Feel Effects: Enriching Storytelling with Haptic Feedback

ALI ISRAR, SIYAN ZHAO, and KAITLYN SCHWALJE, Disney Research
ROBERTA KLATZKY, Carnegie Mellon University
JILL LEHMAN, Disney Research

Despite a long history of use in communication, haptic feedback is a relatively new addition to the toolbox of special effects. Unlike artists who use sound or vision, haptic designers cannot simply access libraries of effects that map cleanly to media content, and they lack even guiding principles for creating such effects. In this article, we make progress toward both capabilities: we generate a foundational library of usable haptic vocabulary and do so with a methodology that allows ongoing additions to the library in a principled and effective way. We define a *feel effect* as an explicit pairing between a meaningful linguistic phrase and a rendered haptic pattern. Our initial experiment demonstrates that users who have only their intrinsic language capacities, and no haptic expertise, can generate a core set of feel effects that lend themselves via semantic inference to the design of additional effects. The resulting collection of more than 40 effects covers a wide range of situations (including precipitation, animal locomotion, striking, and pulsating events) and is empirically shown to produce the named sensation for the majority of our test users in a second experiment. Our experiments demonstrate a unique and systematic approach to designing a vocabulary of haptic sensations that are related in both the semantic and parametric spaces.

Categories and Subject Descriptors: H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Artificial, augmented, and virtual realities*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Haptic I/O*

General Terms: Human Factors

Additional Key Words and Phrases: Haptics media, haptic vocabulary, feel effect, authoring tools, vibrotactile feedback

ACM Reference Format:

Ali Israr, Siyan Zhao, Kaitlyn Schwalje, Roberta Klatzky, and Jill Lehman. 2014. Feel effects: Enriching storytelling with haptic feedback. ACM Trans. Appl. Percept. 11, 3, Article 11 (September 2014), 17 pages.
DOI: http://dx.doi.org/10.1145/2641570

1. INTRODUCTION

Special effects are an important part of storytelling and enrich user experiences in movies, shows, games, rides, virtual simulations, and social and educational media. In particular, sound and visual effects are frequently used to embellish content and emphasize events. In recent years, haptic technologies have been introduced to further enhance the user experience with dynamic feedback across the body [Danieau et al. 2012; Israr et al. 2011; Sodhi et al. 2013]. Key challenges in using haptic technologies in storytelling, however, are the lack of both haptic vocabulary and authoring methods for

Author's address: A. Israr, S. Zhao, K. Schwalje, and J. Lehman, 4720 Forbes Ave. Suite 110, Pittsburgh, PA 15213; email: {israr, siyan.zhao, jill.lehman}@disneyresearch.com, kaischwa@gmail.com; R. Klatzky, 354K Baker Hall, 5000 Forbes Avenue, Pittsburgh, PA; email: klatzky@cmu.edu.

DOI: http://dx.doi.org/10.1145/2641570

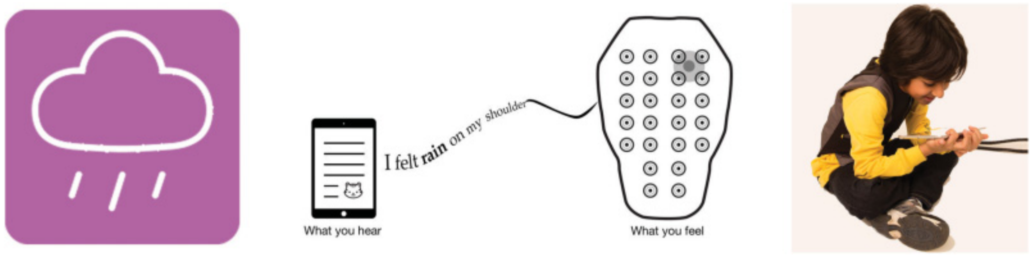


Fig. 1. A young user feels expressive haptic feedback while listening to a story.

creating meaningful and realistic haptic representations of content events. As a result, our work has two goals. The first is to create a library of *feel effects* (FEs) that extends the richness of an interaction by engaging the haptic senses in the same way that libraries of sound and visual effects are used to engage the auditory and visual senses. The second goal is to devise authoring methods for adding expressive, convincing, and natural haptic representations to the library on an ongoing basis.

In essence, an FE is an artificially generated haptic pattern that, by virtue of its connection to a meaningful event, enhances media content through sensations on the user's skin. We concentrate on the vibrotactile array as the source of sensation and the back as the surface for stimulation. Although the back has a low density of receptors, the area is large, naturally accessible in gaming chairs, theater seats, ride vehicles, gaming vests, and so forth, and typically otherwise unused during interactions. Use of a haptic vest is depicted in Figure 1.

A key feature of an FE is that it correlates the semantic interpretation of an event as judged by human users with the parametric composition of the sensation in terms of physical variables (such as intensity, duration, temporal onsets, etc.). Just as the phrases describing events are associated with each other by semantic logic, the corresponding FEs are associated with each other by a parametric relationship. For example, the relationship between "hop" and "jump" can be simulated by two patterns of haptic stimulation that are similar in rhythmic timing but differ in the number of contact points. In this article, we test and find confirming evidence for two hypotheses: (1) semantically similar FEs lie in close proximity in haptic parameter space, and (2) semantic reasoning for relating events can be applied to the haptic space to derive new FEs. As a practical consequence, we populate our library with an initial set of more than 40 FEs, covering six distinct types of sensations.

We also demonstrate a method for eliciting initial FEs by priming sense memory through language, as well as methods for automatically extending the vocabulary when the parametric loci of a small number of FEs in a semantic family are already known. Next, we briefly examine the history of haptic technology most relevant to linking sensations to meaningful events. Subsequent sections lay out our framework and describe the methods for empirically developing an initial corpus of FEs and automatically extending them. We conclude with a discussion of limitations and applications.

2. BACKGROUND AND RELATED WORK

Using haptic feedback to convey meaningful information to a user has a long history. Early efforts were directed toward artificial aids that would assist the blind to see and the deaf to hear. Primarily, these technologies were based on two principles: pictorial mapping and frequency-to-place mapping [Tan and Pentland 2001]. An early pictorial mapping device was the Optacon, which yoked the input from an optical sensor to an array of vibrating pins, enabling a user to successfully read written text [Linville and Bliss 1966]. The same principle was used in Tactile Television [Collins 1970], which mapped dynamic imagery from a camera to an array of vibrating points on the back, thus succeeding in rendering crude shapes and motion but not detailed image features.

Frequency-to-place mapping schemes were extensively used to communicate speech and environmental sounds to the deaf via touch. In an early exemplar, Teletactor [Gault 1927], incoming speech was processed through five band-pass filters that were yoked to five vibrators attached to the fingers and thumb of the user's hand, with the goal of allowing the user to decode speech from its spectral shape. Similar systems, so-called tactile vocoders, were subsequently developed and explored, including the Felix system and Tactaid VII [Tan and Pentland 2001]. These mapping techniques, though successful with impaired users, have characteristics that are undesirable for media enhancement of users with intact sensory systems, such as the necessity for extensive training (hours to days), nonintuitive mapping, and perceptual confusion between haptically conveyed patterns.

To develop usable haptic feedback for neurotypical users, researchers have attempted to design a haptic vocabulary that is perceptually differentiable and easily interpreted and learned. Such vocabularies have been constructed for haptic tones [Sahami et al. 2008], haptic messages [Brewster and Brown 2004], haptic icons [Maclean and Enriquez 2003], haptic phonemes [Enriquez et al. 2006] and haptic metaphors [Brunet et al. 2013]. The purpose of these vocabularies was to establish building blocks for more complex haptic signals related to urgency, warning, navigation, and guidance. Other related work aims to determine the underlying semantic or aesthetic contents of haptic sensations [O'Sullivan and Chan 2006; Gunther and O'Modhrain 2003] and to link hedonic haptic responses to visually perceived object features [Klatzky and Peck 2012].

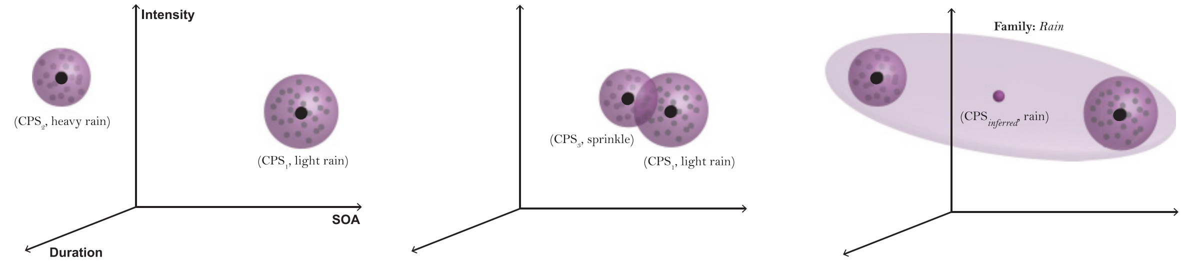
New haptic rendering algorithms have been developed that use tactile illusions to create effective and dynamic sensations on the skin. Kim et al. [2013] showed a systematic process to render 3D tactile features, such as bumps, on flat surfaces. Israr and Poupyrev [2011] developed a Tactile Brush that used a coarse grid of vibrating actuators to render high-resolution and moving haptic patterns on the back. Currently, commercial haptic devices are available that enhance the realism and excitement in games and movies. For example, D-box (D-box, Quebec, Canada) is a motion-platform–based technology that uses motion cues from movie scenes to move the user's body. The Marvel Avengers Vybe Haptic Gaming Pad (Comfort Research, Grand Rapids, Michigan) allows users to feel game content by using sound cues that are loosely associated with events in the game to drive dynamic haptic patterns.

Despite the availability of a broad range of haptic technologies, there are no guiding principles for creating haptic effects that are explicitly associated with the content watched, heard, or read by a user. In this article, we develop a framework for associating haptic patterns to the mental interpretation of events and create a library of FEs that can be used by designers and artists to generate expressive and meaningful haptic content for their media.

3. THE HAPTIC-SEMANTIC FRAMEWORK

We define a feel effect (FE) as a pair of components drawn from two spaces. The first component is drawn from the *haptic space* that can be produced by a particular array of vibrotactile actuators. The haptic component specifies how the sensation will unfold over time and location on the skin via parameter settings for SOA (stimulus onset asynchrony, i.e., the interval between two actuations), duration, intensity, and ramp-up for each actuator in the array. The second component of an FE is drawn from *semantic space*. The semantic component describes what experience the sensation feels like, and is specified using a common language phrase (LP).

In a useful haptic library, the sensation produced by an FE must be reliably experienced as an instance of the language phrase. In other words, we want to be able to look up the parameter settings that feel like being jabbed or rained on, much as we can look up the waveform of a car horn honking or a robin's song in a sound effects library. Given individual variability, as well as the relative ambiguity



(a) (b) (c)

Fig. 2. Location and distribution of the parametric settings for FEs in a 3D haptic parametric space. Each light dot is the setting of one subject; the heavy dot represents the weighted mean value (CPS). (a) Parametric settings of FEs for antonyms are widely separated in the space, while (b) settings of semantically similar or synonymous FEs are in close vicinity with overlapping volumes, and (c) a family of semantically related FEs define a larger volume within which additional FEs (purple dot) can be inferred.

of touch in general, we expect a given language phrase to correspond more or less well to a range of parameter settings. Conceptually, the range of parameter settings defines a volume in *haptic-semantic* space where the surface of the volume represents the boundary between “feels like” and “doesn’t feel like” the named sensation, and the centroid value for the volume constitute the most recognizable instance for the most users. Figure 2(a) illustrates this idea, showing one range of parameter settings that are perceived as heavy rain and a separate, somewhat larger range of values that are perceived as light rain. Within each cluster there is a single {SOA, intensity, duration} tuple—the canonical parameter settings (CPSs)—that best represents the set.

We also expect that a range of parameter settings may correspond more or less well to multiple language phrases, a many-to-many mapping that can be understood as overlapping volumes. Such overlaps may occur when language phrases are essentially synonymous, as illustrated in Figure 2(b), but overlaps may also occur when semantically distinct events are haptically indistinguishable, resulting in potential confusions. Settings associated with the sensation of being jabbed by a small stick, for example, may also be associated with being jabbed by a medium-sized stick, or even with being poked by a finger, despite the obvious differences in reality.

The problem of creating a haptic library is essentially the problem of searching haptic-semantic space for the subset of (CPS, LP) pairs representing settings that are experienced by the majority of people as the named sensations. Given the likely sparseness of such pairs in the space, we proceed heuristically, by using semantic relationships already reflected in language to generate candidate FEs. The relationships among the parameters then predict when overlap and potential confusability will occur.

Central to the heuristic approach is the notion of a *family*—a set of semantically-related LPs partially defined by at least one common haptic parameter and differentiated by others. “Knock,” “jab,” “tap,” and “poke,” for example, are semantically related language phrases that share the hypernym “strike.”[1] The *Strike* family (which includes the LP “strike”) is defined by a quick contact between an object and a surface; family members are differentiated primarily by the force of contact and/or the

[1] A hypernym is a word that names a broad category used to organize other words, a convention in linguistics similar to the IS-A link in AI. Ontologies are typically realized as a forest of trees representing the similarities and distinctions among classes of concepts. Different ontologies, for example, WordNet [Miller 1995] or Verbnet [Schuler 2005], cover somewhat different portions

degree to which repetition of contact is expected. True to their semantic nature, families do not have to be denoted by simple verbs: the *Rain* family includes all of "sprinkle," "light rain," and "rain," as shown in Figure 2(c), as well as "rain shower," "heavy rain," "raining cats and dogs," "deluge," and "downpour."

We hypothesize that a family defines a natural volume in haptic-semantic space where semantic distinctiveness predicts haptic distance. In other words, the weaker the linguistic difference between two LPs with respect to a haptic dimension, the greater the overlap in the ranges of the parameter values tied to that dimension and the more confusable the sensations will be. One person's "tap" is unlikely to be another person's "pound," but one's "poke" may well be another's "jab." If the mapping between semantic distinctness and haptic distance holds within a family, then language pairs that are synonymous (poking stick vs. twig) should produce essentially the same parameter settings.

Antonyms, on the other hand, should reveal parameter settings at the boundaries of the family: "tap" versus "pound," or "sprinkle" versus "downpour." The ranges of parameter settings associated with antonyms should be the least overlapping, their CPSs the most distant, and the sensations the least confusable. Further, if the parameter settings for antonyms define the endpoints of a haptic dimension, we should be able to infer values for language terms that lie between them. The CPS for the *Rain* family, for example, should be equivalent to the CPS for "rain" and lie about midway between the centroids for "light rain" and "heavy rain" (see Figure 2(c)).

We expect the distinctiveness-distance mapping to apply between families as well. Semantic classes may be more or less distinct; hence, family volumes may be far apart in haptic-semantic space or quite close together. "Strike" and "rain" belong to disjoint verb groups; accordingly, members of the *Strike* family should have largely nonoverlapping volumes with members of the *Rain* family, and *Strike* FEs should be difficult to confuse with *Rain* FEs.[2] In contrast, "rain" and "snow" belong to the same verb group, so their parameter settings should place them relatively close together.

A useful haptic library will have easy-to-access contents that are applicable in a wide range of contexts. The power to search and reason in language greatly enhances the user's ability to generalize old content to the authorship of new effects. In essence, the language phrases of known FEs act as "titles" for the associated haptic parameters, offering starting points for finding the parameter values of semantically related terms. The underlying semantic relationships can be used to infer haptic distances between known and novel LPs and suggest modifications of the corresponding CPSs.

Before we can infer new FEs, however, we must have an initial set of established values. In the next section, we describe a user study through which we acquired 23 candidate FEs in six families. To these 23, we added 31 more derived via synonymy and inference, and, in a second study (Section 5), we evaluated the degree to which each candidate was experienced as the named sensation and tested the distinction-distance hypothesis in relation to the entire set.

4. STUDY 1: FINDING CANDIDATE FES

An FE has two parts: the language phrase (LP) that describes a sensation and the canonical parameter settings (CPSs) that produce the named sensation for the majority of people. The purpose of the first study was to elicit data from which we could derive CPS values to test our hypotheses about family relationships.

of the world and tend to have different names for their organizing relations. For convenience, we use WordNet's relations throughout.

[2]Where the distinction fails, family overlap can occur: a single drop of water dripping on a surface is likely indistinguishable from tapping.

4.1 Method

Participants (N = 22, 8 female, 14 male), ages 18 to 61 (M = 25.14 years, SD = 9.38 years), were recruited via methods approved by our institutional review board. Participants were compensated 10 dollars for their time.

4.1.1 *Materials and Design.*

Stimuli consisted of 23 LPs grouped into six semantic families that potentially differ in haptic features. Most of the LPs in each family were specifically chosen to be "antonyms" in the sense of being highly contrastive pairs on one or more haptic-semantic dimensions (e.g., "walking" vs. "running" on the SOA-speed dimension and "tapping" vs. "pounding" on the intensity-hardness dimension).[3] The families and their members were:

- *Rain*: water droplets experienced as multiple simultaneous points of contact in random locations—**4 LPs** ("light rain," "sprinkle," "heavy rain," "downpour") including two sets of synonyms that contrast size, force, and rate of contact.
- *Travel* (multilegged locomotion): experienced as multiple sequential points of contact in a recognizable pattern of locations—**6 LPs** ("bird walking," "bird running," "cat walking," "cat running," "lizard walking," "lizard running") that contrast speed, size, and force of contact and bipedal/quadrupedal patterns.
- *Strike*: experienced as a single point of contact or multiple points of contact over time in the same location—**6 LPs** ("finger tapping," "hand tapping," "stick poking," "teddy bear poking," "knocking," "pounding") that contrast number, size, and force of contacts.
- *Brush*: experienced as an area of contact that moves—**4 LPs** ("cat paw swiping," "lion paw swiping," "feather stroking," "leaves stroking") that contrast size, force, and speed of contact.
- *Pulse*: experienced as alternating points of contact in a specific rhythmic pattern—**2 LPs** ("calm heart beating," "racing heart beating") that contrast speed and force.
- *Motor sound*: experienced as a stationary area of contact in a recognizable rhythmic pattern—**1 LP** ("cat purring").[4]

To elicit parameter settings for these LPs, we rely on the ability of an LP to prime sense memory: participants read an LP, then match the sensation it evokes against what they feel and, using an experimental interface as shown in Figure 3, specify how the sensation must change to fit the LP better. Note that the LP (a) is embedded in a sentence template and concretely describes the evoked experience. Similarly, the five-button scales by which participants modify the sensation (b–d) are labeled semantically, allowing them to reason in relative terms about the experience's qualities rather than the underlying physical parameters actually being controlled. The interface also includes a "Try it" button (e) to test the current sensation and an additional five-button scale (f) for judging its match to the description.

Because the procedure is iterative, receptor fatigue is a real concern. For this reason, the set of actuators used (the location on the back where sensation was felt) was predetermined, and the parameters controlled by participants were limited to the two or three dimensions that distinguished the chosen antonym pairs. The mapping from button to specific parameter value was also predetermined, under the constraints of incorporating likely good representations and extending either beyond them or until the limits of the actuator were reached. The resulting range was divided into four equal intervals,

[3]A list of words, phrases, and events common in children storybooks was first compiled to select potential LPs and families for this study. A subset of the list was chosen and contrastive terms added to produce the final stimuli.

[4]"Cat purring" was included because the FE was needed for an application of the technology to children's story listening. It was the only LP where a parameter's values were programmed by sound and the only one where we were unable to create a satisfactory range of parameter values for a candidate antonym (e.g., a train rumbling) with our actuator array.

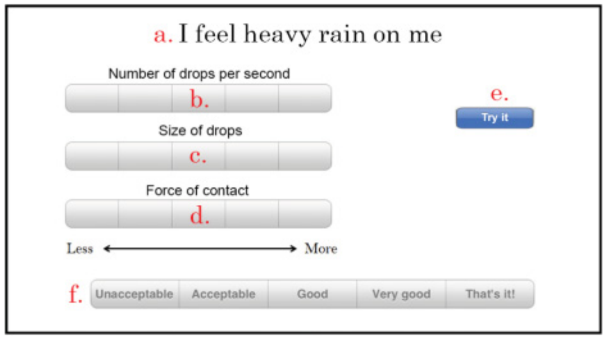


Fig. 3. The interface for eliciting values for SOA, duration, and intensity for "heavy rain" in the *Rain* family.

yielding five values. Thus, the data acquired in this study are a function of both the salience of the haptic-semantic differences in memory and constraints of the haptic device. Parameter values can be read from Figure 4.

4.1.2 *Apparatus.* Stimuli were presented on a custom made haptic pad, embedded with 24 high quality vibrotactile transducers, called *tactors* (type C-2, Engineering Acoustics Inc., Casselberry, Florida). The tactors were arranged in an array grid: five rows of four tactors across the broad portion of the back, and two rows of two closer to the waist (see Figure 1). The actuator grid was embedded in a 14×21-inch pad that slipped over a straight-backed chair. A custom microprocessor-based driver and API were developed to control the parameters of each tactor at 1ms resolution through the USB port of a laptop. The intensity of the actuators was calibrated to match 10V to about 45dB sensation level (SL) of the nominal tactile sensation on the back. The laptop communicated with an iPad mini application via UDP protocol to allow participants to enter their responses on a small handheld device. Noise produced by the vibrotactile array was masked by pink noise played through V-MODA headphones.

4.1.3 *Procedure.* Participants were seated and their posture checked for good contact with the pad. A brief tutorial introduced both vibratory sensation and the method of using a scale of buttons to control it. Specifically, participants were asked to systematically change the intensity of a sensation produced by four actuators in the middle of the back. When participants felt they understood the mechanism, the experimenter guided them step by step through a training trial to familiarize them with the interface and task. Participants were told to read the description ("I feel a cat's paw swiping me" during training) and try different combinations of button selections to find the most realistic sensation. They were encouraged to explore many combinations before rating one, an action that advanced them to the next trial. After training, participants proceeded independently through the remaining 22 LPs presented in random order.

4.2 Results and Discussion

Because CPSs are intended to capture the sensation that is most likely to be recognizable as the phrase by the majority of users, we derive them from a biased subset of the data. Twenty participants produced "Good" or better ratings for more than half the stimuli, so two participants were discarded as outliers (one produced no "Good" or better ratings, and the other produced "Good" or better ratings for only a quarter of the sentences).

With the remaining data, we calculated the CPS as the weighted mean SOA, duration, and intensity for each language phrase using only those parameter values that were rated "Good" or better by more

Table I. CPSs for 23 FEs means calculated over parameter values rated "Good" or better by more than one participant. The centroid is the n-tuple formed by the weighted means. SOA scale is linear in milliseconds except for "cat purring," which is the percentage of compression or stretching of the sound wave; Duration scale is linear; Intensity scale is logarithmic.

Family	Language Phrase	SOA (ms) [% tempo]	Duration (ms)	Intensity log(volts)
Pulse	calm heart beating	505.00	—	2.02
	racing heart beating	248.89	—	2.71
Rain	light rain	390.00	57.06	1.74
	sprinkle	320.00	70.00	1.98
	heavy rain	78.47	104.44	3.01
	downpour	82.50	120.00	2.91
Strike	finger tapping	417.06	73.97	2.45
	hand tapping	415.63	178.33	2.69
	stick poking	—	131.76	3.10
	teddy bear poking	—	279.50	2.24
	knocking	43.82	—	2.76
	pounding	58.33	—	3.06
Motor	cat purring	[13.33]	—	2.64
Travel	bird walking	399.41	135.00	2.19
	bird running	145.94	110.00	2.42
	lizard walking	310.00	194.17	1.95
	lizard running	126.81	156.39	2.46
	cat walking	608.82	497.22	2.75
	cat running	179.41	397.92	2.63
Brush	feather stroking	96.56	150.74	2.13
	leaves stroking	71.92	175.00	2.08
	cat paw swiping	74.47	104.44	2.80
	lion paw swiping	50.83	325.00	3.15

than one participant. Every CPS was based on ratings from at least 13 people. Table I summarizes the results, organized by family. Variability inevitably reflects the range of acceptable choices relative to the options offered. Here the options were intended to span but not exceed a range of reasonable representations, but 40% of participants or more still agreed on the modal value of a parameter for about two thirds of the language phrases. Figure 4 shows the distribution of participants' responses of all three parameters.

We note that the expected synonyms ("light rain" and "sprinkle," "heavy rain" and "downpour") have CPSs that are close together, whereas antonyms and contrastive phrases tend to have values that are quite different for at least one parameter. The clear exception is the contrast between bipedal and quadrupedal *Travel* for small animals. Whether walking or running, the bird and the lizard seem to occupy volumes that are distinct from the cat but not from each other, given our actuator array. The values in Table I are candidates; no participant rated these exact (CPS, LP) pairs, nor was the confusability of the values within or across family members explicitly tested. A second study examines these issues.

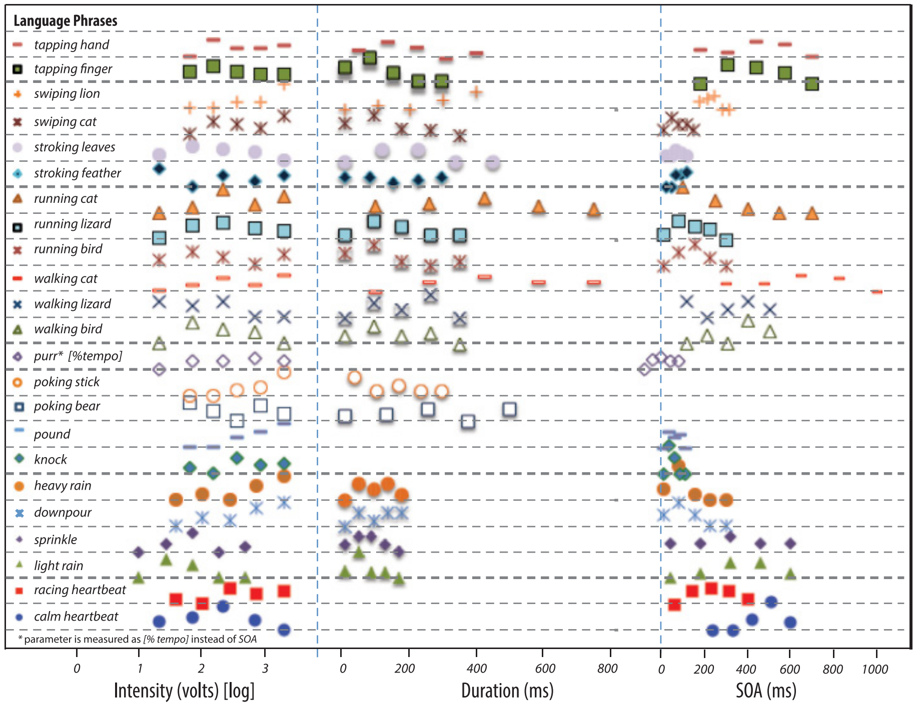


Fig. 4. Distribution of parameters selected by participants in Study 1. The horizontal grids are 10 counts apart.

5. STUDY 2: TESTING FAMILIES

There were two goals for this study. The first was to validate the candidate canonical values in Table I by demonstrating that the majority of users would judge the sensation created from the CPS to be an instance of the associated LP. The second goal was to test whether semantic relationships between families and family members map into relative haptic distances. In particular, it should be the case that FEs within a family are more confusable than FEs drawn from different families. Moreover, we should be able to predict parameter settings for some new FEs based on semantic reasoning alone, with no need for trial-and-error exploration or the explicit data gathering phase of Study 1.

5.1 Method

New participants (N = 63, 36 female, 27 male), aged 18–55 (M = 25.81, SD = 8.44) were recruited as in the first study. Participants were compensated 10 dollars.

5.1.1 *Materials and Design.* Participants used the same vibrotactile array and noise-isolating headphones as in the previous study. Stimuli consisted of four types of FEs to be rated:

Core: 19 of the 23 FEs from Study 1, with "sprinkle," and "downpour," moved to *Synonyms* and "bird walking," and "bird running" moved to *Inferences*, as explained in the following text.

Mismatches: 153 FEs that paired the language phrase from one *Core* FE with the CPSs from a different *Core* FE. The set was exhaustive within-family (i.e., the full cross-product of (CPS, LP) pairs was included). Across-family, every LP in *family1* was paired with one CPS of *family2* (see Table III for the exact pairings). For ease of communication, going forward, we continue to use quotes to refer to the prompting language and put the same phrase in boldface to refer to the sensation produced by the

CPS. So, "light rain": **light rain** refers to the *Core* FE, whereas "light rain": **heavy rain** refers to the *mismatch* in which the LP "light rain" is paired with the canonical sensation for heavy rain.

Synonyms: 10 FEs pairing a synonym for a *Core* LP with the associated CPS: "downpour" paired with the CPS for "heavy rain," "feather brushing," with "feather stroking," "lizard skittering" with "lizard running," "sprinkle" with "light rain," "branches poking" with "stick poking," "leaves brushing" with "leaves stroking," "rapping" with "knocking," "nervous heart beating" with "racing heart beating," "excited heart beating" with "racing heart beating," and "relaxed heart beating" with "calm heart beating."

Inferences: 21 FEs with CPSs that were inferred from the CPSs of *Core* FEs. Four inferences paired nonsynonymous nouns directly with *Core* CPSs, inferring a distance change of zero given the range that could be represented with the array: ("bird running" with "lizard running," "bird walking" with "lizard walking," "hamster running" with "lizard running," and "hand stroking" with "leaves stroking"). The remaining 17 used semantic reasoning to derive new CPS values based on the family as a whole. In the *Rain* family, "drizzle," "rain," and "rain shower"; in the *Travel* family, "bird darting," "lizard dashing," and "cat creeping"; in the *Strike* family, "bird pecking," "finger jabbing," "joystick knob poking," "paw tapping," "pen tapping," "thumb poking," "elbow poking," and "someone throwing seeds at me"; in the *Brush* family, "something as I squeeze into a cave," and "something as I squeeze into a cockpit"; and in the *Motor* family, "the engine of a spaceship humming."[5]

5.1.2 *Procedure.* Participants were seated and familiarized with vibratory sensation as in Study 1, followed by a new step-by-step training trial to teach the task and interface. Task instructions asked participants to decide whether and how well a sensation corresponded to a description. The training trial used the sentence, "I feel a ball bouncing on me," paired with a brief repeated sensation in the middle of the back.

LPs were again embedded in a sentence template designed to prime sense memory. To avoid sensory fatigue, participants could play a sensation at most five times within a trial. Rating occurred in two phases to allow for finer distinctions on a five-point scale than we had in the previous study. First, the participant simply judged whether the sensation felt like the description by choosing either "yes" or "no." Clicking "no" led immediately to the next trial. Clicking "yes," popped up a five-point scale with endpoints of "Acceptable" and "That's it," a midpoint of "Good," and two unlabeled values subdividing the intervals. Participants could change their goodness rating within a trial; the interface recorded only the value that was chosen when the participant explicitly pressed the button to move to the next trial.

Every participant rated every *Core*, *Synonym*, and *Inference* FE, but, to avoid sensory fatigue, only a third of the *Mismatches*. Types were intermixed and presented in random order.

5.2 Results and Discussion

Data for three of the participants were discarded because they had difficulty understanding the stimulus sentences. Of the remaining 60, all but one gave at least 20% "no" responses, indicating a willingness to differentiate between acceptable and unacceptable (CPS, LP) pairs. In addition, all but two participants gave ratings from the lower, middle, and upper portions of the scale, indicating a willingness to distinguish among degrees of acceptability. We note, however, that 32 of 60 participants used only the labeled buttons on the scale (values of 1, 3, and 5), possibly reducing the sensitivity of the instrument overall.

[5]Like "cat purring" in the *Core* set, some of the *Inference* LPs were chosen for use in a story listening application.

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*	*R*	*S*
A light rain	2.72	1.55			1.25		0.55								0.60		0.70		0.75
B heavy rain	0.55	3.13	1.50						0.60				0.40					0.70	0.50
C lizard walk	1.55		2.22	2.40	1.85	3.00		0.35						0.85			0.40		0.25
D lizard run	1.05		2.35	2.75	1.75	2.50					0.30				0.65		0.60		0.55
E cat walk		1.95	1.45	2.25	2.27	1.40				0.40						0.35		0.45	0.30
F cat run		1.30	1.35	1.35	0.90	2.12						0.20				0.60		0.05	0.20
G finger tap		0.65			0.65		2.32	2.25	1.10	1.25	2.95	2.45				0.20		1.70	0.15
H hand tap		0.85		0.60			2.00	1.47	0.70	0.35	3.10	3.10	0.05					1.20	0.15
I stick poke	0.30			0.90			1.90	1.90	1.53	1.75	1.95	2.70	0.35				0.55		0.20
J teddy poke		0.85				0.50	1.00	0.60	0.65	0.88	0.90	1.05		0.50				0.25	0.35
K knock	0.70		0.25				1.85	1.85	0.75	0.50	4.10	4.45			0.45		0.85		0.20
L pound		0.30			0.20		0.40	1.30	0.75	0.25	3.10	3.28				0.45		0.90	0.05
M feather stroke		0.20	0.85									0.05	2.10	2.35	1.20	0.30		0.15	0.30
N leaves stroke	0.25				0.35		0.05			0.35			2.40	1.58	0.55	0.40	0.50		0.10
O cat swipe	0.35			1.30						0.75			1.50	2.25	1.92	1.20	0.50		1.45
P lion swipe		0.90				0.35					0.40		0.85	0.85	0.80	1.35		0.15	0.45
Q calm heart		0.25			0.00			0.45						0.30			3.48	1.75	0.80
R racing heart	0.90		0.35						0.05						0.55		0.90	3.65	0.70
S cat purr	1.10		0.85				0.45						1.80				0.45		3.07

Fig. 5. Confusion matrix for 19 *Core* FEs with cell values corresponding to the mean rating given when the row LP was paired with the column CPS. Families are indicated with boxes and cells are shaded in proportion to the rating.

Ratings for *Core* FEs address validity: if our method for choosing a CPS from the data elicited in Study 1 is valid, then *Core* FEs should be recognized as acceptable instances of the named sensation by the majority of participants. The results in Table II show that, with the exception of "teddy bear poking," the majority of participants judged each FE to be at least "Acceptable," with six FEs receiving a mean rating of "Good," or better. With respect to the five lowest scorers—"hand tapping," "teddy bear poking," "stick poking," "leaves stroking," and "lion paw swiping"—we note that they were all chosen to be contrastive with other family members primarily with respect to the concept of size (as implemented by the duration parameter). In each pair, the LP receiving the lower rating is the less typical instrument for the action. The lower ratings, then, may reflect more uncertainty or variability in size when people imagine how the less typical sensation feels. The inclusion of "stick poking" in the low scoring set was surprising but might be explained along the same lines if participants had different size sticks in mind.

Mismatch FEs address the distinction-distance relationship across and within families by testing confusability. We specifically chose families whose definitions differed in haptic terms, expecting those distinctions to translate into nonoverlapping volumes. Figure 5 shows the confusion matrix of mean ratings for the 19 *Core* LPs, with families outlined in boxes. Each cell represents the mean rating when the LP in the row was paired with the CPS in the column. Values along the diagonal are the ratings for the *Core* FE itself, reiterating the third column of Table II. Cells are shaded in proportion to their values making it easy to see that there is little overlap between families. The saturation of color in rows and columns aligned with, but outside of, a set of boxed values are almost always much lighter than the values along the diagonal within the box. In other words, most *Core* FEs have parameter settings that make them difficult to perceive as belonging to another family. Further, darker shades tend to occur within boxes rather than outside them; members within a family are more likely to be confused

Table II. Ratings of *Core* FEs:
The percentage of users rating the CPS as at least "Acceptable" for the named sensation and the mean rating on a five-point scale (1 = "Acceptable," 3 = "Good," and 5 = "That's it!").

Language Phrase	% "yes"	Mean Rating
light rain	0.87	2.72
heavy rain	0.98	3.13
lizard walking	0.78	2.22
lizard running	0.90	2.75
cat walking	0.78	2.27
cat running	0.83	2.12
finger tapping	0.73	2.32
hand tapping	0.62	1.47
stick poking	0.67	1.53
teddy bear poking	0.40	0.88
knocking	0.98	4.10
pounding	0.93	3.28
feather stroking	0.77	2.10
leaves stroking	0.62	1.58
cat paw swiping	0.83	1.92
lion paw swiping	0.67	1.35
calm heart beating	0.92	3.48
racing heart beating	0.95	3.65
cat purring	0.90	3.07

with each other than with outsiders. Where exceptions do occur, they involve the five lower-rated FEs identified earlier.

Mismatches within a family were also expected to be revealing. Recall that most family members were designed as pairs at opposite ends of a haptically relevant scale to elicit family boundaries: light versus heavy rain, walking versus running, tapping versus pounding, and so on. The more semantically distinct two family members are, the less confusable they should be. For most pairs, the expectation holds—"light rain": **heavy rain** scores poorly (i.e., the pair is rated as mismatched), as do "cat running": **cat walking**, "finger tapping": **stick poking**, "feather stroking": **lion paw swiping**, and "calm heart beating": **racing heart beating**, to name just a few. But there are pairs that, despite being semantic endpoints, are nevertheless confusable. The FEs for the lizard walking and running, for example, are easily confused despite the SOA and intensity being quite different. In this case, the problem probably lies with the size of the actuators relative to the size of the array—the lizard's footfalls are distant enough even when walking to give the impression that ground is covered quickly. "Lizard walking": **cat running** and "cat walking": **lizard running** are probably confusable for similar reasons. Overall, it seems that the vibrotactile array we used was simply not big enough to capture the gait distinctions inherent in the semantics.

Although most mismatches produce ratings that are lower than the original *Core* FE, the mismatch pairs are not perceived symmetrically. **Heavy rain** is a more acceptable sensation for "light rain" than the sensation of **light rain** is for "heavy rain." A **racing heart** is more acceptable when prompted to expect a "calm heart" than a **calm heart** is when you are expecting "racing," and even a **cat running** is a more acceptable version of a "lizard walking" than vice versa. These violations of symmetry may simply reflect well-known asymmetries in semantic scales more generally. Semantic scales often have

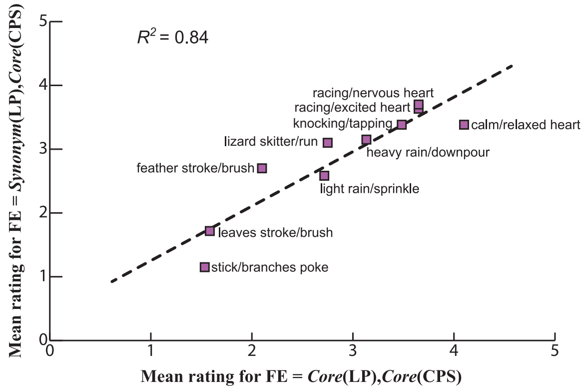


Fig. 6. Average ratings for *Core* versus *Synonym* FEs.

a marked end and a normative, unmarked end, which is typically used in the question form (absent other context we ask, "How hard is it raining?" and "How fast is your heart beating?" rather than "How softly?" or "How slowly?" [Lyons 1977]). Similarly, in our results, it is the sensation associated with the normative, unmarked end of the spectrum that is more likely to be accepted in the opposite context. In a few cases, the CPS for the unmarked end even received a higher rating for the language phrase at the marked end then the CPS for the marked end did. For example, the sensation of a **cat running** is not just rated higher for "lizard walking" than **lizard walking** is rated for "cat running," **cat running** is actually rated higher for "lizard walking" than **lizard walking** is. The same holds true for the extremes in the *Strike* family. We note here that this confusability appears as a function of lexical-semantic processing *absent other context*, but we examine the issue further in the next section.

The third type of stimuli, *Synonym* FEs, tests the simplest method for authoring new values based on semantic reasoning alone. They are created by pairing a synonym for the noun or verb in a *Core* LP with the *Core's* CPS. In terms of the distinction-distance hypothesis, if LPs have the same meaning then their associated parameter ranges occupy the same volume in haptic space. Thus, ratings for *Synonym* FEs should be predicted by their associated *Core* values. Figure 6 plots the mean ratings for each of the 10 pairs tested. The $\mathbf{R^2}$ value indicates that the perception of a *Synonym* FE is extremely well predicted by the perception of its corresponding *Core* value.

Finally, an *Inference* FE has parameter settings that are derived analogically, with distinctions in semantic space mapping to distances in one or more parameters in haptic space. For example, "rain" is treated as the representative LP for the *Rain* family volume whose boundaries occur at "heavy rain" and "light rain." Accordingly, the CPS for "rain" is mapped to the midpoints between the boundaries' CPSs for each parameter. Inferences can be inherited through synonymy: the CPS values for "rain shower" were set to the same values as "rain." Inferences can also chain: the CPS values for "drizzle" were defined to be consistent with both empirically derived and inferred values in the *Rain* family—about half way between "light rain" and "rain."

Four of the inferred FEs were created with the exact CPSs of a *Core* member of the family under the assumption that the semantic differences between the LPs were not haptically meaningful. "Bird running" and "bird walking" used the lizard values, as suggested by the closeness of the means from Study 1. "Hamster running" also used the lizard running values. "Hand stroking" used the values from "leaves stroking" in an effort to find a more natural noun to contrast with "feather."

Table III. Ratings of *Inferred* FEs:
The percentage rating the CPS as at least "Acceptable" for the named sensation and the mean rating on the five-point scale. Starred values (*) paired non-synonymous nouns with exact CPS values for a *Core* FE.

Language Phrase	% "yes"	Mean Rating
drizzle	0.97	3.40
rain	0.92	3.35
rain shower	0.97	3.43
lizard dashing	0.92	3.05
bird running*	0.83	2.40
bird darting	0.80	2.00
bird walking*	0.75	2.25
hamster running*	0.92	2.68
cat creeping	0.68	1.73
pen tapping	0.72	1.88
paw tapping	0.58	1.35
bird pecking	0.77	2.28
someone throwing seeds	0.62	1.43
finger jabbing	0.70	1.82
thumb poking	0.78	2.05
elbow poking	0.50	1.17
joystick poking	0.68	1.62
hand stroking*	0.80	2.23
squeeze into a cave	0.55	1.13
squeeze into a cockpit	0.50	0.88
engine humming	0.75	2.30

As is evident from Table III, inference worked well overall, with 19 of 21 FEs rated "Acceptable" or better by more than 50% of participants. The three additions to the *Rain* family were particularly successful, and "hand stroking" does seem to be a more natural label than "leaves stroking" for that CPS in the *Core*. The poor performance of "elbow poking" was a surprise. It has much stronger force of contact (intensity) than the more successful jabs and pokes of smaller objects, but only a bit more size (duration); increasing the size may improve it.

6. CONCLUSIONS AND APPLICATIONS

We had two goals in this work: to generate the beginnings of a library of usable haptic vocabulary and to do so with a methodology that would allow ongoing additions to the library in a principled and effective way. By reframing the problem as a search through haptic-semantic space, we were able to create a task that was easy for people with no haptic background to do (Study 1) and which produced verifiable results that could be extended automatically (Study 2). In both studies, we relied on cueing participants' memory for haptic sensation via language. The advantage of this approach is that it provides enough context to be evocative without involving cues from other sensory systems that may or may not be usable in specific applications. The situation is analogous to the creation of a sound effects library, where it is assumed that accuracy in the stand-alone form will produce synergy when combined with other media.

The use of lexical prompts was not without consequences, however. We saw a number of instances in which word choice mattered, including cases where atypical word combinations (“leaves stroking,” “teddy bear poking”) led to marginally acceptable ratings, and cases where a more judicious choice of LP improved ratings for the same CPS (“feather brushing,” “hand stroking”). In retrospect, it seems clear that if the goal is to evoke a maximally uniform response via language then common, even trite, phrases are the best semantic cues.

Lexical prompts also seem to have introduced a bias effect in tests for confusability that apparently violates the distinctness-distance hypothesis in the exact circumstances meant to expose family boundaries. The problem here may be more with the test than the underlying regularity, however. Although there were cases where the phrase associated with the unmarked end of the scale elicited higher ratings for the CPS at the marked end, the congruent phrase also elicited good ratings. Even if our intent were for FEs to always stand alone, storytelling requires gradation: the gentle tap at the door becomes insistent pounding when left unanswered. But our intent has never been for FEs to stand alone in practice, and the confusability introduced by our minimalist context is unlikely to occur in applications where additional sensory co-referents—seeing the finger tapping or hearing the fist pounding—will modulate the user’s expectations.

The use of FEs in real applications creates a host of issues that we have not begun to address: reproducibility of (or improvement on) an FE given a different array configuration, form factor or body surface, and coordination of vibrotactile sensation with other modalities, to name just a few. Nevertheless, we look forward to seeing the library used to enrich:

Entertainment content: probably the most direct use of FEs is in augmenting the entertainment experience in games, movies, music, shows, and rides, where realistic representation of events all around the user’s body may lead to a deeper sense of immersion and believability. Imagine a swipe or stroke on the back that follows an eerie wail just before a ghost comes into view, or the feeling of debris raining down after an explosion. Perhaps viewers will identify more strongly with a character when they can feel his or her heartbeat quicken at the approach of a loved one, or gradually slow after a chase.

Education and training: another motivation for establishing a haptic vocabulary is to supplement children’s story listening. Story listening is a valuable part of early literacy curricula, but it can be difficult to keep children who are used to visual stimuli engaged; it may be possible to make the experience more exciting by introducing FEs. More importantly, we are testing whether FEs can improve comprehension and memory for key story ideas and relations. In the same way, the use of FEs in introducing new vocabulary may facilitate children’s story reading and comprehension.

Social interactions: sensation affects emotion. The FE library may also be of use to researchers interested in designing expressive and affective haptic social content [Tsetserukou et al. 2009; Yohanan and Maclean 2012]. An FE experienced while reading email might add emotional information to words of sympathy, anger, and excitement, or, like an emoticon [Haans and Ijsselsteijn 2006], help to disambiguate intent.

We expect content creators in these areas to identify new FE families that are important to their work as they explore how FEs might enhance their craft. Thus, in addition to continuing to explore haptic-semantic space ourselves, we are looking at designing intuitive and user-friendly haptic authoring tools [Enriquez et al. 2003; Jonas 2008] based on the tools and interfaces developed for our studies. One such design, using interconnected, multiple views of family information, is shown in Figure 7. Our goal is to give media artists who are not familiar with the parametric composition of an FE the ability to design an FE as our participants did, using slider displacements that map lexical semantics to the haptic parameter space.

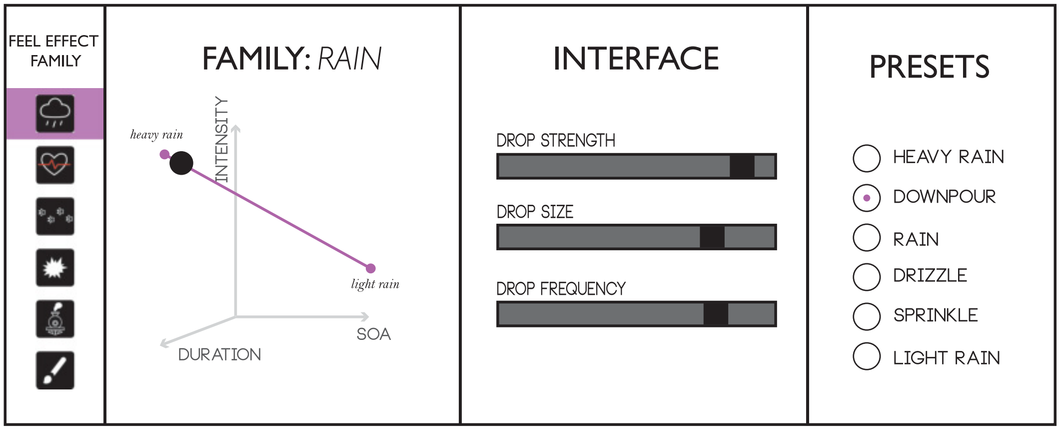


Fig. 7. A multiview authoring tool. Presets for "heavy rain" and "light rain" are shown in parametric space in the leftmost view, with the heavy dot indicating the current parameter settings as controlled by sliders in the middle view. The available library of presets for the family is listed in a meaningful order in the rightmost view.

ACKNOWLEDGMENTS

We wish to thank Kurt Kilpela and Jacklin Wu for their help in determining initial scale values for the FEs and implementing the data collection software. We are also thankful to Ivan Poupyrev and Disney Research to support this research.

REFERENCES

S. Brewster and L. Brown. 2004. Tactons: Structured tactile messages for non-visual information display. In *Proceedings of the 5th ACM Conference on Australasian User Interfaces*. 15–23.

L. Brunet, C. Megard, S. Paneels, G. Changeon, J. Lozada, M. P. Daniel, and F. Darses. 2013. Invitation to the voyage: The design of tactile metaphors to fulfill occasional travelers needs in transportation networks. In *Proceedings of the IEEE World Haptics Conference*. 259–264.

C. Collins. 1970. Tactile television: Mechanical and electrical image projection. *IEEE Transactions on Man-Machine Systems*. 65–71.

F. Danieau, J. Fleureau, P. Guillotel, N. Mollet, A. Lecuyer, and M. Christie. 2012. HapSeat: Producing motion sensation with multiple force-feedback devices embedded in a seat. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*. 69–76.

M. Enriquez, K. Maclean, and C. Chita. 2003. The hapticon editor: A tool in support of haptic communication research. In *Proceedings of the 11th IEEE Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. 356.

M. Enriquez, K. Maclean, and C. Chita. 2006. Haptic phonemes: Basic building blocks of haptic communication. In *Proceedings of the 8th International Conference on Multimodel Interfaces*.

R. Gault. 1927. Hearing through the sense of organs of touch and vibration. *Journal of the Franklin Institute* 204 (1927), 329–358.

E. Gunther and S. O'Modhrain. 2003. Cutaneous grooves: Composing for the sense of touch. *Journal of New Music Research* 32 (2003), 369–381.

A. Haans and W. Ijsselsteijn. 2006. Mediated social touch: A review of current research and future directions. *Virtual Reality* 9 (2006), 149–159.

A. Israr and I. Poupyrev. 2011. Tactile brush: Drawing on skin with a tactile grid display. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 2019–2028.

A. Israr, I. Poupyrev, C. Ioffreda, J. Cox, N. Gouveia, H. Bowles, A. Brakis, B. Knight, K. Mitchell, and T. Williams. 2011. Surround haptics: Sending shivers down your spine. In *ACM SIGGRAPH 2011 Emerging Technologies*. 14:1–14:1.

M. Jonas. 2008. *Tactile Editor: A Prototyping Tool to Design and Test Vibrotactile Patterns*. Technical Report. RWTH Aachen University.

S. Kim, A. Israr, and I. Poupyrev. 2013. Tactile rendering of 3D features on Touch Surfaces. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. 531–538.

R. Klatzky and J. Peck. 2012. Please touch: Object properties that invite touch. *IEEE Transactions on Haptics* 5 (2012), 139–147.

J. Linville and J. Bliss. 1966. A direct translation reading aid for the blind. *Proceedings of the IEEE* 54 (1966), 40–51.

J. Lyons. 1977. *Semantics*. Cambridge University Press.

K. Maclean and M. Enriquez. 2003. Perceptual design of haptic icons. In *Eurohaptics*.

G. Miller. 1995. WordNet: A lexical database for english. *Communications of the ACM* 38, 11 (1995), 39–41.

C. O'Sullivan and A. Chan. 2006. An activity classification for vibrotactile phenomena. In *Haptic and Audio Interaction Design*. Lecture Notes in Computer Science. Vol. 1429. 145–156.

A. Sahami, P. Holleis, A. Schmidt, and J. Häkkilä. 2008. Rich tactile output on mobile devices. In *Proceedings of the European Conference on Ambient Intelligence*. Springer Verlag, Berlin, 210–221.

K. Schuler. 2005. *Verbnet: A Broad-Coverage, Comprehensive Verb Lexicon*. Ph.D. dissertation. University of Pennsylvania, Philadelphia, PA.

R. Sodhi, I. Poupyrev, M. Glisson, and A. Israr. 2013. AIREAL: Interactive tactile experiences in free air. *ACM Transactions on Graphics* 32, 134 (2013).

H. Tan and A. Pentland. 2001. Tactual displays for sensory substitution and wearable computers. In *Fundamentals of Wearable Computers and Augmented Reality*. W. Barfield and T. Caudell, Eds. Lawrence Erlbaum Associates, 579–598.

D. Tsetserukou, A. Neviarouskaya, H. Prendinger, N. Kawakami, and S. Tachi. 2009. Affective haptics in emotional communication. In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. 1–6.

S. Yohanan and K. Maclean. 2012. The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature. *International Journal of Social Robotics* 4 (2012), 163–180.

Received June 2014; accepted June 2014