2009

Orality-Grounded HCID: Understanding the Oral User

Roni Rosenfeld

Jahanzeb Sherwani

Nosheen Ali

Carolyn P. Rosé

Published In
Information Technologies and International Development,, 5, 4, 37-49.

Research Article

Orality-Grounded HCID: Understanding the Oral User

Jahanzeb Sherwani
jsherwan@cs.cmu.edu
Adjunct Faculty
Language Technologies Institute
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213 USA
(412) 567-1204

Nosheen Ali
nosheen.ali@cornell.edu
Post-Doctoral Researcher
Department of Development Sociology
Cornell University
119 Warren Hall
Ithaca, NY 14853 USA
(650) 996-7122

Carolyn Penstein Rosé
cprose@cs.cmu.edu
Assistant Professor
Language Technologies Institute and Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213 USA
(412) 268-7130

Roni Rosenfeld
roni@cs.cmu.edu
Professor
School of Computer Science
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213 USA
(412) 414-3701

Abstract

While human-computer interaction (HCI) methodologies are designed to be general, they have most often been applied in the context of literate end users in the West. These methodologies may, however, need rethinking for application in HCI for the developing world (HCID) contexts, where many of the basic assumptions that underpin the methods may not always hold true. In this article, we present an overview of one factor that is significantly different in the HCID context—the literacy of the end user—by drawing on the literature of orality, and we offer a framework for HCID methodology that we argue is more appropriate for the HCID context. Based on this framework, we then present guidelines for design and user research methodologies in such contexts, highlighting seminal HCID research that corroborates these guidelines.

Introduction

One of the most valuable contributions from the field of human-computer interaction (HCI) to date is the collection of methodologies for user-centered design that have evolved through generations of design work that focus on the needs of users within their context of work or life. While, at the most abstract level, these methodologies have been designed to be general, they have most frequently been applied to populations within the developed world. The danger in such a population bias is that it may lead to certain blind spots that can result in naïve applications of these practices in the developing world if practitioners do not deeply consider how the practices must be parameterized in order to generalize and be effective with these very different populations. The contribution of this article is to raise awareness to these blind spots by contrasting common assumptions underlying HCI as it is practiced within the developed world with recent findings from work in HCI for the developing world (HCID).

In HCID literature, users unfamiliar with the tools of literacy—reading and writing—are commonly described as illiterate, non-literate, low-literate, or semi-literate, and it is these users that most HCID interventions focus on. We suggest that a better understanding of these users can be achieved through the concept of orality, which has been notably theorized by Walter Ong in his seminal work *Orality and Literacy* (1982). In this article, we use this framework as a lens with which to re-examine common practices from mainstream HCI.

Orality describes how people think, communicate, and learn in a culture where writing has not become internalized. Orality theory argues that writing has so fundamentally transformed consciousness in literate cultures that we—literate researchers—are unable to grasp how oral cultures and people operate. The very categories through which such people are described is a case in point: They are described in terms of what they

are not, such as "illiterate," rather than what they are, that is, oral. Terms like "illiterate" devalue the identity and knowledge of oral cultures by implicitly suggesting that lack of literacy is equivalent to backwardness. Orality theory emphasizes that we need to understand such communities in their own terms, rather than from the perspective and biases of literate users.

Ong discussed orality in the context of oral cultures, how they contrast with literate cultures, and how modern media is creating a "secondary orality" for literate cultures. While many HCI researchers have referenced Ong's work, the focus has been on the implications of orality for literate societies, especially on the notion of secondary orality in such contexts, and there has been little exploration of Ong's work with respect to oral cultures. In this article, we attempt to begin this exploration by leveraging Ong's concept of orality as a foundation on which to build a model for oral users in HCID contexts, and to use this model to develop recommendations and predictions for user interface design, as well as user-study methods in these contexts.

Related Work

Ong (1982) presented an extensive contrast between what he referred to as oral cultures and what he referred to as literate cultures, a contrast on which we base much of our proposal and through which we interpret findings from our and other people's fieldwork in the developing world. Ong himself did not mean his synthesis primarily to be used in the context of work in the developing world, but rather painted a picture of how modern media was creating a secondary orality for literate cultures. In that vein, Ong's work has already played an influential role within the field of HCI, facilitating the application of insights from oral cultures to design within the developed world, increasing so recently in the light of conversational Web 2.0 technologies. Sometimes Ong's insights have been applied at a general level (e.g., Mathur & Karaholios, 2009; Conrad, 2003), simply taking a single quote from his work and using it as inspiration for an otherwise bottom-up, user-centered design process. Several researchers in the past decade have discussed the emergence of Web 2.0 technologies (e.g., Stahl, 1999; Fischer, 2006; Wright, 2008; Bergstrom & Karahalios, 2009), pointing to the conversational natural of those media, and framing those media within the vision for a secondary orality. Other times, his framework has been adopted or scrutinized at a more detailed level (e.g., Davis, 1997; Grudin, 2007; Wright, 2008), sometimes challenging or extending Ong's ideas. In much of this work, Ong's vision has not been adopted wholesale, but rather has been used as a catalyst for discussion, often about the future of technology. For example, Grudin (2007) added a third distinction to Ong's two-way distinction between oral cultures and literate, or digital, cultures, and used this to make predictions about trends in computer usage by modern youth. Ackermann & Decortis (2007) similarly explored Ong's ideas in connection with e-literacy and youth, looking at opportunities for youth to engage in storytelling behaviors on the Web. Davis (1997) reinterpreted some of Ong's ideas in terms of what they might mean for video as a form of communication. In contrast to this body of work, which took insights from the developing world and used them to provide new insights into design for the developed world, our work brings these insights full circle, applying them within their context of origin, the oral cultures themselves.

One of the initial efforts to create design recommendations for illiterate users drew primarily from literature on screen-readers for the visually impaired and from sighted users working without visual displays, since "human computer interaction and user-interface design literature provided little specific resources for designers interested in building software for illiterate users or members of the developing world" (Huenerfauth, 2002). However, both visually impaired users in the developed world and users working without visual displays are literate groups of users. Furthermore, these strategies were never tested on actual oral end users.

Within the field of evidence-based HCID research, there has been much work on the specifics of individual projects and artifacts, both in terms of interface design and evaluation, such as HealthLine (Sherwani et al., 2007; Sherwani et al., 2009) and Digital Green (Gandhi et al., 2007). There is also a substantial body of work that goes beyond the specifics of one artifact, and aims to provide generalized guidelines for interface design and evaluation grounded on empirical, field-based research, including the following: design guidelines for text-free interfaces (Medhi et al., 2006), user interface (UI) strategies for low-literacy users (Parikh et al., 2005), full-context video (Medhi & Toyama, 2007), and the

Bollywood and auto-rick radio measurement methods (Chavan, 2007). While these research contributions do not provide a theoretical framework that explains why these methods work well, they do provide valuable insights that can be used to build such a theoretical framework. We will now summarize some of the findings from these projects to provide context for the discussion in the remainder of this article.

Digital Green

In the Digital Green project (Gandhi et al., 2007) for agricultural extension, researchers at Microsoft Research in India designed, developed, and evaluated a mechanism for generating and disseminating video-based instructional content for teaching agricultural best practices to farmers. While watching a video did not always lead to behavioral change on its own, the researchers' "finding was that mediation [through a local facilitator] is essential to the process of extension that farmers were most convinced by appropriately targeted and pitched content." Also, "village mediators encourage these farmers to share their personal experiences to motivate their peer groups." Finally, during video screenings, "viewers frequently ask for the names and villages of recorded farmers."

Full-Context Video

The "full-context video" approach to providing tutorials for an interface (Medhi & Toyama, 2007) has been shown to lead to remarkably improved task success rates; one out of 17 participants was able to complete the job-finding task without the help of a full-context video, while 18 out of 18 participants who were shown the video were able to complete the same task—an improvement from 6% task success to 100%. In that study, the full-context video showed not only the mechanics of how to complete the job-search task, but more importantly, also presented a dramatization of the context of the task: An employer uses an interface to post a job, a job-seeker is told by a friend how to use the interface to find the job, and the story ends with the job-seeker finding the employer and getting the job. Medhi and Toyama concluded:

> Overall, the video appeared to instill a great amount of confidence among the test participants. Subjects were seen to gesture in agreement, to smile, and to laugh at various points in the video. It was as if the video provided a shift in concern, from anxiety about how to use the device to concerns about the content itself. . . . Subjects who saw the video first appeared more determined—eyebrows knit, body leaning into the monitor—when moving onto the application itself, as compared with the control group . . . More than one subject explained that they had never understood what computers were for, but that the video showed why people use them. The dramatized video made this clear (p. 2).

HealthLine

In our own work (Sherwani et al., 2009), we have been designing, developing, and evaluating speech-based health information access interfaces for low-literate community health workers. Our research has been conducted in rural parts of the Sindh province in Pakistan, where we have conducted user studies with up to 20 participants in each, all of them community health workers. These community health workers are often not able to read at all, or do so with great difficulty.

We have seen how human-guided tutorials—where the participant uses the interface with the active guidance of the facilitator, as a training tool—outperform verbal instruction, written instruction, and even video-based instruction (the video we tested was not full-context video). Further, we have seen how oral participants have greater difficulty comprehending the same material even when they speak the language as fluently as their literate counterparts, and how content designed for literate consumption needs to be significantly adapted to be used by oral users. It was our experience in these studies, and the difficulty in applying the tools used in standard research and design practice, that led us toward the work on orality.

We now turn to a discussion of orality, to present a framework for oral users, after which we will present our recommendations and predictions, in the context of existing work.

Orality versus Literacy

In this paper, we argue that HCID research can be substantially improved by grounding it in an understanding of orality. While there is a continuum between completely oral users and completely literate ones, in this paper we will use the term "oral user" to refer to aspects of completely oral users that are present even in somewhat literate people.

To design for oral users, we must first understand their cultural practices of community knowledge-building and transmission, including how they (a) organize and transmit information, (b) learn infor-

mation, and (c) remember information. Additionally, we need to grasp what Ong (1982) referred to as the "psychodynamics of oral thought" to understand the fundamental differences between oral and literate users as individuals within that cultural context. Below, we expand on each of these aspects.

Cultural Practices of Knowledge Building and Transmission

1. How information is organized and transmitted

A key difference between oral and literate cultures is how knowledge is organized and transmitted. The sum of an oral community's knowledge—its history, its identity, its culture, and its whole way of life—is represented and remembered in the oral tradition, often referred to as a culture's folklore. Among other dimensions, this oral tradition comprises stories, myths, proverbs, riddles, poems, and songs. These forms provide a pool of wisdom that individuals can draw on as a means of daily survival, as well as deal with new situations.

Notably, this knowledge is not statically stored; rather, it is kept alive through continual verbal, extempore performance. Storytellers dynamically render different parts of this knowledge at different occasions, altering the content based on the current political scenario, social sentiment, and immediate audience. Since this knowledge is never written down, it is fixed neither in form or content, nor is it repeated word for word. Hence, knowledge is dynamic and evolving, fluid and creative, and it is composed iteratively over generations.

One of the striking characteristics of oral knowledge is that it "knows no lists or charts or figures" (Ong, p. 97). The emphasis is on narrating events in time, with a correspondence between these narratives and human experience.

2. How information is learned

According to Ong (1982), "human beings in primary oral cultures . . . learn by apprenticeship . . . by discipleship, by listening, by repeating what they hear, by mastering proverbs and ways of combining and recombining them, by assimilating other formulary materials, by participation in a kind of corporate introspection—not by study in the strict sense." What is especially relevant to HCID practitioners is what is not mentioned in the above quotation: Oral people do not learn from a neutral, stand-alone object, such as a book, or automated system, which contains a set of abstract instructions to be applied across situations; rather, they learn "in situ," embedded in concrete situations and practical experience.

3. How information is remembered

Since information in oral cultures is transmitted through oral delivery—and hence it is recited and heard—it needs to be presented in a form that is conducive for aural-oral reception and retention. Ong elaborated:

> In a primarily oral culture, to solve effectively the problem of retaining and retrieving carefully articulated thought, you have to do your thinking in mnemonic patterns, shaped for ready oral recurrence. Your thought must come into being in heavily rhythmic, balanced patterns, in repetitions or antitheses, in alliterations or assonances, in epithetic and other formulary expressions, in standard thematic settings . . . in proverbs which are constantly heard by everyone so that they come to mind readily and which themselves are patterned for retention and ready recall, or in other mnemonic form. Serious thought is intertwined with memory systems. Mnemonic needs determine even syntax. (Ong, p. 34)

In short, the usage of rhythm, repetition, locally known idioms, and dramatized settings are key to keeping orally constructed knowledge alive and remembered. Even though these mechanisms can aid memory in literate contexts, the ability to store information externally through the technology of writing means that such mechanisms are not as crucial in literate contexts as they are in oral contexts.

Psychodynamics of Oral Thought (Contrasted with Literate Thought)

1. Additive, not subordinative

While the preferred grammatical form of literate writing is subordinative, oral thought prefers additive forms. Ong's example from an older translation of the Bible—one considered to have a large amount of oral residue—contrasts neatly with a modern translation. The older translation reads:

> In the beginning, God created heaven and earth. And the earth was void and empty,

> and darkness was upon the face of the deep; and the spirit of God moved over the waters. And God said: Be light made. And light was made. And God saw the light that it was good; and . . . (Ong, p. 37)

The word "and" is used to join phrases and sentences liberally, which is not the case with a modern translation:

> In the beginning, when God created the heavens and the earth, the earth was a formless wasteland, and darkness covered the abyss, while a mighty wind swept over the waters. Then God said, "Let there be light", and there was light. God then separated the light from the darkness . . . (Ong, p. 37)

Here, the use of conjunctions—when, while, then—to join phrases is the common case, with a more hierarchical and subordinative organization than the additive pattern shown in the previous passage. Ong argues that the former example is as natural to the oral mind as the latter example is to the literate mind.

2. Aggregative, not Analytic

Oral tradition favors the use of reusable formulas. Thus, it is not just "the soldier," but "the brave soldier"; not just "the princess," but "the beautiful princess"; and not just "the oak," but "the sturdy oak." These clusters aid memory by creating archetypal concepts that are repeatedly used in narratives. Furthermore, according to Ong:

> Traditional expressions in oral cultures must not be dismantled: it has been hard work getting them together over the generations, and there is nowhere outside the mind to store them. . . . Once a formulary expression has been crystallized, it had best be kept intact. Without a writing system, breaking up thought—that is, analysis—is a high risk procedure. (p. 39)

These reusable formulae are, in some ways, the fundamental building blocks of oral tradition.

3. Redundancy

All spoken information is ephemeral, and if information were never repeated, even the slightest loss of concentration would mean a complete misunderstanding on the part of the listener. Redundancy is necessary to ensure that the speaker and listener both remain on track. When reading, however, it is not necessary to repeat information, as the reader can always backtrack to re-read earlier material—to the point where this becomes second nature. The differing economies of hearing and reading result in a significant difference between the outputs of oral and literary content in terms of the redundancy inherent in the content. Even though redundancy may be a natural part of the thought process, "with writing, the mind is forced into a slowed-down pattern that affords it the opportunity to interfere with and reorganize its more normal, redundant processes" (ibid., p. 40).[1]

4. Conservative/Traditionalist

A corollary of redundancy is conservatism. Since oral knowledge vanishes unless repeated again and again, oral cultures place a premium on repeating previously held knowledge rather than experimenting and discovering new knowledge. To the technologically minded, this can best be explained with an analogy to hard drive capacities. When hard drive space was expensive, one had to keep only a limited amount of data, which was most likely static; however, once capacity became cheap, it was easier to keep storing newer and newer data, as there was no need to choose between the old and the new. Oral cultures function with the former mindset.

This is not to say that oral cultures never update their knowledge store or that they lack originality:

> Narrative originality lodges not in making up new stories but in managing a particular interaction with this audience at this time—at every telling the story has to be introduced uniquely into a unique situation, for in oral

1. The content used by NGOs for information dissemination is often in a highly summarized format in an attempt to crystallize the salient points that need to be conveyed. However, this is not the optimal format to present information to oral users. This has been our experience in the HealthLine project, and researchers at the Meraka Institute in South Africa faced the same issue in the OpenPhone HIV Information Line project (Aditi Sharma Grover, personal communication).

> cultures an audience must be brought to respond, often vigorously. But narrators also introduce new elements into old stories." (ibid., p. 41)

5. Close to the Human Lifeworld

> In the absence of elaborate analytic categories that depend on writing to structure knowledge at a distance from lived experience, oral cultures must conceptualize and verbalize all their knowledge with more or less close reference to the human lifeworld, assimilating the alien, objective world to the more immediate, familiar interaction of human beings. . . . An oral culture has no vehicle as neutral as a list. . . . [It] likewise has nothing corresponding to how-to manuals for the trades. (ibid., p. 42)

Trades are learned primarily through apprenticeship, involving observation and practice, with minimal verbalized knowledge, if any. Furthermore, oral cultures do not preserve knowledge of skills in abstract, self-contained corpora. Finally, oral people do not treat all human lifeworld-based information as equally important; it is their specific lifeworld that is relevant, and hence, important. For instance, they give specialized names to details that matter in their specific lives, but generalize others that do not (e.g., "that is merely a flying animal").

6. Agonistically Toned

To literate individuals, oral people might appear to be extraordinarily agonistic or argumentative. This is because stories and proverbs—the basis of oral tradition—are not meant to be just a store of knowledge, but also a means to engage a dialectic dialog. It is this argumentation that Ong considers to be the predecessor of the dialectic method of Socrates and Plato. Thus, oral thought is often instantiated as an interplay of competing ideas.

7. Empathetic and Participatory

With literacy, knowledge is disengaged and is supposedly "objective." In an oral culture, according to Ong, "learning means achieving close, empathetic, communal identification with the known." Oral people usually do not memorize, other than for rituals; instead, learning involves an amalgamation of the new with the self.

> Literates are usually surprised to learn that the bard planning to retell the story he has heard only once wants often to wait a day or so after he has heard the story before he himself repeats it. In memorizing a written text, postponing its recitation generally weakens recall. An oral poet is not working with texts or in a textual framework. He needs time to let the story sink into his own store of themes and formulas, time to 'get with' the story. In recalling and retelling the story, he has not in any sense 'memorized' its metrical rendition from the version of the other singer. (ibid., p. 59)

8. Homeostatic

On the order of generations, oral cultures eliminate memories that are obsolete, simply by not repeating that information. As a consequence, oral cultures never need dictionaries, since all words in usage are commonly known and understood by everyone in the culture. The number of words never grows unmanageably, as obsolete words are pruned away. Literate cultures, on the other hand, can access all words from the present day back thousands of years ago, and require dictionaries to store and access these words.

9. Situational, not Abstract

Perhaps the most fundamental difference between oral and literate cultures is that of situational thinking versus abstract thinking. Because oral knowledge is rooted in the human lifeworld, oral people are most comfortable in thinking and learning situationally, instead of in the abstract. This does not mean that oral people cannot think categorically; it is just that they make categories differently than literate people. For example, when asked to categorize a set of objects, oral people categorize based on features important to the human lifeworld (e.g., usefulness) as opposed to abstract features (e.g., function). For instance, Ong wrote that:

> Subjects were presented with drawings of four objects, three belonging to one category and the fourth to another, and were asked to group together those that were similar and could be placed in one group or designated by one word. One series consisted of drawings of the objects hammer, saw, log, hatchet. Illiterate subjects consistently

> thought of the group not in categorical terms (three tools, the log not a tool) but in terms of practical situations—"situational thinking"—without adverting at all to the classification "tool" as applying to all but the log. If you are a workman with tools and see a log, you think of applying the tool to it, not of keeping the tool away from what it was made for—in some weird intellectual game. (Luria, 1976 in Ong, 1982)

Further, any form of thought other than grounded "operational thinking" is likely to be considered "not important, uninteresting, trivializing." (Ong, 1982)

Operational thinking is also the mode in which oral individuals interpret reality. For example, Luria found that oral people interpreted circles as a plate, sieve, bucket, watch, or moon, although school-going children readily identified them as circles (Luria, 1976).

Likewise, oral people do not think in formal, syllogistic logic. For instance, when given the query, "precious metals do not rust, gold is a precious metal—does it rust or not?" one oral respondent said: "Do precious metals rust or not? Does gold rust or not?" Another said, "Precious metal rusts. Precious gold rusts." Syllogisms are special riddles where the conclusions are derived from the given premises alone. Ong said:

> Persons not academically educated are not acquainted with this special ground rule but tend rather in their interpretation of given statements, in a syllogism as elsewhere, to go beyond the statements, as one normally does in real life or in riddles (common in all oral cultures). (p. 53)

Furthermore, requests for definitions are resisted in oral cultures. When asked to explain what a tree is, one oral peasant replied, "Why should I? Everyone knows what a tree is, they don't need me telling them." Ong's response to this is,

> Why define, when a real-life setting is infinitely more satisfactory than a definition. . . . The peasant was right, there is no way to refute the world of primary orality. All you can do is walk away from it into literacy. (p. 53)

One might argue that the above questions were not asked in the correct context. However, it appears that there is no conceivably "correct" way to ask oral people such questions. In fact, in oral cultures, intelligence is something that is measured based on a person's actions or skills, not by their response to verbal questions in a test: "Written examination questions came into general use (in the West) only well after print had worked its effects on consciousness, thousands of years after the invention of writing" (ibid., p. 55). Thus, asking abstract questions to test someone's knowledge without any social context is an alien concept in an oral culture.

It is important to note that just because oral people think situationally or concretely does not mean they are incapable of abstract thought. Indeed, there are complex layers of meaning in oral knowledge that are captured in idioms, folklore, mythology, riddles, and historical narratives.

Orality-grounded HCID

Drawing on Ong's conceptualization of orality, we suggest the framework of orality-grounded HCID. By drawing on a specific focus on oral end users, this framework calls for a fundamental rethinking of HCID both in terms of design principles and user research methodology. While the corollaries of the framework echo the sentiments of many user-centric methodologies popular in HCID research—such as user-centered design, participatory design, and participatory action research—there is still a new story to tell here in that the proposed framework attempts to explicitly take into consideration how such users think, learn, and process information. Orality-grounded HCID is based on a testable, and falsifiable, model of oral users, and thus enables HCID researchers to engage in more grounded research and design.

Below, we outline our proposal in terms of principles of design and methodological considerations. These guidelines make explicit the fundamental differences between mainstream HCI practice and orality-grounded HCID at both levels.

Principles for Interface and Information Design

When designing interfaces for information access by oral users, several dimensions of orality are useful to consider:

Information needs to be rooted in common experience with specific examples. Abstract descriptions, such as "breastfeeding should be con-

tinued even when the child has diarrhea" is not as effective as describing a specific example of a mother who is faced with this particular issue. Ideally, new information should be described in terms of familiar cultural memes, and preferably using the culture's own oral formulae. Thus, instead of using a generic "mother," the information should draw on existing characters—perhaps a widely known maternal character in this case—in the community's folklore. Furthermore, drawing on a culture's oral formulae is best left to local members of the culture. IDE hired local storytellers to write and perform a play centered on a treadle pump (Polak, 2008), which was an excellent strategy since local storytellers understand their culture's oral formulae better than anyone else.

Narrative stories are more memorable and more effective at conveying information than neutrally listed bullet points. Oral cultures do not have neutral lists. Ong wrote, "Even genealogies out of such orally framed tradition are in effect commonly narrative. Instead of a recitation of names, we find a sequence of 'begats,' of statements of what someone did: 'Irad begat Mehajael, Mehjael begat Methusael, Methusael begat Lamech' (Genesis 4:18)" (Ong, p. 98). Thus, listing bullet points of information, such as how to improve cow yield, will not work as well as telling the story of a farmer who used a specific method and how he was able to increase his yield (Gandhi et al., 2007). Moreover, dramatic descriptions will work better than neutral ones. IDE created a full-length movie that showed the value of a treadle pump through a dramatic storyline (Polak, 2008), which was more effective than a neutral description of how such a pump could improve a farmer's bottom line. The full-context video method (Medhi & Toyoma, 2007) leveraged the same principle.

Rhythm aids recall. IDE hired local troubadours to compose a song about a treadle pump and to perform it at farmers' markets and fairs in Bangladesh (Polak, 2008). Even without the creation of an entire song, content with rhyme and alliteration is likely to be both understood and remembered by oral users more effectively than prose.

Linguistic style should be structured additively, not hierarchically. While it is common in literate material, using subordinative conjunctions such as "while," "then," "since," and "although" is uncommon in oral users' communication. Complicated sentence structures impose a cognitive load on the user. Instead, it is preferable to use coordinating conjunctions, such as "and," "or," and "so," which do not create hierarchy.

Redundancy needs to be embedded in the content. Redundancy is an important part of oral communication, mainly because of the ephemeral nature of speech. The user should also be given ample opportunity to request repeated presentations of content that has been given before. It should be noted, however, that explicitly requesting repetition may be less natural for the oral user than having the correct amount of redundancy already embedded within the content. In the Digital Green project, researchers found that farmers would ask for the video content to be repeated multiple times; our prediction is that a more optimal solution would be to design the video content to have redundancy "baked in" the content itself.

Each and every word needs to be understood. In an oral community that speaks one common language, there are no unfamiliar words, and oral users never face unknown words in daily life. Thus, even one unfamiliar word can confuse the user completely, and care should be taken to ensure that no such words exist in the system's content. This is common when content comes from experts, who are literate and have a more diverse linguistic background than an oral person. This strongly suggests that locally generated content—as in the case of the Digital Green project—is likely to be the most effective option in many cases, in terms of its relevance and understandability.

Abstract categories should be avoided. Oral users do not categorize the same way as literate users, and do not think in terms of abstract categories (Luria, 1976). The use of categories should be minimized and if their use is essential, it should be kept in mind that the designer's choice of categories will most likely not match the expectations of the user. Also, hierarchies in information architecture should be avoided. Browsing multiple depths of information (e.g., as in a Web page, or when using navigation metaphors of "up a level") is difficult for oral users (Deo et al., 2004).

Requiring adherence to specific spoken words or phrases is less likely to succeed. Oral users perceive speech as a continuous stream, rather than

as discrete words. In our work with speech recognition interfaces, one of the most common complaints of oral users is that they are afraid that they will "say the wrong thing"; they understand they have to speak in individual words, an action that is unnatural for them, and their comments reflect their discomfort with this requirement. While it is technically difficult, it may be necessary to widen the speech recognizer's grammar to allow for more natural utterances, rather than forcing the user to speak in single word commands.

Oral people do not internalize new information the same way as literate people. Since internalizing new information is comparatively expensive for oral people—they cannot offload their memory requirements onto the technology of writing—it appears that they are more selective when choosing whether to internalize new information. In our research, we have seen many times that even when users understand each word in the content perfectly, when asked a question on that content, users will respond based on prior knowledge. In a study we conducted with health workers of varying literacy levels, even after reading a specific paragraph in their manual fluently, their response to a question on that content was the opposite of what it said, based on prior knowledge (Sherwani et al., 2007). Similar effects have been found in other research on oral people (Luria, 1976) and poor readers (Doak, 1995), and have also been anecdotally reported by other HCID researchers (Etienne Barnard, personal communication). This might happen because the formal content is not seen as relevant, or it isn't organized in the same linguistic style that users expect (e.g. narrative, concrete examples rather than neutral and abstract information), or for other reasons; however, it is important in the design of an interface to know that this is a problem that needs to be engaged.

Oral people give more importance to the source of information than literate people. Writing establishes "context-free language" (Ong, 1982), where information is not linked to any particular source. For oral people, however, all information is social and traceable to a person. When farmers were shown videos of other farmers demonstrating agricultural best practices, the most common question the viewers asked was what the demonstrating farmer's name was, and which village he was from (Gandhi et al., 2007). This may also explain why source-neutral information—of the form presented in a book or in a persona-less system—is not internalized: It is missing the essential feature of social context, not only of the information's relevance to the real world, but also its traceability to a trustworthy human being. Thus, systems, such as Digital Green, that provide information through people perceived to be trustworthy are likely to be more successful than systems that transmit information without this feature.

Given all the recommendations listed above, it becomes apparent that information designed for literate users—virtually all written material—is not appropriate for oral consumption. Oral and literate users require content with different organization, presentation, and context.

Furthermore, end users may be the best resource for content creation and content adaptation. Literate people produce both written and spoken content optimized for literate people (Ong, 1982, p. 56). Even "human access points" (Marsden et al., 2008), or literate technologists from the local community, may not understand how to effectively alter content for optimal consumption by oral users, although they would do a better job than a naïve, non-local designer. However, it may be the case that involving users is essential for content creation and content adaptation. The success of the Digital Green project is a case in point; it doesn't just involve users in the creation of the content, but features them as the star performers in the "Farmer Idol" videos that form the backbone of the system.

Methodologies for User Research

Even more basic than principles for design, methodologies for conducting user research must also be adapted for application within oral cultures. We can learn from the significant number of successful evaluations of designed artifacts in the context of oral cultures, projects such as Digital Green (Gandhi et al., 2007), eSagu (Reddy et al., 2007) and Warana Unwired (Veeraraghavan et al., 2007). In these and other related efforts, the common experience has been that standard user studies have various layers of problems. Ultimately, it may be the case that a fundamentally different method of evaluation needs to be explored for the HCID context. In this section, we explore the issues with existing methodologies and make some suggestions for how the methodologies can be adapted in light of the differences between oral and literate cultures.

Overall, standard user study practice needs to be considerably rethought for application in oral contexts. Even though literate participants may find user studies a novel concept, they are likely to have experienced the individual components of a user study in school, and would likely consider them somewhat familiar. For oral people, a user study is an alien experience. The various steps involved in a user study—the facilitator reading out instructions to the participant, teaching the steps in the task, and then asking the participant to perform it themselves to answer examination-style questions—is a clinical abstraction that is alien to the lifeworld of a typical oral person. Oral people do not think and learn in the way that user studies expect them to, and they are not used to being asked abstract, context-free questions. Thus, it is arguable whether results from such studies are of much analytical value in oral contexts. Further work is needed to explore alternative methods for user studies adapted to the HCID context.

The experimenter always has a higher socio-economic status than the participant and is usually foreign to the local community, leading to feelings of intimidation and "performance anxiety." This is often an overlooked feature of HCID user studies and deserves a fundamental rethinking of user study roles. Recruiting and training local user study facilitators from the community have worked well in our experience. Moreover, involving the facilitator in the motivation and design of the user study is useful for a number of reasons, as it increases their ownership of the process, and provides them the necessary context for why each step of the user study is important, in addition to listing the actual steps they need to perform.

When presenting a novel interface, the context in which it is to be used and the motivation for the system should be presented as concretely and vividly as possible. Given oral users' preference for situationally grounded examples relevant to their lifeworld, it is imperative to present the system as a solution to a problem that users deem important and for which users perceive current solutions as inadequate. Whenever possible, the system should be introduced as a solution to a widely remembered, specific instance of a local problem. For example, in health care, there may be a recent story of a person affected by an illness that no one knew how to cure, which everyone remembers vividly, perhaps because of its tragic nature. By asking locals about such examples beforehand and then portraying a health information access system within a narrative that ties into the local example of a problem, as the potential solution, users may see the social context of the system more clearly than they would if it were to be given as a solution to the problem in general (e.g., "Imagine if a person in your village became ill . . ."). That approach is still better than a generic solution in the abstract (without a narrative, e.g., "The system can be used to give health information when the need arises"). A motivation presented with a narrative, tied to the specific local folklore of the community, could make the intervention seem more relevant to the needs of the community.

This suggests why the full-context video method (Medhi & Toyama, 2007) works so well: It creates a dramatized and visual example of how the interface could be useful, instead of introducing it in abstract. Given the need to localize all content, it is unclear whether a video would be more cost-effective at providing context than other alternatives (e.g., face-to-face interaction) with a narrative that is customized to the specific experiences and memory of the user or community. Given the high success rate of the full-context video method (100% of the users exposed to the video were able to successfully complete the task, compared to 6% in the control group), full-context video is clearly a winning method, and may be the optimal choice.

Oral participants do not remember neutral user study tasks, and do not actively engage with them. It is a known issue that artificial scenarios and fictitious needs are unlikely to be internalized by participants. For oral users, tasks cannot be presented textually—and so cannot be reviewed at will—and furthermore, if the task is abstract and neutral, which is standard HCI methodology, it is unlikely that the task will be remembered, much less internalized. Engaging with a task that one does not remember or believe in is also unlikely. The Bollywood method (Chavan in Shaffer, 2004) is an excellent solution to this problem. In this method, user study tasks are couched in a dramatized narrative, which leads to more involved and interactive participants, who would otherwise not contribute any feedback when faced with neutrally presented tasks. Thus, oral users are more likely to remember and engage with a dramatic narrative than with an

abstract and neutral request for information. It is unclear how much of the difference is attributable to (a) participants' improved memory of the dramatized tasks, (b) their appreciation for these tasks' wider social context and relevance to the human lifeworld (an issue of motivation), and (c) their perceived freedom to critique when in role play (what Chavan attributes its success to), or other factors. Further work is needed to explore what aspects of this method work well, and why.

Giving user study participants descriptive instructions on how to use the system is not sufficient. Oral people do not learn by verbalized instruction—rather, they learn through apprenticeship and practice. This is neatly analogous to the difference between declarative and procedural knowledge in the field of artificial intelligence. In our own work, we have found that a mentor-apprentice model, in which a facilitator does a guided walkthrough of a sample task with a participant, works well. Also, participants' understanding of the system improves further with an "incrementally-removed training wheels"-approach with multiple practice tasks, where explicit help is provided on each step for the first sample task, and is gradually reduced for each subsequent practice task, until no help is provided on the final sample task unless required. The approach used in the full-context video method is different, and we hypothesize that oral users would learn better with a personal, mentored learning experience than by watching a video demonstrating another person's use of a system. Further work is needed to ascertain the best strategy for teaching an interface to the user—both for the design of the system and for evaluating the system in a user study.

Likert Scales are an inappropriate measurement tool for oral people. A Likert Scale question consists of an abstract statement, such as "I found this system easy to use," and a list of numbered options, usually 1 through 7, with 1 representing "strongly disagree" and the highest number (e.g., 7) representing "strongly agree." Participants are asked to select the option that most closely describes their level of agreement with the given statement. When used with oral people, this tool is problematic for a number of reasons:

- The Likert Scales method requires that the options to be written down and presented textually; however, oral people cannot read and verbal presentation of Likert Scales has not been rigorously proven as a valid methodology, even though they might be used in practice.
- The concept of "context-free" abstract statements is a literate construct. Oral people do not conceive of statements in this form—there has to be someone who said any given statement. Moreover, who this someone is makes a difference as to how likely they are to agree with the statement; issues of credibility, trust, authority, politeness, political considerations, etc., are involved.
- Categorizing agreement into discrete chunks is also a literate construct. Oral people do not think in terms of such categories, and we have seen countless examples of oral user study participants confused at such categorization.
- Praise and criticism are often given indirectly in oral cultures, and thus asking participants to directly praise or critique an artifact with their subjective evaluation might not align with participants' expectations.

In our own work, we have had significant problems with Likert Scales, where participants often gave polite, positive answers, which on deeper examination and triangulation were not representative of their true opinions. In one user study we conducted in rural Pakistan, oral participants were unable to speak the language of the speech interface they were using and could not correctly answer any of the health information questions they were asked to learn from the system. Yet, when asked to rate the system in terms of various dimensions—including ease-of-use and understandability—they rated it very highly (Sherwani et al., 2009). Similar experiences have been reported by other researchers in South Africa (Aditi Grover, personal communication).

The "auto-rick radio" method (Chavan, 2007) provides an interesting alternative to Likert Scales. Instead of representing a subjective evaluation on a linear scale, participants are asked to twist a knob, shaped like a radio's volume knob, to represent their feelings on a particular dimension. Chavan wrote that:

> For these users, the concept of a difference in degree (moving from negative to positive) being represented by a horizontal straight line seemed very conflicting. The feeling was that if the different

> points in the scale represented different degrees of an attribute, then they could not appear to be on the same level, as they did with the straight horizontal line. . . . Hence a knob control was devised which resembled the volume control knob of the radio that all users were very familiar with. (Chavan, p. 13)

However, this explanation confounds various issues: the difficulty in quantifying emotion for oral users, the difficulty in relating emotion to spatial terms, and the difficulty of expressing their choice using a novel interface. By providing a familiar interface (the radio volume knob), participants may have found it easier to express themselves, but whether their expression correctly represented their subjective emotions remains to be seen. Rigorous methodological research is required to explore these issues in detail.

One potential alternative to Likert Scales in general is to instead solicit open-ended answers and have the experimenter decide on the appropriate category—after all, what is the point of asking a participant to do something, like quantify an emotion, that that person really cannot do? Chavan suggested "using storytelling to find dissonances" as a means of asking users about their experience with a particular product (ibid.); however, this method could be extended to subjective measurement in which participants could be asked, for example, to narrate instances where they had trouble performing the user study task, as an alternative to asking them to subjectively rate its difficulty. The experimenter could use the quantity and quality of these instances to determine a quantification of the user's experience, if desired. However, if the participant and experimenter are both present in the same room during the study, the participant would likely wonder why the experimenter is asking such questions, since surely the experimenter saw the participant's performance and should obviously then know the answers. As such, care would need to be taken to set up the scenario appropriately—perhaps by having a different experimenter conduct the subjective evaluation—to ensure that the participant views these questions as valid and worthy of a real response.

Thus, it can be seen that standard user studies have various layers of problems, and while the above discussion points out some suggestions for how the methodologies can be adapted, it may be the case that a fundamentally different method of evaluation needs to be explored for the HCID context. Alternatively, testing deployed interfaces "in the wild" should yield more representative results than user studies would, and more so than in literate contexts. There are a significant number of research projects designed this way, such as Digital Green (Gandhi et al., 2007), eSagu (Reddy et al., 2007) and Warana Unwired (Veeraragavan et al., 2007), and this may be a better, though in most cases more difficult, method of evaluation.

Conclusion and Future Work

In this paper, we have argued that because a disproportionate amount of HCI practice to date has focused on literate users, it is important to understand that standard HCI practice is not always applicable in oral contexts. Orality theory provides a unique lens with which to understand oral users, and we have attempted to synthesize a framework that provides guidelines and testable predictions for design and evaluation with such contexts. We have also discussed best practices in HCID based on existing research in the context of the framework. Much work is needed, however, to substantiate, qualify, verify, refute, and build on this work. ■

References

Ackermann, E. K., & Decortis, F. (2007). Kids story "writers": POGO, tell-tale, sprite. *Proceedings of the 6th International Conference on Interaction Design and Children,* pp. 161–162, Aalborg, Denmark.

Bergstrom, T., & Karahalios, K. (2009). Conversation clusters: Grouping conversation topics through human-computer dialogue. *Proceedings of the 27th International Conference on Human Factors in Computing Systems,* pp. 2349–2352, Boston, MA, USA.

Chavan, A. (2007). Around the world with 14 methods. *Human Factors International.* White paper retrieved on October 10, 2008, from http://www.humanfactors.com/downloads/whitepapers.asp#bollywood

Conrad, E. (2003). Aether: Reading with tactile vision. *ACM SIGGRAPH 2003 Sketches & Applications,* p. 1, San Diego, CA, USA.

Davis, M. (1997). Garage cinema and the future of media technology. *Communications of the ACM,* 40(2), pp. 42–48, New York, NY, USA.

Deo, S., Nichols, D., Cunningham, S., & Witten, I.

(2004). Digital library access for illiterate users. *Proceedings of 2004 International Research Conference on Innovations in Information Technology.,* pp. 506–516, Dubai, UAE.

Doak, C., Doak, L., & Root, J. (1996). Teaching patients with low literacy skills. Retrieved on September 10, 2008, from http://www.hsph.harvard.edu/healthliteracy/doak.html

Fischer, G. (2006). Distributed intelligence: Extending the power of the unaided, individual human mind. *Proceedings of The Working Conference on Advanced Visual Interfaces,* pp. 7–14, Venezia, Italy.

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

Gandhi, R., Veeraraghavan, R., Toyama, K., & Ramprasad, V. (2007). Digital green: Participatory video for agricultural extension. *Conference proceedings of ICTD 2007,* pp. 21–30, Bangalore, India.

Grudin, J. (2007). Living without parental controls: The future of HCI. *Interactions, 14*(2), 48–52, New York, NY, USA.

Luria, A. (1976). *Cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Marsden, G., Maunder, A., & Parker, M. (2008). People are people, but technology is not technology. *Philosophical transactions of the Royal Society,* A:1–10, London, UK.

Mathur, P., & Karahalios, K. (2009). Visualizing remote voice conversations. *Proceedings from the 27th International Conference on Human Factors in Computing Systems,* pp. 4675–4680, Boston, MA, USA.

Medhi, I., Sagar, A., & Toyama, K. (2006). Text-free user interfaces for illiterate and semi-literate users. *Proceedings from ICTD 2006: International Conference on Information and Communication Technologies and Development,* pp. 72–82, Berkeley, CA, USA.

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

Ong, W. (1982). *Orality and literacy: The technologizing of the word.* London: Routledge.

Parikh, T., Ghosh, K., & Chavan, A. (2003). Design studies for a financial management system for micro-credit groups in rural India. *Proceedings of the 2003 Conference on Universal Usability,* pp. 15–22, Vancouver, BC, Canada.

Polak, P. (2008). *Out of poverty.* Berrett-Koehler Publishers. San Francisco, CA, USA.

Reddy, P., Ramaraju, G., & Reddy, G. (2007). *eSagu™: A data warehouse–enabled personalized agricultural advisory system.* Paper presented at 2007 ACM SIGMOD International Conference on Management of Data, pp. 910–914, Beijing, China.

Shaffer, E. (2004). *Institutionalization of usability.* Addison Wesley Professional. Boston, MA, USA.

Sherwani, J., Ali, N., Mirza, S., Fatma, A., Memon, Y., Karim, M., et al. (2007). HealthLine: Speech-based access to health information by low-literate users. *Proceedings for 2007 IEEE/ACM International Conference on Information and Communication Technologies and Development.,* pp. 131–139, Bangalore, India.

Sherwani, J., Palijo, S., Mirza, S., Ahmed, T., Ali, N., & Rosenfeld, R. (2009). Speech vs. touch-tone: Telephony interfaces for information access by low literate users. *Proceedings for ICTD 2009: Information and Communications Technologies and Development,* pp. 447–457, Doha, Qatar.

Stahl, G. (1999). Reflections on WebGuide: Seven issues for the next generation of collaborative knowledge-building environments. *Proceedings of the 1999 Conference on Computer Support for Collaborative Learning,* Article 73, Palo Alto, CA, USA.

Veeraraghavan, R., Yasodhar, N., & Toyama, K. (2007). *Warana unwired: Replacing PCs with mobile phones in a rural sugarcane cooperative.* Paper presented at ICTD 2007: Information and Communication Technologies and Development, pp. 89–98, Bangalore, India.

Wright, A. (2008). Primal Actions, *Interactions, 15*(1), 11–12, New York, NY, USA.