Foundations and Trends® in Human-Computer Interaction
Vol. 8, No. 1 (2014) 1–72

DOI: 10.1561/1100000047

User Interface Design for Low-literate and Novice Users: Past, Present and Future

Indrani Medhi Thies
Microsoft Research India,
Bangalore, India
indranim@microsoft.com

Contents

Abstract

One of the greatest challenges in providing information and communication technology access is that about 775 million people in the world are completely non-literate and many are able to read only with great difficulty and effort. Even though mobile phone penetration is growing very fast, people with low levels of literacy have been found to avoid complex functions, and primarily use mobile phones for voice communication only. This monograph looks at how we can design ICT user interfaces (UIs) such that novice and low-literate users can access a broad range of services and utilities, increasingly available to them, with minimal training and external assistance. It begins by observing the challenges in designing for low-literate users, and in conducting user studies among low-income communities in the developing world. It discusses techniques used by researchers in overcoming some of these challenges. It presents a review of existing examples of UIs for novice and low-literate users, which have used various combinations of input–output modalities other than text. It goes on to discuss strong trends that are starting to emerge in this design space and concludes with opportunities and future directions for research and design of UIs targeted at populations with low-literacy.

I. Medhi Thies. *User Interface Design for Low-literate and Novice Users: Past, Present and Future*. Foundations and Trends® in Human-Computer Interaction, vol. 8, no. 1, pp. 1–72, 2014.
DOI: 10.1561/1100000047.

1

Introduction

Today Information and Communication Technologies (ICTs) are becoming increasingly available to people around the globe. The prices of devices are rapidly dropping and people across income groups are getting access to a range of technologies, from mobile phones and PCs to handheld tablets. If we consider mobile phones, as of 2014 there were about 6.9 billion of them in the world and 78% of the subscribers lived in developing countries [ITU, 2014]. Beyond facilitating communication, ICTs have transformed the way people send money, manage health, check market prices, engage with government, manage emergency response, and many other things. Therefore, many entities with a global development focus have turned to ICT as a potential platform for delivering services related to international development.

PC penetration may not be as high as mobile phones. But in the past decade there has also been a surge in public kiosks and telecenters [Best and Kumar, 2008, Heeks and Kanashiro, 2009, Kuriyan et al., 2006]. A telecenter is a public place where people have shared access to PCs and the internet so they can create, learn, and communicate with others while developing digital skills. The primary objective of telecenters is to contribute to the development of a community by bridging the digital divide, promoting health issues, creating economic opportunities, and reaching out to youth (Telecentre.org retrieved July 7, 2014).

However, one of the greatest challenges in providing any kind of services on ICTs — mobile phones, PCs, and others — is that 775 million people in the world are completely non-literate, and even more are able to read only with great difficulty and effort. Of the non-literate population 85% live in 41 countries, most of which are between developing to least developed [UN News Centre, 2012]. And among poor populations in these countries even the literate typically are novice users of ICTs. Research shows that non-literate populations avoid complex functions, and primarily use phones for synchronous voice communication only [Chipchase, 2005]. For many low-literate people, even the contact function on their phones is too difficult to use, so they dial numbers from scratch every time they need to make a voice call [Medhi et al., 2011].

Indeed there is reason to believe that UI design should be different for low-literate users because they have a different set of cognitive skills. Studies by cognitive science researchers in developed countries (e.g., the U & Netherlands) have shown that low-literate people with limited formal education differ from people with good educations in their performance on a variety of cognitive skills: language processing, visual organization and visual memory, mental spatial orientation, speed of cognitive processing, vigilance, divided attention and perceived self-efficacy [Van Linden and Cremers, 2008, Manly et al., 2003]. These skills have been said to be essential for realizing interaction on ICTs. Indeed one conception of "digital literacy" proposes that it requires a variety of complex cognitive, motor, sociological, and emotional skills, and a mature understanding of cyberspace "rules" so as to function effectively in digital environments [Eshet-Alkalai, 2004].

All this brings us to the question: how can we design ICT user interfaces (UIs) such that novice and low-literate users can access a broad range of services and utilities that are increasingly available to them? Researchers in the domain of Human–Computer Interaction for International Development (HCI4D) have been investigating this question, exploring how UIs can be made more user-friendly for low-literate users. HCI4D is an emerging domain of research in the field of Information and Communication Technologies and Development (ICTD). In this monograph, we first examine the various challenges that have to be taken into account while designing these interfaces. We discuss issues of low formal education and low textual literacy, limited exposure to ICTs, and other cognitive challenges associated with low education. We also discuss challenges in conducting user studies in low-income, low-literate communities, and techniques researchers in HCI4D have used to overcome the same. We then go on to review different examples of UI research done over the last several years that have proposed non-textual designs. We review research conducted in the context of non-literate adults; work done with children is outside of the scope of this monograph. Note that many of the studies reviewed are from India, as much of the research has been conducted there and we are more familiar with this work. Of course, other regions and cultures have their own unique context, but we believe many of the findings also apply to these areas.

Studies in HCI4D have used various input–output modalities other than text. Researchers have looked at speech or voice and touch as input mostly because it is a natural means of expression than typing [Patel et al., 2010, Sherwani et al., 2007]. Other forms of input modalities studied are pen-based and touch-tone input [Raza et al., 2013, Underwood et al., 2013]. In output or information display a number of research studies have explored the use of graphics and imagery [Grisedale et al., 1997, Huenerfauth, 2002, Medhi et al., 2011]. Researchers have also studied the use of audio output combined with graphics or text and on Interactive Voice Response systems (IVR) [Friscira et al., 2012, Koradia et al., 2013, Parikh et al., 2003]. Video output has been used to help low-literate users overcome the inability to read text [Medhi and Toyama, 2007, Smyth et al., 2010b]. Other design features have included consistent help, no scroll bars, use of numbers, ultra-simple navigation, fewer menus and dedicated buttons [Jones and Marsden, 2005, Kurvers, 2002, Medhi et al., 2011]. In addition to the above there have also been comparative studies examining the trade-offs between various combinations of input–output modalities going from inflexible to flexible in input (such as typing, structured speech and freeform speech) and lean to rich output (such as text, audio and graphics+audio) [Chakraborty et al., 2013, Cuendet et al., 2013, Patnaik et al., 2009]. We make design recommendations for UIs for low-literate users based on current work.

Above and beyond these examples we review strong research trends in HCI4D that are currently starting to emerge in the design of UIs for low-literate users. Researchers are beginning to look at how non-literacy is not just about the inability to read text, but is correlated with cognitive skills such as the ability to transfer learning in video-based skills training [Medhi et al., 2012b], and the ability to navigate hierarchical organization of information architectures [Medhi et al., 2013a,b]. These studies have offered implications for design for both PC and mobile UIs. Another emerging trend we review is of UIs for production of content by low-literate users. Most of the examples we review leverage IVR systems to overcome issues of non-literacy. These are in citizen journalism [Mudliar et al., 2013], agriculture Q&A [Patel et al., 2010], a virally spread voice manipulation and forwarding system [Raza et al., 2013], and a community moderated talent competition [Koradia et al., 2013]. In addition, we examine a trend for natural UIs, but for literate and tech savvy users, that use gestures, speech, touch and other forms of natural interaction [Windows Phone Cortana, 2014, Google, 2014, Xbox 360 + Kinect, 2014, Google now, 2014, Apple Siri, 2014] (retrieved July 18, 2014). Through these examples, we discuss how design principles for literate users and non-literate users could be starting to converge.

Research in design for low-literate users presents interesting challenges because of the environment and ecosystem that these users live in. Urban environments could be vast informal settlements found in a rundown area of a city characterized by substandard housing, and commonly referred to as ‘slum areas’ [UN-HABITAT, 2007]. Average areas of households are small (<200 sq.ft.) for about 6–7 family members per household. And infrastructure in water, electricity and sanitation are constrained. Rural environments could have comparatively larger household areas, but the infrastructure conditions are similar. In both these rural and urban environments exposure to ICTs also tends to be constrained. Given this, as a follow-up we discuss conversations that have started emerging in designing UIs for low-literate users that goes beyond usability. These issues relate to socio-cultural and socio-psychological concerns such as shared and mediated uses of technology [Parikh and Ghosh, 2006, Sambasivan et al., 2010] and intimidation caused by technology [Medhi et al., 2010].

Finally, we present opportunities in low-literate UI design research for future work. We discuss directions for studying UIs for training effects and how learning occurs through longitudinal field deployments. We also discuss opportunities in designing for multiple users for both collaborative and competitive scenarios. In addition, we think there is scope for designing for the uptake of technology in mediated and assisted scenarios, such as when a technology is seeded into a community through a human mediator. We identify opportunities in UI and product design research for lowering intimidation among low-income, low-literate users, for both formal evaluations and pilot deployments. We discuss opportunities for studying how cost-consciousness can impact user experience. We conclude with some thoughts on how research in this area may impact livelihoods among low-income, low-literate communities.

2

Challenges in Designing for Low-Literacy

Most studies in HCI4D have looked at two main challenges in designing for adult users with low-literacy: (a) low levels of formal education and textual non-literacy, (b) difference in cognitive abilities. We discuss each of them in detail here:

Low levels of formal education and textual non-literacy

One of the biggest challenges in designing for low-literacy is that people lack formal education and are unable to read text. A number of studies have thus used formal education as a proxy for literacy [Grisedale et al., 1997, Huenerfauth, 2002, Medhi et al., 2011, Parikh et al., 2003]. In these studies, target users had formal education less than Grade VII, with a majority having never attended a formal school. And even among people who had been to school, said that they were "out of practice" since the time they left school. They "rarely" read and wrote in their day-to-day lives. Many of these people are reported to have functional non-literacy: not able to read real-world print, e.g., road signs, bus schedules, etc. [Medhi et al., 2012a]. People spoke various local languages but did not have any working knowledge of English.

While studies in HCI4D and cognitive science have used formal education as the proxy for textual literacy, researchers have also shown that textual literacy might not necessarily be correlated with the level of formal education. As the Vai Project conducted on a small Liberian population that home-schools its children suggests some cognitive skills are linked with textual literacy, but not necessarily with formal education [Scribner and Cole, 1981].

The overall education status of an individual could depend on a number of factors, which includes what school the individual attended, quality of teaching, role of parents and home environment, amount of effort invested, school attendance, nutrition conditions, genetics, etc. [Becker, 1993, Ermisch and Francesconi, 1997, 2000]. However, these factors interact in multiple ways and their complex interaction is not fully understood, and in any case, separating and measuring the cause and effect of each of them accurately is impractical. To overcome this some studies in cognitive science have used textual literacy — the ability to read and write at the time of study — as a proxy for literacy, instead of level of formal education. Examining memory decline among ethnically diverse elders [Manly et al., 2003], researchers found that textual literacy more accurately reflects the quality of the educational experience at the time of testing. In addition, textual literacy could be a more accurate reflection of native ability because it does not assume that all individuals get the same amount of learning from a certain grade level.

Following up on studies in cognitive science, HCI4D researchers have also used degree of textual literacy at the time of the study as a proxy for overall quality of education of study participants [Medhi et al., 2013a,b]. These studies have found test takers to have reading — writing skills equivalent to or less than those of Grade VII text books from local government-run schools. People who passed the Grade VII level tests could read short passages in the local language, but people who did not pass the Grade I test were not able to identify even the basic alphabet. Many of the people who passed the intermediate levels were able to read isolated words but only with great difficult and effort. However, most people had some form of numeracy: were able to read up to 3-digit Indo-Arabic numerals, e.g., 0,1,2,3, etc.

(a) Difference in cognitive abilities:

Most studies in HCI4D have focused on non-literacy solely as the inability to read. And it is only recently that researchers have started taking into account differences in cognitive skills while designing UIs for low-literate users [Medhi et al., 2012a,b, 2013a,b]. But we discuss this later in the emerging trends section. In this section, we review studies in cognitive science in developed countries (e.g. United States and Netherlands) that have examined differences in cognitive abilities among low-literate populations.

There have been studies in the cognitive sciences that support the hypothesis that formal education is correlated with general cognitive skill development. Studies have suggested that much of the causal pathway between IQ and schooling points in the direction of the importance of the quantity of schooling one attains (highest grade successfully completed) [Ceci, 1991].

One study looks at how literacy when introduced to a "primary oral culture" (which has never known writing), can have extremely wide-ranging impacts in all areas of life of that culture, including economics, politics, art, and more (Ong, 2002). Oral cultures require strategies for preserving information in the absence of writing, which include a reliance on proverbs, epic poetry, and stylized culture heroes. Writing makes these features no longer necessary, and introduces new strategies of remembering cultural material, which itself now changes.

In addition to the skills of reading and writing, educated people seem to acquire cognitive skills and strategies for efficient processing of information such as skills for: language processing, listening, fluency, understanding instructions, learning capabilities, visual organization and visual memory, mental spatial orientation, mental alertness, divided attention and general self-efficacy [Van Linden and Cremers, 2008]. Several other behavioural studies have demonstrated through empirical research that education level influences various cognitive skills by comparing literate and lower literate/non-literate test participants:

(a) language tasks (such as repeating pseudowords, memorizing pairs of phonologically related words compared to pairs of semantically related words, generating words according to a formal criterion) [Abadzi, 2003, Castro-Caldas, 2004, Castro-Caldas et al., 1998, Kurvers, 2002, Manly et al., 2003, Morais et al., 1979, Reis and Castro-Caldas, 1997].

(b) general self-efficacy (the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations) [Bandura, 1977, 2005, Czaja et al., 2006].

(c) visuospatial and visual organization tasks that required abstraction, specifically the comparison, discrimination, and grouping of objects by picking their features according to abstract semantic categories [Luria, 1974]. Other visuospatial skills such as figure copy of a cube, house, Rey-Osterrieth complex figure, etc., and construction of figures with varying degrees of complexity related to rotation, distortion and disarticulation. [Ardila et al., 1989, Matute et al., 2000, Reis et al., 2001]. There is also naming of two-dimensional representations of common, everyday objects in terms of accuracy and reaction times [Reis et al., 2001]. Abstract icons have been known to be less recognized by participants with limited education — they possibly have difficulty integrating details of 2D line drawings into meaningful wholes [Castro-Caldas, 2004].

Most of the above work was undertaken in developed regions — North America and Western Europe — and therefore, is subject to caveats of cultural bias that may differ in other geographies. Nevertheless, the strength of the evidence suggests that formal education can shape cognitive skills beyond the mere ability to read and write. If anything, in environments where standards of education are even poorer, one might expect differences in cognitive skill arising from educational quality to be even more pronounced.

There has been limited work among developing world communities, to understand the impact of literacy on cognitive processes [Akinnaso, 1981, Bernardo, 2000]. In his essay *Literacy and Individual Consciousness,* Akinnaso used first-hand experiences growing up in his native small village in Nigeria to discuss how literacy impacts not just reading and writing, but also cultural traditions, linguistic behavior, socio-economic organization, cognitive processes, and child development. He discusses two main views of literacy: causal and facilitating. Literacy as a causal agent means that literacy itself creates new cognitive capacities. Literacy as a facilitating agent means that literacy makes it easier to acquire these cognitive skills. Meanwhile, Bernardo studies five rural and urban low-income communities in the Philippines and compares their performance on a number of cognitive tasks involving conceptual organization, conceptual categorization, conceptual comparison, deductive reasoning and explanation. This is done through quasi-experimental approaches and by drawing upon an ethnographic study.

Finally, relevant to this monograph one study [Sherwani et al., 2009a] finds the previously mentioned orality theory (Ong, 2002) to provide a unique lens to understand low-literate users. The study discusses various examples of HCI projects from the developing world, particularly India and Pakistan, and finds that low-literate users and cultures do not learn as well from a neutral, stand-alone object, such as a book, or an automated system, which contain a set of abstract instructions to be applied across situations. Rather they learn better in situ, embedded in concrete situations and practical experience. The study then synthesizes and recommends a framework that provides guidelines for design and evaluation of HCI interventions in developing world contexts. The methodological recommendations include avoiding Likert scales, which can be abstract for low-literate users, and motivating and contextualizing the system using concrete examples.

In addition to the challenges of low-levels of formal education, textual non-literacy and differences in cognitive abilities discussed above, there are a number of socio-cultural challenges that researchers in HCI4D have encountered while doing user studies among low-income, low-literate communities. These challenges may or may not be directly attributable to low-literacy per se, but literature in the design of UIs for low-literate users has discussed these at length. We elaborate on a few recurrent themes in the section below. We also discuss some of the techniques used in addressing these challenges.

2.1 Challenges in conducting user studies and techniques to overcome

Beyond lack of formal education, textual non-literacy and differences in cognition, there are a host of nuanced socio-cultural issues that seem to mediate how low-literate users interact with an ICT. Most low-literate people studied in HCI4D typically live in low-technology contexts without much exposure to computing technology in their living environments. Researchers have studied populations in both the rural [Cuendet et al., 2013, Mudliar et al., 2013, Parikh et al., 2003]) and urban context [Chakraborty et al., 2013, Medhi et al., 2013a, 2011, 2013b]. The household income of populations studied in a number of projects has typically been between USD 80–200 per month. People in these communities had informal sector jobs: domestic workers, farmers, daily wage labourers, plumbers, carpenters, construction workers, mechanics, fishermen, fruit and vegetable vendors. Their living environments are characteristic of housing built of low-cost building material. Living spaces are small, of about 100–150 sq. ft. for a household size of 6 or 7 people. General infrastructure like roads and sanitation conditions are poor and resources like water and electricity are in short supply. Nearly all households in these communities had television sets, and over half of them had some video playback device (VCDs and DVDs). A large section of men in the communities owned low-end limited functionality phones, although "China phones" with larger screens and higher multimedia capability, and some low-end Android phones was becoming more common. Compared to men, relatively fewer women owned mobile phones. None of the people had any previous experience using computers.

In addition to income and education, there could be differences that stem from variations in ethnicity, age and other socio-demographic characteristics between researchers and the communities they work with. All of this gives rise to a number of challenges while conducting user studies and we highlight a few here. While some of these are a problem in standard user studies even in developed world contexts, they seem to be much more pronounced in the developing world:

(a) Social intimidation and low self-esteem

In user studies researchers have observed that in spite of ensuring that participants feel comfortable they are often intimidated that it is they that are being tested and not the technology. This causes nervousness and anxiety during usability tests [Medhi et al., 2010]. Or the fear that they would break the technology if not handled properly and be taken to task for it. The more expensive the device appears, higher is the level of discomfort in using that technology. In fact it has been found that users often blame themselves for problems while using a technology (than blame it on the technology itself). This is especially true of older subjects in India, who appear to have low confidence on a technology that is new to them.

Usability studies in standard HCI are typically conducted in controlled lab environments. In HCI4D to help overcome social intimidation among participants, researchers have conducted usability studies in physical settings which are routine for the participants, e.g., the participants' own homes, a community space, in the village or slum area that participants are familiar with, offices of non-profit organizations that work among these communities, etc.

To make study participants comfortable with ICTs that they would be ultimately tested on, researchers have adopted various field improvisations. In one study, women participants were shy to hold the stylus of a Tablet PC being used to test an employment search application [Medhi et al., 2005]. To help overcome the reluctance, before using the app, participants were asked to draw something they were very familiar with on the Tablet — a rangoli using MS Paint. Rangoli is a traditional Indian decoration and pattern made with ground rice and other colors at the entrance of a house particularly during festivals. The women participants warmed up to this idea and created their designs, the researchers were thus successful in removing some bit of the fear and reluctance in using the Tablet PC.

(b) Participant response bias

Given the increasingly common situation in which researchers have higher social status and social power than the people they investigate, study participants may be particularly susceptible to a type of response bias known as demand characteristics. Demand characteristics refer to aspects of a study that may convey the researcher's hypothesis to participants who then adjust their behavior in relation to what they perceive to be the researcher's expectations [Orne, 1962]. In an HCI4D setting one quantitative analysis of participant response bias due to interviewer demand characteristics found that respondents are about 2.5x more likely to prefer a technological artifact they believe to be developed by the interviewer, even when the alternative is identical [Dell et al., 2012]. When the interviewer is a foreign researcher requiring a translator, the bias towards the interviewer's artifact increases to 5x. In fact, the interviewer's artifact is preferred even when it is degraded to be obviously inferior to the alternative. This study was done through 450 interviews in Bangalore, India.

Apart from this study there are others that have discussed anecdotal evidence for demand characteristics in similar settings. In one field study reflections paper some users were eager to please researchers, both because of their perceived difference in status and because of their curiosity about the various gadgets tested [Anokwa et al., 2009]. The study noted that regardless of whether the users were actually listening to the foreign researcher, understanding what she was saying or anything, they would always tell her exactly what they thought she wanted to hear, which was generally a "yes".

Researchers have recommended paying attention to the types of response bias that might result from working with any participant population and dissociating themselves as much as possible from any particular design or solution [Dell et al., 2012]. Obtaining factual, rather than subjective information has been recommended for user studies [Bernhart et al., 1999]. It has been said that using implicit metrics [Czerwinski et al., 2001] or triangulation [Mackay, 1998] to validate the data collected could further increase confidence in the results of a study.

Particularly for the HCI4D context, researchers have used the "Bollywood" method [Chavan, 2005] to motivate test participants toward desired tasks and encourage honest feedback. In this method usability tasks are embedded in dramatized stories involving the test participant. Particularly in a South Asian context, where participants tend to be reserved about giving feedback to people they perceive to be in authority, this has been found to be a useful method.

(c) Communication barriers: language and culture

More often than not, researchers in HCI4D do not speak the language of the communities they work with. To bridge this gap local translators and facilitators are used during field interviews. Researchers have noted that if not trained properly, a lot might be lost during translation as translators often tend to rephrase and summarize participants' comments [Squires, 2009, Twinn, 2008].

Problems due to language may also extend to the ICT intervention itself. Researchers have observed how the official language of a region was not the optimal choice for an IVR health information system [Sherwani et al., 2009b], as study participants spoke other regional languages. Other studies have found that even languages suggested by the local partner NGOs for the intervention turned out to be sub-optimal [Cuendet et al., 2013, Sherwani et al., 2009a]. Participants either spoke a local dialect or used colloquial terms that the partner NGOs were not cognizant of while co-designing content for the application with the researchers.

In addition to language studies have noted various socio-cultural barriers in conducting research in low-income, low-literate settings. In one study in Pakistan researchers noted early on that all user studies would be conducted with at least one female facilitator to help female users feel comfortable in the presence of male researchers. [Anokwa et al., 2009, Sherwani et al., 2007]. In another study researchers noted how they were being perceived as "novelties from the big city", thus creating communication barriers with study participants [Parikh et al., 2003]. The researchers were most often seated in chairs in the center of a large circle, with the women sitting around on the ground. This made it a little bit difficult to have an equal exchange of ideas. It was observed that a deep immersion in the community would have to be done to make the target community feel at ease. Researchers in HCI4D have typically relied on intermediaries, such as non-profit organizations, to gain access into communities and overcome communication barriers with them.

(d) Social and peer pressure

Low-income, low-literate community members, especially in urban contexts, reside in densely populated areas. Much of the field studies are conducted in open, public spaces in these areas, where every spectator might be free to participate. This creates challenges for the researcher in conducting interviews without interruptions. One field study reflections paper notes how when interviews were being conducted, outspoken onlookers demanded to know what the study was about, what the motivations were, or what was in it for them or their community [Anokwa et al., 2009]. These kind of situations not only cause interruptions during the interviews, but also could influence the behavior of the participant being interviewed.

In another study female users of a graphics+speech multimodal video search phone system were shy, and did not dare to speak into the phone while using speech input. And when they did, they did not speak loudly [Cuendet et al., 2013]. One woman who had used the system noted that she felt shy of speaking because it was her in-laws' house where the interviews were being conducted, and they were listening to her speak while waiting outside the room.

In such situations it is best to conduct the interviews one-on-one with no onlookers around. While it is important to interview study participants in a physical setting they are familiar with, it is important to do so in a place that is secluded and free from interruption from other members of the community.

In the above section, we observed a number of socio-cultural challenges that researchers in HCI4D have experienced while designing UIs for low-literate, low-income communities. We also discussed techniques that have been used to overcome some of these challenges. In the following sections, we will look at how existing work in UI design for low-literate users are helping address the challenges of limited formal education, low-textual literacy and difference in cognitive abilities.

3

Examples of UIs for Low-Literate Users

To counter problems associated with limited education and non-literacy, HCI4D researchers have proposed doing away with text, and have instead advocated non-textual designs. It has been found that common text-based interfaces were completely unusable for low-literate users [Medhi et al., 2007], and were error- prone for literate, but novice, users [Medhi et al., 2011]. And to counter these problems, UIs that use voice, graphics, and video have been proposed. The different modalities of these non-textual UIs offer different affordances to the low-literate user. Researchers have looked at the following input/output modalities:

Input

Voice/Speech

Researchers have studied speech or voice mostly because it is a natural means of expression well-suited to input, and avoids some of the issues related to non-literacy. Speaking comes naturally to people and thus studies have leveraged the use of local language in speech. Speech interfaces also avoid many typed language-related complications, such as the absence of keyboard standards and unique fonts/scripts (or, in the case of some languages, a lack of any written script) [Boyera, 2007, Huenerfauth, 2002]. However, using speech as an input mechanism is technically challenging and costly because of the large amount of data needed to train an automatic speech recognizer (ASR) with an acceptable word error rate [Moore, 2003]. These data typically are not available for languages in developing countries, and collecting them would be a huge effort with small impact because of the variations in dialect and accent [Plauché et al., 2006]. Training a single ASR for a given language/dialect/accent requires many hours of manually annotated speech.

Research has shown that by limiting the vocabulary to less than 100 words, one can develop a reasonable speech-based system for some applications [Qiao et al., 2010, Sherwani et al., 2007]. These researchers developed SALAAM, a method that requires only a fraction of the training data as traditional ASRs to create a speech recognizer. It allows small-vocabulary (about 100 words or fewer) recognition by using the acoustic model of any existing ASR and performing cross-language phoneme mapping between the language of the ASR and the target language. IVR systems using SALAAM have been designed, deployed and tested for health data collection among community health workers [Sherwani et al., 2009b] and a voice manipulation and forwarding system for spreading job-information through entertainment [Raza et al., 2013]. While ASRs are much cheaper and easier to create using SALAAM, the method does have some limitations: it only works for small vocabulary and only allows one vocabulary item to be recognized at a time. These limitations do impact the natural interaction promised by speech; free speech is not possible, and one must shape the interaction for short utterances.

Given this constraint, researchers have studied how speech could be used in combination with graphics and touch for a better experience for novice or non-literate users. VideoKheti, a multimodal interface combines SALAAM ASR with graphics and touch on a smart phone designed for agriculture video search [Cuendet et al., 2013]. The study was conducted among a novice rural population using the SALAAM ASR for local language/dialects. Results showed that although the speech interface was not able to help low- literate participants overcome non-literacy related problems in using the system, it worked well for cases where there was a long list of choices and selections comprised of short and familiar words or expressions. Researchers also did a related study to show how with very little training data VideoKheti was able to achieve >90% accuracies for test and field deployments [Bali et al., 2013]. In another study, the original, pure speech interface was augmented with touch, graphics, audio and a small amount of text to accommodate unsophisticated or new users [Plauché et al., 2006]. It was found that the though the speech recognition was poor, participants reported that the interface was easy to use. Participants with the lowest literacy preferred to listen to the system for several minutes before speaking to it.

Then there are other speech systems that exploit the ubiquity of inexpensive mobile phones. Avaaj Otalo, for example, allows users to ask questions and browse others' questions and answers on agricultural topics through a simple mobile-phone call [Patel et al., 2010]. The system is controlled by simple speech prompts (or by numeric inputs discussed later) and is used to collect information about farmers' harvests in the state of Gujarat, India. The ASR was adapted from American English to be usable by Gujarati speakers. Another example is Spoken Web that attempts to create a secondary audio version of the web, accessible by any phone [Kumar et al., 2010]. The framework presented allows users to create "voice-sites" by means of voice interaction. These voice-sites can then be accessed by a phone call. There have been several demonstration applications on Spoken Web, one of which was to provide farmers with crop and market information. These systems demonstrate the utility of voice-based systems for providing information to low-literate farmers in rural India.

Finally there are Wizard-of-Oz studies where a speech interface was compared with a graphical+audio interface. It was shown that although the graphical interface led to a higher completion rate, users who understood the speech system were able to complete tasks faster [Medhi et al., 2011]. Words on the speech interface were short, and some users found it easier to determine which word to speak on the speech interface than identifying what button to press on the graphical UI. These users had prior experience using IVRs for mobile phone recharge.

Touch

Touch in this section refers to direct interaction that enables computers, phones and tablets to respond to the touch of a finger on the screen. Researchers in HCI4D have explored the use of touch to enable a better and more natural experience for first-time, low-literate users. Previous research showed that low-literate users experience challenges when faced with soft-key mapping on devices with keypads [Medhi et al., 2011]. Soft keys are usually unlabeled keys, often appearing directly below the screen on mobile phones, that have different functions in different contexts. They could also be numeric keys when used to choose from an enumerated list on the screen. Soft keys were difficult to understand because they require mapping to the changing functions displayed on screen.

To overcome issues with mapping, researchers have explored the use of touch that enables direct interaction on the screen. Researchers on VideoKheti used touch on smartphones, in combination with graphics and speech, to enable easy interaction among farmers to do agriculture video search [Cuendet et al., 2013]. On every page, the system asked users what information they wanted to know about and explicitly named all choices available. A choice uttered by the system was simultaneously highlighted on the graphical interface. Once all prompts were done playing, a first touch on an item highlighted it and played the corresponding audio. A second touch validated the choice and the system navigated to the corresponding page. Results showed that more than half of attempted interactions and three quarters of successful interactions among low-literate users were via touch compared to speech.

Researchers have also used touch input in comparing the performance of non-literate users on a search task between a hierarchical interface and a multipage list design. The hierarchy was four levels deep and the multipage list design was spread across seven pages [Medhi et al., 2013a]. Results show that participants using the multi-page list performed better even when the list UI design required browsing through multiple pages of items on the phone. It was further found that test participants, despite having no previous experience using touch-screen phones, seemed fairly comfortable moving about the pages using the forward and back buttons.

One study looked at enabling non-literate users to use text-messaging in conjunction with audio, text, and visuals through a touch screen interface [Friscira et al., 2012]. It showed that using touch-screen phones did not represent a cognitive problem for non-literate users, but only a problem in terms of lacking confidence or technological literacy. And this can be overcome and users are able to use the system with initial training.

Besides the above, there are examples focused on using touch interaction in the education context, eg. the new XO-4 OLPC laptop (http://one.laptop.org/ retrieved July 11, 2014) [One Laptop Per Child], the computer used in the "Hole-in-the-wall" studies [Mitra, 2003]. But these are targeted at children and are thus outside of the scope of this monograph.

Others (Digital pen, DTMF)

Researchers in HCI4D have explored other input techniques that leverage natural interaction people already practice in their everyday lives, and have mapped these on to new technologies. One such example is using pen as input device. The Milpa project leverages the familiar pen–paper notebook interaction and maps that on to digital slate technology [Ratan, et al., 2010]. The slate accepts handwritten input from a digital ballpoint pen on ordinary paper notebooks placed on the digitizing pad. Immediate electronic feedback is then provided on the interactive touch screen display. The back of the pen also serves as a stylus for touch screen input. This allows for a seamless transition from older, paper-based systems to digitally recorded information. Milpa has been piloted in the context of microfinance [Ratan, et al., 2010] and health data collection [Medhi et al., 2012a,b] among frontline workers in rural India.

PartoPen, an interactive digital pen-based system, was designed to work with an existing paper-based labor monitoring system, the partograph [Underwood et al., 2013]. It uses customizable software written for the digital pen to capture and synchronize audio and handwritten text. The handwritten notes are then digitized into searchable and printable documents. By adding new technology to an older system, PartoPen improved maternal and fetal health outcomes while maintaining the continuity of an older paper system.

Other examples of leveraging familiar interactions include using DTMF or touch-tone as input modalities. Low-literate users are numerate and are comfortable using the numeric keypad on their mobile phones [Medhi et al., 2011]. DTMF or touch-tone input allows users to press numbers on a simple phone call so as to access and navigate an IVR system. Examples of IVRs aimed at low-literate users include: (a) Avaaj Otalo, a Q&A forum for small-scale farmers to ask and listen to agriculture queries [Patel et al., 2010], (b) CGNet Swara, a citizen journalism portal where rural users can report and listen to news stories of local interest [Mudliar et al., 2013], (c) Polly, a voice manipulation and forwarding system to virally spread job information through entertainment [Raza et al., 2013, 2012], and Gurgaon Idol, a talent competition in which community members can call to record their songs, and vote to select the best songs [Koradia et al., 2013]. All of these are discussed more in detail in Section 4.

Output:

Graphics

A number of early UI studies aimed at low-literate users recognized the value of imagery and advocated extensive use of graphics to help overcome the inability to read text [Grisedale et al., 1997, Huenerfauth, 2002, Medhi and Toyama, 2007, Parikh et al., 2003]. While all of these papers emphasize the importance of user-centeredness in graphics, the first one uses stick-figure style icons in health information data collection among rural health workers [Grisedale et al., 1997], the second discusses the trade-off of different styles of graphics [Huenerfauth, 2002], and the third 'text-free UIs' uses hand-drawn style images in a job information dissemination application for domestic helpers [Medhi et al., 2007]. The fourth study also uses representational identifiers such as icons and images but in a financial management system for rural microfinance [Parikh et al., 2003].

Text-Free UI studies have compared five different representational styles — local language text, static hand-drawn graphics, photographs, video and animation — all with and without audio output in the domain of health information dissemination [Medhi et al., 2007]. And through a 200 people usability study established that static, hand-drawn representations are better understood than photographs or icons by low-literate users. It was found that photographs have extraneous details that can confuse users and icons can be too abstract for these users. But at the same time there is also work that talks about how graphics can have more photorealism with deeper interaction, as the specificity of the information increases [Medhi et al., 2007].

More recent work has also reinforced the use of graphical icons and pictures for low-literate users across a variety of domains: 'Easy-Texting' for enabling use of text-messaging in conjunction with touch, text and audio [Friscira et al., 2012]; 'Igwana' to navigate large and complex data sets found on the Web [Bhattacharya and Feldman, 2012]; smartphone applications in agriculture [Agrawal et al., 2013]; 'Karaoke', an assistive technology to help learn the alphabet and early literacy skills [Dew et al., 2013]; 'Parichaya' for medication adherence among tuberculosis patients [Seth and Sorathia, 2013]; 'WATER alert!', a water delivery alert and quality reporting system for better citizen involvement [Brown et al., 2012]. Graphical icons and pictures have also been used in comparison studies of GUI widgets and navigation styles in fluid and nutrition monitoring among chronically-ill patients [Chaudry et al., 2012].

Within graphics for UIs low-literate users, there is work that has looked at the use of color in imagery. Researchers have designed an icon and color based visual phonebook, 'Rangoli', for non-literate people and have showed that colors could help them sort and identify contacts, though within a limited range [Joshi et al., 2008]. Another group of researchers also designed a phonebook with colors and icons and got preliminary success with novice users in managing their contacts [Wiedenbeck, 1999].

Audio

In UIs for low-literate users, audio as output has been used either in and of itself, or in combination with another modality, such as graphics or text. A number of previous works have underlined the benefits in combining audio annotation with graphics, where the pre-recorded audio is used to explain what the graphic depicts. 'Text-Free UI' studies have discussed how low-literate domestic helpers were thrilled to hear the computer 'speak' in their native language while using a text-free job-information application [**?**]. Other examples of using a pre-recorded local language audio in combination with graphics include applications such as: a financial management system for micro-credit groups [Parikh et al., 2003]; 'VideoKheti': a mobile video search system for agriculture best practices [Cuendet et al., 2013]; 'Parichaya': medication adherence system among tuberculosis patients [Seth and Sorathia, 2013].

Audio output has also been used in combination with text to help low-literate users do text-messaging [Friscira et al., 2012]. The 'Igwana' system allowed non-literate users to listen to received SMS and compose text messages by augmenting words with touch-initiated text-to-speech support. Audio as output has also been used in assistive technology in combination with text, to help learn the alphabet and early literacy skills [Dew et al., 2013]. The on-demand playback uses text-to-speech and can be controlled by gestures, leaving their use to the user's discretion so as to avoid social stigma.

In addition to the above, there are a number of examples of IVR systems that use prerecorded or dynamically generated audio to direct users on how to proceed within the system. Examples of IVRs aimed at low-literate users include: Avaaj Otalo, a Q&A forum for small-scale farmers [Patel et al., 2010]; CGNet Swara, a citizen journalism portal [Mudliar et al., 2013]; Polly, a voice manipulation and forwarding system [Raza et al., 2013, 2012], and Gurgaon Idol, a community talent competition [Koradia et al., 2013]. All of these are discussed more in detail in the next section in Section 4.

Video

Video output in UIs aimed at low-literate users has been used for both information delivery as well as exchange. Researchers have looked at 'Full-Context video' for information on how to use text-free UIs [Medhi and Toyama, 2007]. Following previous work it was observed that in spite of users' understanding of UI mechanics, they experienced barriers beyond non-literacy in interacting with the computer: lack of awareness of what the PC could deliver, fear and mistrust of the technology, and lack of comprehension about how information relevant to them was embedded in the PC. All these challenges were addressed with full-context video, which included soap opera-style dramatizations of how a user might use the application and how relevant information comes to be contained in the computer, in addition to a tutorial of the UI.

Video output in information delivery also featured in projects such as VideoKheti, a mobile video-search system [Cuendet et al., 2013], and Digital Green, mediated video instructions for small and marginal farmers [Gandhi et al., 2007]. However in VideoKheti, video features only as content and is not integral to the interaction of the UI. And Digital Green uses video for information delivery on agriculture best practices, but only in a broadcast mode through handheld pico-projectors. Also, direct interaction with the video is not aimed at low-literate users per se, but for relatively literate agriculture mediators.

Video has also been used for asynchronous exchange of information in UIs aimed at low-literate users. The 'MOSES' project looked at an interactive computer kiosk system, for use in Liberia's post-conflict reconciliation effort [Smyth et al., 2010a]. The system allows users to browse a collection of videos recorded by previous users, watch them, and then record their own videos, to be stored and viewed by subsequent users. The kiosk is housed in a rugged, self-contained booth such that it can be easily disassembled and transported.

Another study looked at the design and evaluation of an asynchronous peer-to-peer communication application [Prasad et al., 2008]. After considering different message formats it was found that video+audio was the most viable for low-literate users. The other formats included text, freeform ink, and audio. User studies showed that users were able to grasp the basics of the application and with the help of an onscreen audio assistant, complete the given tasks, such as record a video message for a family member. Most importantly, they were able to do so even after a 10-day break from the initial demonstration by the experimenter.

Other Design Features:

Studies have proposed a host of other features for making UIs for low-literate users more accessible. Non-literate users, as previously mentioned, have been known to be numerate. A number of studies have therefore proposed the use of Indo-Arabic numerals, "O", "1", "2", etc. even in non-textual UIs. Numerals have been used to denote wages in a domestic job-information dissemination application [Medhi et al., 2007], account information in rural microfinance applications [Ratan, et al., 2010, Parikh et al., 2003], etc.

Secondly, it has been found that the use of voice-based help instructions allows an application to be more autonomously used [Medhi et al., 2007]. An on-screen character could be optionally placed so that users can put a visual figure to voice playback. Another study used an onscreen audio assistant, in the form of a woman dressed as an Indian post-person, to provide context specific voice help in an asynchronous peer-to-peer communication application [Prasad et al., 2008]. This audio assistant remained on screen at all times to provide a consistent place where users could turn to for help at any time. The help was found to be a constant source of reassurance to users; it was observed that users were significantly more confident and needed less prompting. Congratulatory audio messages on task accomplishment also seemed to produce excitement and encouragement.

Researchers have also suggested minimizing use of scroll bars, as non-literate users did not realize that there were functions "beneath" what was displayed. Explicit demonstrations were required to teach these users what scrollbars were and how to use them [Medhi et al., 2011]. One study proposes the replacement of vertical scrollbars with much larger up/down arrow icons at the top/bottom of lists as required [Prasad et al., 2008].

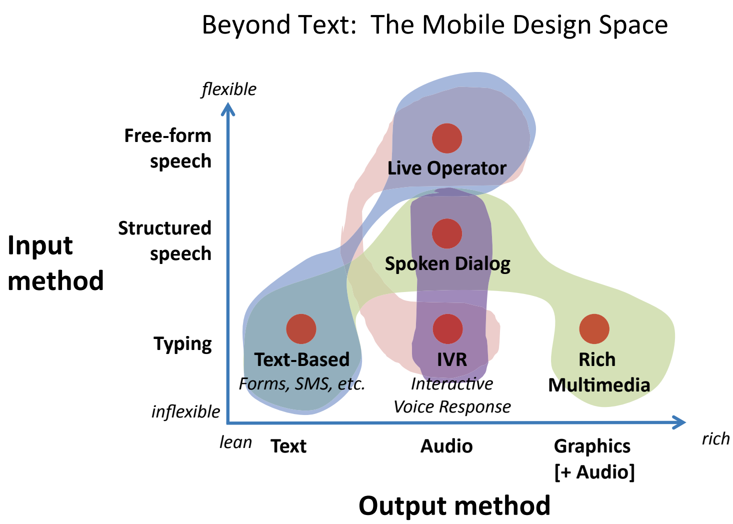

Figure 3.1: Beyond text: the mobile design space.

Finally there are studies that have questioned the suitability of menu-based navigation for novice users [Jones et al., 2000, Jones and Marsden, 2005], and have discussed designs that advocate fewer menus and dedicated buttons for this target group [Kurvers, 2002].

Apart from the above studies, there are others that compare various *combinations* of input–output modalities for exploring alternative UIs for low-literate users. Classifying mobile interfaces along two axes, according to the flexibility of their input and output modalities, one can arrive at Figure 3.1. In the recent years, researchers have started exploring rigorous comparisons between various points in this space.

Spoken dialog vs. IVR

Researchers have examined the tradeoffs between IVRs (spoken menu output with keypad input navigation) and spoken dialogue systems (SDS) (spoken menu output with speech input navigation) and have reached varying conclusions regarding the benefits of typed versus spoken inputs. The studies have varying results: one reports that task completion is higher with keypad inputs than with speech inputs [Patel et al., 2009]; one reports comparable task completion but a preference for dialed inputs over speech inputs [Grover et al., 2009]; and one reports that speech inputs significantly out-perform dialed inputs [Sherwani, 2009]. The first study [Patel et al., 2009] was conducted in the context of Avaaj Otalo, a Q&A forum for small-scale farmers in India to ask and listen to agriculture queries. Users expressed discomfort speaking single word commands on the SDS and "talking to the computer" was perceived as unnatural. Additionally, there was difficulty in recovering from speech recognition errors made by the system or bad or no input by the user. The second study [Grover et al., 2009] was designed for provision of health information to caregivers of HIV positive children in Botswana. Users reported that they found the IVR system with dialed input to: have clearer instructions, be faster to use and be a more private experience. It was also found that experience loading airtime on a mobile phone was the sole significant factor in user preference of IVR over SDS. The third study was conducted in the context of Healthline, an information access system to be used by Community Health Workers in Pakistan [Sherwani et al., 2007]. Results showed that a well-designed SDS could significantly outperform an IVR system for both low-literate and literate users. For "in-grammar utterances" (i.e., when the user said something that the recognizer should have been able to recognize), recognition accuracy was 91%. Literate participants reported that they had to remember less with the SDS interface and preferred it over IVR. While low-literate participants reported that the SDS interface was harder to use, and that they preferred the IVR, in reality *they performed better with the SDS interface on average.* Turned out that when participants found the system's language difficult to understand, they were hesitant to speak at all on the SDS interface after few attempts for fear of exposing their confusion publicly. Though when given the IVR interface, they still attempted to press buttons since it could be done more discreetly.

Text-based form vs. SMS vs. Live operator

Researchers have also experimented with another very interactive interface: that of a live operator, in which users accomplish tasks by talking to an actual human over the phone. This interface was compared in one study to textual interfaces, based on either electronic forms or SMS, conducted in the domain of mobile health data collection [Medhi et al., 2011, Patnaik et al., 2009]. The education level of the participant health workers was between 10 and 12 years though they had limited exposure to technology. Unlike many other studies, results were reported in terms of the error rate in data entered, rather than the rate of task completion. It was found that the live operator was 10x more accurate than the electronic forms and SMS; there was only one error for the live operator during the whole trial. Though while the operator interface offered the lowest error rates, it also led to the longest entry times; 1.43x higher than forms and SMS. Another result was that the overall error rates observed for SMS were not significantly higher than that of electronic forms.

In spite of the live operator interface outperforming the forms and SMS interfaces, health workers reported in follow-up qualitative interviews that they preferred the forms and SMS. This was perhaps due to poor phone connections experienced during the trial, the study explains. Many people found conversations with the operator to be frustrating due to the bad quality of the call.

The above results suggest that for domains where the accuracy of information is critical, and at locations like India where setting up call centers is affordable, a live operator interface might be viable provided the quality of connection is good. The study also suggested that it might be viable to consider a simple SMS reporting system in cases where it is too complex or costly to develop an electronic forms solution.

Text-based app vs. Spoken dialog vs. Graphics+audio UI

Researchers have experimented with media rich interfaces, comparing trade-offs with cost-effective UIs [Medhi et al., 2009, 2011]. Rich interfaces can display graphics and audio, making them user-friendly for low-literate users, but they also come with their own drawbacks, including greater complexity, greater cost, and less platform universality. This particular study explored how low-literate participants in India reacted to three different UIs that made tradeoffs between cost and richness:

(1) a local language text-based UI, (2) a Wizard-of-Oz spoken dialogue system, and (3) a text-free graphical UI, in which users press keys to navigate voice-annotated images on their mobile phone. The study was conducted in the context of a simulated mobile banking system involving money transfers and account balance inquiry.

Results confirmed that the text-based UI was unusable by low-literate users, and that task completion was highest with a graphical interface. But those who understood the SDS could use it more quickly due to their comfort and familiarity with voice interfaces doing airtime loading previously. On the graphical UI, the graphics+audio output were easy to understand, but typing as input was scary to some users. On the SDS, spoken directions as output were sometimes difficult to understand, but speech as input came naturally to these users. Based on this, as part of future work, the study suggested a combination of speech input and graphics+audio output for UIs for low-literate users. The VideoKheti, mobile video search system was built on these recommendations [Cuendet et al., 2013].

Touch+Graphics+Audio vs. Speech+Touch+Graphics+Audio

Following up on recommendations from the study mentioned in the previous sub-section, researchers have experimented with Videokheti, an interface with graphics+audio output combined with speech input on touch screen phones. Speech input was done for a limited vocabulary (less than 100 words) using the SALAAM system's [Chan and Rosenfeld, 2012, Qiao et al., 2010, Sherwani, 2009] cross language phoneme mapping. This interface was compared with a system with graphics+audio output, without the speech part, again on touch screen phones. The study was conducted in the context of mobile search for agricultural videos for small and marginal farmers in India. It was found that while farmers could use the VideoKheti system, success greatly depended on their education level. Younger farmers with about Grade X education used the system much more successfully than farmers with less than Grade III education. While participants were enthusiastic to talk to the system, the speech part of the system was not able to help overcome issues related to low-literacy. The difficulties encountered by low-literate users included problems with understanding and remembering task scenarios; vocabulary comprehension and reproduction; and even understanding the hierarchical organization of information. It seemed that the potential benefit of adding speech as input were outnumbered by the cognitive overload it brought along for users to deal with. But speech was useful where there was a long list of items in the UI and user selections comprised of short, familiar words or expressions, e.g., at the first level on the UI when the user chose one out of 22 crop names. This result, obtained with a touch interface, is in line with what Avaaj Otalo [Patel et al., 2010] had hypothesized without being able to demonstrate, when comparing speech input on SDS with touchtone input on an IVR system.

IVR vs. Live operator

Finally, researchers have also compared IVRs with a live operator system, in terms of accuracy, speed, and cost of conducting a job survey among low-income job seekers in India [Chakraborty et al., 2013]. The IVR interview included a brief introduction by a live operator, to provide context for the call. This was then followed by prompts for questions and confirmations of responses. It was found that out of the people who completed both surveys, overall the IVR system compared favorably with the live operator, incurring only 4.0% error rate. The errors on the IVR system were incurred during free-form oral responses and multi-digit responses. There were no errors observed during the multiple-choice questions. The 4% figure was comparable to error rates reported previously for mobile data collection via SMS (4.5%) and electronic forms (4.2%) [Medhi et al., 2011, Patnaik et al., 2009]. However, unlike the previous study, the IVR system did not require training of participants, making it suitable for a large-scale deployment. With regard to speed it was observed that the IVR system required 2.5 times longer than the live operator interface, though this was due to the conservative slow pacing of prompts on the IVR system.

Given current costs in India, this implied that the IVR system could offer modest reductions (1.5x), not taking into account the cost

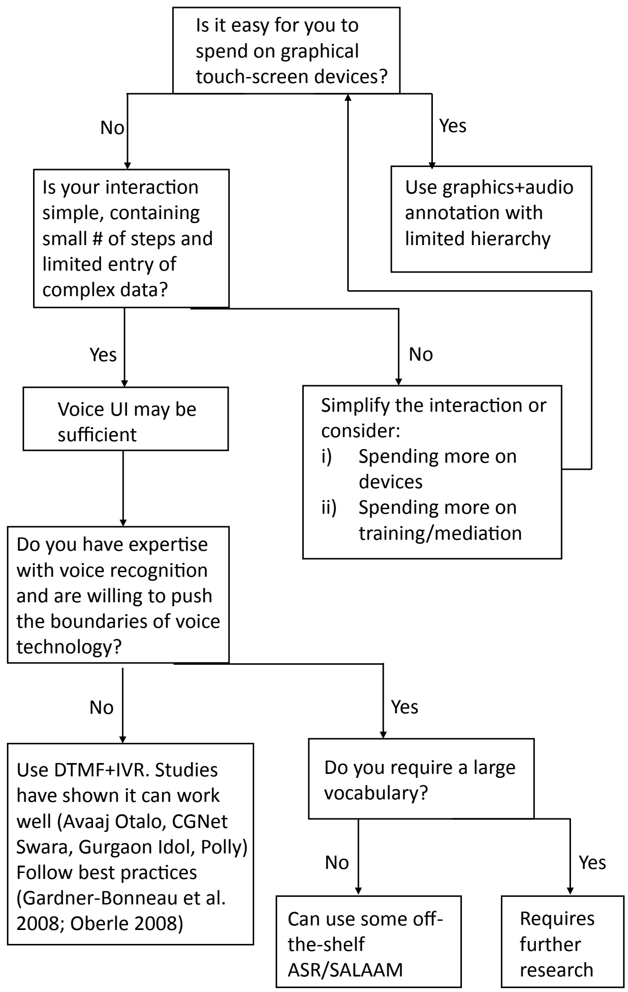

Figure 3.2: Flow chart of overall design recommendations.

of technology setup and maintenance. Furthermore the study indicated that the IVR prompts could be recorded in many languages and dialects, making it more accessible to people who might have trouble understanding the accent of the operator.

In this section, we looked at existing examples of UIs from the perspective of different input/output modalities and the various affordances each of them allows to low-literate users. We also looked at comparative studies, studying different points of input/output in the mobile design space. Based on these current examples of work, we make overall design recommendations for UIs for low-literate users. We present this in the form of a flowchart in Figure 3.2.

4

Current Emerging Trends

In addition to helping low-literate users get around problems associated with inability to read, and realize useful interaction, currently there are other strong trends emerging in design of UIs. Some of these look beyond UI mechanics, and explore greater participation for the user than just consumption of content. We discuss a few of the trends here:

(a) Non-literacy is not just about the inability to read text, it is also correlated with cognitive skills that have implications for design

A number of previous studies in HCI4D considered non-literacy as the inability to read text and hence, as elaborated in the previous sections, recommended non-textual UIs that use a combination of voice, video and graphics. But as described earlier, research in cognitive science has shown that non-literacy is correlated with a host of cognitive skills, including those that are related to language, self-efficacy and visuo-spatial and visual organization skills [Abadzi, 2003, Ardila et al., 1989, Bandura, 1977, 2005].

One previous study shows that low-literate users have "less developed cognitive structures and linguistic sequential memory" when compared to educated users. And calls for attention to these "unorganized" structures when doing instructional design for rural e-learning applications [Katre, 2006]. This study, however, is a small-sample qualitative analysis that does not offer concrete evidence for difference in cognitive skills. Another work discusses how non-literate users might have difficulty in hierarchical navigation and understanding abstract categorization of information [Sherwani, 2009]. However, these observations are anecdotal and the claims are not investigated systematically.

More recently researchers in HCI4D have started systematic investigations on how non-literacy may be correlated with certain cognitive skills relevant for ICT interactions, and how these have implications for UI design [Medhi et al., 2013a,b]. The two specific cognitive skills studied were the ability to transfer learning from video-based skills training and the ability to understand and navigate hierarchical information architectures (hereafter referred to as IAs) [Medhi et al., 2012a, 2013a, 2012b, 2013b]. These skills were chosen for investigation because there was previous anecdotal evidence for low-literate users having difficulty with navigating hierarchies and while transferring learning from instructions to actual usage [Jones and Marsden, 2005, Medhi et al., 2011, Sherwani, 2009]. The correlational studies were conducted for concrete confirmation of previous anecdotes. They were carried out among people in India with less than Grade VIII formal education, although literacy tests were also conducted to assess reading–writing ability of test takers at the time of testing. We discuss these two correlational studies separately below:

Non-literacy and the ability to transfer learning from video-based skills training

Instructional videos are an increasingly popular medium for teaching people a wide range of skills and tasks. Examples of popular websites of instructional videos include howcast.com, e-how.com, and youtube.com. In ICTD, Digital Green has been successful in using video to teach agricultural techniques to farmers in rural India [Gandhi et al., 2007]. Also, full-context video instructions, with dramatized storylines, seem to work reasonably well for imparting information on how to use UIs to populations with limited education [Medhi and Toyama, 2007]. Another research reinforces this by showing that storytelling and narratives are indeed effective in video-based instructions for low-literate users [Ladeira and Cutrell, 2010]. However, even though video-based instruction is not dependent on reading, other cognitive skills may be required to comprehend the instructions demonstrated in the video, and to then transfer that learning to actual implementation in similar real-world tasks.

Transfer of learning in this study [Medhi et al., 2012a,b] was defined as the ability to transfer learning from specific examples of a task demonstrated in instructional video to actual implementation in circumstances similar to (in attribute or relationship), but not necessarily identical with, that shown in the video. Like the previous study [Ladeira and Cutrell, 2010], the video used was for operating vacuum cleaners for various cleaning and maintenance tasks. A controlled experiment compared 56 participants from low-income communities in India, split into groups of 28 based on a test of textual literacy used as a proxy for assessing overall educational level. 'Literate' participants included those who passed the cut-off condition on the test of literacy; 'non-literate' participants included those who did not. Participants were then rated for their ability to generalize video instructions on how to use a vacuum cleaner to similar, but not necessarily identical tasks. Results confirmed that both groups did worse on abstracted transfer learning tasks where the task to be implemented was different from that demonstrated in the video. On specific learning tasks, where the tasks in the video and the required implementation were the same, they performed better. Results also showed that literate participants did better than non-literate participants all-around on all tasks. In addition, it was found that diversification/generalization within instructions, i.e., providing more than one example of a vacuum cleaner, helped literate participants in transfer of learning, but did not measurably help non-literate participants.

In the case of the non-literate participants, the improvement due to the diversified examples was non-significant. The study suspected that more than one example in the video could be cognitive overload for some non-literate participants. Since another vacuum cleaner was shown in the video (but not used in the tasks) and yet had to be retained in memory for any potential use. Furthermore, it was observed that both literate and non-literate participants required less assistance for accomplishing functions with more similarities compared to those with fewer similarities, between the demonstrated video and actual tasks.

This study recommended that for users with limited education instructional video should demonstrate examples that are as close as possible to actual instances of the task. In other words, while designing how-to video manuals for the use of a mobile phone, a camera, or any other device, the instructions should be, as much as possible, for the same set of design features, keys and buttons layout and other specifications as the phone or the camera under consideration.

Non-literacy and the ability to understand and navigate hierarchical UIs

The same group of researchers have also looked at how non-literacy may be correlated with the ability to understand and navigate hierarchical IAs, and how that has implications for UI design [Medhi et al., 2013a,b]. Traditional computing software is structured in the form of hierarchical IAs to enable structured navigation of large information systems by concentrating on a few issues at a time. One of the principle benefits of hierarchical IAs is that space needed for navigation can be reduced by nesting. However, previous research has shown that hierarchies can be difficult to use [Allen, 1983, Parush and Yuviler-Gavish, 2004, Tullis, 1985], particularly for low-literate people [Jones et al., 2000, Sherwani, 2009]. The latter two studies conducted in the context of low-literate users however were anecdotal, without any concrete evidence for differences in ability to navigate hierarchies. There were also no concrete design implications identified.

This particular study [Medhi et al., 2013a,b] investigates how limited education impacts the ability to navigate hierarchical UIs, even when there is no text. Textual literacy tests were used to score 60 participants from low-income communities in Bangalore, India as “low”, “medium”, and “high” literate. These participants were then asked to perform search tasks on any of the following text-free designs: a list UI with 40 items organized in a grid layout, a shallow hierarchy with the same 40 items organized 2 levels deep with average 8 branching factor, and a deep hierarchy with again the same 40 items organized 4 levels deep and average 3 branching factor. It was found that participants with “high” literacy completed more correct tasks and required less time to navigate on the deep hierarchy, than both groups of participants with “low” literacy, and “medium” literacy. The low literacy participants performed the best on the list UI design where all items of interest were visible all at once on the PC screen; in half the time and had more than 2x accurate responses compared to the deep hierarchy. High literacy participants did about the same on the list in terms of correct responses as compared to the deep hierarchy, but in terms of time took one tenth of the time on the list.

Participants who could not complete tasks correctly or took longer on the hierarchical UIs did not seem to understand the concept of nesting. When asked to do a search task, they randomly selected all unrelated graphics on any given page. They had watched instructional videos on how to navigate the UIs. But it seemed like they did not understand how selecting items corresponded to movement within the hierarchy. Some of them did not remember the use of the “back” button and thus could not remember how to navigate back to higher levels once they had gone down the incorrect path in the hierarchy. Follow-up conversations revealed that some participants had not understood that they had to apply what they had learnt in the instructional video to actual usage during the UI tests. This reinforces the findings from another study where low-literate participants experienced challenges transferring learning from instructional video [Medhi et al., 2012a,b].

Based on the above results it was recommended to minimize hierarchical depth in UIs and keep navigation linear as much as possible. If all the search items in an application could be accommodated on a single screen, it was recommended that the items be laid out in a grid pattern loosely grouping them together.

Another study compared linear, hierarchical and cross-linked navigation performance on mobile UIs for constrained populations in a Western context. The results indicated that users performed best when navigating a linear structure, but preferred the features of cross-linked navigation, e.g., provision to go back to the ‘home’ page. However, this study was with a relatively higher literate group (having between nine and twelve years of formal education), half of who had previous experience playing computer games or browsing the internet occasionally [Chaudry et al., 2012].

While keeping navigation linear seems well-suited both for low-literate and constrained users, for any real world design screen real estate may be limited in certain devices. And it may not be possible to display all search items at once, e.g., on mobile phone screens. So as a follow-up, the same group of researchers conducted a study to investigate the trade-offs of paging through multiple pages of a list UI on a phone, compared to a deep hierarchy where all of the items at a given level were all visible at once [Medhi et al., 2013a]. In the multi-page list design the 40 items were organized across seven pages in a 3×2 matrix on each page; six items per page up to the 6th page and the remaining four items (in a 2×2 matrix) on the 7th page. The deep hierarchy UI design was the same as the previous study, four levels deep with average three branching factor.

A controlled experiment compared the performance of 10 non-literate participants on the multiple-page phone list with another 10 non-literate participants on a deep hierarchy phone UI. Results showed that the multiple-page phone list performed better both in terms of time taken for navigation as well as % correct tasks, compared to the deep hierarchy phone UI.

Like in the previous PC hierarchy study, participants had difficulty understanding the concept of nesting and how selecting items corresponded to movement within the hierarchy. On the multipage list UI, it was observed that even though not all of the items were visible at once, participants did not hesitate to move about quickly through the pages. There was better recall for the use of the "back" button compared to (the hierarchies in) the previous PC study, perhaps due to its placement next to the "forward" button. The study suspected that whether on the phone or on the PC, hierarchies are difficult because the user has to remember they are on a hierarchy and hold it from the root in their thinking. Whereas list navigation really does not require a user to remember much — it is just moving back and forth and knowing where he/she is along a single line.

Comparing results from the single page PC list and multipage phone list, a surprising result was that the multipage phone list performed better, by requiring less time to navigate and completing more correct tasks. This, the study suspected, was perhaps because on the multipage phone list it was not overwhelming to see all of the items all at once like on a single-page PC list. Also, participants using the phone list seemed relatively more comfortable using the device, than the participants who used the single-page PC list. This despite not having previous experience using a touch screen phone. The study suspected that this could be because of general familiarity with the form factor of a phone, even though the interaction was through the touch screen, which users had no experience with.

(b) UIs not just for consumption of content but also for production by low-literate users

One emerging trend in HCI4D is UI work to enable low-literate users to participate in production of content, and not just remain silent consumers. Production of content by low-literate users has been made possible through innovations in voice UIs that have helped transcend issues of textual literacy, local language and device constraints. Having low-literate populations produce their own content also helps generate more locally relevant material. In the following sections we discuss some relevant examples of these UIs for content production:

CGNet Swara

CGNet Swara is a citizen journalism portal where rural users can report and listen to news stories of local interest through an IVR system [Mudliar et al., 2013, CGNET Swara, 2014]. The goal of the project is to extend civic participation beyond the Internet's reach by making the system available through ordinary phones. Users can record and listen to spoken menu outputs by navigating the system through keypad inputs.

Currently CGNet Swara mainly targets low-literate tribal populations in the Indian state of Chhattisgarh, which still remain underserved by mainstream media. Due to lack of trained journalists, there are no established news sources in the local tribal languages. This situation is further worsened by regulations on community radio in India where it is illegal to broadcast or discuss news. While there are newspapers and television stations, they do not offer a voice to the tribal populations. CGNet Swara addresses this need by allowing local community members to record and listen to local stories that are meaningful.

Recordings are less than 3 minutes long and these are moderated to ensure they are clear, audible and appropriate for dissemination. The user sends a missed call to the server, which then calls them back. Then through a few keypad presses, they are able to record or listen to voice messages. After recordings have been moderated by a human mediator (a senior former BCC journalist), they are made available for listening on the phone and the internet website. To keep the phone lines available, only the four most recent posts are available for playback on the phone. Currently there is no search or browse feature for posts on the phone, though the text summaries made available on the internet website are searchable. The various categories of posts include local news, grievances, appeals, interviews, and also song and poetry performances!

The study reports that CGNet Swara allows its users different affordances in a variety of contexts. Users have been known to find the voice interface convenient and easy to understand. In addition, the phone's portability has allowed users to act as "broadcasters" and "infomediaries", generating and sharing messages with others present. Finally, broadcasters and infomediaries appreciate that the playing of voice messages on speaker lets others listen in. As of now CGNet Swara has received 300,000 phone calls and 5,000 recordings. Majority of the callers only listen to content; about 200 calls are received in a day, whereas only 3 new posts are generated.

Avaaj Otalo

Avaaj Otalo is a Q&A forum for small-scale farmers in Gujarat, India, to ask and listen to queries on a wide range of agricultural topics through a simple mobile phone call [Patel et al., 2009, 2010]. The goal of the system is to let farmers learn about best practices through recorded experiences of other farmers, and advice from experts. The service also includes an announcement board that allows a local NGO to broadcast information on topics of general interest, agriculture, animal husbandry, government programs, market prices, etc.

Avaaj Otalo was originally conceived as an interactive, on-demand informational resource that would complement a popular weekly Gujarati radio program targeting small and marginal farmers. The program was produced by a local NGO, Development Support Center (http://www.dscindia.org/ retrieved July 28, 2014). Avaaj Otalo was developed as a system that the NGO could use to incorporate listener feedback into the radio program, in addition to effectively communicating with low-literate listeners.

Farmers on the system are able to record, browse and respond to Q&A by dialing a phone number and navigating through speech prompts or by numeric inputs. As has been discussed earlier, research was conducted to identify what was better suited to low-literate users: numeric inputs or speech inputs [Patel et al., 2009]. The ASR for speech inputs was adapted from American English to be usable by Gujarati speakers. Users would be prompted to either say the given keyword or press the numeric key corresponding to the option they wanted. It was observed that users preferred numeric keypad input significantly over speech input, largely because they experienced discomfort speaking single word commands and "talking to the computer" was perceived as unnatural.

During user studies it was observed that no users expressed major difficulty in understanding how to operate the system through numeric keypad input, including several fully non-literate participants. However, one challenge across both modes of input was navigating voice command-driven menus and knowing when to provide input. Every voice prompt had to be followed with a beep to indicate that input was requested.

During the pilot study in 2009, it was found that more than 60% of Avaaj Otalo's traffic was in Q&A. Topics of questions included a wide range of agricultural subjects from crop planning to crop marketing. The Avaaj Otalo project led to the founding of Awaaz.De, a company in India that provides a hosted solution for deploying voice-based social media (http://awaaz.de/ retrieved July 17, 2014).

Polly

Polly is a voice manipulation and forwarding system on IVR to virally spread job information through entertainment ([Raza et al., 2013, 2012]; Polly CMU http://www.cs.cmu.edu/∼./Polly/ retrieved July 28, 2014). It allows any caller to record a short message, choose from several entertaining voice manipulations, forward the manipulated voice to their friends, provide feedback or listen to the latest job ads. The caller usually initiates by giving a "missed call" and is then called back by the server. The goal of the project is to explore if it is possible to virally spread awareness and train low-literate populations in speech-based services using entertainment as a motivation. The aim also is to study how this voice service can be used as a conduit for introducing more core development-related services. A user who receives a call from Polly with a forwarded job ad first receives the information of the sender and is given Polly's phone number for future use. After listening to the forwarded message they can choose to browse the job ad list.

Polly is currently deployed in Pakistan. The second version of Polly was seeded with five users in 2012 and as of one year later it had 164,807 users who had taken part in 636,536 calls. Of the 33,682 people who chose to listen to a job option, forwarded them 33,484 times to their friends. A follow-up survey through 200 calls revealed that 77% of the users of Polly had less than 10 years of formal education [Raza et al., 2013]. Most respondents described their primary use of the service as "fun" and reported using it for making prank calls to friends, sending messages to say hello, to share poetry and browse job ads as a pastime. A subset of users said that they used the service to send birthday/holiday greetings, know a friend's whereabouts or to browse and apply for jobs.

More rigorous user studies remain to be conducted, but among the feedback from the survey there have been requests to increase recording time, prioritize the unchanged voice recording and have the ability to store messages for later listening. As far as job ads are concerned, the current system uses a push model where job ads from local newspapers are audio recorded and made available for browsing as part of the Polly menu. But the survey has revealed that users of Polly were interested in posting job ads from their end as well.

Gurgaon Idol

Gurgaon idol is a talent competition in collaboration with a Community Radio station in Gurgaon, India, in which community members could call an IVR system to record their songs, and vote to select the best songs [Koradia et al., 2013, Gram Vaani http://www.gramvaani.org/?p=1048 retrieved July 28, 2014]. The Community Radio station is called ‘Gurgaon ki Awaaz’ (GKA) (http://www.trfindia.org/community_radio.php retrieved July 28, 2014) and its primary listener base is migrant workers who have moved to Gurgaon from several Indian states. Of the 22 hours of broadcast in a day programs include folk music, non-folk music, song request programs, and topical programs on health and employment. While songs are sourced through various means, continued curation remains a challenge. The goal of the Gurgaon Idol program was to help GKA enhance its song bank by having a singing competition that listeners could easily access and participate in. Interest was also enhanced by involving people in judging the best songs.

Participants could call in on the IVR system, record their name, age, and the song they wanted to enlist. There were two age groups to select from: below 30 years and above 30 years. Participants could call repeatedly to re-record their song until they were satisfied with their performance. Most of the people who were recruited for the usability test had about 10th–12th Grade education, though there were also people with education higher than that. Even then in the usability test it was found that words like “record” were hard to relate to for first-time IVR users. Users had difficulty remembering a long sequence of instructions, and it was observed that limiting choices through specific instructions could remove anxiety.

All songs recorded by participants were assigned an entry number for voting. Three different voting methods were tested: (a) playing one song and having participants choose among the options of ‘thumbs up’, ‘thumbs down’ or ‘neutral’, (b) playing two songs, and having people select the better among the two options, (c) playing four songs and having people choose the best among the four. In the usability test it was found that even though both thumbs up-thumbs down and best-of-two were easy to use, there were several cultural factors that impacted preference. Best of four method was hard to use because of the difficulty in remembering earlier songs, but that did not impact voting results.

Finally, it was observed that task completion could be improved by training people over the phone and in person. Though keys presses for terminating audio recording were hard to learn for some participants in spite of repeated training calls.

Other than voice UIs, there is one recent study that looks at content production by low-literate users not through voice UIs, but through a graphical touch-screen UI. We describe the system below:

KrishiPustak

KrishiPustak is an audio-visual social networking mobile application for low-literate farmers in rural India [Medhi et al., 2015]. It allows low-literate farmers to create photo or audio posts through their own accounts on a shared smartphone. Farmers can reply to posts using photos or audio as well. There was no text in the UI, though numerals were used to denote the number of replies to a post. Since the system did not have the concept of ‘friends,’ anybody who registered on the system could view content from and reply on the posts of all other users.

In the pilot study, eight mediators from a partner organization BAIF, who worked as para-veterinary workers, were used to seed the system. Each of the mediators was given a Nokia Lumia 820 mobile phone with the application installed. The mediators were from the same village as the farmers, though their education level was between Grade 10 and 12. The literate mediators were to register low-literate farmers with little or no experience with mobile phone uses beyond voice calls, and no existing SN accounts. The mediators were to help these farmers use the application throughout the pilot study.

Over a four-month deployment, 306 farmers registered through the phones of the eight agricultural mediators making 514 posts and 180 replies. The context of mediated use and agricultural framing had a powerful impact on system understanding and usage. During follow-up interviews mediators largely described the system as being for sharing agricultural information and farming best practices, and most posts reflected this. However, despite a strong professional bent, other uses also showed through, with posts related to families, local grievances, and aspirational content. Overall, KrishiPustak was useful and usable, but nevertheless the study identified a number of design recommendations for similar SN systems.

(c) UIs for the literate, tech-savvy, and those for low-literate are converging

Another encouraging theme is that increasingly there are these UIs designed for higher-income, literate users, which use various natural user interfaces (NUIs) that may well be very relevant for non-literate users. Examples include Siri, Cortana, Google Now, Xbox and Kinect, and Google Glass [Apple Siri, Google, Xbox 360 + Kinect, Google now, Windows Phone Cortana, retrieved July 18, 2014]. In these UIs natural dialog systems use spoken language, user modelling through machine learning (for making interactions easier and more intuitive), natural gestures and novel output. Siri, Cortana and Google Now have been positioned as "intelligent personal assistants" that use spoken language to answer questions, make recommendations, and perform actions on smartphones. Xbox and Kinect use natural gestures and spoken language enabling users to interact with their gaming console/computer without a game controller. Google Glass is a wearable technology with an optical head-mounted display that uses spoken language and displays information in a smartphone-like hands-free format. The idea behind these innovations is to leverage interaction mechanisms that come naturally to people. While all of these NUIs are currently aimed at higher income groups, the design principles could also apply to non-literate users. But this needs to be empirically tested. Also, it is not clear that the cost of many of these technologies are within reach for low-income, low-literate users in the near term. But as smartphone prices continue to drop, they could well be within the purview of these users.

(d) Looking beyond immediate human-machine interaction to the larger ecosystem

One emerging theme in the design of UIs for low-literate users are conversations around designing with respect to a host of nuanced socio-cultural and socio-psychological issues, beyond strict usability [Medhi et al., 2010, Parikh and Ghosh, 2006, Sambasivan et al., 2009, Sambasivan and Smyth, 2010]. Most low-literate users live in low-technology contexts without much exposure to computing technology in their living environments and this ecosystem gives rise to interesting design challenges. In the following section, we discuss a few of these conversations that are starting to emerge.

Minimizing intimidation caused by technology

During usability tests in Text-Free UIs on PCs for example, in spite of ensuring that participants felt comfortable, researchers have observed that older participants usually feared that they would "break the device" if not used "correctly" [**?**]. This has been ascribed to the implicit class hierarchy between test takers in low-income contexts and experimenters. It has also been suspected that sometimes the physical appearance of the device used in usability studies can be intimidating to users. The more expensive the device appears higher is the level of discomfort. As such studies have recommended designs that can minimize intimidation caused by technology [Medhi et al., 2010].

Providing relevant, contextual information

It has been observed that the lack of knowledge of application context also seems to influence how low-literate users interact with various technologies [Medhi et al., 2010]. To register for an existing mobile money-transfer service, users were required to fill up details such as their mother's maiden name. Most of the participants in the usability test of this service did not understand what the term "maiden name" meant. Information such as mother's maiden name is usually required for banking protocols and most low-literate test participants being unbanked had no previous context for this information field. Researchers have recommended providing relevant, contextual information in the applications targeting low-literate users [Medhi et al., 2011].

Checking for user motivation for given application

Most current research work in UI design for low-literate users ignores the centrality of user motivation in determining adoption of services. Motivation indeed could be a powerful force that can influence how a user interacts with technology in general. Though it is often contingent on what the technology application has to offer. One study found that the motivation to adopt, if powerful enough, trumps the obstacles in the path of adoption of a new technology [Smyth et al., 2010b]. Through an ethnographic study on sharing and consumption of entertainment media on low-cost mobile phones in urban India, it was found that novice technology users will traverse multiple levels of complex UI navigation, if the motivation exists. In this particular case, the motivation was for the desire to be entertained. It was observed that users traversed as many as 19 steps to do Bluetooth transfer of entertaining content consisting of music, music videos, film dialogs, and comedy clips. This was despite minimum technical knowledge. This observation suggested that while designing services for low-income, low-literate populations, in addition to focusing on users' inability to read, it is important to pay attention to the level of user motivation for the concerned application.

Designing for multiple user scenarios

ICTs have been traditionally designed for individual usage scenarios. However in the case of low-income, low-literate populations, it has been observed that during single-user tests, users usually appear nervous and uncomfortable when asked to perform a task. A group, on the other hand, seems more comfortable, while interacting with a computer [Medhi et al., 2007]. For a health information dissemination application at the public hospital waiting lobby, the researchers had instructed patients to use the system, one by one. These patients were really anxious and hesitant to even come near the kiosk. However, at one point a group of patients at the hospital began playing with the application between the formal tests. It was surprising to see that these patients seemed more confident, suggesting ideas to one another, discussing the purpose of the application, watching over each other's shoulders, advising and learning from one another, and interacting more boldly with the computer. Studies in HCI4D have discussed how technologies in low-income contexts are often shared among members of a household [Sambasivan et al., 2010]. Male members of a household might own a mobile phone and female members might share and use the phone as and when required.

Collaborative usage scenarios have been studied in the context of education amongst children in resource constrained environments [Pal et al., 2006, Pawar et al., 2006]. However, in our knowledge, in the context of adult users with low-literacy levels, group usage scenarios have not been examined so far. But there are conversations beginning to emerge for UI design research taking into account a collaborative user model [Medhi et al., 2010].

Designing for mediated, assisted scenarios

Low-literate users usually reside in communities with high population density and variation in literacy and digital literacy levels. One research work discusses how non-literate members of slum communities very often seek help from proximate users, to interact with a technology owned personally or by the household. A proximate could be termed a technology aide for semi- or non-digitally literate members. They help realize a technological interaction, by aiding to overcome its navigational, functional, or UI complexity [Sambasivan and Smyth, 2010]. People get help from younger family members and friends to carry out functions such as: receiving and interpreting text messages, carrying out mobile money-transfer transactions through sharing secret passwords, and other sophisticated functions on the phone. This seems to be due to the proximity, approachability and trusted relationship with the proximates.

Again there is a study that discusses how the proxy user's filtering and funneling decisions limit the low-literate users' information-seeking behavior [Parikh and Ghosh, 2006]. The low-literate user might have an unequal power relationship with the proxy. Therefore, he or she might never know the full scope of actions and knowledge available to them. The paper thus recommends developing technologies that explicitly support intermediated tasks by recognizing the needs and aspirations of all users, including those without direct access to the UI. Other researchers have also argued that technological interventions will be more effective if they take into consideration the underlying human infrastructures such as intermediation that are embedded in low-income communities [Sambasivan et al., 2010].

While most current ICTs are still designed and tested for individual usage scenarios, one recent study looks at using agricultural field workers as intermediaries for low-literate farmers to access and use a mobile social-networking system, KrishiPustak [Medhi et al., 2015]. The role of the literate mediators was to provide technology intermediation — that is access to the smartphone technology, support, encouragement and on-the-spot training — to the low-literate farmers who were the end-users of KrishiPustak.

(c) Increasing literacy levels among younger populations in low-income communities

Apart from all of the above themes that are emerging in design of UIs for low-literate users, another relevant trend is the increase of literacy levels among younger members in low-income communities. Researchers in HCI4D have observed during usability experiments that more literate participants tended to be younger [Medhi et al., 2013a,b]. In countries like India this is likely due to the increase in school enrollment and quality of education in recent years, owing to the government's efforts towards universal elementary education [Sarva Shiksha Abhiyan retrieved November 28, 2014]. Younger members of low-income communities have been reported to use mobile phones for texting [Medhi et al., 2013a,b]. One study conducted among low-income youth in slum communities in India found that they were accessing social networking sites such as Facebook within the first month of mobile internet usage [Rangaswamy and Cutrell, 2012]. With more and more younger people in these communities gaining literacy and familiarity with technology it would be interesting to note how the design principles for UI design for low-literate users will hold in the context of the younger, literate users.

5

Opportunities

In the previous sections, we saw that currently a number of innovative UI designs are making use of different combinations of input and output modalities to enable easy ICT access to low-literate users. The numbers of such designs are only increasing, and going forward there seem to be even more directions to push forward in the research in UI design for low-literate users. We discuss a few of these design opportunities here:

(a) UI studies for training effects

A number of UI studies targeted at low-literate users have focused on first-time usage [Cuendet et al., 2013, Grisedale et al., 1997, Medhi et al., 2011, Parikh et al., 2003]. First-time usage is when target community participants use an application for the very first instance in a formal evaluation. Participants may have used low-fidelity prototypes of the application informally through the iterative cycle, but most observations are often based on the formal evaluation of the final prototype. As we have seen, usability test participants may be anxious due to the contrived nature of the testing environment: their first encounter with a new technology, the implicit class hierarchy with experimenters, a task-oriented and rigorously timed usage scenario, and other such issues [Medhi et al., 2010]. This may not be a fair representation of how the ICT gets adopted had it been more carefully situated into the natural ecosystem that users were familiar with. Systems deployed over a longer period have been appropriated by low-literate target communities to varying degrees of success. Examples of such projects include CGNet Swara, Polly, Avaaj Otalo, and Gurgaon Idol. Previous research has also shown that the outcome of usability studies can be strongly impacted by providing focused training to users before usability tests. In a study exploring data collection with novice health workers in India, it was seen that there was only one error during the whole study testing a live operator interface [Patnaik et al., 2009]. Another study proposed human-guided instructions in which users learned to use a speech-based health information system Healthline with a human mentor [Sherwani, 2009]. Compared with the authors' own prior work using video tutorials, it was shown that the interactivity and individually tailored nature of the cooperative human-guided tutorials made it easier to learn for low literate users.

All of this brings up opportunities for assessing learning on UIs for low-literate users over a longer period of time. More longitudinal field deployments could be conducted to study and measure training effects and how UIs are learned and appropriated over time. Which socio-cultural issues need to be taken into account and what UI design innovations could there be to enable better learning and easier access.

(b) Designing for multiple user scenarios

Most UI design studies still focus on individual usage scenarios. That is, the system is designed to take and process user input from only single users at any given time. However, studies have observed that in the real world context low-literate, low-income households often share technological artifacts [Medhi et al., 2011, Sambasivan and Smyth, 2010]. Younger members of families help older members to realize useful interactions with ICTs like mobile phones. Female members share and use mobile phones owned by male members of a family. This is often to meet constraints in literacy and the affordability of devices. Researchers have also observed how using a technology by oneself in a test setting could be very intimidating to first-time low-literate users [Medhi et al., 2007]. But in the same setting when used by a group of users with one person taking the lead, people are far more confident. Studies have observed how in a group usage scenario, people could be suggesting ideas to the person taking the lead, watching over each other's shoulders and learning from one another [Medhi et al., 2007]. The CGNet Swara project for citizen journalism showed how "infomediaries" use the speaker of a mobile phone to play voice messages on the IVR for the consumption of a group of users [Mudliar et al., 2013].

In the context of educational technologies there have been studies that have explored collaborative usage scenarios for resource constrained environments [Pal et al., 2006, Pawar et al., 2006]. In one study researchers developed 'multipoint' software that allowed up to five children to use one computer through the use of five mice and five individual cursors on a single screen. This was done to circumvent the problem of high student-to-computer ratio in resource constrained schools. Taking this forward 'mouse mischief' was developed so as to allow a classroom size of up to 32 children to use individual mice and cursors to interact with a single large display [Moraveji et al., 2009].

There could be similar thrust for exploring collaborative usage paradigms among low-literate adult users. ICTs could be designed to take input from multiple users at a given time. There could be various collaborative and competitive usage scenarios to design for, which need to be studied. There could also be research opportunity in exploring shared usage where users have multiple accounts on a single device but use it individually at different times. Researchers have previously studied various modes of mobile phone sharing and the related deep processes of human organization, though this is in the domain of ethnography [Steenson and Donner, 2009]. Taking this forward there are opportunities for research in designing around similar shared usage scenarios.

(c) Designing for mediated, assisted scenarios

We have previously discussed how in low-income communities, non-literate members often seek help from someone who has access to technology and is more digitally literate [Sambasivan et al., 2010]. This mediated use of technology can amplify its use to many people in the community who would otherwise not be able to use it. Research can be conducted to explore system designs around mediated usage, or to support existing mediated usage. For such studies the collective digital-literacy of the social group will need to be taken into account. These systems could be studied in the natural environment of the community so as to observe and understand real-world challenges. For doing such studies it would be important to identify the human mediators through whom to seed the system. This could be done through partner organizations who work closely with target communities.

Some notable projects which rely heavily on human infrastructure include: DakNet (human transport networks — busses, motorcycles, ox carts) [Pentland et al., 2004]; MOSES (Groups of kiosk users) [Smyth et al., 2010a]; Digital Green (farmers and villagers in close-knit communities) [Gandhi et al., 2007]; KrishiPustak (agricultural mediators from the village) [Medhi et al., 2015]; data entry accuracy for health using forms, SMS, and voice (Human data entry operations) [Patnaik et al., 2009]; and rural mobile health (community health workers and patients) [Medhi et al., 2012a,b, Ramachandran et al., 2009].

(d) Designing for lowering intimidation

We have noted how many low-literate users experience trepidation when using an ICT for the very first time or in formal evaluations. Research could be conducted in exploring user experience and product design opportunities for lowering intimidation. What form factor of the ICT might help users feel more comfortable using it? There could also be opportunities in identifying methods and techniques to help lower intimidation while introducing a technology in a formal evaluation or a pilot deployment.

The ‘Bollywood Method’ [Chavan, 2005] has been found to be useful in making novice users more comfortable in going about a usability test. In this method tasks are embedded in dramatized stories involving the test participant. This has been found to be better at motivating participants toward the desired tasks. Methods similar to the Bollywood method could be useful in contexts where there might be participant response bias, due to the social distance between test participants and experimenters [Dell et al., 2012].

(e) Studying how cost-consciousness impacts user experience

Low-literate, low-income communities in HCI4D contexts have been known to be judicious with cost of technology. Researchers have noted how a large section of people in low-income communities still owned low-end limited functionality phones [Medhi et al., 2011]. "China phones" with larger screens and higher multimedia capability, and some low-end Android phones were becoming more common, but they are still not owned by majority of users. As discussed earlier, this could be why low-literate users experience intimidation when using an expensive device for the very first time or in formal evaluations.

The above cost consciousness is not limited to devices alone, it extends to the use of phone talk time [Donner, 2007] and mobile internet [Donner et al., 2011]. Industry estimates suggest 83% users in emerging economies are "prepay" or pay-as-you-go users [Almanac, 2011]. Consumers get to pay as much or as little as they can afford each day/week/month. Researchers have observed how consumers in emerging economies are cost sensitive and will continue to be particularly acute around mobile video [Oeldorf-Hirsch et al., 2012]. Many users do not download a movie to their handset, lest it "eat their cap" or "drain their airtime" [Chetty et al., 2012]. Given this context one study looks at mobile video quality trade-offs for bandwidth-constrained consumers [Oeldorf-Hirsch et al., 2012]. A series of online studies testing the effects of manipulating a video's content, bit rate, frame rate, and audio quality showed that video quality can be greatly reduced with relatively little decrease in outcomes of enjoyment. A field experiment with low-income users in urban India suggested that offering lower-quality videos to bandwidth-constrained users could provide monetary savings with only minimal reduction in consumer satisfaction.

There is opportunity for more research in user experience of mobile and other ICT experiences given the cost-consciousness among users in HCI4D contexts. How could overall experience be preserved or enhanced even if there was bandwidth or other monetary constraints?

(f) Studying how application design impacts livelihoods

There is also an opportunity to study how a given application design impacts livelihoods of low-income, low-literate communities. Among current innovations, paid crowdsourcing platforms have been publicized as a potential way to help generate income for low-income workers in the developing world. About one third of the workers on crowdsourcing platforms like Mechanical Turk are based in India [Ross et al., 2010]. However, in practice, it has been found that most of these workers are college-educated and have an income that is more than double the Indian average [Khanna et al., 2010]. One recent study presented, mClerk, a platform that sends and receives tasks via SMS, making it accessible to anyone with a low-end mobile phone [Gupta et al., 2012]. This system leveraged a little-known protocol to send small images via ordinary SMS for distribution of graphical crowdsourced tasks. There could be opportunities to extend innovations like mClerk to low-literate users. In addition to graphics, other modalities such as audio and video could be explored for making these applications accessible. Audio on IVR could be a potential direction for UIs for low-literate users that can impact livelihood.

Another application that currently offers livelihood services to low-income community members is Babajob.com (babajob.com retrieved July 24, 2014). Babajob is a job portal that connects registered job seekers with potential employers. Job seekers are typically from the unorganized labor sector where jobs are found through informal social networks, e.g., drivers, domestic helpers, cooks, security guards, data entry operators, etc. Due to low levels of literacy among job seekers, 'missed calls' and SMSs with numeric information are leveraged to disseminate information about matching jobs.

There is opportunity for research in applications similar to Babajob that allow access to livelihood services. Various non-textual interfaces with graphics, audio and video could be leveraged so low-literate users can use these applications more independently.

6

Concluding Summary

One of the greatest challenges in designing for low-literate users is that they are unable to read text. Textual interfaces are therefore unusable. Research in cognitive science also shows that beyond the inability to read, there are certain cognitive skills relevant for ICT use that low-literate users experience difficulties with. Despite this, currently there is widespread penetration of ICTs, especially mobile phones, among low-income, low-literate communities. Though it has been observed that low-literate users use mobile phones for very limited functions, like voice calls only.

The above context provides tremendous opportunity for design and research in UIs for low-literate populations. The goal would be to make UIs usable and accessible for a broad range of services across agriculture, healthcare, governance, livelihoods, money management and many other domains. For more than a decade researchers in HCI4D have experimented with non-textual UIs for low-literate users, and this remains a promising area of work. In this monograph, we review a number of these examples from this past decade. We examine studies that have used modalities more natural than text entry: voice/speech, touch and other forms of input; and graphics, audio, and video as output. There have been other studies that have compared various combinations of input and output modalities in the mobile design space, going from inflexible to flexible and lean to rich, respectively. We make recommendations for UI design for low-literate users based on these examples.

In addition to the above work, currently there are some strong trends that are emerging in the research for UI design for low-literate users. Researchers have studied how non-literacy is not just about the inability to read text, but correlated with cognitive skills that have implications for UI design. One of the skills examined was the ability to transfer learning in video-based skills training, and the related implications for design of instructional video. It was also shown that non-literacy is correlated with the ability to navigate hierarchical organization of information architectures. The study closed with design recommendations for UI navigation for both PCs and mobile phones.

Another strongly emerging trend we discuss is of UIs for not just consumption of content by low-literate users, but also production. Most of these examples are in IVR systems that use spoken menu output with keypad input navigation. Current examples are in the domains of citizen journalism, agriculture Q&A, a virally spread voice manipulation and forwarding system, and a community moderated talent competition. We follow this up with a recent example of an audio-visual social-networking system that allows low-literate farmers to make and reply to photo and audio posts through shared smartphones.

Also, in the recent years, there have been natural UIs for the literate and tech-savvy, with design principles that could very well apply to UIs for low-literate users. This new generation of natural UIs use gestures, speech, touch and other forms of natural interaction, instead of text. It is a relatively new domain that presents exciting possibilities for using technologies in spoken dialogs, agents with personality (like Microsoft's Cortana, Apple's Siri), proactive intelligent user modelling and gestures for the next wave of research for UIs for low-literate users.

Conversations have also begun in looking beyond immediate human–machine interaction to the larger ecosystem. Researchers have talked about designing for a host of nuanced socio-cultural and socio-psychological issues, beyond strict usability, which emerge due to the interesting ecosystem low-literate users are situated in. These range from minimizing intimidation caused by technology, providing contextual information, to designing for multiple and mediated usage scenarios. It would be interesting to see how the increasing literacy levels among younger populations in low-income communities will impact the UI design principles established for low-literate users.

Going forward there seem to be even more directions to push forward in UI design for low-literate users. There are opportunities in UI research for studying training effects and learning over a longer period of time. These can be done through longitudinal field deployments. Research can also be conducted in designing ICTs for multiple user scenarios. There are both collaborative and competitive scenarios to think of. Designing for mediated and assisted scenarios, where a technology is seeded into a community through a human mediator, is another direction to push forward. It would be interesting to study usage and uptake in the mediated context. There is also opportunity in UI and product design research for lowering intimidation among low-income, low-literate users. It would be exciting to identify well-suited methods and techniques to help low-literate users feel more comfortable while using a technology. This is both for formal evaluations as well for pilot deployments. There is also opportunity in studying how cost-consciousness among low-income, low-literate users impacts user experience. Finally, there is potential for research in understanding and designing UI applications for generating income. It would be interesting to study how UI applications could lead to the impact on livelihoods among low-income, low-literate communities.

Acknowledgments

Thanks to my mentor Kentaro Toyama for the encouragement to write this monograph. I thank my manager Ed Cutrell, and colleague and husband Bill Thies for many useful suggestions throughout the writing of this monograph. Thanks to all my co-authors and collaborators on various papers that are covered in this monograph.

References

H. Abadzi. *Improving Adult Literacy Outcomes: Lessons from Cognitive Research for Developing Countries*. World Bank, Washington, DC, 2003.

R. Agrawal, M. Atray, and S. K. Sundari. Exploring suitable interfaces for agriculture based smartphone apps in India. In *Proceedings of the Asia Pacific Conference on Computer Human Interaction (APCHI 2013)*. 2013.

N. Akinnaso. The consequences of literacy in pragmatic and theoretical perspectives. *Anthropology & Education Quarterly*, 12(3):163–200, 1981.

R. B. Allen. Cognitive factors in the use of menus and trees: An experiment. *IEEE Journal on Selected Areas in Communications*, SAC-1(2):333–336, 1983.

T. Ahonen Almanac. Mobile telecoms industry review. 2011, 2011.

Y. Anokwa, T. N. Smyth, D. Ramachandran, J. Sherwani, Y. Schwartzman, R. Luk, M. Ho, N. Moraveji, and B. DeRenzi. Stories from the field: Reflections on HCI4D experiences. Information Technologies & International Development, 2009.

Apple Siri. https://www.apple.com/ios/siri/. Retrieved July 28, 2014.

A. Ardila, M. Rosselli, and P. Rosas. Neuropsychological assessment in non-literates: Visuospatial and memory abilities. *Brain Cognition*, 11(2):147–166, 1989.

Babajob. http://www.babajob.com/. Retrieved July 28, 2014.

K. Bali, S. Sitaram, S. Cuendet, and I. Medhi. A hindi speech recognizer for an agricultural video search application. ACM Symposium on Computing for Development (ACM DEV), January, 2013.

A. Bandura. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84:191–215, 1977.

A. Bandura. Self-efficacy. In V. S. Ramachaudran, editor, *Encyclopedia of Human Behavior*, pages 71–81. Academic Press, New York, 2005.

G. Becker. "Human Capital Revisited*," Human Capital*. University of Chicago Press, Chicago, 1993.

A. Bernardo. *Literacy and the Mind: The Contexts and Cognitive Consequence of Literacy Practice*. Luzac Oriental, 2000.

M. Bernhart, I. Wiadnyana, H. Wihardjo, and I. Pohan. Patient satisfaction in developing countries. *Social Science and Medicine*, 48, 1999.

M. L. Best and R. Kumar. Sustainability failure of rural telecenters: The sustainable access in rural india project. *Information Technologies and International Development*, 4(4):31–45, 2008.

S. Bhattacharya and L. Feldman. Igwana: a text-free search interface. In *Proceedings of the Australian Computer–Human Interaction Conference*, pages 34–37. Melbourne, Australia, November 26–30 2012.

S. Boyera. The mobile web to bridge the digital divide? Paper presented at the IST-Africa Conference 2007, Maputo, Mozambique, 2007.

D. Brown, G. Marsden, and U. Rivett. Water alert!: using mobile phones to improve community perspective on drinking water quality in South Africa. In *Proceedings of the International Conference on Information and Communication Technologies and Development*. Atlanta, Georgia, March 12–15 2012.

A. Castro-Caldas. Targeting regions of interest for the study of the non-literate brain. *International Journal of Psychology*, 39:5–17, 2004.

A. Castro-Caldas, K. M. Petersson, A. Reis, S. Stone-Elander, and M. Ingvar. The illiterate brain: learning to read and write during childhood influences the functional organisation of the adult brain. *Brain*, 121:1053–1063, 1998.

S. J. Ceci. How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Development Psychology*, 27(5):703–722, 1991.

CGNET Swara. http://cgnetswara.org/. Retrieved July 28, 2014.

D. Chakraborty, I. Medhi, E. Cutrell, and W. Thies. Man versus machine: Evaluating IVR versus a live operator for phone surveys in India. In *ACM Symposium on Computing for Development (ACM DEV)*. January 2013.

H. Y. Chan and R. Rosenfeld. Discriminative pronunciation learning for speech recognition for resource scarce languages. In *Proceedings of ACM DEV 2012, Annual ACM Symposium on Computing for Development*. Atlanta, GA, March 2012.

B. Chaudry, K. Connelly, K. A. Siek, , and J. L. Welch. Mobile interface design for low-literacy populations. In *Proceedings of the 2012 ACM SIGHIT International Health Informatics Symposium*. 2012.

A. L. Chavan. Another culture, another method. In *Proceedings of the International Conference on Human–Computer Interaction, HCII2005*. 2005.

M. Chetty, R. Banks, A. J. B. Brush, J. Donner, and R. E. Grinter. "you're capped!" Understanding the effects of bandwidth caps on broadband use in the home. In *Proceedings of CHI 2012*. Austin, TX, USA, May 5–10, 2010 2012.

J. Chipchase. Understanding non-literacy as a barrier to mobile phone communication. http://www.kiwanja.net/database/document/report_literacy_barrier.pdf, 2005.

S. Cuendet, I. Medhi, K. Bali, and E. Cutrell. VideoKheti: Making video content accessible to low-literate and novice users. In *ACM Conference on Human Factors in Computing Systems*. 2013.

S. J. Czaja, N. Charness, A. D. Fisk, C. Hertzog, S. N. Nair, and W. A. Rogers. Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (CREATE). *Psychology and Aging*, 21:333–353, 2006.

M. Czerwinski, E. Horvitz, and E. Cutrell. Subjective duration assessment: An implicit probe for software usability. In *IHM-HCI Conference*. 2001.

N. Dell, V. Vaidyanathan, I. Medhi, E. Cutrell, and W. Thies. "Yours is Better!" participant response bias in HCI. In *The ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)*. 2012.

K. Dew, M. Haddadin, C. Fishel, and A. Dawale. Karaoke: an assistive alternative interface for illiterate users. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Paris, France, April 27–May 02 2013.

J. Donner. The rules of beeping: Exchanging messages via intentional "missed calls" on mobile phones. *Journal of Computer-Mediated Communication*, 13(1), 2007.

J. Donner, S. Gitau, and G. Marsden. Exploring mobile-only internet use: Results of a training study in urban South Africa. *International Journal of Communication*, 5:574–597, 2011.

J. Ermisch and M. Francesconi. *Family Matters*. Institute for Social and Economic Research, 1997.

J. Ermisch and M. Francesconi. *The Effect of Parents' Employment on Children's Educational Attainment*. United Kingdom, University of Essex UP, 2000.

Y. Eshet-Alkalai. Digital literacy: A conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia*, 13 (1):93–106, 2004.

E. Friscira, H. Knoche, and J. Huang. Getting in touch with text: Designing a mobile phone application for illiterate users to harness SMS. In *ACM DEV*. 2012.

R. Gandhi, R. Veeraraghavan, K. Toyama, and V. Ramprasad. *Digital Green: Participatory video for Agricultural Extension*. ICTD, 2007.

D. Gardner-Bonneau, H. E. Blanchard, and B. Suhm. Ivr usability engineering using guidelines and analyses of end-to-end calls. In *Human Factors and Voice Interactive Systems, Signals and Communication Technology*. 2008.

Google. http://www.google.com, Retrieved July 28 2014.

Google now. http://www.google.com/landing/now/, Retrieved July 28 2014.

Gram Vaani and Gurgaon Idol. http://www.gramvaani.org/?p=1048. Retrieved July 28, 2014.

S. Grisedale, M. Graves, and A. Grunsteidl. Designing a graphical user interface for healthcare workers in rural India. In *ACM Conference on Human Factors in Computing Systems*. New York, 1997.

A. S. Grover, M. Plauché, E. Barnard, and C. Kuun. Hiv health information access using spoken dialogue systems: touchtone vs. speech. In *ICTD*. 2009.

A. Gupta, W. Thies, E. Cutrell, and R. Balakrishnan. mclerk: Enabling mobile crowdsourcing in developing regions. In *ACM Conference on Human Factors in Computing Systems (CHI 2012)*. 2012.

Gurgaon ki Awaaz. http://www.trfindia.org/community_radio.php. Retrieved July 28, 2014.

R. Heeks and L. L. Kanashiro. Telecentres in mountain regions: A peruvian case study of the impact of information and communication technologies on remoteness and exclusion. *Journal of Mountain Science*, 6(4):320–330, 2009.

M. Huenerfauth. Developing design recommendations for computer interfaces accessible to illiterate users. Master's thesis, University College Dublin, 2002.

ITU. The world in 2014 : ICT facts and figures, 2014.

M. Jones and G. Marsden. *Mobile Interaction Design*. John Wiley & Sons, 2005.

M. Jones, G. Buchanan, H. Thimbleby, and G. Marsden. User interfaces for mobile web devices. [www9 mobile workshop position paper]. In *Proceedings of the International World Wide Web Conference*. Amsterdam, 2000.

A. Joshi, N. Welankar, B. L. Naveen, K. Kanitkar, and R. Sheikh. Rangoli: A visual phonebook for low-literate users. In *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services*, pages 217–223. MobileHCI '08. ACM, New York, NY, Amsterdam, The Netherlands, September 02–05, 2008 2008.

D. S. Katre. Unorganized cognitive structures of non-literate as the key factor in rural e-learning design. *I-Manager's Journal of Education Technology*, 2 (4):67–72, 2006.

S. Khanna, A. Ratan, J. Davis, and W. Thies. Evaluating and improving the usability of mechanical turk for low-income workers in India. In *ACM Symposium on Computing for Development (DEV 2010)*. 2010.

Z. Koradia, P. Aggarwal, A. Seth, and G. Luthra. Gurgaon idol: A singing competition over community radio and IVRS. *ACM DEV*, 3:6, 2013.

A. Kumar, S. K. Agarwal, and P. Manwani. The spoken web application framework. In *Proceedings of W4A*. 2010.

R. Kuriyan, K. Toyama, and I. Ray. Integrating social development and financial sustainability: The challenges of rural computer kiosks in Kerala. In *Proceedings of International Conference on Information Technology and Communcation and Development*. Berkeley, May 2006.

J. Kurvers. Met ongeletterde ogen. kennis van taal en schrift van analfabeten. [with nonliterate eyes. knowledge of language and writing of non-literates]. (Thesis, Tilburg University), 2002.

I. Ladeira and E. Cutrell. *ACM/IEEE International Conference on Information and Communication Technologies and Development*. London, December 2010.

A. R. Luria. *Cognitive Development: Its Cultural and Social Foundations*. Harvard University Press, Cambridge, MA, 1974.

W. Mackay. Triangulation within and across HCI disciplines. *Human-Computer Interaction*, 13(3), 1998.

J. J. Manly, P. Touradji, M. X. Tang, and Y. Stern. Literacy and memory decline among ethnically diverse elders. *Journal of Clinical and Experimental Neuropsychology*, 25:680–690, 2003.

E. Matute, F. Leal, D. Zarabozo, A. Robles, and C. Cedillo. Does literacy have an effect on stick construction tasks? *Journal of the International Neuropsychological Society*, 6:668–672, 2000.

I. Medhi and K. Toyama. Full-context videos for first-time, non-literate PC users. In *IEEE/ACM International Conference on Information and Communication Technologies and Development*. Bangalore, India, 2007.

I. Medhi, B. Pitti, and K. Toyama. Text-free UI for employment search. In *Asian Applied Computing Conference*. 2005.

I. Medhi, M. Sagar, and K. Toyama. Text-free user interfaces for illiterate and semi-literate users. *Information Technologies and International Development*, 4(1):37–50, 2007.

I. Medhi, S. N. N. Gautama, and K. Toyama. A comparison of mobile money-transfer UIs for non-literate and semi-literate users. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. 2009.

I. Medhi, E. Cutrell, and K. Toyama. It's not just illiteracy. India HCI/Interaction Design & International Development, Mumbai, March 20–24 2010.

I. Medhi, S. Patnaik, E. Brunskill, S. N. N. Gautama, W. Thies, and K. Toyama. Designing mobile interfaces for novice and low-literacy users. *ACM Transactions on Computer-Human Interaction*, 18:1, 2011.

I. Medhi, M. Jain, A. Tewari, M. Bhavsar, M. Matheke-Fischer, and E. Cutrell. Combating rural child malnutrition through inexpensive mobile phones. In *Nordic Conference on Human-Computer Interaction*. 2012a.

I. Medhi, S. R. Menon, E. Cutrell, and K. Toyama. Correlation between limited education and transfer of learning. *Information Technologies and International Development*, 8:51–65, June 2012b.

I. Medhi, M. Lakshmanan, K. Toyama, and E. Cutrell. Some evidence for impact of limited education on hierarchical user interface navigation. In *ACM Conference on Human Factors in Computing Systems*. 2013a.

I. Medhi, K. Toyama, A. Joshi, U. Athavankar, and E. Cutrell. A comparison of list vs. hierarchical UIs on mobile phones for non-literate users. In *Interact. IFIP TC13 Conference on Human–Computer Interaction*. 2013b.

I. Medhi, P. Ferreira, N. Gupta, J. O'Neill, and E. Cutrell. KrishiPustak: A social networking system for low-literate farmers, CSCW, 2015.

S. Mitra. Minimally invasive education: a progress report on the "hole in the wall" experiments". *British Journal of Education Technology*, 34(3): 367–371, 2003.

R. K. Moore. A comparison of the data requirements of automatic speech recognition systems and human listeners. In *Proceedings of Eurospeech*, pages 2582–2584. Geneva, 2003.

J. Morais, L. Cary, J. Alegria, and P. Bertelson. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7:323–331, 1979.

N. Moraveji, K. Inkpen, E. Cutrell, and R. Balakrishnan. A mischief of mice: Examining children's performance in single display groupware systems with 1 to 32 mice. In *CHI 2009*, pages 2157–2166. 2009.

P. Mudliar, J. Donner, and W. Thies. Emergent practices around CGNet swara: A voice forum for citizen journalism in rural India. *Information Technologies and International Development (ITID)*, 9(2 (ICTD 2012 Special Issue)), 2013.

F. Oberle. Who why and how often? key elements for the design of a successful speech application taking account of the target groups. In *Usability of Speech Dialog Systems, Signals and Commmunication Technologies*. Springer, 2008.

A. Oeldorf-Hirsch, J. Donner, and E. Cutrell. How bad is good enough? exploring mobile video quality trade-offs for bandwidth-constrained consumers. In *Proceedings of the Nordic Conference on Human-Computer Interaction: Making Sense through Design (NordiCHI '12)*, pages 49–58. New York, NY, USA, 2012.

One Laptop Per Child. Xo-4 touch specifications. www.one.laptop.org.

M. Orne. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 1962.

J. Pal, U. S. Pawar, E. A. Brewer, and K. Toyama. The case for multi-user design for computer aided learning in developing regions. In *Proceedings of the International Conference on World Wide Web*, pages 781–789. Edinburgh, Scotland, New York, NY, May 23–26 2006. WWW '06.

T. Parikh, K. Ghosh, and A. Chavan. Design studies for a financial management system for micro-credit groups in rural India. ACM Conference on Universal Usability, 2003.

T. S. Parikh and K. Ghosh. Understanding and designing for intermediated information tasks in india. *Pervasive Computing, IEEE*, 5(2):32–39, 2006.

A. Parush and N. Yuviler-Gavish. Web navigation structures in cellular phones: The depth/breadth trade-off issue. *International Journal of Human-Computer Studies*, 60:753–770, 2004.

N. Patel, S. Agarwal, N. Rajput, A. Nanavati, P. Dave, and T. S. Parikh. A comparative study of speech and dialed input voice interfaces in rural India. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. 2009.

N. Patel, D. Chittamuru, N. Jain, P. Dave, and T. S. Parikh. Avaaj otalo — a field study of an interactive voice forum for small farmers in rural India. In *ACM Conference on Human Factors in Computing Systems*. 2010.

S. Patnaik, E. Brunskill, and W. Thies. Evaluating the accuracy of data collection on mobile phones: A study of forms, sms, and voice. In *Proceedings of the International Conference on Information and Communication Technologies and Development*. 2009.

U. S. Pawar, J. Pal, and K. Toyama. Multi-mouse for rural education. In *Proceedings of International Conference on Information Technology and Communication and Development*. Berkeley, May 2006.

A. Pentland, R. Fletcher, and A. Hasson. Daknet: Rethinking connectivity in developing nations. *Computer*, 37(1):78–83, 2004.

M. Plauché, U. Nallasamy, J. Pal, C. Wooters, and D. Ramachandran. Speech recognition for illiterate access to information and technology. In *ICTD*, pages 83–92. 2006.

A. Prasad, I. Medhi, K. Toyama, and R. Balakrishnan. Exploring the feasibility of video mail for illiterate users. In *Proceedings of AVI 2008 — the Conference on Advanced Visual Interfaces*, pages 103–110. 2008.

F. Qiao, J. Sherwani, and R. Rosenfeld. Small-vocabulary speech recognition for resource-scarce languages. In *Proceedings of ACM DEV 2010, Annual ACM Symposium on Computing for Development*. London, UK, December 2010.

D. Ramachandran, J. Canny, P. D. Das, and E. Cutrell. Mobileizing health workers in rural India. In *ACM Conference on Human Factors in Computing Systems*. 2009.

N. Rangaswamy and E. Cutrell. Anthropology, development and ICTs: Slums, youth and the mobile internet in Urban India. In *ICTD*. 2012.

A. L. Ratan, et al. Managing microfinance with paper, pen and digital slate. In *Proceedings of ICTD 2010*. 2010.

A. Raza, F. U. Haq, Z. Tariq, M. Pervaiz, S. Razaq, U. Saif, and R. Rosenfeld. Job opportunities through entertainment: Virally spread speech-based services for low-literate users. In *ACM Conference on Human Factors in Computing Systems*. Paris, France, 2013.

A. A. Raza, M. Pervaiz, C. Milo, S. Razaq, G. Alster, J. Sherwani, U. Saif, and R. Rosenfeld. Viral entertainment as a vehicle for disseminating speech-based services to low-literate users. In *ICTD*. 2012.

A. Reis and A. Castro-Caldas. Illiteracy: A bias for cognitive development. *Journal of International Neuropsychological Society*, 3:444–450, 1997.

A. Reis, K. M. Petersson, A. Castro-Caldas, and M. Ingvar. Formal schooling influences two- but not three-dimensional naming skills. *Brain and Cognition*, 47(3):397–411, 2001.

J. Ross, L. Irani, M. Silberman, A. Zaldivar, and B. Tomlinson. Who are the crowdworkers? In *Shifting Demographics in Mechanical Turk*. CHI EA, 2010.

N. Sambasivan and T. Smyth. The human infrastructure in ICTD. In *Proceedings of ACM/IEEE Interactional Conference on Information and Commmunication Technologies for Development (ICTD)*. London, UK, 2010.

N. Sambasivan, N. Rangaswamy, E. Cutrell, and B. Nardi. Ubicomp4D: Interaction and infrastructure for international development-the case of Urban Indian slums. In *Proceedings of ACM Conference on Ubiquitous Computing (UBICOMP)*, pages 155–164. Orlando, Florida, 2009.

N. Sambasivan, E. Cutrell, K. Toyama, and B. Nardi. Intermediated technology use in developing communities. In *Proceedings of ACM conference on Human Factors in Computing Systems (SIGCHI)*, pages 2583–2592. Atlanta, GA, 2010.

Sarva Shiksha Abhiyan. http://ssa.nic.in/, Retrieved November 28 2014.

S. Scribner and M. Cole. *The Psychology of Literacy*. Harvard University Press, 1981.

H. Seth and K. Sorathia. Parichaya — a low-cost device to increase adherence among tuberculosis patients in rural Assam. In *Proceedings of the Asia Pacific Conference on Computer Human Interaction*. Bangalore, India, September 24–27 2013.

J. Sherwani. Speech interface for information access by low-literate users in the developing world. PhD thesis, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA, May, 2009. (Also published as technical report CMU-CS-09-131.).

J. Sherwani, M. S. Ali, A. Fatma, Y. Memon, M. Karim, R. Tongia, and R. Rosenfeld. Healthline: Speech-based access to health information by low-literate users. In *IEEE/ACM Int'l Conference on Information and Communication Technologies and Development, Bangalore, India*. 2007.

J. Sherwani, N. Ali, C. P. Rose, and R. Rosenfeld. Orality-grounded HCID: Understanding the oral user. *Information Technologies & Internatinal Development*, 5(4):35–48, 2009a.

J. Sherwani, S. Palijo, S. Mirza, T. Ahmed, N. Ali, and R. Rosenfeld. Speech vs. touch-tone: Telephony interfaces for information access by low literate users. In *Proceedings of Information & Communications Technologies and Development*. Doha, Qatar, April 2009b.

T. Smyth, J. Etherton, and M. L. Best. MOSES: Exploring new ground in media and post-conflict reconciliation. Presented at CHI, Atlanta, GA, 2010a.

T. N. Smyth, S. Kumar, I. Medhi, and K. Toyama. Where there's a will there's a way: Mobile media sharing in urban India. In *Proceedings of the International Conference Extended Abstracts on Human Factors in Computing Systems*. New York, NY, 2010b.

A. Squires. Methodological challenges in cross-language qualitative research: A research review. *International Journal of Nursing Studies*, 46(2):277–287, February 2009.

M. Steenson and J. Donner. Beyond the personal and private: Modes of mobile phone sharing in urban India. In S. W. Campbell and R. Ling, editors, *The Reconstruction of Space and Time: Mobile Communication Practices*, volume 1, pages 231–250. Transaction Books, Piscataway, NJ, 2009.

Telecentre.org. Foundation www.telecentre.org.

T. S. Tullis. Designing a menu-based interface to an operating system. In *Proceedings of CHI'85*, pages 79–84. 1985.

S. Twinn. An exploratory study examining the influence of translation on the validity and reliability of qualitative data in nursing research. *Journal of Advanced Nursing*, 2008.

UN-HABITAT. Press release. the challenge of slums: Global report on human settlements, 2007.

UN News Centre. For international literacy day un flags key role of reading and writing in global peace. http://www.un.org/apps/news, 2012.

H. Underwood, R. Sterling, and J. Bennett. The design and implementation of the partopen maternal health monitoring system. *ACM DEV*, 2013.

S. Van Linden and A. H. M. Cremers. Cognitive abilities of functionally non-literate persons relevant to ict use. In *International Conference on Computers Helping People with Special Needs*, pages 705–712. Linz, Austria, July 9–11 2008.

S. Wiedenbeck. The use of icons and labels in an end user application program: An empirical study of learning and retention. *Behaviour & Information Technology*, 18(2), 1999.

Windows Phone Cortana. http://www.windowsphone.com/en-us/how-to/wp8/cortana/meet-cortana, Retrieved July 28 2014.

Xbox 360 + Kinect. http://www.xbox.com/en-in/xbox360/index, Retrieved July 28 2014.