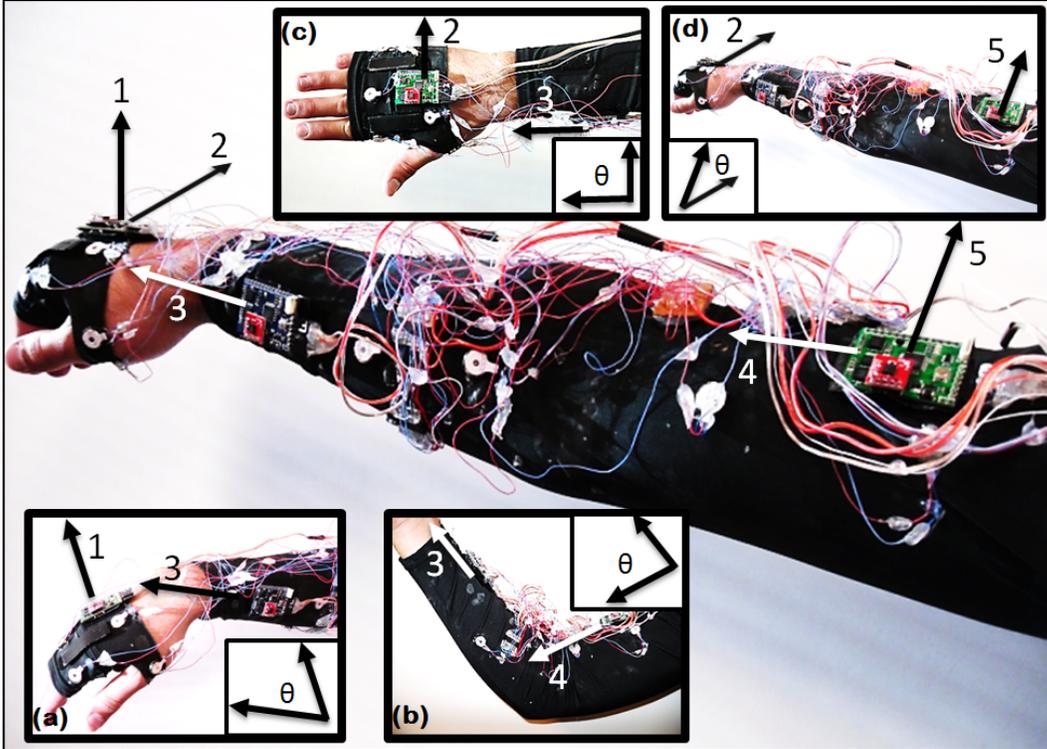


Figure 78. Sensors and vectors (numbered) involved in relative angle calculations for vibrotactile feedback: (a) wrist flexion/extension; (b) elbow flexion/extension; (c) wrist abduction/adduction; and (d) forearm pronation/supination.

The relative positions of limbs are used to drive feedback signals. Angle estimates of a fundamental movement are used to calculate the user's current speed (with respect to this fundamental movement only). The system recognizes that a movement is being made when this speed surpasses a threshold; in this implementation, 15°/s worked well, helping to avoid detection of movements that could simply be jitter. A speed of 15°/s had to be maintained for at least 3 samples for the system to classify the current movement as valid. If classified as valid, when speed falls below the threshold (i.e., ceases), feedback is initiated.

For feedback related to speed, the speed of the completed movement (computed from the median value of speed samples to mitigate effects of acceleration and deceleration at the beginning and end of the movement, respectively) is compared to a

pre-defined target speed. If the current speed is less than the target speed (outside an acceptable amount of error), the “speed up” rhythm is displayed; if the current speed is more than the target speed, the “slow down” rhythm is displayed; and finally, if the speeds match, then a feedback signal indicating correctness (vibration running up the length of the arm) is displayed.

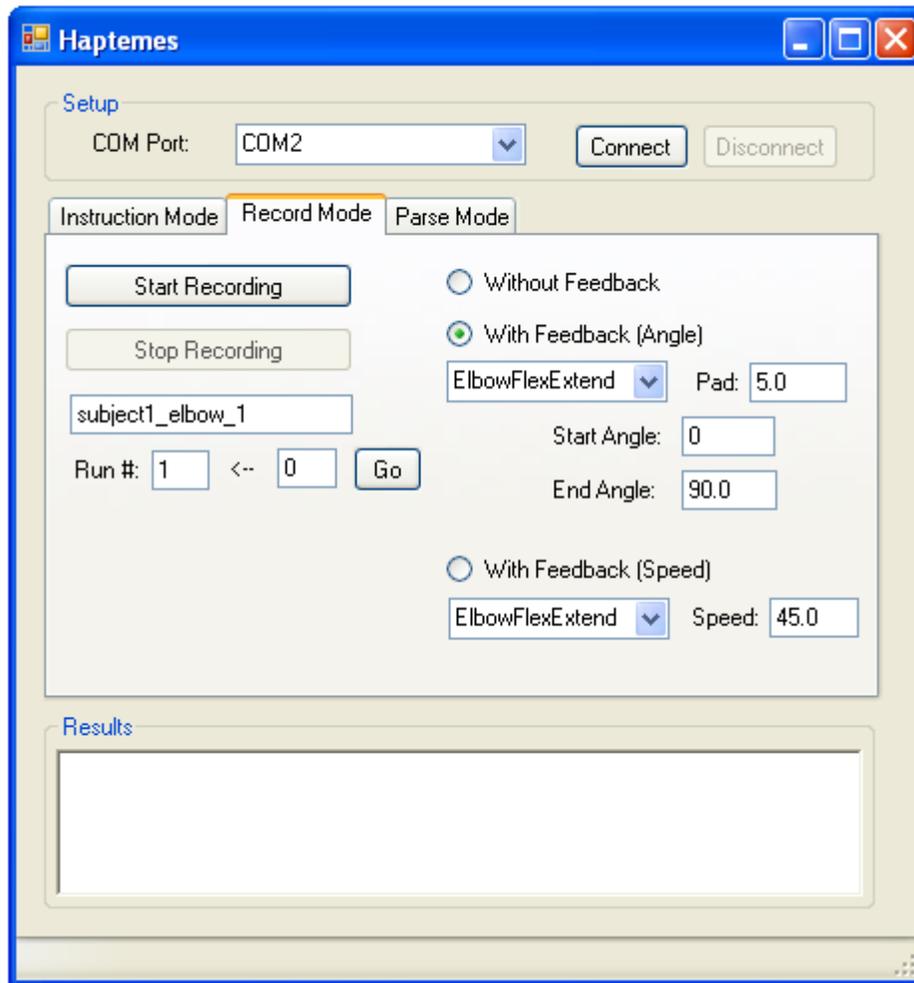
This simple feedback scheme received mixed comments during pilot testing; an approach that was deemed as more usable discretized the range of speed. Speed values were classified as “very slow” ( $< 45^\circ/s$ ), “slow” ( $\geq 45^\circ/s$  and  $< 75^\circ/s$ ) and “moderate” ( $\geq 75^\circ/s$ ). If the current speed falls into the classification containing the target speed, it is found correct; otherwise, users were informed to speed up or slow down. This feedback scheme was more lenient, reducing frustrations for users during pilot testing.

Positioning feedback compared the current angle to a pre-defined target angle. If the current angle is less than or more than the target angle (outside an acceptable amount of error), vibrotactile stimulation “nudges” the user in a direction to reduce this error. When the user reaches an angle that is close to the target angle—that is, within an acceptable amount an error—a feedback signal indicating correctness is displayed (after this angle is held for about 500 ms). Pilot test results for both speed and positioning feedback were positive; participants found the feedback signals easy to learn and intuitive, and appreciated the interactivity of the system.

*Software description.* The graphical user interface previously described for sending instructions to the haptic suit, was augmented to provide vibrotactile feedback for correcting positioning or speed errors. Two new modes were implemented: recording and parsing. Under recording (figure 79), a trainer has the option of recording movement without delivering feedback; recording movement while displaying real-time feedback for positioning errors by entering a start and end angle (feedback is based only on the end

angle); or recording movement while displaying near real-time feedback for speed errors by entering a target speed in degrees per second. Trainers can enter custom file names, which are appended with a “run” or trial number to keep track of the number of times a movement has been practiced.

The parsing function enables users to select and load a previously recorded movement file for segmentation. The raw measurements contained in a file are parsed, and fundamental movements are extracted. These are movements that meet the aforementioned criteria of a “valid” movement—those movements that reach and maintain  $15^\circ/\text{s}$ . Along with the start and end angle of each fundamental movement contained in a recording, its start time, duration, and median speed are also given. The segmentation is outputted to a new file for storage.



*Figure 79.* Graphical user interface for recording movements with options for recording movement without feedback or with feedback (position or speed). A target angle (degree) or speed (degrees/second) can be directly entered for feedback functions ('Pad' value represents the acceptable amount of error in degrees).

**Confirm: vibrotactile feedback.** This section presents the formal evaluation of the proposed somatic language extensions for vibrotactile feedback related to errors in positioning and speed of movements. The structure of the study is similar to our assessment of vibrotactile motor instructions in that it investigates psychophysical responses, rather than motor learning, to shed light on the language's distinctness and naturalness.

***Aim.*** The purpose of this study was to separately assess the proposed vibrotactile feedback designs for position and speed adjustments through two experiments, respectively. The proposed feedback signals were assessed in terms of distinctness, naturalness and usability through objective (number of learning trials, recognition accuracy and response time) and subjective (post-experiment questionnaire) measures. Each participant goes through both experiments, the order of which is counterbalanced across participants to eliminate order effects. The study was approved by Arizona State University's Institutional Review Board.

***Subjects.*** Sixteen participants (8 males and 8 females) completed the user study. The average age was 24 (SD: 8). Only one participant (male) was involved in the previous study (vibrotactile motor instructions) described earlier. No learning effects were found. Each participant was randomly assigned to an experiment order, *position->speed* or *speed->position*, where each order was completed 8 times. No participants had any known motor or tactile impairments that would bias the results of this study.

***Apparatus.*** The haptic feedback suit was used to deliver the proposed vibrotactile feedback for position and speed corrections. The hardware previously described remained the same, but some modifications were made to the suit's firmware and user interface. As only responses were of interest, participants were not asked to reach target angles or target speeds—rather, they simply responded to feedback with the system recording their movement responses in real-time. The new software design also accounted for the structure of the experiment, enabling the experimenter to customize test cases (in support of randomized experiments). For example, the experimenter may select an experiment (position or speed), fundamental movement, and feedback type, 'increase' or 'decrease', to be delivered once a user stopped moving (detected when a threshold of 15°/s is exceeded, maintained for at least three samples, and then drops below threshold). By

delivering feedback immediately following an initial movement, we may assess psychophysical response to feedback signals comparable to situations found in motor learning applications of the proposed system.

Since only responses are of concern, feedback is not linked to motor performance. When in positioning feedback mode, ‘increase’ refers to elbow flexion, wrist flexion, wrist abduction or forearm supination, depending on the fundamental movement selected; and vice versa. The ‘increase’ or ‘decrease’ feedback signal continues until the experimenter stops the recording, after which the captured angles and timestamps are written to a file based on the run number. When in speed feedback mode, ‘increase’ refers to feedback to speed up, whereas ‘decrease’ refers to feedback to slow down. The feedback signal is presented once a movement is completed, after which recording continues until the experimenter stops the recording. No vibrotactile stimulations representing correctness are displayed. Selections for initial movement direction, ‘left’, ‘right’, ‘up’ or ‘down’, are counterbalanced across pattern presentations for each participant (during testing only). The firmware on the LilyPad was updated to accommodate these new modes of operation.

File output was modified to simplify extraction of results (recognition accuracy, response time, etc.) and provide a clear understanding of the movements recorded within each file—critical since synchronized video was not captured. Each output file has two parts: the first part represents the initial movement made by the participant (segmented using the aforementioned threshold) before feedback began. The file is automatically annotated (using the parsing algorithm previously described) to indicate the start of the movement (angle and timestamp), the end of the movement (angle and timestamp), and all samples between these (captured every 120 ms). If the initial movement does not surpass this threshold, feedback will not begin. The second part consists of samples

recorded immediately after feedback began in the case of positioning feedback; or immediately following a vibrotactile alert for cueing a follow-up, speed-adjusted movement in the case of speed feedback. Positioning feedback responses that were very slow were segmented using a threshold of 5°/s and verified to be valid movements by examining their range of motion. Speed feedback responses to slow down and speed up were segmented with a threshold of 5°/s and 15°/s, respectively, and also verified.

**Procedure.** Participants first read and signed their consent, and then completed a subject information form requesting age, sex, height, weight, and descriptions of any known tactile or motor impairments. Each participant was then randomly assigned an experiment order. A brief introduction to the study was given, and the haptic feedback suit was introduced. Participants were shown how to put on the shirt, and help was provided when requested. To ensure that the motors are as close to the skin as possible, participants were told ahead of time to wear a thin short sleeve shirt. During the study, participants performed movements while standing and facing the experimenter. Layman terminology was used to describe movements with respect to specific postures: elbow flexion/extension was termed “elbow up”/“elbow down” (arm held out in front of body with palm facing up); wrist flexion/extension was termed “wrist down”/“wrist up” (arm held out in front of body with palm facing down); wrist abduction/adduction was termed “wrist right”/“wrist left” (arm held out in front of body with palm facing down); forearm supination/pronation was termed “rotate right or clockwise”/“rotate left or counterclockwise” (arm held out in front of body with palm facing any direction).

After participants were acclimated to the movements involved in the study, the phases for the first experiment (*position* or *speed*) began, after which the second experiment followed after a brief rest break. Phases of familiarization, training and testing, in that order, involved a sequence of recording sessions where individual

movements and feedback responses were captured and stored. Recording sessions will be referred to as familiarization, training and testing trials, respectively. Before each trial, participants were told which limb they would be moving. During familiarization and training, participants were asked to, for example, “perform an elbow movement, either up or down”. During testing, these instructions were made more specific by considering the direction of participant’s initial movement; for example, “perform an elbow up”, “perform a wrist right”, etc., to counterbalance the direction a participant was moving immediately before they felt the feedback signal. Since the system saturates at high speeds, participants were requested to perform smooth movements at speeds referred to as slow to moderate (15 to 100°/s)—these movement speeds were demonstrated by the experimenter. For either experiment, participants were requested to perform their initial movement at a slow speed within the middle of this range.

During familiarization, participants were acclimated with the operation of the system for providing feedback related to positioning or speed errors. Feedback signals for position/speed adjustments via elbow flexion/extension, wrist flexion/extension, wrist abduction/adduction and forearm supination/pronation, were presented in that order—these eight feedback signals are referred to as a *set*. Repetitions of any of these signals were allowed when requested. Only during the familiarization phase were participants told in advance what feedback signals to expect. For each trial, the system procedure is as follows for real-time positioning error feedback (speed error feedback operation is described next):

- 1) A “start signal” alerts the participant to ready him or herself for performing a movement. It is a brief (1s) vibration delivered to the elbow joint. Participants were requested to begin moving only after the start signal ended.

- 2) The system records the initial movement, which is detected when a speed of  $15^{\circ}/s$  is surpassed and briefly maintained. The end of the movement is detected when speed drops below  $15^{\circ}/s$ . If the participant's speed was too slow for detection, the experimenter provided feedback, and the trial was repeated. Participants were informed to make a full stop in the middle of the range of the motion of the involved joint so that a response to the feedback signal could be accurately recorded.
- 3) Once movement stopped, the system immediately delivered feedback to adjust position. Participants were requested to respond as quickly but as accurate to the feedback as indicated by the position adjustment it conveyed; and continue moving for the full range of motion or until the vibrations ceased. The experimenter manually stopped the feedback signal via the user interface. The type of position adjustment was random in that it was not linked to motor performance.

Or the following system operation (for near real-time speed error feedback):

- 1) A "start signal" alerts the participant to ready him or herself for performing a movement. It is a brief (1s) vibration delivered to the back of the hand. Participants were requested to begin moving only after the start signal ended.
- 2) The system records the initial movement, which is detected when a speed of  $15^{\circ}/s$  is surpassed and briefly maintained. The end of the movement is detected when speed drops below  $15^{\circ}/s$ . If the participant's speed was too slow for detection, the experimenter provided feedback, and the trial was repeated. Participants were informed to move through their full range of motion at a smooth, slow speed that could be slowed down or sped up in response to feedback without saturating the system.

- 3) Once movement stopped, the system immediately delivered feedback to adjust speed. Participants were requested to re-adjust to their start position while the rhythm was displayed, and wait for a second “start signal”. The type of speed adjustment was random in that it was not linked to motor performance.
- 4) The second “start signal” (identical to the first), readies the participant for their updated, speed-adjusted movement based on the feedback they perceived. Participants were requested to respond as quickly but as accurate once the second start signal ended, and continue moving through their full range of motion with a constant speed. During this follow-up movement, the system records samples (movements and timestamps). This recording is manually stopped by the experimenter.

The training phase involved a sequence of *training sets*. As previously described, a *set* is the presentation of the eight feedback patterns. Unlike familiarization, the patterns within each set were presented randomly, and participants had to recognize specific adjustments, as indicated by the feedback stimulations, on their own—although experimenter feedback was provided to identify wrong responses or confirm correct guesses. To move on to the testing phase, participants had to score 80% or better (at least 7 out of 8 patterns guessed correctly) during a single training set. Since responses are analyzed offline through a parsing algorithm, feedback was provided manually by the experimenter through careful visually analysis. Responses were visually observed and documented as correct, incorrect or corrected—the latter response is initially incorrect but immediately corrected without feedback (corrected responses apply only to positioning error feedback since speed error feedback is near real-time). Corrected responses in which the correction occurred at about 500 ms or less, were not counted wrong toward

the 80% performance threshold, whereas those that took more time to correct were. Observational accuracy was later confirmed by offline analysis via automated parsing functions. No more than three training sets were given per participant.

The testing phase involved the random presentation of 32 patterns generated from four presentations of each of the eight patterns, where for each of the four presentations, half were movements in one direction, and half were movements in the opposite direction. For each of the eight patterns, participants were told to perform their initial movement in a specific direction; for example, if the feedback signal involved movements related to elbow flexion/extension, the participant was told to perform an “elbow up” or “elbow down” as their initial movement. For any movement, participants were requested to begin at the extent of their range of motion; revisiting the previous example, participants were told to start fully extended or fully flexed, respectively; but as before, stop in the middle of their range of motion (for *position*) or move all the way through their range of motion (for *speed*). No feedback was given during the testing phase. Once completed, participants were asked to complete a post-experiment questionnaire. The questionnaire considered of three sets of questions related to general usability, positioning feedback usability and speed feedback usability. Responses were recorded through Likert scales with the exception of two open ended questions—(10) and (14)—where comments and suggestions could be written.

**Results.** The mean number of training sets for *position* and *speed* experiments were 1.25 (SD: 0.57) and 1.12 (SD: 0.34), respectively. Of the 160 training trials for *position*, 7 trial recordings were corrupted (e.g., due to sensor saturations from fast movements), and hence, omitted from analysis. Of the 144 training trials for *speed*, 12 trial recordings were corrupted, and also omitted from analysis. Offline analysis was used to verify the accuracy of the experimenter’s manual feedback regarding correct or

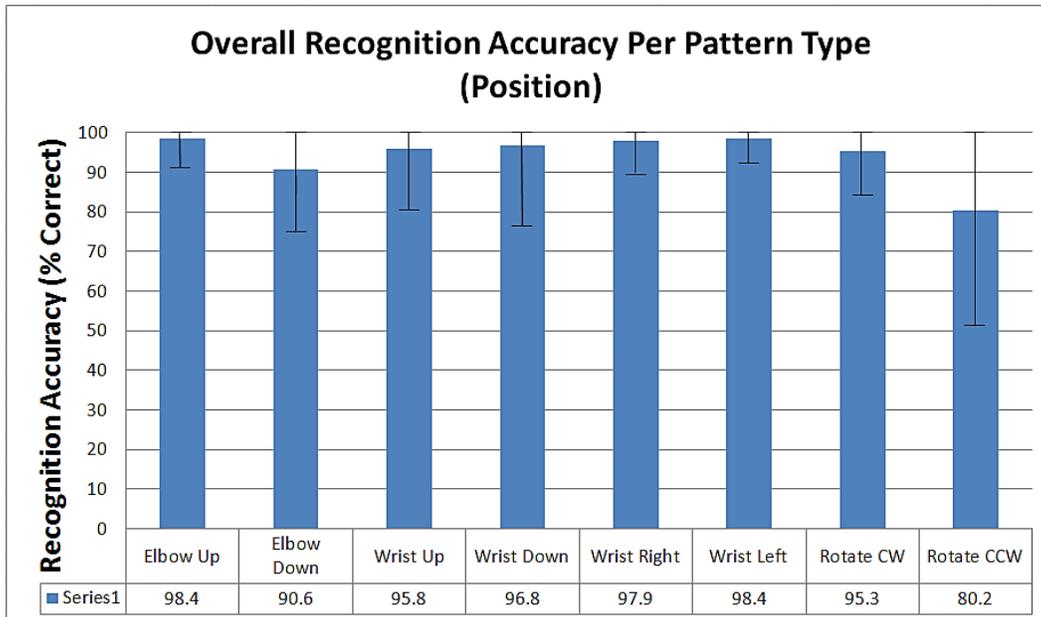
incorrect responses via visual observation: Out of 153 trials for *position* that were successfully recorded, no inconsistencies were found; and out of 132 trials for *speed* that were successfully recorded, only four inconsistencies were found.

For the position experiment, we differentiated between two types of measures of recognition accuracy: (1) Response accuracy is the number of *correct responses* out of the number of presented patterns where *correct responses* do not include *corrected responses*. Recall that *corrected responses* are guesses that are initially incorrect, but eventually corrected. This measure allows us to assess a participant's immediate response and initial interpretation of the feedback signal; (2) Recognition accuracy is the number of *correct* and *corrected responses* out of the number of presented patterns. The overall recognition accuracy for feedback for position adjustments, averaged across participants and patterns, was 94.2% (SD: 6.2%). The overall response accuracy, averaged across participants and patterns, was 91.2% (SD: 7.1%). Out of 512 testing trials, 14 corrected responses were recorded. Overall recognition accuracy and response accuracy per pattern are depicted in figure 80 and 81, respectively. Nine of these responses were corrected in less than a second; three in just above one second; and two in about two seconds. Of the 512 testing trials, 18 trial recordings were corrupted, and hence, omitted from analysis.

For the speed experiment, there was no opportunity to correct responses since feedback was near real-time. Therefore, there is only one measure of recognition accuracy: the number of *correct responses* out of the number of presented patterns. The overall recognition accuracy for feedback for speed adjustments, averaged across participants and patterns, was 90% (SD: 9.7%). Overall recognition accuracy per pattern is depicted in figure 82. Out of 512 testing trials, 20 trial recordings were corrupted, and hence, omitted from analysis.

Response time was recorded for each trial throughout position and speed experiments. For the position experiment, response time is defined as the delay between the presentation of the positioning feedback signal, and the instance when adjustment begins. For speed, response time is defined as the delay between the end of the second start signal, and the instance when the updated movement begins. The overall response time for positioning error feedback, averaged across participants and patterns, was 847 ms (SD: 202 ms) for the training phase, and 881 ms (SD: 205 ms) for the testing phase. The overall response time for speed error feedback, averaged across participants and patterns, was 198 ms (SD: 214 ms) for the training phase, and 247 ms (SD: 182 ms) for the testing phase. Overall response time per pattern between training and testing is depicted in figure 83 and 84 for position and speed experiments, respectively.

Mean responses to the questionnaire are shown in tables 7, 8 and 9. Responses were recorded using a Likert scale ranging from 1 through 5.



*Figure 80.* Mean recognition accuracy per pattern (position experiment) with error bars representing standard deviations. Recognition accuracies have been averaged across participants.

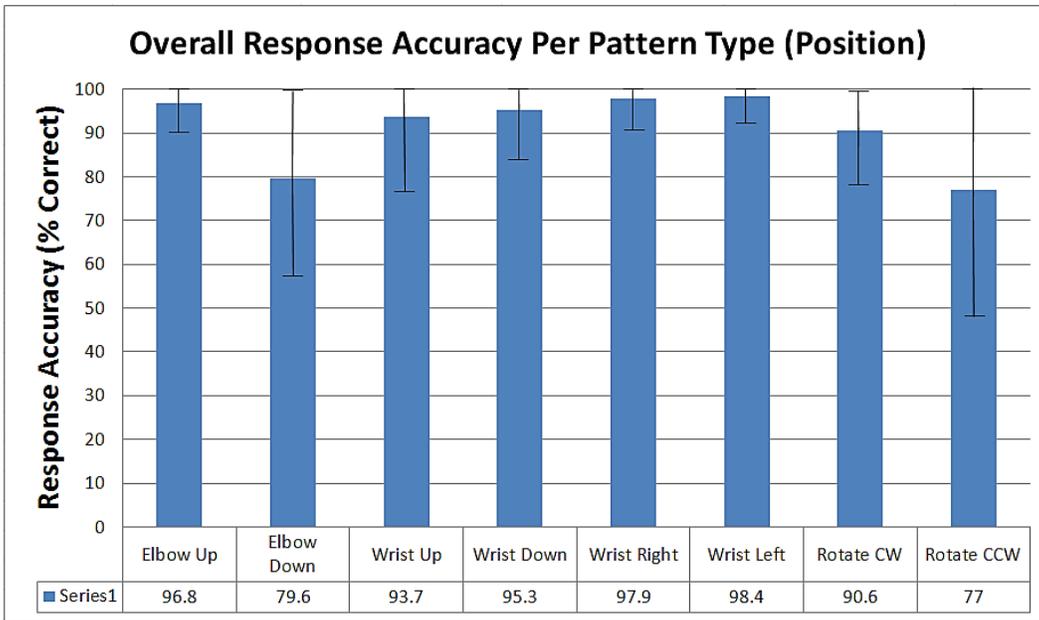


Figure 81. Mean response accuracy per pattern (position experiment) with error bars representing standard deviations. Recognition accuracies have been averaged across participants.

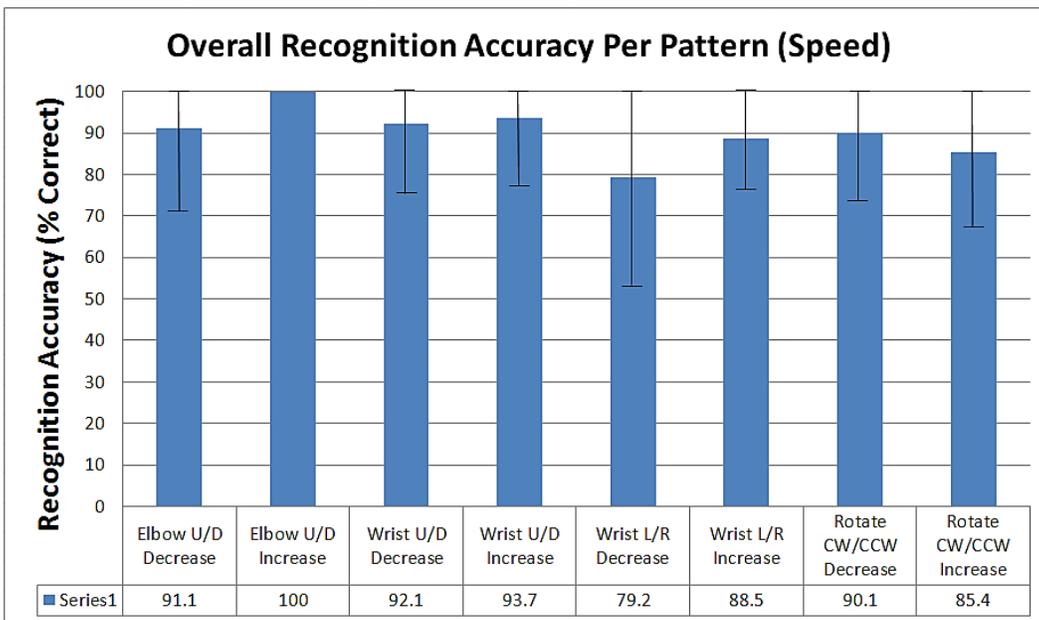


Figure 82. Mean recognition accuracy per pattern (speed experiment) with error bars representing standard deviations. Recognition accuracies have been averaged across participants.

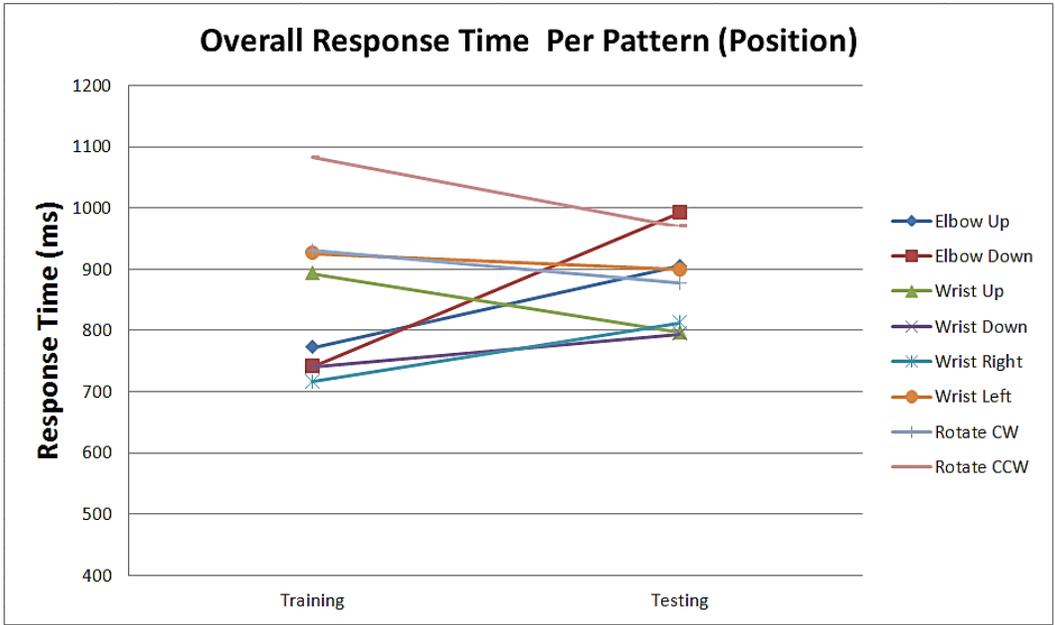


Figure 83. Mean response time per pattern for training and testing phase (position experiment). Response times have been averaged across participants.

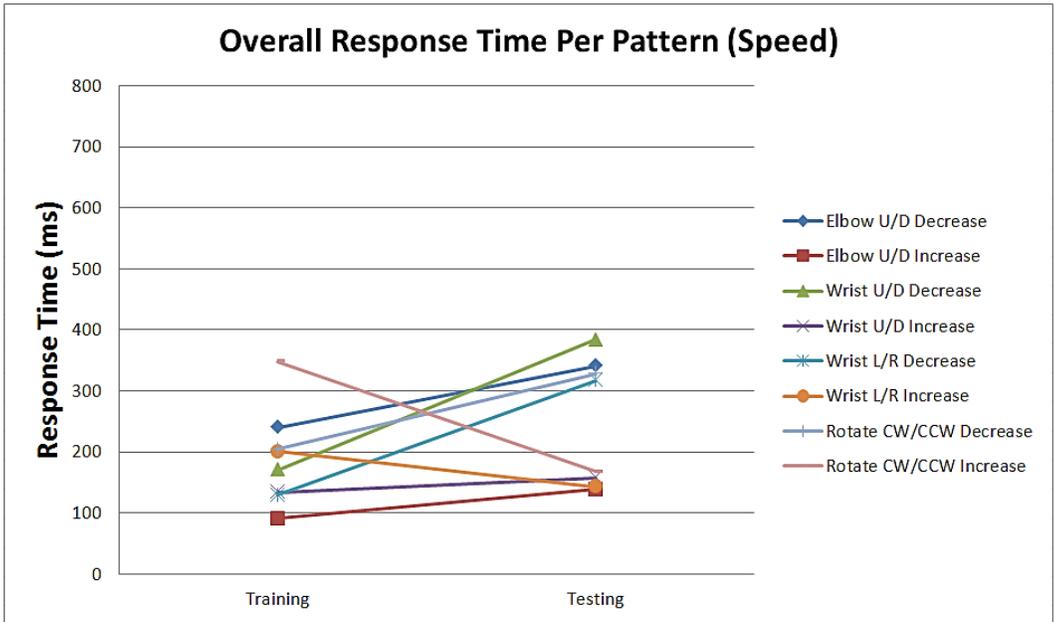


Figure 84. Mean response time per pattern for training and testing phase (speed experiment). Response times have been averaged across participants. In the legend, 'U' is up, 'D' is down, 'L' is left, 'R' is right, 'CW' is clockwise and 'CCW' is counterclockwise.

Table 7

*Mean Responses to General Usability Questions*

Questions	Mean	SD
1) How easy was it to put on the suit?	3.68	0.87
2) How easy was it to take off the suit?	3.43	1.15
3) How easy was it to perform the movements?	4.43	0.62
4) How comfortable was the suit?	3.87	0.95
5) How lightweight was the suit?	4.75	0.57
6) How silent were the suit's vibration motors?	3.37	0.95

*Note.* Questions (1)-(3) were answered using a Likert scale with a range of '1' (very difficult) to '5' (very easy); question (4) used a Likert scale of '1' (very uncomfortable) to '5' (very comfortable); question (5) used a Likert scale of '1' (very heavy) to '5' (very light); and question (6) used a Likert scale of '1' (very loud) to '5' (very quiet).

Table 8

*Mean Responses to Position Feedback Questions*

Questions	Mean	SD
<b>7) How easy was it to recognize the following feedback signals for changing position:</b>		
• a) Elbow 'up' adjustment:	4.37	0.88
• b) Elbow 'down' adjustment:	4.00	1.09
• c) Wrist 'up' adjustment:	4.87	0.34
• d) Wrist 'down' adjustment:	4.68	0.47
• e) Wrist 'right' adjustment:	4.56	0.62
• f) Wrist 'left' adjustment:	4.62	0.50
• g) Rotate 'right' (CW) adjustment:	3.68	1.19
• h) Rotate 'left' (CCW) adjustment:	3.75	1.18
<b>8) How easy was it to learn the following feedback signals for changing position:</b>		
• a) Elbow 'up' adjustment:	4.50	0.73
• b) Elbow 'down' adjustment:	4.12	1.02
• c) Wrist 'up' adjustment:	4.93	0.25
• d) Wrist 'down' adjustment:	4.93	0.25
• e) Wrist 'right' adjustment:	4.87	0.34
• f) Wrist 'left' adjustment:	4.87	0.34
• g) Rotate 'right' (CW) adjustment:	3.87	1.14
• h) Rotate 'left' (CCW) adjustment:	3.93	1.12
<b>9) How natural (intuitive) were the following feedback signals for changing position:</b>		
• a) Elbow 'up' adjustment:	4.37	0.80
• b) Elbow 'down' adjustment:	3.81	1.10
• c) Wrist 'up' adjustment:	4.68	0.70
• d) Wrist 'down' adjustment:	4.68	0.70
• e) Wrist 'right' adjustment:	4.68	0.70
• f) Wrist 'left' adjustment:	4.68	0.70
• g) Rotate 'right' (CW) adjustment:	3.62	1.14
• h) Rotate 'left' (CCW) adjustment:	3.68	1.13

*Note.* Questions (7)-(8) were answered using a Likert scale with a range of '1' (very difficult) to '5' (very easy); and question (9) was answered using a Likert scale with a range of '1' (not natural) to '5' (natural).

Table 9

*Mean Responses to Speed Feedback Questions*

Questions	Mean	SD
<b>11) How easy was it to recognize the following feedback signals for changing speed:</b>		
• a) Decrease speed for elbow... ('up'/'down'):	4.56	0.81
• b) Increase speed for elbow... ('up'/'down'):	4.81	0.54
• c) Decrease speed for wrist... ('up'/'down'):	4.62	0.80
• d) Increase speed for wrist... ('up'/'down'):	4.87	0.34
• e) Decrease speed for wrist... ('left'/'right'):	4.62	0.80
• f) Increase speed for wrist... ('left'/'right'):	4.87	0.34
• g) Decrease speed for forearm rotation ('left'/'right'):	4.62	0.80
• h) Increase speed for forearm rotation ('left'/'right'):	4.87	0.34
<b>12) How easy was it to learn the following feedback signals for changing speed:</b>		
• a) Decrease speed for elbow... ('up'/'down'):	4.68	0.70
• b) Increase speed for elbow... ('up'/'down'):	4.87	0.34
• c) Decrease speed for wrist... ('up'/'down'):	4.56	0.81
• d) Increase speed for wrist... ('up'/'down'):	4.75	0.57
• e) Decrease speed for wrist... ('left'/'right'):	4.50	0.96
• f) Increase speed for wrist... ('left'/'right'):	4.68	0.79
• g) Decrease speed for forearm rotation ('left'/'right'):	4.68	0.70
• h) Increase speed for forearm rotation ('left'/'right'):	4.87	0.34
<b>13) How natural (intuitive) were the following feedback signals for changing speed:</b>		
• a) Decrease speed for elbow... ('up'/'down'):	4.75	0.77
• b) Increase speed for elbow... ('up'/'down'):	4.81	0.54
• c) Decrease speed for wrist... ('up'/'down'):	4.75	0.77
• d) Increase speed for wrist... ('up'/'down'):	4.87	0.34
• e) Decrease speed for wrist... ('left'/'right'):	4.75	0.77
• f) Increase speed for wrist... ('left'/'right'):	4.87	0.34
• g) Decrease speed for forearm rotation ('left'/'right'):	4.75	0.57
• h) Increase speed for forearm rotation ('left'/'right'):	4.81	0.40

*Note.* Questions (11)-(12) were answered using a Likert scale with a range of '1' (very difficult) to '5' (very easy); and question (13) was answered using a Likert scale with a range of '1' (not natural) to '5' (natural).

**Discussion.** This section discusses the aforementioned objective and subjective results as they relate to distinctness and naturalness. In particular, learning rates, recognition accuracies, response times and questionnaire responses for both position and speed experiments are discussed.

*Learning rate (position and speed).* The mean number of training sets for both position and speed experiments are impressive at 1.25 and 1.125, respectively. Most participants achieved 80% recognition accuracy (or better) with just one training set. For position, two participants needed two training sets, and one needed three training sets. For speed, two participants needed two training sets. These results correlate with subjective results: For both position and speed, participants gave high ratings on questions related to the ease of learning feedback signals—see table 8 (question 8) and table 9 (question 12), respectively. The mean learnability ratings (averaged across participants and patterns) for position and speed were 4.5 and 4.69, respectively. The distinctness and naturalness of the feedback signals for correcting positioning and speed errors clearly influenced learnability. These attributes are described next.

*Recognition accuracy (position).* Given the short training times, the mean recognition and response accuracies depicted in figure 80 and 81, respectively, are impressive. For recognition accuracy, all individual accuracies, with the exception of ‘Rotate CCW’ (80.2%), are above 90%. For response accuracy, most individual accuracies are above 90% with exceptions to ‘Elbow Down’ (79.6%) and ‘Rotate CCW’ (77.0%). These results correlate with questionnaire responses—see question (7) of table 8; overall, participants found the patterns easy to recognize, providing an overall rating of 4.31 (averaged across participants and patterns) with ‘Elbow Down’ and ‘Rotate CW’/‘Rotate CCW’ receiving lower, but still satisfactory scores compared to other patterns. Recall that recognition accuracy counts *corrected* movements as correct (that is,

a participant's immediate response was incorrect, but soon corrected—usually in less than a second); whereas response accuracy considers only immediate responses. The latter measure provides a better estimate of participant's initial interpretation and reaction to the feedback signal; however, the former measure is still useful to learn whether participants eventually recognized the feedback signals even when initial reactions were incorrect. If many corrections are needed, this may hint that more time, or even movement, is needed to accurately sense and perceive the proposed feedback signals.

As previously described, out of 512 testing trials, only 14 corrections were made, and the overall recognition accuracy was comparable to the overall response accuracy: 94.2% (SD: 6.2%) to 91.2% (SD: 7.1%), respectively. This shows that initial reactions were accurate, but it is interesting to note that of these 14 corrections, half occurred during 'Elbow Down' adjustments (compare figure 80 with 81)—this finding is discussed below.

A Friedman's analysis of variance by ranks revealed a significant difference between both recognition accuracies of individual patterns,  $\chi^2(7) = 20$ ,  $p < 0.05$ , and response accuracies of individual patterns,  $\chi^2(7) = 18.2$ ,  $p < 0.05$ . Indeed, figure 80 and 81 suggest that feedback signals for 'Elbow Down' were sometimes confused with those for 'Elbow Up', and feedback signals for 'Rotate CCW' were sometimes confused with those for 'Rotate CW'. Subjective feedback, described next, provides insight into these results.

Regarding feedback signals for elbow flexion/extension adjustments, half of the participants commented that the push/pull metaphor was difficult to use for these movements, particularly 'Elbow Down', as all the other movements used the "follow me" conceptual mapping. Hence, there were clear difficulties with switching between conceptual mappings within the same system with many participants commenting that

they would prefer the “follow me” conceptual mapping for elbow flexion/extension adjustments. This confusion also affected participants’ perceived naturalness of the signals (table 8, question 9). But differences between the mean recognition and response accuracy of ‘Elbow Down’ compared to ‘Elbow Up’ require further discussion.

If the “follow me” conceptual mapping was used for elbow flexion/extension adjustments, then a vibration on the volar aspects of the forearm would cue elbow extension (rather than flexion), which is often how participants initially responded since they usually expected “follow me” patterns. However, they often quickly corrected these initial mistakes, suggesting that with additional training, performance may be improved. The conceptual mapping of push/pull for vibrotactile feedback for ‘Elbow Up’ adjustments was more difficult to confuse with the “follow me” conceptual mapping since stimuli is delivered to the bicep rather than the dorsal forearm. Less confusion improved the distinctness and naturalness of feedback signals for ‘Elbow Up’. Indeed, higher recognition accuracies were found for ‘Elbow Up’; as were higher subjective ratings for both ease of recognition (table 8, question 7) and naturalness (table 8, question 9). Lastly, some participants commented that the vibration motors for ‘Elbow Up’ and ‘Elbow Down’ were too closely spaced when fully flexed. As part of future work, the “follow me” conceptual mapping will be used for elbow flexion/extension adjustments, realizing that vibrotactile motor instructions for this movement will take additional training, to simplify recognition through use of a common conceptual mapping. This modification will also increase the spacing between motors to improve distinctness, eliminating confusion between signals when fully flexed.

Participants found wrist movements (flexion/extension and abduction/adduction) very easy to recognize (table 8, question 7). For these movements, the “follow me” conceptual mapping was very intuitive (table 8, question 9), and the spacing between

motors helped increase distinctness. Moreover, recognition results (figure 80 and 81) correlated well with subjective ratings and participant comments. However, rotations also used the “follow me” conceptual mapping, but ‘Rotate CCW’ seemed more difficult to recognize, which also correlated with the lower subjective ratings for both ‘Rotate CW’ and ‘Rotate CCW’—question (7) of table 8. Six of the 16 participants commented that feedback signals for rotational adjustments were harder to distinguish and less natural compared to other patterns (table 8, question 9). We speculate that this discrepancy is related to the arm moving *within* the compression sleeve during rotations, thereby altering the positions of vibration motors depending on the degree of rotation. As part of future work, vibrotactile feedback for forearm pronation/supination adjustments will be moved off the forearm and onto either the hand or upper arm. Differences between recognition accuracy for ‘Rotate CW’ and ‘Rotate CCW’ could be from participants guessing ‘Rotate CW’ when confused.

The challenge with recognizing and using vibrotactile feedback signals related to forearm pronation/supination has been encountered previously. Lieberman and Breazeal (2007) explored vibrotactile feedback to correct positioning errors related to fundamental movements of the right arm during motor learning tasks ranging from simple to complex movements. For the wrist joint and other hinge joints, the intensity of vibrotactile stimulation increased as joint errors increased, intended to *push* a limb back toward its correct position. For these joint types, they found that visual+vibrotactile feedback significantly reduced errors at all times during a trial, and over time through multiple repetitions of trials, compared to using only visual feedback. However, this performance improvement was not found for rotations (forearm and shoulder rotations) in which saltation patterns were used for adjustments as opposed to localized stimulation. Although this work explored two rotational movements as opposed to one, these results

still suggest that novel designs for distinct and natural rotary feedback need to be explored.

*Recognition accuracy (speed).* Recognition accuracy of feedback signals for correcting speed errors, depicted in figure 82, is impressive; most of the individual recognition accuracies are above 90% with the exception of ‘Wrist L/R Decrease’ (79.2%), ‘Wrist L/R Increase’ (88.5%), and ‘Rotate CW/CCW Increase’ (85.4%), which are still satisfactory given the short training times (one to two training sets). A Friedman’s analysis of variance by ranks revealed a significant difference between recognition accuracies of individual patterns,  $\chi^2(7) = 17.1, p < 0.05$ . This difference was likely introduced by some difficulties experienced while recognizing the aforementioned feedback signals for ‘Wrist L/R Decrease’, ‘Wrist L/R Increase’, and/or ‘Rotate CW/CCW Increase’. Note, however, that these accuracies are not only related to recognition, but also, a participant’s ability to *perform* the correct speed adjustment. To mitigate the influence of performance on recognition accuracy, during familiarization and training, participants were instructed to perform speed adjustments that were perceptually distinct from the initial movement. The experimenter provided feedback if speed adjustments were too similar to initial movements. In any case, subjective feedback can help unravel recognition errors from performance errors.

Participants found all the feedback signals for speed corrections very easy to recognize; a summary of their subjective ratings for each pattern is shown in table 9, question 11 (mean: 4.73). This was expected since we are using a single base rhythm with stimulation variations (and no contextual variations)—that is, a vibrotactile rhythm displayed at the same body site (independent of movement type) with two tempo variations (slowing down or speeding up). Participants enjoyed the consistency and naturalness of the rhythms, as detailed in comments, “I liked the consistency between all

of the signals. This made it easy to learn and pick up quickly”, “Very intuitive. I knew exactly what to do after the first time I felt them” and “The speed of the pulses of the motors was very easy and intuitive to figure out”. These comments correlated with the very high subjective ratings for naturalness as shown in table 9, question 13 (mean: 4.79). As shown, the rhythms were found to intuitively represent their respective speed adjustments through tempo variations that noticeably slow down or speed up. As one participant explained, “When resetting my position and feeling the rise or drop in tempo, I could feel my recent arm motion either speeding up or slowing down respectively”. Only two participants commented that they had difficulty distinguishing the rhythm for slow speed; although they said the rhythm for fast speed was distinct and clear.

Based on comparisons between subjective results and recognition accuracies, most misclassifications are likely related to the ability to perform a speed adjustment. It is likely that the short range of motion involved in wrist abduction/adduction and forearm supination/pronation, made speed adjustments more difficult for these fundamental movements compared to movements with a wider range of motion, such as elbow flexion/extension. Indeed, the experimenter observed greater effort, at least physically, when participants were attempting speed adjustments within these shorter ranges. Even so, the accuracy demonstrated within such a short training period is impressive, and will likely improve with continued training and practice.

This work is novel in that it is the first psychophysical evaluation of vibrotactile feedback for speed adjustments of fundamental movements. Lindeman, Yanagida, Hosaka and Abe’s TactaPack (2006) for physical therapy also explored vibrotactile feedback for speed adjustments, but within the scope of warning “nudges” when a patient exceeded an acceleration, or was yet to reach a target acceleration. Moreover, no formal evaluation was presented.

*Response time (position).* The mean response time, as found during testing, to react to a feedback signal for correcting position, is impressive at less than one second (and above 700 ms) for each pattern (figure 83). The mean response time found during testing ( $M = 881$  ms,  $SD = 205$  ms) was higher than the mean found during training ( $M = 847$  ms,  $SD = 202$  ms) resulting in a mean increase ( $M = 33$  ms,  $SD = 186$  ms) in response time per participant. This increase was not significant,  $t(15) = 0.689$ ,  $p = 0.5$ , two-tailed (data normalized using  $\log_{10}$ ). This suggests that participants quickly acclimated to the system, and were able to respond aptly (between half a second to one full second, on average) and consistently to feedback signals between training and testing. We speculate that with continued use and practice, a significant difference would eventually be found as participants became more experienced with using the system and responding to its feedback.

A Friedman's analysis of variance by ranks revealed a significant difference between response times of the testing phase for individual patterns,  $\chi^2(7) = 15.3$ ,  $p < 0.05$ . Indeed, figure 83 suggests that the response times for 'Elbow Down' and 'Rotate CCW' were often higher than other patterns. These results correlated with both recognition/response accuracy and subjective results (ease of recognition and naturalness), for 'Elbow Down' and 'Rotate CCW', showing that recognition difficulties introduced response latencies due to hesitation and/or incorrect initial reactions.

*Response time (speed).* The mean response time, as found during testing, to react to a feedback signal for correcting speed, is impressive at less than 400 ms for each pattern (figure 84). These response times are considerably lower compared to those of position given the near real-time nature of speed feedback. For positioning feedback, participants must respond in real-time and on-the-fly; whereas for speed feedback, participants respond only after the presentation of the feedback and a second start signal.

Therefore, participants have ample time (a several seconds) to recognize and understand how to adjust their speed, and prepare to make this speed-adjusted movement.

The mean response time found during testing ( $M = 247$  ms,  $SD = 182$  ms) was higher than the mean found during training ( $M = 198$  ms,  $SD = 214$  ms) resulting in a mean increase ( $M = 49$  ms,  $SD = 183$  ms) in response time per participant. This increase was not significant,  $t(15) = -1.045$ ,  $p = 0.312$ , two-tailed (data normalized using  $\log_{10}$ ). As with positioning feedback, this suggests that participants quickly acclimated to the system, and were able to respond aptly to feedback immediately following the second start signal.

In contrast to positioning feedback, however, no significant difference was found between mean response times of the testing phase for patterns,  $\chi^2(7) = 8.42$ ,  $p = 0.297$ . This reveals that participants did not hesitate at the start of the follow-up movement, but rather, were confident with their recognition regardless of pattern. Indeed, the impressive recognition accuracy and high subjective ratings in terms of ease of recognition and naturalness, confirm this. This result was expected, however, since the proposed rhythms are consistent across fundamental movements; the base rhythm varies only with respect to stimulation variations that participants were able to easily recognize through a relative comparison between the rhythm's base and tempo change. Future work will explore adding body context to enrich this vibrotactile word—that is, moving this vibrotactile rhythm to different joints based on the fundamental movement it is intended for. This will enable speed feedback for multiple joints at once. As previously described, however, participants enjoyed the consistency of the centralized presentation; one participant commented, “If the speed signal was moved to the respective joint location, it could get confusing. It appears to be correct this [the proposed] way.” The proposed method could also provide a means of speed feedback for more complex movements: the median speed

of multiple fundamental movements could be averaged, and feedback could be delivered related to the speed adjustment of the movement as a whole as opposed to specific, individual joints.

*Post-experiment questionnaire.* Subjective ratings for usability are shown in table 7. Overall, participants found the system reasonably easy to put on (3.68), easy to take off (3.43) and comfortable (3.87). Participants perceived the noise level of vibration motors as being reasonably quiet (3.37). The system was found to be very lightweight (4.75), and very easy to perform movements in while being worn (4.43).

For the majority of feedback signals, the subjective ratings for learnability, detectability and naturalness of positioning and speed error feedback, described throughout the previous sections, show that participants found the patterns very easy to learn; distinct in terms of ease of recognition; and naturally related to their intended movement corrections. The feedback received for position corrections for ‘Elbow Down’ will be improved through use of the “follow me” conceptual mapping rather than the push/pull metaphor, which created confusion for elbow flexion/extension (being the only fundamental movement that used a different conceptual mapping). Novel designs for rotational movements will be explored to learn how vibrotactile positioning feedback may be made more natural and easy to recognize.

Overall, the positive usability feedback, together with positive feedback related to distinctness and naturalness, suggests that the proposed system has potential for real-world motor learning applications. As future work, we will conduct a study exploring motor learning, as opposed to psychophysics, to understand if the proposed system enhances motor learning compared to environments that lack vibrotactile positioning and speed error feedback.

**Conclusion and future work.** Somatic ABC's was applied to design, development and evaluation a language extension for the previously proposed vibrotactile motor instructions. This work bridges the divide between instruction and feedback by proposing a vibrotactile feedback design that is compatible with the motor configurations and conceptual mappings previously presented; in particular, the optimal motor configurations and conceptual mappings for instructions were used when designing the proposed feedback signals to allow for concurrent use and/or separate use depending on the application. The proposed extension was evaluated through a psychophysical study to investigate the distinctness and naturalness of the proposed vibrotactile word and its stimulation variations. Overall, participants were pleased with the design of the vibrations in terms of their distinctness and naturalness, and commented on how much they liked the proposed system for vibrotactile feedback. The relative comparisons of stimulation variations were effective at enriching the proposed somatic language while maintaining low learning curves. Further research objectives related to vibrotactile feedback include:

- Explore perceptual and cognitive differences between the use of one conceptual mapping to describe all movements, and the use of multiple conceptual mappings for movements.
- Evaluate alternative designs for vibrotactile positioning feedback for correcting rotational errors; in particular, vibration motors for this type of feedback will be taken off the forearm, and placed on either the wrist or upper arm for experimentation.
- A longitudinal study should be conducted to assess the proposed system within the context of a motor learning application, such as physical therapy, in terms of error performance and recall over time. Rather than provide simultaneous feedback signals for multiple movements within

complex movements, it is recommended that the system recognize, and provide feedback for, the movement most in error. We hypothesize that this approach will reduce confusion and lower cognitive load. The effects of faded feedback should also be explored—that is, reducing feedback over time so that users become used to performing movements on their own without continually being guided by feedback.

## CONCLUSION AND FUTURE WORK

Today's electronics devices and displays largely engage our vision and hearing. These sensory modalities have become overloaded within our current information-rich, technology-driven lifestyles and careers brought about by the digital revolution. The adverse effects of sensory overload are well known: distractions, confusion and high cognitive load—all increasing the chances of life threatening situations (e.g., texting while driving). It is therefore surprising how little the alternative senses, particular touch, have been investigated toward their use as information delivery channels to ease the burden on sight and hearing. Although some approaches have explored touch-based information delivery, they are limited in terms of high learning curves, applicability and/or expressiveness. The integral component missing from current approaches is a versatile, comprehensive design theory for the building blocks of touch-based information delivery focusing on expandable, efficient, rich, robust, easy-to-learn and easy-to-use languages for somatic (body) communication.

To achieve these objectives, we proposed a novel theoretical framework, inspired by natural, spoken language, called *Somatic ABC's*. This proposed framework guides **A**rticulating (designing), **B**uilding (developing) and **C**onfirming (evaluating), touch-based information delivery languages (somatic languages). The proposed design theory of Somatic ABC's guides the formulation of a somatic language through identifying building blocks similar to those of natural, spoken language; in particular, phonemes, words and sentences. Concepts including body context and stimulation variations enhance somatic word vocabularies to create rich languages. The proposed implementation theory of Somatic ABC's guides language implementation and system construction through design and performance criteria related to functionality, system

performance and general usability. Lastly, the proposed evaluation theory of Somatic ABC's defines a procedure for accurately assessing somatic languages through perceptual distinctness and naturalness—two key criteria that influence a language's learnability and usability.

The usefulness of the proposed theoretical framework was evaluated through two applications: audio-haptic described movies; and vibrotactile motor instructions and feedback. For either application, a somatic language was straightforward to design and enrich—all aspects of the proposed design theory were explored including somatic phoneme, somatic word, somatic sentence and somatic alert creation; and vocabulary enrichment through either body context or stimulation variations. The implementation theory helped guide the successful development of useful and practical systems with sufficient functionality and performance requirements to complete the proposed studies. Lastly, the evaluation theory was followed to design experimental procedures to evaluate the proposed system for each application. Through these evaluations, the communication units and enrichments (context and/or stimulation variations) of both of the proposed somatic languages were shown to be, overall, distinct and natural, which supported learnability and usability. Moreover, general system usability, such as comfort and ease of wearability, received satisfactory to high scores from participants; and participants were excited over both applications, often leaving very positive and enthusiastic comments.

These results clearly demonstrate the effectiveness of Somatic ABC's for designing, developing and evaluating somatic languages that are versatile, rich and easy to learn and use. This work has opened several new vistas for promising directions of future work, described below.

To increase the bandwidth of touch-based information delivery, high-dimensional, multimodal somatic languages must be explored. The proposed applications focus on one modality: vibrations. Increasing dimensionality risks perceptual distinctness; but when coupled with the use of multiple modalities, rich languages could be achieved. As part of future work, Somatic ABC's should be extended to include guidelines for designing multimodal somatic languages, developing these complex languages, and bandwidth evaluation to accurately assess information transfer (*IT*) and information transfer rate (*IT rate*).

Formal grammars for defining somatic languages must be explored. The recursive notation of formal grammars could lend itself well to the structure of somatic languages. Productions could help define the structure of words and sentences, and clearly identify alphabets. This, in turn, would hasten the development of parsers for somatic language-based formal grammars. Since somatic languages are not context-free when body context and stimulation variations are utilized, strategies for overcoming semantic ambiguities need to be investigated. One strategy is to assume that all contextual cues are identifiable in advance, and therefore, can be defined—albeit possibly complicating formal grammars, which may not be acceptable. This may be possible for simple somatic languages, but more complex somatic languages need to be explored.

Potential neurological bases for Somatic ABC's must be explored. Specifically, articulation should take into account existing neuronal circuits in the brain (for tactile and kinesthetic perception) to enhance intuitiveness and simplify learning. As an analogy, consider the development of written language where scribes reinvented and fine-tuned characters to accommodate the perceptual abilities of readers. Our ability to read is not innate, but requires adapting existing neuronal circuits (edge detectors, corner detectors, etc.) to recognize characters (Dehaene, 2010). Likewise, the design of a somatic language

should be adjusted to best fit our neuronal circuits with minimal adaptation to ensure intuitiveness, ease of learning and ease of use. This design strategy not only applies to low level signal parameters, but also the conceptual interpretation of stimuli.

Toward this goal, literature on nurturing touch and its importance in child development must be explored. Studies and anecdotes have hinted at innate tactile sensibilities such as a fetus' ability to recognize its mother's caresses (Field, 2001); but how can these innate perceptual abilities be exploited in somatic language articulation? Moreover, literature on socio-haptic communication (Lahtinen, 2008), haptic idioms (Lemmens, Cromptvoets, Brokken, van den Eerenbeemd, & de Vries, 2009) and haptic perceptual illusions must be explored to learn how natural interactions of social touching and intuitive metaphorical representations, respectively, can be leveraged in the design of conceptual mappings for somatic languages. Ideally, a neurological basis should guide articulation toward literal stimulations by leveraging both innate tactile sensibilities and those naturally developed from societal and cultural environments.

## REFERENCES

- (2010). *Distracted Driving 2009 NHTSA Traffic Safety Facts Research Note*. NHTSA's National Center for Statistics and Analysis, U.S. Department of Transportation National Highway Traffic Safety Administration.
- (2011). *State of the Media: Consumer Usage Report*. The Nielson Company.
- Bach-y-Rita, P. (1972). *Brain Mechanisms in Sensory Substitution*. New York, New York: Academic Press.
- Bach-y-Rita, P. (2004). Tactile sensory substitution studies. *Annals New York Academic Sciences*, 1013, 83-91.
- Bach-y-Rita, P., & Kerchel, S. W. (2003, December). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7(12), 541-546.
- Bach-y-Rita, P., Collins, C. C., Saunders, F. A., White, B., & Scadden, L. (1969, March 8). Vision substitution by tactile image projection. *Nature*, 221, 963-964.
- Bach-y-Rita, P., Kaczmarek, K. A., Tyler, M. E., & Garcia-Lara, J. (1998, October). Form perception with a 49-point electrotactile stimulus array on the tongue: A technical note. *Journal of Rehabilitation Research and Development*, 35(4), 427-430.
- Barralon, P., Ng, G., Dumont, G., Schwarz, S. K., & Ansermino, M. (2007). Development and evaluation of multidimensional tactons for a wearable tactile display. *9th International Conference on Human Computer Interaction with Mobile Devices and Services*, (pp. 186-189).
- Behnke, R. S. (2006). *Kinetic Anatomy*. Champagne, IL: Human Kinetics.
- Benecke, B. (2004). Audio-Description. *Translator's Journal*, 49, 78-80.
- Bonanni, L., Vaucelle, C., Lieberman, J., & Zuckerman, O. (2006). TapTap: A haptic wearable for asynchronous distributed touch therapy. *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, (pp. 580-585).
- Borst, C. W., & Baiyya, V. B. (2007). Enhancing VR-based visualization with a 2D vibrotactile array. *ACM Symposium on Virtual Reality Software and Technology*, (pp. 183-186).
- Borst, C. W., & Baiyya, V. B. (2009). A 2D haptic glyph method for tactile arrays: Design and evaluation. *Third Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 599 - 604).
- Brave, S., & Dahley, A. (1997). inTouch: A medium for haptic interpersonal communication. *CHI '97 Extended Abstracts on Human Factors in Computing Systems: Looking to the Future*, (pp. 363-364).

- Brewster, S., & Brown, L. M. (2004). Tactons: Structured tactile messages for non-visual information display. *Fifth Conference on Australasian User Interface*, 28, (pp. 15-23).
- Brewster, S., Chohan, F., & Brown, L. (2007). Tactile feedback for mobile interactions. *CHI '07 SIGCHI Conference on Human Factors in Computing Systems*, (pp. 159-162).
- Brown, L. M., & Kaaresoja, T. (2006). Feel who's talking: Using tactons for mobile phone alerts. *CHI 2006 Extended Abstracts on Human Factors in Computing Systems*, (pp. 604-609).
- Brown, L. M., Brewster, S. A., & Purchase, H. C. (2005). A first investigation into the effectiveness of tactons. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 167-176).
- Brown, L. M., Brewster, S. A., & Purchase, H. C. (2006a). Multidimensional tactons for non-visual information presentation in mobile devices. *Eighth Conference on Human-Computer Interaction with Mobile Devices and Services*, (pp. 231-238).
- Brown, L. M., Brewster, S. A., & Purchase, H. C. (2006b). Tactile crescendos and sforzandos: Applying musical techniques to tactile icon design. *CHI 2006 Extended Abstracts on Human Factors in Computing Systems*, (pp. 610-615).
- Brown, L. M., Sellen, A., Krishna, R., & Harper, R. (2009). Exploring the potential of audio-tactile messaging for remote interpersonal communication. *27th International Conference on Human Factors in Computing Systems*, (pp. 1527-1530).
- Brown, L., & Williamson, J. (2007). Shake2Talk: Multimodal messaging for interpersonal communication. In I. Oakley, & S. Brewster (Eds.), *Haptic and Audio Interaction Design* (Vol. 4813, pp. 44-55). Springer.
- Cha, J., Oakley, I., Ho, Y.-S., Kim, Y., & Ryu, J. (2009, July-September). A Framework for Haptic Broadcasting. *IEEE Multimedia*, 16(3), 477-491.
- Chan, A., MacLean, K., & McGrenere, J. (2005). Learning and identifying haptic icons under workload. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 432-439).
- Chang, A., O'Modhrain, S., Jacob, R., Gunther, E., & Ishii, H. (2002). ComTouch: Design of a vibrotactile communication device. *4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, (pp. 312-320).
- Chen, H.-Y., Santos, J., Graves, M., Kim, K., & Tan, H. Z. (2008). Tactor localization at the wrist. In M. Ferre (Ed.), *Haptics: Perception, Devices and Scenarios* (Vol. 5024, pp. 209-218). Berlin: Springer Berlin Heidelberg.

- Cholewiak, R. W. (1979). Spatial factors in the perceived intensity of vibrotactile patterns. *Sensory Processes*, 3(2), 141-156.
- Cholewiak, R. W., & Collins, A. A. (2003). Vibrotactile localization on the arm: Effects of place, space, and age. *Perception & Psychophysics*, 65(7), 1058 -1077.
- Cholewiak, R. W., Brill, J. C., & Schwab, A. (2004). Vibrotactile Localization on the Abdomen: Effects of Place and Space. *Perception and Psychophysics*, 66(6), 970-987.
- Chua, P. T., Crivella, E., Daly, B., Hu, N., Schaaf, R., Ventura, D., et al. (2003). Training for physical tasks in virtual environments: Tai Chi. *IEEE Virtual Reality*, (pp. 87-94).
- Craig, J. C., & Sherrick, C. E. (1982). Dynamic tactile displays. In W. Schiff, & E. Foulke (Eds.), *Tactual Perception: a sourcebook* (pp. 209-233). New York, New York, USA: Cambridge University Press.
- Crystal, D. (2007). *How Language Works*. New York: Avery Trade.
- Dehaene, S. (2010). *Reading in the Brain: The New Science of How We Read*. Penguin.
- DiSalvo, C., Gemperle, F., Forlizzi, J., & Montgomery, E. (2003). The Hug: An exploration of robotic form for intimate communication. *12th IEEE International Workshop on Robot and Human Interactive Communication*, (pp. 403-408).
- Distraction.gov. (2011, March 4). Retrieved September 26, 2011, from <http://www.distraction.gov/>: <http://www.distraction.gov/files/for-media/2011/2011-03-04-distracted-driving-brochure-print.pdf>
- Drobny, D., Weiss, M., & Borchers, J. (2009). Saltate!: A sensor-based system to support dance beginners. *27th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 3943-3948).
- Edwards, N., Rosenthal, J., Moberly, D., Lindsay, J., Blair, K., Krishna, S., et al. (2009). A Pragmatic Approach to the Design and Implementation of a Vibrotactile Belt and its Applications. *IEEE International Workshop on Haptic Audio visual Environments and Games*, (pp. 13-18). Lecco.
- Effenberg, A. O. (2005, April-June). Movement sonification: Effects on perception and action. *IEEE Multimedia*, 12(2), 53-59.
- Enriquez, M. J., & MacLean, K. E. (2003). The hapticon editor: A tool in support of haptic communication research. *11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 356-362).
- Enriquez, M., MacLean, K., & Chita, C. (2006). Haptic phonemes: Basic building blocks of haptic communication. *8th International Conference on Multimodal Interfaces*, (pp. 302-309).

- Erp, J. B., Veen, H. A., Jansen, C., & Dobbins, T. (2005, April). Waypoint Navigation with a Vibrotactile Waist Belt. *ACM Transactions on Applied Perception*, 2(2), 106-117.
- Ferscha, A., Emsenhuber, B., Riener, A., Holzmann, C., Hechinger, M., Hochreiter, D., et al. (2008). Vibro-Tactile Space-Awareness. *Adjunct Proceedings of the 10th International Conference on Ubiquitous Computing*.
- Field, T. (2001). *Touch*. Cambridge, MA: MIT Press.
- Fogg, B., Cutler, L. D., Arnold, P., & Eisbach, C. (1998). HandJive: A device for interpersonal haptic entertainment. *SIGCHI Conference on Human Factors in Computing Systems*, (pp. 57-64).
- Förster, K., Bächlin, M., & Tröster, G. (2009). Non-interrupting user interfaces for electronic body-worn swim devices. *2nd International Conference on Pervasive Technologies Related to Assistive Environments*, 38, pp. 1-4.
- Foulke, E. (1982). Reading braille. In W. Schiff, & E. Foulke (Eds.), *Tactual Perception* (pp. 168-208). Cambridge University Press.
- Friskin-Gibson, S. F., Bach-y-Rita, P., Tompkins, W. J., & Webster, J. G. (1987, December). A 64-solenoid, four-level fingertip search display for the blind. *IEEE Transactions on Biomedical Engineering*, BME-34(12), 963-965.
- Gagnon, L., Foucher, S., Heritier, M., Lalonde, M., Byrns, D., Chapdelaine, C., et al. (2009). Towards computer-vision software tools to increase production and accessibility of video description for people with vision loss. *Univ Access Inf Soc*, 8, pp. 199-218.
- Gaw, D., Morris, D., & Salisbury, K. (2006). Haptically annotated movies: Reaching out and touching the silver screen. *14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 287-288).
- Geldard, F. A. (1957). Adventures in tactile literacy. *American Psychologist*, 12(3), 115-124.
- Geldard, F. A. (1960). Some neglected possibilities of communication. *Science*, 131(3413), 1583-1588.
- Geldard, F. A. (1966). Cutaneous coding of optical signals: The optohapt. *Perception & Psychophysics*, 1, 377-381.
- Geldard, F. A., & Sherrick, C. E. (1972, October). The cutaneous 'rabbit': A perceptual illusion. *Science*, 178(4057), pp. 178 -179.
- Grindlay, G. (2008). Haptic guidance benefits musical motor learning. *Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, (pp. 397-404).

- Gunther, E. (2001). Skinscape : a tool for composition in the tactile modality. *Unpublished master's thesis*. Massachusetts Institute of Technology, Cambridge, MA.
- Hedgpeth, T., Black, J., & Panchanathan, S. (2006). A demonstration of the iCARE portable reader. *Eighth International ACM SIGACCESS Conference on Computer & Accessibility* (pp. 279-280). Portland: Oregon.
- Hesse, S., Schmidt, H., Werner, C., & Bardeleben, A. (2003). Upper and lower extremity robotic devices for rehabilitation and for studying motor control. *Current Opinion in Neurology*, 16, 705-710.
- Holden, M. K. (2005, June). Virtual environments for motor rehabilitation: Review. *CyberPsychology & Behavior*, 8(3), 187-211.
- Huang, K., Starner, T., Do, E., Weinberg, G., Kohlsdorf, D., Ahlrichs, C., et al. (2010). Mobile Music Touch: Mobile tactile stimulation for passive learning. *28th International Conference on Human Factors in Computing Systems*, (pp. 791-800).
- Johansson, R. S. (1978). Tactile sensibility in the human hand: Receptive field characteristics of mechanoreceptive units in the glabrous skin area. *Journal of Physiology*, 281, 101-123.
- Johansson, R. S., & Vallbo, A. B. (1979). Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *Journal of Physiology*, 286, 283-300.
- Johansson, R. S., & Vallbo, A. B. (1983, January). Tactile sensory coding in the glabrous skin of the human hand. *Trends in Neurosciences*, 6, 27-32.
- Jones, L. A., & Lederman, S. J. (2006). *Human Hand Function*. New York, New York: Oxford University Press, Inc.
- Jones, L. A., & Ray, K. (2008). Localization and pattern recognition with tactile displays. *Symposium on Haptic Interfaces for Haptic interfaces for Virtual Environment and Teleoperator Systems*, (pp. 33-39).
- Jones, L. A., & Sarter, N. B. (2008). Tactile displays: Guidance for their design and application. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(1), 90-111.
- Jones, L. A., Held, D., & Hunter, I. (2010). Surface waves and spatial localization in vibrotactile displays. *IEEE Haptics Symposium*, (pp. 91-94).
- Jones, L. A., Kunkel, J., & Piatetski, E. (2009). Vibrotactile pattern recognition on the arm and back. *Perception*, 38, 52-68.

- Jones, L. A., Kunkel, J., & Torres, E. (2007). Tactile vocabulary for tactile displays. *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 574-575).
- Jones, L. A., Lockyer, B., & Piatetski, E. (2006). Tactile display and vibrotactile pattern recognition on the torso. *Advanced Robotics*, 20(12), 1359-1374.
- Kaaresoja, T., & Linjama, J. (2005). Perception of short tactile pulses generated by a vibration motor in a mobile phone. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 471-472).
- Kaczmarek, K. A., Webster, J. G., Bach-y-Rita, P., & Tompkins, W. J. (1991, January). Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Biomedical Engineering*, 38(1), 1-16.
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of Neural Science* (fourth ed.). New York, New York, USA: McGraw-Hill Medical.
- Kane, L. A. (2004). *Martial Arts Instruction*. Boston, MA: YMAA Publication Center.
- Kapur, P., Jensen, M., Buxbaum, L. J., Jax, S. A., & Kuchenbecker, K. J. (2010). Spatially distributed tactile feedback for kinesthetic motion guidance. *IEEE Haptics Symposium*, (pp. 519-526).
- Kennedy, C. A., & Yoke, M. M. (2009). *Methods of Group Exercise Instruction*. Champaign, IL: Human Kinetics.
- Kim, Y., Cha, J., Oakley, I., & Ryu, J. (2009, September). Exploring Tactile Movies: An Initial Tactile Glove Design and Concept Evaluation. *IEEE Multimedia*, PP(99).
- King, T. W. (2000). *Modern Morse Code in Rehabilitation and Education: New Applications in Assistive Technology*. Needham Heights, MA, USA: Allyn & Bacon.
- Knapp, M. L. (1978). *Nonverbal Communication in Human Interaction*. New York: Holt, Rinehart and Winston.
- Kosuge, K., Hayashi, T., Hirata, Y., & Tobiyama, R. (2003). Dance Partner Robot -Ms DanceR-. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, (pp. 3459-3464).
- Krishna, S., Bala, S., & Panchanathan, S. (2010). Exploring the dorsal surface of the fingers for visuo-haptic sensory substitution. *IEEE International Symposium on Haptic Audio-Visual Environments and Games*, (pp. 1-6).
- Krishna, S., Bala, S., McDaniel, T., McGuire, S., & Panchanathan, S. (2010). VibroGlove: An assistive technology aid for conveying facial expressions. *28th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 3637-3642).

- Lahtinen, R. M. (2008). *Haptices and Haptemes: A case study of developmental process in social-haptic communication of acquired deafblind people*. Frinton-on-Sea, England: A1 Management UK.
- Lakritz, J., & Salway, A. (2006). *The Semi-Automatic Generation of Audio Description from Screenplays*. University of Surrey, Guildford.
- Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, *19*, 342-368.
- Lemmens, P., Cromptvoets, F., Brokken, D., van den Eerenbeemd, J., & de Vries, G. J. (2009). A body-conforming tactile jacket to enrich movie viewing. *Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 7-12).
- Lieberman, J., & Breazeal, C. (2007, October). TIKL: Development of a wearable vibrotactile feedback suit for improved human motor learning. *IEEE Transactions on Robotics*, *23*(5), 919-926.
- Lin, M.-W., & Cheng, Y.-M. (2008). Using tactons to provide navigation cues in pedestrian situations. *5th Nordic Conference on Human-Computer Interaction: Building Bridges*, (pp. 507-510).
- Lindeman, R. W., & Yanagida, Y. (2003). Empirical studies for effective near-field haptics in virtual environments. *IEEE Virtual Reality*, (pp. 287-288).
- Lindeman, R. W., Page, R., Yanagida, Y., & Sibert, J. L. (2004). Towards full-body haptic feedback: The design and deployment of a spatialized vibrotactile feedback system. *ACM Symposium on Virtual Reality Software and Technology*, (pp. 146-149).
- Lindeman, R. W., Yanagida, Y., Hosaka, K., & Abe, S. (2006). The TactaPack: A wireless sensor/actuator package for physical therapy applications. *14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 337-341).
- Lindeman, R., Yanagida, Y., Noma, H., & Hosaka, K. (2006). Wearable vibrotactile systems for virtual contact and information display. *Virtual Reality*, *9*(2), 203-213.
- Linville, J. G., & Bliss, J. C. (1966, January). A direct translation reading aid for the blind. *Proceedings of the IEEE*, *54*(1), 40-51.
- MacLean, K., & Enriques, M. (2003). Perceptual design of haptic icons. *EuroHaptics*.
- Magana, M., & Velazquez, R. (2008). On-shoe tactile display. *IEEE International Workshop on Haptic Audio Visual Environments and Games*, (pp. 114-119).

- McDaniel, T. L., Krishna, S., Colbry, D., & Panchanathan, S. (2009). Using tactile rhythm to convey interpersonal distances to individuals who are blind. *27th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 4669-4674).
- McDaniel, T. L., Villanueva, D., Krishna, S., Colbry, D., & Panchanathan, S. (2010). Heartbeats: A methodology to convey interpersonal distance through touch. *28th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 3985-3990).
- McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., & Panchanathan, S. (2011). Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements. *19th ACM international conference on Multimedia*, (pp. 543-552).
- McDaniel, T., Krishna, S., Balasubramanian, V., Colbry, D., & Panchanathan, S. (2008). Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. *IEEE International Workshop on Haptic Audio Visual Environments and Games*, (pp. 13-18).
- Montagu, A. (1986). *Touching: The Human Significance of the Skin* (3rd ed.). New York, New York: Harper & Row, Publishers, Inc.
- Morley, J. W., & Rowe, M. J. (1990). Perceived pitch of vibrotactile stimuli: Effects of vibration amplitude, and implications for vibration frequency coding. *The Journal of Physiology*, *431*(1), 403 -416.
- Nielson. (2010). Retrieved September 26, 2011, from Nielsonwire:  
<http://blog.nielson.com/nielsenwire/press/nielsen-fact-sheet-2010.pdf>
- O'Modhrain, S., & Oakley, I. (2004). Adding Interactivity: Active Touch in Broadcast Media. *Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 293-294).
- Oakley, I., & O'Modhrain, S. (2002). Contact IM: Exploring Asynchronous Touch Over Distance. *CSCW*.
- Oakley, I., Kim, Y., Lee, J., & Ryu, J. (2006). Determining the feasibility of forearm mounted vibrotactile displays. *International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 27-34).
- O'Modhrain, S., & Oakley, I. (2003). Touch TV: Adding feeling to broadcast media. *European Conference on Interactive Television: From Viewers to Actors?*, (pp. 41-47).
- Osawa, N. (2006). Tactile glyphs for palpation of relationships. *International Conference on Information Visualisation*, (pp. 575-584).
- Penfield, W., & Rasmussen, T. (1950). *The Cerebral Cortex of Man*. New York: The Macmillan Company.

- Piateski, E., & Jones, L. (2005). Vibrotactile pattern recognition on the arm and torso. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 90-95).
- Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., et al. (Eds.). (2008). *Neuroscience* (fourth ed.). Sunderland, Massachusetts, USA: Sinauer Associates, Inc.
- Rahman, M. A., Alkhaldi, A., Cha, J., & Saddik, A. E. (2010). Adding Haptic Feature to YouTube. *ACM Multimedia*. Firenze, Italy.
- Reed, C. M., & Delhorne, L. A. (1995, November). Current results of a field study of adult users of tactile aids. *Seminars in Hearing*, 16(4), 305-315.
- Reed, C. M., Delhorne, L. A., Durlach, N. I., & Fischer, S. D. (1990). A study of the tactual and visual reception of fingerspelling. *Journal of Speech and Hearing Research*, 33, 786-797.
- Reed, C. M., Delhorne, L. A., Durlach, N. I., & Fischer, S. D. (1995). A study of the tactual reception of sign language. *Journal of Speech and Hearing Research*, 38, 477-489.
- Reed, C. M., Durlach, N. I., & Delhorne, L. A. (1992). Natural methods of tactual communication. In I. R. Summers (Ed.), *Tactile Aids for the Hearing Impaired* (pp. 218-230). London, England: Whurr Publishers Ltd.
- Reed, C. M., Rabinowitz, W. M., Durlach, N. I., & Braida, L. D. (1985). Research on the Tadoma method of speech communication. *Journal of Acoustical Society of America*, 77(1), 247-257.
- Réhman, S. u., & Liu, L. (2008). Vibrotactile rendering of human emotions on the manifold of facial expressions. *Journal of Multimedia*, 3(3), 18-25.
- Roberts, J. C., & Franklin, K. (2005). Haptic Glyphs (Hlyphs) - Structured haptic objects for haptic visualization. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 369-374).
- Rosenthal, J., Edwards, N., Villanueva, D., Krishna, S., McDaniel, T., & Panchanathan, S. (2011). Design, implementation, and case study of a pragmatic vibrotactile belt. *IEEE Transactions on Instrumentation and Measurement*, 60(1), 114-125.
- Rovers, A. F., & van Essen, H. A. (2004). HIM: A framework for haptic instant messaging. *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, (pp. 1313-1316).
- Rupert, A. H. (2000). An instrumentation solution for reducing spatial disorientation mishaps. *IEEE Engineering in Medicine and Biology Magazine*, 19(2), 71-80.

- Salway, A. (2007). A Corpus-based Analysis of Audio Description. In J. Díaz Cintas, P. Orero, & A. Remael, *Media for All*. Radopi.
- Saunders, F. A., Hill, W. A., & Franklin, B. (1981). A wearable tactile sensory aid for profoundly deaf children. *Journal of Medical Systems*, 5(4), 265-270.
- Schmidt, R. A., & Wrisberg, C. A. (2000). *Motor Learning and Performance: A Problem-Based Learning Approach* (2nd Edition ed.). Champaign, IL: Human Kinetics.
- Shieh, M.-D., & Wu, Z.-B. (2008). Tactile icon design using a vibration actuator in mobile devices. In S. Lee, H. Choo, S. Ha, & I. C. Shin (Eds.), *Computer-Human Interaction* (Vol. 5068, pp. 240-247). Berlin: Springer Berlin Heidelberg.
- Snyder, J. (2005). Audio description: The visual made verbal. *International Congress Series*, 1282, 935-939.
- Spelmezan, D., Jacobs, M., Hilgers, A., & Borchers, J. (2009). Tactile motion instructions for physical activities. *27th International Conference on Human Factors in Computing Systems*, (pp. 2243-2252).
- Takahata, M., Shiraki, K., Sakane, Y., & Takebayashi, Y. (2004). Sound feedback for powerful karate training. *Conference on New Interfaces for Musical Expression*, (pp. 13-18).
- Tan, H. Z., & Pentland, A. (2001). Tactual displays for sensory substitution and wearable computers. In W. Barfield, & T. Caudell, *Fundamentals of Wearable Computers and Augmented Reality* (pp. 579-598). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Tan, H. Z., Durlach, N. I., Rabinowitz, W. M., Reed, C. M., & Santos, J. R. (1997). Reception of Morse code through motional, vibrotactile, and auditory stimulation. *Perception & Psychophysics*, 59(7), 1004-1017.
- Vallbo, A. B., & Johansson, R. S. (1984). Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. *Human Neurobiology*, 3, 3-14.
- Vallbo, A. B., Olausson, H., Wessberg, J., & Kakuda, N. (1995). Receptive field characteristics of tactile units with myelinated afferents in hairy skin of human subjects. *Journal of Physiology*, 483.3, 783-795.
- van der Linden, J., Johnson, R., Bird, J., Rogers, Y., & Schoonderwaldt, E. (2011). Buzzing to play: Lessons learned from an in the wild study of real-time vibrotactile feedback. *2011 Annual Conference on Human Factors in Computing Systems*, (pp. 533-542).
- van der Linden, J., Schoonderwaldt, E., & Bird, J. (2009). Towards a real-time system for teaching novices correct violin bowing technique. *IEEE International Workshop on Haptic Audio Visual Environments and Games*, (pp. 81-86).

- van Erp, J. B. (2005). Vibrotactile spatial acuity on the torso: Effects of location and timing parameters. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, (pp. 80-85).
- van Erp, J. B., & van Veen, H. A. (2001). Vibro-tactile information presentation in automobiles. *Eurohaptics*, (pp. 99-104).
- van Erp, J. B., & van Veen, H. A. (2003). A multi-purpose tactile vest for astronauts in the international space station. *Eurohaptics*, (pp. 405-408).
- van Erp, J. B., van Veen, H. A., Jansen, C., & Dobbins, T. (2005). Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception*, 2(2), 106-117.
- van Veen, H. A. H. C., & van Erp, J. B. F. (2001). Tactile information presentation in the cockpit. In S. Brewster, & R. Murray-Smith (Eds.), *Haptic Human-Computer Interaction, LNCS 2058* (pp. 174-181). Springer-Verlag Berlin Heidelberg.
- Verrillo, R. T. (1963). Effect of contactor area on the vibrotactile threshold. *The Journal of the Acoustical Society of America*, 35(12), 1962-1966.
- Verrillo, R. T., & Chamberlain, S. C. (1972). The effect of neural density and contactor surround on vibrotactile sensation magnitude. *Perception and Psychophysics*, 11(1B), 117-120.
- Verrillo, R. T., & Gescheider, G. A. (1992). Perception via the sense of touch. In I. R. Summers, *Tactile Aids for the Hearing Impaired* (pp. 1-36). London: Whurr Publishers.
- Verrillo, R. T., Fraioli, A. J., & Smith, R. L. (1969). Sensation magnitude of vibrotactile stimuli. *Perception and Psychophysics*, 6(6A), 367-372.
- Viswanathan, L. N. (2011). *Enhancing Movie Comprehension For Individuals Who Are Visually Impaired Or Blind*. Thesis (M.S.), Arizona State University.
- Viswanathan, L., McDaniel, T., & Panchanathan, S. (2011). Audio-haptic description in movies. In C. Stephanidis (Ed.), *HCI International 2011 – Posters' Extended Abstracts* (pp. 414-148). Berlin, Heidelberg: Springer Berlin Heidelberg.
- von Békésy, G. (1962). Can we feel nervous discharges of the end organs during vibratory stimulation of the skin? *Journal of the Acoustical Society of America*, 34(6), 850-856.
- Wall III, C., & Weinberg, M. S. (2003, March/April). Balance prostheses for postural control. *IEEE Engineering in Medicine and Biology Magazine*, pp. 84-90.
- Weber, E. H. (1834/1996). *E.H. Weber on the Tactile Senses* (2nd ed.). (H. E. Ross, & D. J. Murray, Eds.) East Sussex: Erlbaum (UK) Taylor & Francis.

- Weinstein, S. (1968). Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality. In D. R. Kenshalo, *The Skin Senses* (pp. 195-222). Springfield, Illinois, USA: Charles C Thomas.
- Weisenberger, J. M., & Percy, M. E. (1994, November). Use of the tactaid II and Tactaid VII with children. *The Volta Review*, 96(5), 41-57.
- White, B. W., Saunders, F. A., Scadden, L., Bach-y-Rita, P., & Collins, C. C. (1970). Seeing with the skin. *Perception & Psychophysics*, 7(1), 23-27.
- Whitehead, J. (2005). What is audio description. *International Congress Series*, 1282, 960-963.
- Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, 31, 285-289.
- Yang, U., & Kim, G. J. (2002, June). Implementation and evaluation of "Just Follow Me": An immersive, vr-based, motion-training system. *PRESENCE*, 11(3), 304-323.