Shake2Talk: Multimodal Messaging for Interpersonal Communication

Lorna M. Brown[1] and John Williamson[2]

[1] Microsoft Research, 7 JJ Thomson Ave, Cambridge CB3 0FB
lornab@microsoft.com
[2] Department of Computing Science, University of Glasgow, G12 8QQ
jhw@dcs.gla.ac.uk

Abstract. This paper explores the possibilities of using audio and haptics for interpersonal communication via mobile devices. Drawing on the literature on current messaging practises, a new concept for multimodal messaging has been designed and developed. The *Shake2Talk* system allows users to construct audio-tactile messages through simple gesture interactions, and send these messages to other people. Such messages could be used to communicate a range of meanings, from the practical (e.g. "home safely", represented by the sound and sensation of a key turning in a lock) to the emotional (e.g. "thinking of you" represented by a heartbeat). This paper presents the background to this work, the system design and implementation and a plan for evaluation.

Keywords: haptics, audio, vibrotactile, multimodal interaction, mobile phones, messaging, remote communication, gesture recognition.

1 Introduction

It has been reported that there may be a need for new genres of communication [1, 2]. Harper [1] observed that people wish to use a mix of communication channels rather than one single channel, giving the example that paper mail is still in use despite the introduction of email and instant messaging. It seems that there may still be more new ways in which people could communicate, and this research aims to explore new possibilities for mobile communication. Currently the main forms of mobile communication are voice calls, text messaging (SMS) and multimedia messaging (MMS). This research investigates the area of non-visual messaging, in particular messaging using non-speech audio and vibrotactile display.

Non-speech sound and touch are ubiquitous in our everyday lives, but their potential in remote technology-mediated communication has not yet been realized. Using non-visual modalities such as sound and touch offers new opportunities for mobile messaging. A user's eyes are often engaged in other tasks when a message arrives, and they cannot look away from their current activity to attend to it. In addition if they are engaged in a physical task (e.g. driving, cooking), they are unable to pick up the phone and interact with it to receive a message. If the message was a non-speech audio or a vibrotactile message, instead of a visual message, and was presented to the user upon arrival (rather than requiring that the user open the inbox to

I. Oakley and S. Brewster (Eds.): HAID 2007, LNCS 4813, pp. 44–55, 2007.
© Springer-Verlag Berlin Heidelberg 2007

retrieve it), the user would be able to receive the information peripherally without having to disengage from their current activity to interact with the device.

Modalities such as non-speech audio and vibrotactile displays might offer new ways of communicating over and above text, speech, or picture messaging. It is said that a picture is worth a thousand words; could the same be true for a sound or a touch? A simple touch can have a much stronger impact than words, and sounds can evoke strong emotions and associations. It would be interesting to explore how people might communicate if they could send sounds or vibrotactile messages in place of, or in addition to, text or multimedia messages.

Based on these ideas, this paper presents a new concept for mobile messaging. The *Shake2Talk* system allows a user to create multimodal audio-tactile messages through simple gesture interactions with a mobile device, and then send these messages to another user. The long term aim of this research is to explore how such a non-visual, multimodal communication system might be used for interpersonal communication, alongside current messaging genres, when deployed with users. The paper presents a review of current messaging genres (Section 2) and related work on non-visual communication (Section 3), and then describes the Shake2Talk system, along with a plan for evaluation (Section 4).

2 Current Mobile Messaging Genres

In designing a new genre for communication it is important to consider the ways in which people currently communicate using mobile phones. The two main messaging genres are text messaging (SMS) and multimedia messaging (MMS) and a number of studies have investigated how people, particularly teenagers, use these services [3-7].

Ling, Julsrud and Ytrri [5] provide a categorization of the uses of text messaging. They state that the main uses are: co-ordination of events, questions, grooming (compliments/small talk), short one-word answers, commands/requests, information, personal news, invitations, jokes, thank you notes and apologies. Kopomaa [4] said that the main uses of SMS are for setting up meetings, exchanging gossip, giving info/reminders, and coordinating shared activities. Kasesniemi and Rautiainen [3] found that teenagers' use of SMS differs when they text their peers from when they text their family members. Between peers, teenagers use text messaging to express emotions, to gossip, to express longing, and to say things that they might not say in person, whereas, within a family group, they use text messaging for practical matters. SMS is also used within families to reinforce the family unit. The authors report the delight of a mother who received a message saying "mommy, I love you, I took out the garbage" from her 11 year old daughter in the middle of a work day [3].

Kindberg *et al.* [7] report a study of camera-phone use. They found that photo messaging was used for both emotional and functional purposes. The main uses were: extending an experience to absent friends, embodying personal common ground (e.g. sharing something that reminds you of a person, and which they will understand because of a shared history), as part of a conversation, to complete a task (e.g. sending images of an item of clothing to an absent person while shopping to ask if this is the item they want), conveying news, and providing evidence of an event (e.g. sending a photo when you arrive somewhere to show that you are there safely).

It will be interesting to see where a new form of messaging, such as audio-tactile messaging, might fit into these genres. Would it mainly be used for playful messaging, or might it also be used to provide information or commands/requests? Might it reveal new types of messaging that have previously not been used?

The literature on SMS shows that text messages are often treated like gifts [6]. They are carefully crafted by senders and often saved and treasured by recipients [2, 3]. Some of the literature on multimedia messaging indicates that the gift-like quality of messages is dependent on the effort put into crafting or creating the message by the sender [5]. This raises interesting questions for audio-tactile messaging. Will such messages also be treated as gifts? For them to be treated as gifts do they need to be crafted or constructed by the user? If so, does this mean that the user needs to record their own sounds and send them? Or would it be enough "craftsmanship" if the user created their own unique messages by interacting with a set of sounds and tactile sensations to create an audio-tactile message?

Another element that comes out of the literature on mobile phone use is that fiddling with objects, like beads, cigarettes, keys, etc seems to be a fundamental part of human nature, and a means of obtaining pleasure, and that fiddling with a mobile phone may also provide pleasure [4]. It might be possible to exploit this by using the fiddling with the phone itself to create messages, for example through gesture interaction, so that the fiddling itself actually becomes the means by which the message is created. This physical creation of the messages might also enable the craftsmanship required to create a message which would be considered to be a gift.

3 Non-visual Communication

In addition to considering the use of commercial messaging services, it is useful to consider related research in non-visual communication. This section discusses the use of touch and sound in real world and technology-mediated communication.

3.1 Communication Via Touch

Touch is widely used in social communication to enhance other forms of communication, and can "emphasize, qualify or contradict spoken words" [8]. Thayer [8] states that touch "will be trusted more by the person touched as a genuine reflection of feelings than all other forms of human communication". The literature also reports that "touching another's body generates an immediate demand for a response" [8], and that a lack of response may imply rejection [9]. Jones and Yarborough [9] grouped touch in social interaction into six categories: positive affect touches (support, appreciation, inclusion, sexual, affection), playful touches (playful affection, playful aggression), control touches (compliance, attention getting, announcing a response), ritualistic touches (greeting, departure), task related touches (reference to appearance, touch during a task) and accidental touches.

Considering the functions of touch in social communication may provide inspiration for the design of audio-tactile messages. The sense of touch is particularly interesting in remote communication since it is something that, currently, can only occur in face to face communication and not remotely. A touch of the hand can give reassurance; what if someone could send a "touch" to someone else's mobile phone to reassure them remotely? This does not necessarily mean that touch in face to face communication should be replaced in remote communication by mechanical touch but, rather, that the functions of social touching should be enabled through remote messaging. Therefore, the above categories of social touching may provide inspiration for the types of non-visual messages that people might want to send.

Vibrotactile displays, such as pager motors and other similar actuators are low-cost and widely available. Such displays are appropriate for communication systems as they are private to the user, and are attention grabbing. The disadvantage is that they need to be in contact with the skin for messages to be felt. A range of work has been conducted in the field of mediated social touch, both using vibrotactile display and tangible interfaces. A full review of this work is available from [10]. One such system is The Hug [11]: a robotic device which allows physical communication between two people. Hugging, stroking and squeezing a Hug device will send heat and vibrations to another Hug device. Other systems have been built to allow people to send vibrotactile messages via mobile phones [12] and instant messaging applications [13]. In recent work, Smith and MacLean [14] have explored the communication of emotion through a virtual hand stroke using a single degree of freedom haptic knob, with very promising results. Research has also been conducted into how to design complex vibrotactile messages for communicating rich data [15, 16]. However, it has been shown that, when these messages are abstract, training is needed, and the recognition rates are quite low [15]. Recognition rates are higher when tactile messages that use a metaphorical mapping to real world concepts are used, but the set of such messages is limited [16].

3.2 Communication Via Non-speech Audio

In the real world, we use sounds to understand what is going on around us, e.g. the sounds of doors opening and closing in our home or workplace indicate that people are arriving or leaving, the sound of pots and pans indicates that someone is cooking dinner, etc. In general, such peripheral awareness through sound is only available through co-location. If people could send sounds, then perhaps they could use this as a means to provide people with remote awareness of their actions. Audio display is widely available in mobile devices, with all mobile phones having audio output capabilities and many featuring high quality audio for their MP3 player functionality. The disadvantage of audio, compared to tactile display, is that it can be heard by other people, unless headphones are worn.

Two main forms of non-speech audio have been used in computer interfaces, Earcons and Auditory Icons. Earcons [17] are structured, abstract non-speech sounds, whereas Auditory Icons [18] use real world sounds to communicate information. Auditory Icons are of particular interest to this research since they use real world sounds with which users will already be familiar and, therefore, offer more opportunities for immediate expression without training to learn meanings.

Much of the research on non-speech audio display has focused on feedback on user actions and notification of system state rather than on communication. For example, Gaver's SonicFinder [18] used auditory icons to provide feedback on user interface events such as dragging, copying, opening or deleting files/folders. Earcons have also been used for feedback on system state and actions [17]. Gaver's [19] EARS (Environmental Audio Reminders) system used non-speech audio to support collaboration and awareness in a work environment. Auditory icons were presented to offices and common areas to remind people of meetings (the sound of murmuring voices) or to announce a trip to the pub (the sound of voices and of a beer being poured). In addition sounds were used to indicate emails arriving or people connecting to a video camera. All of this research has shown is that it is possible for people to attend to, and understand, non-speech sounds, while engaged in other tasks.

Some research has investigated the use of non speech audio for communication. For example, the Hubbub system [20] allowed people to send Sound Instant Messages (SIMs), in the form of Earcons, alongside text instant messages. In addition the system played sounds to indicate when contacts signed in or out, to provide peripheral awareness. Users could choose from a set of 14 pre-defined messages to send as SIMs. These SIMs were used quite regularly, but people found it hard to remember many different sounds. This might be improved by using auditory icons instead of Earcons, as people can then use their own real world associations to remember the messages. In addition, users might be more creative and expressive if they could interpret the sounds in their own way rather than having the meanings pre-assigned.

3.3 Discussion of Related Work on Non-visual Communication

This review has shown that both sound and touch are used regularly in real world communication and awareness, and there has been a range of work using these modalities for remote communication. However, many of these systems have a very limited vocabulary for communication. For example, squeezing or stroking an object to send a "hug" is very literal and offers little scope for new interpretations or new types of expression. Harper and Hodges [2] note that many remote communication applications "are popular at first, but soon wither: their value turning out to be essentially gimmicky and short-lived. Moreover, it also appears that they wither in part because what is communicated (and to some degree how) is sometimes too literal from the user perspective". They go on to say that, for a genre to succeed, it needs to be expressive rather than constrained, allowing users to communicate in rich ways [2]. This needs to be considered when designing any new communication genre. Whereas many of the systems described above have a single, literal function, e.g. squeezing an object to send a "hug", a system for richer communication needs to enable a wider range of messages to be sent, and to allow users to interpret these messages in their own ways. In so doing, it may be possible to open up a more open and less literal communication channel that allows new means of expression.

From this review of related work, a number of lessons can be learned for designing a new system for remote, interpersonal communication. The system should allow people to express themselves in rich ways, and should, therefore, offer a wide vocabulary and use messages that are open to interpretation by users rather than using very literal messages. Non-visual modalities, in particular non-speech real world audio and vibration, seem to offer promise for a new genre of communication. By building a non-visual communication system and deploying it with users it will be possible to investigate a number of questions, e.g. how do people communicate when using such a system, and how is it used alongside other forms of communication? One limitation of much of the previous work on non-visual communication is that there have not been any long term evaluations of the systems in use. The long term aim of this research is to deploy *Shake2Talk* in a longitudinal study, to understand how it is used alongside existing communication methods.

4 The Shake2Talk System

Based on the above discussion, a new concept has been generated for remote non-visual communication. *Shake2Talk* is a mobile audio-tactile messaging system. Audio-tactile messages are created by a user through simple gesture interactions with a mobile phone. After the message has been composed, the user presses send and the message is sent across the mobile phone network via SMS to the chosen recipient. Upon arrival the message is presented immediately (Figure 1). The reasons for these choices of input and output modalities are discussed further in the following sections.

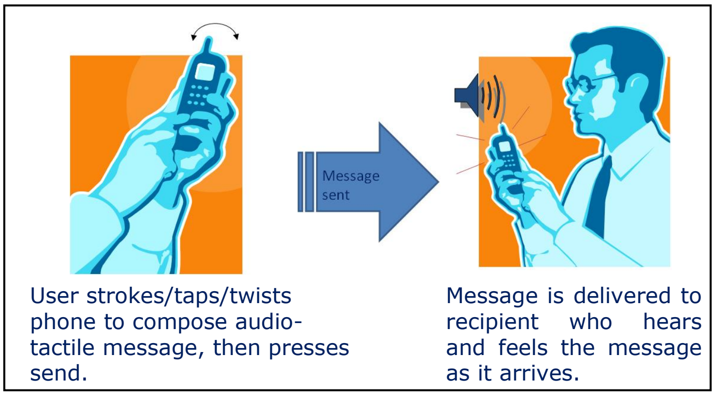

Fig. 1. The interaction with *Shake2Talk*

The *Shake2Talk* system comprises a Windows Smart Phone with a SHAKE device from SAMH Engineering fixed to the back of the phone (Figure 2). The SHAKE device contains inertial sensors (accelerometers, gyroscopes, capacitive sensors), which are used for gesture recognition. It also contains an eccentric-weighted pager motor to provide vibrotactile output (with control over onset, offset and frequency). The *Shake2Talk* application is run on the Smart Phone, which also provides audio output. The system is integrated with a mobile phone as it was felt that people would be more likely to adopt a system which works with their existing phone, than to carry a separate device. Although the SHAKE device is currently a separate unit, this could be integrated into the phone. Some phones contain accelerometers (e.g. Nokia 5500, Samsung SCH-310) and most feature vibration motors.

Fig. 2. The SHAKE device from SAMH Engineering, (left) and the SHAKE device attached to a smart phone and held by a user (right)

4.1 Output Modalities

Section 3 indicated that both audio and vibrotactile modalities offer a number of advantages and disadvantages. Tactile messages are private to the user and would not be overheard by others, but only a limited number of tactile messages can be identified, and training is necessary. By using audio messages it is possible to find a wider range of messages which can easily be interpreted by users. Real world sounds (auditory icons) could be particularly effective since people already have associations with such sounds. Whilst audio messages are not private to the user, it is possible that privacy regarding what is being communicated can be retained. The meaning of the audio message is likely not to be literal and, therefore, is probably only understandable to the parties to the communication. For example, whilst one couple might send a heartbeat sound to indicate that they are thinking of one another, another couple might send a heartbeat sound to indicate that they are nervous. The communication is, thus, dependent on a shared audio vocabulary, which results from a shared experience. Thus, such a system will be most appropriate for those people who know each other best, e.g. couples, close friends or family members. It should also be noted that people already use a range of sounds for personalized ringtones and, thus, there may be less concerns about phones producing sounds than one might expect.

Given the argument that audio offers a richer vocabulary than tactile, it might seem sensible to use audio-only messaging. However, by adding vibrotactile feedback to the audio, the information is redundantly coded in two modalities and the user will still feel the message whilst in a noisy environment. The tactile feedback also adds to the feeling of engagement when creating the message, and enhances the output. If distinctive tactile messages are used, they could be learned implicitly alongside the audio messages and then, eventually, could be used alone for discreet communication. Given these arguments, the final design for *Shake2Talk* uses non-speech everyday sounds, paired with synchronized tactile feedback (see Section 4.3 for further detail).

4.2 Inputs

A number of different input techniques could be used to select or create audio-tactile messages, with the most obvious being to select a preset message from a list. It might be more engaging for the user to interact directly with the audio and tactile feedback to create a message themselves. This can be done through gesture interaction with the device. In addition to providing engagement, it may be quicker for the user to perform a simple gesture (such as a tap, twist or a stroke) than to navigate through a series of menus to find the message. It was discussed in Section 2 that people fiddle with their phone to fill time, and it could be interesting to explore how such "fiddling", through gesture interaction, could be a means by which the messages are created. The element of "gift giving" in messaging may be partly dependent on the fact that the sender has taken time and effort to create the message. By having users create an audio-tactile message dynamically through gestures, these gift-like qualities may be retained more than if the user had simply selected a preset message from a list. Using gesture input may also lead to a very different experience than using menu input. In particular it results in an asymmetry between the experience of the sender and the experience of the recipient. When menu selection input is used, the sender and the recipient both experience the message passively. When gesture input is used, the sender actively interacts with the sounds and vibrations, whereas the recipient's experience is passive. It will be interesting to investigate whether this asymmetry of experience affects the perception of the meaning of the messages. *Shake2Talk* has been designed with both input options (menu selection and gesture input) so that they can be evaluated and compared.

Table 1. The Four Types of Gesture used in Shake2Talk

Gesture	**Recognition**
Stroke: User slides a finger from one capacitive sensor to the other in a "stroking" motion.	Recognition is performed by a simple finite state machine, based on thresholds on the capacitive sensor values and their derivatives. The machine accepts sequences of the form 1-down-1-up-2-down-2-up, within a certain timeout. On reaching the final state, the gesture event is triggered.
Tap: User taps a finger on a single capacitive sensor.	The tap recogniser also uses a state machine, with state changes triggered by threshold crossings from a single capacitive sensor. When the capacitive sensor is quickly activated and deactivated, the appropriate event is generated.
Flick: User moves the device forwards, then backwards, in a quick, sharp motion, like cracking a whip.	Flicking is sensed by accelerometers. The flick recognizer uses a physical model of a point mass anchored via a spring inside a sphere with a simulated viscous fluid. Rapid motions overcome the attraction of the spring and the damping effect of the fluid to strike the wall of the sphere, triggering the gesture event.
Twist: The user turns the device through a 180 degree rotation.	Twisting is sensed by gyroscopes, using a leaky integrator which operates on a single angular axis. The gesture event is triggered when the integrator output crosses a threshold.

Four different gestures are recognised: stroke, tap, flick and twist (Table 1). These were selected as gestures that could easily be distinguished from each other, and which required little learning from users (based on results of informal pilot tests). Recognition of gestures is effected by an ensemble of simple models which are continuously and simultaneously run. These models combine elementary dynamical systems with finite state machines to recognise the movements.

These gestures are associated with audio-tactile messages, e.g. tapping on the device might result in the sound and sensation of tapping or hitting something, whereas twisting the device might result in the sound and feel of wine pouring or a key turning in a lock (additional mappings are presented in Section 4.3). Users can perform a single gesture to generate a short audio-tactile message or sequences of gestures to create a longer message containing multiple audio-tactile messages.

The SHAKE device features a three-way button, and the gesture recognition will only occur when this button is pressed, to avoid accidental triggering of gestures. This three-way button also determines the particular sound and vibration that will be produced, with three possible audio-tactile messages for each gesture (one for each of the up, middle and down positions). This limits the number of sounds per gesture to three; to access a wider palette of sounds, different themes could be created and the user could switch theme to access another sound set.

4.3 Shake2Talk Messages

Once the decision had been made to use audio-tactile messages for output and gesture interaction for input, the types of messages that people might wish to send were considered. It was discussed above that what is needed is a system that allows people to communicate in new expressive ways that are not too literal and which are open to interpretation by the sender and recipient. Therefore, the aim with this design was to create a set of audio-tactile messages but not to impose meanings. Instead, users could appropriate the messages for any function they wished, based on their own interpretation of the sounds. However, in order to select sounds it was beneficial to consider the scenarios in which people might wish to use this system. A brainstorming session was held to generate ideas about the types of scenarios in which the system might be used, and the types of sounds that people might want to send. The current uses of SMS and MMS along with the categories of touch in social communication were used to generate these ideas. In addition, the types of gestures which could be recognized were used as an inspiration, by considering what types of sounds might map to these gestures. Table 2 shows a list of possible scenarios along with the related sounds and gestures. The sounds are short, and synchronized with vibrotactile feedback. The vibrotactile sensations are designed to match the temporal pattern and amplitude contour of the sounds. For example the sound of tapping on wood is accompanied by gentle, short vibrotactile pulses, whereas the sound of hitting metal uses stonger, longer pulses. The sound of a cat purring is accompanied by a series of vibrotactile pulses that match the amplitude contour of the sound, and the resulting sensation feels much like that experienced when stroking a cat. To illustrate the system in use, three scenarios are described below.

Scenario 1: “Call when you can”. Lucy wants to chat with her husband but it is not urgent, so she wants to indicate that he should call her when he has time. She picks up her *Shake2Talk* phone and taps on the device twice with her index finger. She hears and feels a gentle tapping, then presses send. When the message arrives on her husband’s phone, his phone reproduces the same audio-tactile message, making a tapping sound and vibrating with a “tap” sensation. This analogy to someone tapping him on the shoulder indicates that he should contact her when he has a chance. In contrast, if the message were urgent, she might select a different output, such as tapping on a wine glass to indicate that a more urgent reply was needed.

Scenario 2: "Home safely". Bob leaves his parent's house for a long drive home. When he arrives home he picks up his *Shake2Talk* phone and makes a twist gesture, like turning a key in a lock. This sends an audio-tactile message with the sound and feel of a key turning in a lock to his mother. When it arrives on her phone she will hear and feel the message, and immediately know that Bob has arrived home safely. An SMS saying "home safely" would have required her to pick up the phone to look at the message: an audio-tactile message, on the other hand, is heard peripherally thus notifying her immediately even if she is engaged in another task.

Scenario 3: "I'm nervous". Mary is feeling nervous before an exam. By stroking her *Shake2Talk* phone, a heartbeat sound is generated. Mary strokes the device faster to speed up the heartbeat and then sends the message to her friend, Barbara. Barbara receives the sound and sensation of a heartbeat and is aware that Mary is nervous. She twists her phone to generate the sound of a glass of wine being poured and sends this back to Mary to indicate that she should relax and not worry.

Table 2. Possible scenarios, with corresponding sounds and gestures

Scenario	Sound	Gesture
"Call when you can"	Gentle tapping	Tap
"I need to talk to you"	Tapping on a wine glass	Tap
"Call me now (angry)"	Banging on metal	Tap
"Fancy a drink?"	Beer pouring	Twist
"Relax!"	Wine pouring	Twist
"Home Safely"	Key in lock	Twist
"Thinking of you"	Regular heartbeat	Stroke (regular speed)
"I'm nervous"	Racing heartbeat	Stroke (fast)
"I'm bored"	Snore	Stroke
"happy"	Cat purring	Stroke
"I'm rushing"	Fast footsteps	Twist back and forth
"I've put the dinner on"	Rattling of pots and pans	Twist back and forth
"Angry"	Plates smashing	Flick
"Playful slap"	Slap	Flick
"Hurry Up"	Whip crack	Flick

4.4 Evaluation

The next stage of this work is to deploy the *Shake2Talk* system with users to evaluate how it might be used for communication. In the first instance, the system will be deployed with four pairs of users (couples or close family members), over a four week period, and the use of the system will be monitored through data logging, interviews, and diaries to see how people use the system. In addition to recording their use of the system (and reflecting upon the data logged by the system), users will be asked to record their use of other communication methods during this period. During the four weeks users will try the system with the gesture interaction input method for two weeks and the menu selection input method for two weeks (the order will be counterbalanced between pairs), to investigate how the methods compare. A number of research questions will be addressed by this study:

1. When and where do people send audio-tactile messages? What kinds of messages do they send? What kinds of messages would they like to send (that are not currently available in the system)?
2. How is *Shake2Talk* used in combination with other forms of communication?
3. How does the use, and perception, of *Shake2Talk* differ when messages are created through gestures rather than menu selection?
4. Does the gift giving element of text messaging transfer to audio-tactile messaging? Is this dependent on the sender creating messages themselves?
5. What are the implications of the fact that the sound always announces itself upon arrival? Does this cause annoyance or raise privacy concerns?

This evaluation will provide insights into how people might use an audio-tactile messaging system alongside other communication methods. In addition, it will act as a probe to understand more about how people currently communicate, by prompting them to think about the kinds of messages they currently send, and how these could be replaced by audio-tactile communication. The system will then be refined and a longer term study will be conducted, so as to gain greater insights into long-term use.

5 Conclusions and Future Work

This paper has introduced *Shake2Talk*: a mobile communication system in which users create audio-tactile messages through simple gesture interactions with a mobile phone and send these to other *Shake2Talk* users. The aim of this research was to design a new form of messaging, using non-visual modalities. Audio and tactile modalities have been combined in these messages so as to benefit from the affordances of each. Vibrotactile messages are attention grabbing and will be felt even in noisy environments when an audio message might be missed. Combining these with audio messages means that a richer set of messages can be used, and less training is needed as people can use their real world associations with these sounds.

The problem with many new communication systems is that they are short lived, having novelty value with first use, but failing to be adopted in the long term. To succeed it has been suggested that a communication system needs to offer rich expression and allow users to interpret messages in their own way, rather than being limited to literal communication. *Shake2Talk* has been designed to allow people to express themselves in new ways, without imposing meanings on the types of messages that people can send, but only by long term deployment will we understand whether this has been achieved. Therefore, the *Shake2Talk* system will now be deployed with a number of users in a longitudinal evaluation. This evaluation will investigate how, when and why people use audio-tactile messaging to communicate, and will provide insight into how such a system might be adopted and used.

Acknowledgments. Thanks to the SDS Group at MSRC (in particular Abigail Sellen, Richard Harper and Phil Gosset) and to Dominic Robson for his work on the sound design. This research was conducted while John Williamson was an intern at MSRC.

References

1. Harper, R.: Towards a new Communications Genre. IEEE Computer On-Line Magazine, 99–101 (2005)
2. Harper, R., Hodges, S.: Beyond talk, beyond sound: Emotional expression and the future of mobile connectivity. In: Höflich, J., Hartmann, M. (eds.) Mobile Communication in Everyday Life: An Ethnographic View, Frank & Timme (2006)
3. Kasesniemi, E., Rautiainen, P.: Mobile Culture of Children and Teenagers in Finland. In: Katz, J.E., Aakhus, M.A. (eds.) Perpetual Contact: Mobile Communication, Private Talk, Public Performance, pp. 170–192. Cambridge University Press, Cambridge (2002)
4. Kopomaa, T.: The breakthrough of text messaging in Finland. In: Harper, R., Palen, L., Taylor, A. (eds.) The Inside Text: Social, Cultural and Design Perspectives on SMS, Springer, Heidelberg (2005)
5. Ling, R., Julsrud, T., Yttri, B.: Nascent communication genres within SMS and MMS. In: Harper, R., Palen, L., Taylor, A. (eds.) The Inside Text: Social, Cultural and Design Perspectives on SMS, Springer, Heidelberg (2005)
6. Taylor, A., Harper, R.: Age-old Practises in the 'New World': A study of gift-giving between teenage mobile phone users. In: Proc. CHI 2002, pp. 439–446. ACM Press, New York (2002)
7. Kindberg, T., Spasojevic, M., Fleck, R.: I saw this and thought of you: Some social uses of camera phones. In: Proc. CHI 2005, pp. 1545–1548. ACM Press, New York (2005)
8. Thayer, S.: Social Touching. In: Schiff, W., Foulke, E. (eds.) Tactual Perception: A Sourcebook, Cambridge University Press, Cambridge (1982)
9. Jones, S.E., Yarborough, E.: A Naturalistic Study of The Meanings of Touch. Communication Monographs 52(1), 19–56 (1985)
10. Haans, A., IJsselsteijn, W.: Mediated social touch: a review of current research and future directions. Virtual Reality 9(2), 149–159 (2006)
11. DiSalvo, C., et al.: The Hug: An Exploration of Robotic Form For Intimate Communication. In: ROMAN 2003, pp. 403–408. IEEE, Los Alamitos (2003)
12. Chang, A., et al.: ComTouch: Design of a Vibrotactile Communication Device. In: Proc. DIS 2002, pp. 312–320. ACM press, New York (2002)
13. Rovers, A.F., van Essen, H.A.: HIM: A Framework for Haptic Instant Messaging. In: Ext. Abstracts CHI 2004, pp. 1313–1316. ACM Press, New York (2004)
14. Smith, J., MacLean, K.E.: Communicating Emotion Through a Haptic Link: Design Space and Methodology. International Journal of Human-Computer Studies (IJHCS), Special Issue on Affective Evaluation–Innovative Approaches 65(4), 376–387 (2007)
15. Brown, L.M., Brewster, S.A., Purchase, H.C.: A First Investigation into the Effectiveness of Tactons. In: Proc. World Haptics 2005, pp. 167–176. IEEE, Los Alamitos (2005)
16. Chan, A., MacLean, K., McGrenere, J.: Learning and Identifying Haptic Icons under Workload. In: Proc. World Haptics 2005, pp. 432–439. IEEE, Los Alamitos (2005)
17. Brewster, S.A., Wright, P.C., Edwards, A.D.N.: A detailed investigation into the effectiveness of earcons. In: Proc. ICAD 1992, pp. 471–498. Addison-Wesley, Reading (1992)
18. Gaver, W.W.: The SonicFinder: An Interface that Uses Auditory Icons. Human Computer Interaction 4(1) (1989)
19. Gaver, W.W.: Sound Support For Collaboration. In: Proc. ECSCW 1991, pp. 293–308. Kluwer Academic Publishers, Dordrecht (1991)
20. Isaacs, E., Walendowski, A., Ranganathan, D.: Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. In: Proc. CHI 2002, pp. 179–186. ACM Press, New York (2002)