What's in a Gesture: Exploring the potential of gesture-based interaction for low literacy user groups.

Sukeshini Grandhi
RWTH Aachen University
Human Technology Centre
Natural Media & Engineering
Aachen, Germany
grandhi@humtec.rwth-aachen.de

Gina Joue
RWTH Aachen University
Human Technology Centre
Natural Media & Engineering
Aachen, Germany
joue@humtec.rwth-aachen.de

Irene Mittelberg
RWTH Aachen University
Human Technology Centre
Natural Media & Engineering
Aachen, Germany
mittelberg@humtec.rwth-aachen.de

CHI'12, May 5–10, 2012, Austin, Texas, USA.
ACM 978-1-4503-1016-1/12/05.

Abstract
Gesture-based interfaces are seen as effective means of enabling intuitive and natural ways of interaction with technology and thus can potentially be a great interaction technique to combat challenges of technology use in low literacy user groups. However, these interfaces can be effective and intuitive only if anchored in a deep understanding of how humans use gestures to communicate within the social and cultural context of the users. This position paper presents our preliminary ideas on how we could effectively extract gesture vocabulary for the design of gesture-based applications for low literacy user groups, bringing together insights from HCI, semiotics, and cognitive science.

Keywords
Gestures, cognitive-semiotic principles, metonymy, user acceptance, interaction design, illiteracy.

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces

Introduction

A major barrier in adopting technology for illiterate and limited-literate users is the lack of applications that are designed for their usability and their context. The challenge in designing user friendly applications for this user group stems primarily from not being able to rely on existing tried and tested design features and principles, such as text-based input methods, and mental models, such as sequential or hierarchal navigation of content, of literate users [14][5].

Previous research has shown the ineffectiveness of text-based interaction in mobile applications for low literacy groups [14]. Alternate methods such as voice and graphical displays, however, improved user's task performance and user experience [13][4][23]. Conspicuously absent in previous research is the exploration of gestures as a way to interact with technology for low literacy user groups, although there has been a recent surge in research exploring the effectiveness of gesture-based interaction in mainstream HCI design. The intuitive, natural and pervasive use of hand gestures in human-human communication has set HCI researchers and practitioners in the pursuit of the design of gestural interfaces that inherit these qualities. Given this research is still in its nascent form, researchers have focused on applications development in more developed countries with predominantly literate user groups. However previous research, as cited above, has suggested that simple translations of design principles from literate user groups may not be useful. Thus, in this position paper, we present our preliminary ideas on how to explore the feasibility of using gestures as effective means for interaction with technology specifically for low literacy user groups in developing countries.

Gesture-based interaction

Gestures implemented in HCI can be 1) single or multi-point touch gestures on touch screens, 2) free-form gestures in space involving manipulation of a device such as the Wii remote, or 3) touchless free-form gesture inputs which do not require the manipulation of any input device but are often coupled with computer vision systems and/or motion sensors (e.g. Microsoft Kinect/Natal Project).

A key challenge in the implementation of gesture-based interfaces is developing natural, intuitive and meaningful gesture vocabularies appropriate for the tasks in question [3][17][22]. While less so in touch technology, the gesture vocabulary in touchless interfaces is often an ad-hoc choice made by the designer to trigger certain pre-assigned actions (e.g. the gesture of clapping one's hands to turn on the computer). Mostly chosen for their ease of implementation to facilitate distinctive recognition and segmentation, such gestures have arbitrary mappings that require users to learn a set of gestures and the associated commands they trigger [11]. This often puts a strain on user memory and defeats the purpose of using gestures as a way to facilitate intuitive and natural interaction. With an illiterate user group, which may have more difficulty accessing information such as in printed manuals even if pictorial, for instance, having an intuitive gesture vocabulary could be of immense value. Specifically, using mobile devices as gesture sensors could be an optimal solution in low literacy groups. Not only are mobile phones widely adopted in low literacy user groups in developing countries [14], they are also becoming more commonly equipped with accelerometers and gyroscopes, which have shown to be useful in gesture recognition [21] (also Nintendo Wii). Furthermore, using gestures also relieves users from having to work within small screens on mobile phones and from being forced to work with mental models of literate user groups.

Cognitive and Semiotic Principles

We believe that we can develop more intuitive gesture vocabularies for interaction with technology by considering and understanding the cognitive mechanisms which underlie how we as humans symbolically refer to and represent the information in our surroundings due to our cultural influences as well as innate symbolizing (representation) capabilities. After all, the virtual world of technology is comprised of representations of the physical world around us.

Metaphor and metonymy

One such mechanism of representation draws on *metaphoric* (similarity) relationships, which are commonly used in HCI. For example, computer operating systems have virtual "files" and "folders", some of which you can keep on your "desktop". The metaphor is the mapping of two different experiential domains: here, folders in the real world and folders in the virtual world. Using gestures in the system interface to manipulate these virtual objects as if they were the real counterparts could clarify and reinforce this metaphor for the less technologically literate. However, once appropriate metaphors are identified, how should these manipulation interactions be designed?

One possibility is to allow users to pantomime the actions as if the real counterpart were there (e.g. to open a "folder": hand gripping a flap of the virtual folder and motion to open it). Another is for the person to physically represent the manipulated objects themselves (e.g. the hands are the folder: flat palms together oriented laterally, then supinating to palms up). These options are captured by another mechanism of representation, metonymy, which refers to contiguity (e.g. spatial or temporal) relationships. In the example, the pantomimic action of opening a "real folder" is iconic (i.e. physical, such as visual, resemblance) of the action that would be done, and the hands are external to the object manipulated (*external metonymy* [9][16]). In the hand-as-folder alternative (hand becomes part of the folder, hence *internal metonymy* [9][16]), the hand shape bears some iconic resemblance to the shape of a folder. These metonymic distinctions in gestures are also discussed in neuro- and developmental psychology literature. For transitive actions, *internal metonymy* is termed as *body-part-as-object* in neuropsychology, and is an example of *object substitution* in developmental psychology. In neurological examinations, *body-part-as-object* action is considered an error associated with movement planning disorders [7]. The ability to perform hand actions where the hand holds an imagined object (*external metonymy*), is referred to as *pantomimed action with an imagined object* in neuropsychology, and is considered the proper developmental trend in children [2]. Compared to metaphor, metonymy is seldomly explicitly exploited in interface design, but as illustrated above, it plays an important role for gestural interaction.

Both metaphor and metonymy are conceptual or representational shortcuts that can facilitate communication through both auditory and visual channels. A few researchers have noted the importance in HCI design of understanding these mechanisms [17][8], as well as in general, semiotics [1][8], or the study of sign/symbols (i.e. representational) systems. Moreover, researchers also have found that low literacy users seem to be more likely to benefit from use of metaphors involving concepts used in familiar everyday activities and objects [13][19]. We propose that this cognitive semiotic approach can potentially help towards understanding the potential of gesture-based application design for low literacy user groups in developing countries.

Proposed method for gesture elicitation

In order to tap into culturally specific manifestations and more universal aspects of these cognitive mechanisms, we need a method to elicit gestures to allow us to explore the workings of these mechanisms and thus choose the more meaningful and intuitive alternatives. In our recent work on developing an intuitive and natural gesture vocabulary for touchless gestural interaction [8], we adopted a user-centric approach [17][24] to tap into people's conceptualizations and to understand the metaphoric and metonymic modes reflected in their gestures. A user-centric approach to designing gesture vocabulary is even more imperative in developing countries because while universal cognitive principles may arguably exist, cultural differences can dominate. For example, although in European cultures there is a conceptual metaphor of time in terms of space where the future is in front of us whereas the past is behind us (we see "ahead" into the future and we leave memories "behind" us), in certain cultures like the South American indigenous Aymara people, it is the opposite orientation. Such metaphors for time are also reflected in our gestures [18].

In our study, participants were presented pairs of "before" and "after" pictures and had to perform a gesture to "explain and show" the experimenter, seated in front of them in a position where the pictures could not be seen, what needed to be done to achieve to go from the "before" picture to the "after" picture. These scenarios (Fig 1) reflected simple computer tasks (e.g. cut, erase, close), but were camouflaged as everyday non-computer scenarios to minimize the influence of conceptual models of performing these tasks on pre-existing input devices. Before-and-after picture pairs were presented to place emphasis on the goal of the action and avoid constraining participants to perform the tasks with a particular tool. Analyzing these gestures in terms of semiotic principles then revealed patterns, where the most frequently produced patterns would correspond to the more intuitive aspects that should be considered in gesture vocabulary design.

This methodology of free-form gesture elicitation for guiding interface design can be adopted regardless of the interface of interaction. For example it can be used for deviceless interaction with large screens, which could be used in community information exchange boards or even in specially designed ATMs. It can also be adopted for the elicitation of free-form gestures with a mobile device in hand (as in Nintendo Wii). Furthermore it can be used for one-handed deviceless gesturing to a mobile device, which is held in the other hand and equipped with gesture recognition capabilities.

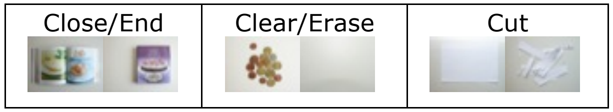

Figure 1 Sample transitive actions used in the study

To illustrate the method for a domain relevant to illiterate user groups, let us consider banking. In the context of mobile banking, designers can look for the style of representations in how users gesture when asked to loan money to a relative, or acquire account summary with a mobile phone in their hand. Stimuli used to elicit these gestures can be 1) before-and-after pictures of artifacts involved such as money and checkbooks and 2) real artifacts depicting before and after change. Use of such different stimuli can shed light on the imagery evoked in user's mind and mental models used to understand or perform these tasks. We believe that this method facilitates choosing gestures which increase imageability, and hence facilitate understanding, of virtual actions.

Broader Research Challenges

While, as HCI researchers, we are eager to bring in gesture-based interaction to mainstream HCI, there are several research challenges that we face before this is achieved for both literate and illiterate user groups. There is a limited understanding of the nature of tasks for which gesture-based interaction will be useful and desired by users. We do not have a clear understanding of applicable and appropriate domains or how we can use gestural interaction for low literacy user groups.

While we have focused on gesture elicitation we acknowledge that several other challenges remain in regarding the feasibility of gesture-based application use for low literacy users in developing countries. Apart from costs of development and implementation, there exist challenges of training and generalizability. Although we propose that gesture interfaces designed with our methodology will facilitate interaction with technology, it remains to be seen how easily users can be trained to use gestures with technology, what technological user resistance we will face, and how well frequently produced gestures translate to users using them for different applications. We hope to create discussion over these issues at the workshop.

Acknowledgements

This work has been funded by the Excellence Initiative of the German Federal and State Governments.

[1] Andersen P.B. What Semiotics Can and Cannot Do for HCI. In *Proc. CHI'2000 Workshop on Semiotic Approaches to User Interface Design*, (2000).

[2] Boyatzis, C.J. and Watson, M.W. Preschool children's symbolic representation of objects through gestures. *Child Development* 64 (1993), 729-735.

[3] Buxton, B. *Sketching User Experiences: Getting the design right and the right design*. Elsevier, Amsterdam, 2007.

[4] Deo, S., Nichols, D.M., Cunningham, S.J., Witten, I.H. and Trujillo, M.F. Digital library access for illiterate users. In the *Proceedings 2004 International Research Conference on Innovations in Information Technology (IIT2004)*, Dubai, U.A.E (2004), pp. 506-516.

[5] Findlater, L., Balakrishnan, R. and Toyama, K. Comparing semiliterate and illiterate users' ability to transition from audio+text to text-only interaction. In *Proc. ACM Conference on Human Factors In Computing Systems (CHI 2009)*, ACM Press (2009), 1751-1760.

[6] Forceville, C.J. and Urios-Aparisi, E. (eds.). *Multimodal metaphor*. Mouton de Gruyter, Berlin/New York, 2009.

[7] Goodglass, H., and Kaplan, E. Disturbance of gesture and pantomine in aphasia. *Brain 86* (1963), 703-720.

[8] Grandhi, S.A., Joue, G., Mittelberg, I. Understanding Naturalness and Intuitiveness in Gesture Production: Insights for Touchless Gestural Interfaces. In *Proc. of the 2011 Annual Conference on Human Factors in Computing Systems*, ACM, New York, NY (2011), 821-824.

[9] Jakobson, R. and Pomorska, K. *Dialogues*. MIT Press, Cambridge, MA, 1983.

[10] Lalji, Z., and Good, J. Designing new technologies for illiterate populations: A study in mobile phone interface design. Interacting with Computers (2008), 20(6), pp.574-586

[11] Lee, J.C. In search of a natural gesture. *XRDS: Crossroads, The ACM Magazine for Students 16*, 4 (2010), 9-12.

[12] McNeill, D. *Gesture and Thought*. Chicago University Press, Chicago, 2005.

[13] Medhi, I. and Ratan, A. Usability test of—bank's biometric-ATM prototype: An evaluation. Tech. rep., Microsoft Research India, 2006.

[14] Medhi, I., Patnaik, S., Brunskill, E., Gautama S.N.N., Thies, W. and Toyama, K. Designing Mobile Interfaces for Novice and Low-Literacy Users, in ACM ToCHI, ACM Transactions on Computer-Human Interaction, April 2011.

[15] Mittelberg, I. Peircean semiotics meets conceptual metaphor: Iconic modes in gestural representations of grammar. In A. Cienki and C. Müller (eds.), *Metaphor and Gesture*, John Benjamins, Amsterdam (2008), 115-154.

[16] Mittelberg, I., and Waugh, L.R. Metonymy first, metaphor second: A cognitive semiotic approach to multimodal figures of thought in co-speech gesture. In C. J. Forceville and E. Urios-Aparisi (Eds.), *Multimodal metaphor*, Mouton de Gruyter, Berlin, New York (2009), 329-358.

[17] Nielsen, M., Moeslund, T., Storring, M. and Granum, E. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In A. Camurri and G. Volpe (eds.). *Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop, GW 2003*, LNCS 2915, Springer, Berlin (2004), 409-420.

[18] Núñez, R. and Sweetser, E. With the Future Behind Them: Convergent Evidence From Aymara Language and Gesture in the Crosslinguistic Comparison of Spatial Construals of Time. *Cognitive Science* (2006), 30, 1-49.

[19] Parikh, T. S., Javid, P., Sasikumar, K., Ghosh, K. and Toyama, K. Mobile phones and paper documents: Evaluating a new approach for capturing microfinance data in rural India. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)* (2006).

[20] Peirce, C.S. Logic as semiotic: The theory of signs (1893-1920). In J. Bucher (ed.), *Philosophical Writings of Peirce*. Dover, New York (1995), 98-119.

[21] Ruiz, J., Li, Y., and Lank, E. User-defined motion gestures for mobile interaction. In *Proc. ACM Conference on Human Factors In Computing Systems (CHI 2011)*, ACM Press (2011), 197-206.

[22] Saffer, D. *Designing Gestural Interfaces*. O'Reilly Media, 2008.

[23] Taoufik, I., Kabaili, H., and Kettani, D. Designing an e-government portal accessible to illiterate citizens. In the *Proceedings of the International Conference on Theory and Practice of Electronic Governance*, (2007).

[24] Wobbrock, J.O., Morris, M.R. and Wilson, A.D. User-defined gestures for surface computing. *Proc. of the 27th Int. Conf. on Human Factors in Computing Systems,* ACM, New York, NY (2009), 1083-1092.