Voice Interfaces for Visually Impaired and Low-Literate Communities in the Developing World

Aditya Vashistha
adityav@cs.washington.edu

Sam Sudar
sudars@cs.washington.edu

INTRODUCTION

According to the World Health Organization, 285 Million people in the world are visually impaired (VI) and 90% of them live in developing countries [28]. VI people in developing countries have limited access to assistive technologies because of various socio-economic, device-related and network related constraints. In India, only 5% of all mobile phone subscribers own a smartphone [19]. The remaining 95% use a basic or feature phone; these typically have only one in-built assistive tool for VI people, a speaking clock. Lack of screen readers on these devices and tactile interfaces force VI users to access information and provide data using voice based communication on their phones.

The literacy rate in most developing countries is quite low. For example, 26% people in India are illiterate and the rate of functional illiteracy is surprisingly high. Even in a developed country like US where the claimed literacy rate is 99%, around 20% of adults are functionally illiterate. Thus, text based interfaces are out of reach of sighted low-literate communities as well. In India, there are millions of speakers of tribal languages like Kui, Kurukh, Gondi. However, the font support for many tribal languages is still non-existent. Thus many literate tribal people also use voice medium for producing and accessing information. The limited usage of computing devices like smartphones, tablets, and desktops limits adoption of graphical user interfaces designed for low-literate low-income communities. As a result, these communities also use their basic or feature phones to access informational services on the voice channel.

In recent years, various researchers and practitioners have used Interactive Voice Response (IVR) systems to provide health information, citizen journalism, educational services, feedback channels, data collection systems, social media platform, etc. for low-literate marginalized communities. IVR systems have also been used as a means to provide information to and collect data from VI people with low socio-economic status. Though there are a lot of success stories of organizations appropriating IVR systems for societal development, there are a few challenges that impede the sustainable usage of such systems. The first challenge is the high cost of calls for accessing information. The second challenge is inefficient audio content management. Hierarchical call trees are used in IVR systems to structure the information much like a menu in graphical user interfaces. However, auditory interfaces are quite different from textual and graphical interfaces. Voice interfaces force users to access information sequentially, which results in frustration, high cost, and low-usability. Moreover, searching and indexing audio recordings on an IVR system are non-trivial operations because of un-availability of high efficacy speech recognition engines for resource constrained languages spoken in developing countries. Because of the above reasons, concentrated efforts are needed for accurate and swift retrieval of information to decrease the time required by callers and to increase usability.

Our aim is to understand ways in which voice interfaces on an IVR platform can be organized so as to optimize information seeking time, information absorption, and call duration. Through this project, we are evaluating expressiveness and effectiveness of various voice interface designs for providing information to low-literate and VI communities. Our research will answer the following research questions:

- Is a hierarchical voice menu better than a linear list?

- How many items are acceptable in a linear list?

- Is a shallow hierarchy better than deep hierarchy?

- How deep or shallow the hierarchy should be?

- How do abstraction capabilities, literacy level, and prior exposure to IVR impact different designs of voice user interfaces?

RELATED WORK

Interfaces for Low-Literate People
A lot of work has been done in designing user interfaces for low-literate low-income communities in the developing world. Most of the work explores usage of graphical interfaces [9, 16, 17, 8] and auditory interfaces [7, 2, 21] for collecting data and providing information to marginalized low-literate communities. Various researchers have explored challenges faced by literate users while accessing hierarchical menus [3, 5]. However, the most relevant work to our project is done by Medhi et. al., where they presented how limited education affects hierarchical user interface navigation on desktops and mobile phones [15, 18]. Medhi measured the textual literacy level and abstract reasoning level for low-income low-literate people in India. The participants were then asked to find items from graphical user interfaces consisting of a linear list, shallow, and deep hierarchy. Medhi

found that low-income low-literate communities performed best when navigating a linear structure on both desktops as well as mobile phones.

Interfaces for VI People

Although VI people can use screen readers to access information on the Internet, this requires a computer/smartphone, expensive screen reader software, and an Internet connection that is beyond the reach of the majority of the low-income VI people in the developing world [14]. Our parallel research (under review in MobileHCI 2014) has demonstrated that many low-income VI people do not use screen reader software, instead using alternative information and communication tools for accessing information. Many participants of the study indicated using IVR systems for accessing educational and entertainment content. Some work has also been done to augment or replace visual menus on desktops, mobile phones and other electronic devices like music players for improved usability and accessibility [29, 30, 12]. However, no work has yet been done to understand the trade-offs between using lists and different kind of hierarchical interfaces to organize information in IVR systems.

IVR Design Guidelines

Various researchers have provided design guidelines to create effective IVR systems and decrease cognitive load on callers [22, 25, 20, 10]. However, a few guidelines focus on the breadth and depth of navigation IVR trees. Various researchers have investigated deep and shallow hierarchies in IVR systems for high-literate users in the developed world [11, 27, 6]. For instance, Commarford et.al. compared deep and shallow IVR hierarchies and concluded that deep hierarchies take longer to navigate and have high cognitive load [6]. However, the participants of the study were undergraduate students with at least 3-months of email experience and were based in the United States.

Researchers have also studies various facets of IVR systems to improve information consumption and production by low-income low-literate populations in the developing world. Chakraborty et.al. has provided design guidelines for creating IVR surveys for low-literate low-income marginalized communities in India [4]. Kummamuru et.al. has analyzed the relationship between cognitive load and various navigational factors like breadth, depth, type of information, and numerical order [13]. They asked semi-literate participants in India to find job information from a synthetic IVR job repository. Each participant had to complete eight tasks. The tasks were designed such that the impact of depth, breadth and other parameters on cognitive load could be measured. They found that the time to choose the correct option decreases as the breadth and depth increases. This is because practice and learning effects impact the time taken to navigate the depth. Moreover, cognitive load is high while listening to the first few options in the breadth and thus the navigation time is lessened for later options in the breadth. The biggest weakness of the paper is the absence of direct comparisons between linear lists, shallow hierarchies and deep hierarchies.

To the best of our knowledge, our work will be the first to directly compare linear list, shallow hierarchy and deep hierarchy for maximizing information absorption by low-literate, low-income communities and visually impaired people in India.

DESIGN

We have conducted 3 between-subjects experiment design to compare three designs of IVR interfaces: linear list, shallow hierarchical interface, and deep hierarchical interface. We have used the *Household items* dataset used by researchers for comparing performance of low-literate communities on list, shallow and deep hierarchy graphical user interfaces [15, 18]. We chose the same dataset in order to directly compare our finding with the findings for graphical visualization. In our experiment, each structure is read aloud by the system. The list user interface is a sequential list of the household items. The shallow hierarchical interface has a branching factor of five and the depth of the tree is two (see figure 1). The branching factor for the deep hierarchical interface is up to four, and the depth of the tree is four (see figure 2). We have implemented the design structures in VoiceXML using Voxeo Prophecy and IVR Junction [26]. We have also logged user's interaction with the IVR system while performing tasks to review the actions user took while performing the tasks.
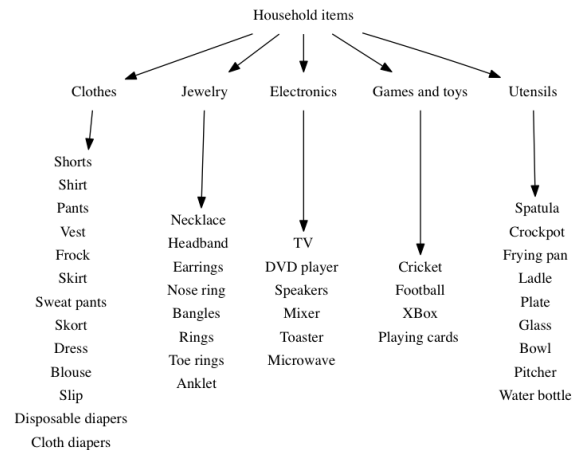

Figure 1. Shallow hierarchy arrangement of household items
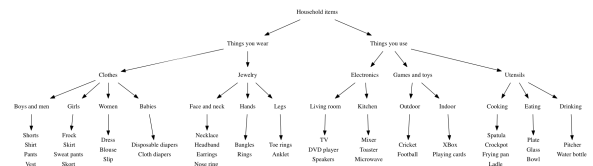

Figure 2. Deep hierarchy arrangement of household items

Participants were randomly divided into 3 groups where each group contained equal number of participants. Each participant performed five tasks, one after another. Each task consisted of selecting a target item from the assigned design structure. All participants searched for five target items (Shirt, Blouse, Rings, Football, and Plate) in the same order. We did not randomize the tasks for participants to ensure that

the learning effects for each participant is same. We selected two easy tasks (*find Shirt, Blouse*), one task of intermediate difficulty (*find Rings*), and two complex tasks (*find Football, Plate*). The comparative complexity of performing a particular task (say *find Rings*) across different hierarchies is same as the number of hops required to reach the target item (excluding navigational prompts) in a brute force manner is same for each design structure.

EVALUATION

We recruited 18 participants (12 Male, 6 Female) to evaluate the three design structures. All the participants were either students or employees of the Computer Science department of the University of Washington. The average age of the participants was 26.7 years (max=52, min=19) and the average education level was Master's. Sixteen participants were heavy users and two were moderate users of IVR systems. None of the participants were visually impaired.

Before beginning the task, participants were shown a demonstration task where they had to find an item from the randomly assigned design structure. We used a different dataset (animals) to illustrate the task and design structure to participants. Participants were allowed to perform demo task at most two times before starting the actual experimental task. We also observed participants while they were performing the task for understanding the challenges they faced while navigating design structures. After completing the task, we interviewed participants to collect demographic information. Participants also completed a short survey where they rated their performance, frustration, and physical, temporal and mental efforts exuded while performing the tasks.

The following criteria were used for evaluating information absorption and swiftness in navigating the design structures:

- Task completion rate

- Task correctness

- Average task completion time

TO DO ~~Tell the overall correctness and time taken for List, Deep and Shallow. Graphs: Box plots of time taken for all tasks in List, Shallow and Deep. Graphs created by Aditya in poster with error bars. Graphs plotted by Sam depicting results of NASA study. explain the graphs~~

Correctness of a task is defined as whether or not the subject correctly selected the target item. For example, if the task was *find Shirt* and the subject made any final selection other than *Shirt*, it was defined as incorrect. The list structure resulted in the lowest correctness, with ~~an average~~ the overall correctness rate of 57% ~~correct~~ across all the tasks. The shallow hierarchy achieved a 93% ~~average across all tasks~~correctness rate, while the deep hierarchy fared slightly better with 97% accuracy ~~across all tasks~~rate.

~~Correctness per task is shown~~

The correctness corresponding to each task for all design structures is depicted in Figure 3. It is clear that the ~~hiearchical~~ hierarchical structures outperformed the ~~simple~~
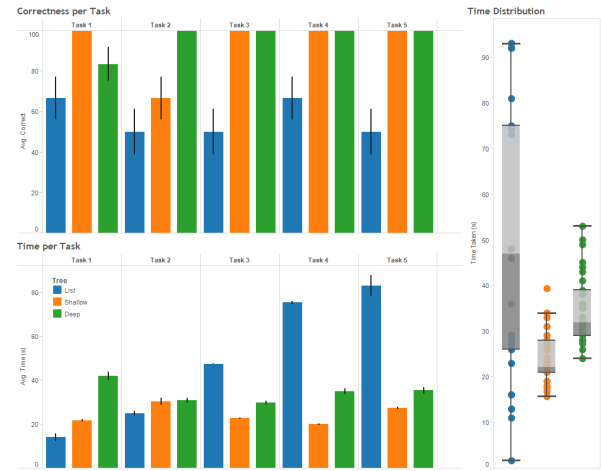

Figure 3. Correctness and time taken

~~list. No task was drastically more difficult than another.~~ linear list. Performance using the list remains poor throughout the experiment, whereas ~~subjects~~ the correctness for participants using the hierarchical structures ~~saw correctness increase~~increased, as evidenced by the fact that no user incorrectly selected an item in the last three tasks. Possible explanations of these effects are discussed in more detail in the ~~Analysis~~ *Analysis* section.

~~Correctness and time taken.~~

The time required to complete a task is equivalent to the amount of time a caller would spend on an IVR system. Lowering this time increases usability. The ~~dstribution~~ distribution of times taken by ~~subjects~~ participants is shown in Figure 3. The box plots for each ~~hierarchy~~ design structure are constructed using 30 data points, as ~~six subjects interacted with each hierarchy and each user~~ we had six participants for each design structure, and all users completed five tasks for the specific design structure they were assigned. The average time taken across all tasks using the linear list structure was 48.9 seconds, the average time taken for the shallow hierarchy was 24.3 seconds, and the average time for the deep hierarchy was 34.5 seconds.

As can be seen in Figure 3, the distribution of times was largest for the list structure, where the time per task ranged from 1.7 to 93 seconds. This ~~variation is in keeping with a list structure, where~~ is because, in the best case an item could be selected almost immediately, whereas in the worst case a user would have to listen to almost all 40 items before making a selection. The shallow hierarchy took the least ~~time on average , with 24.3 seconds~~ average time, and had a range of 15.7 to 39.3 seconds. The deep hierarchy ~~took an average of 34.5 seconds, with a range of~~performed relatively worse than shallow hierarchy and the values were distributed from 24 to 53 seconds. ~~The lower average task time for the hierarchical structures is to be expected, as callers do not have to listen to as many options before reaching their target items.~~

After completing the experiment, user feedback was collected using the NASA TLX Survey [1]. Subjects were asked to
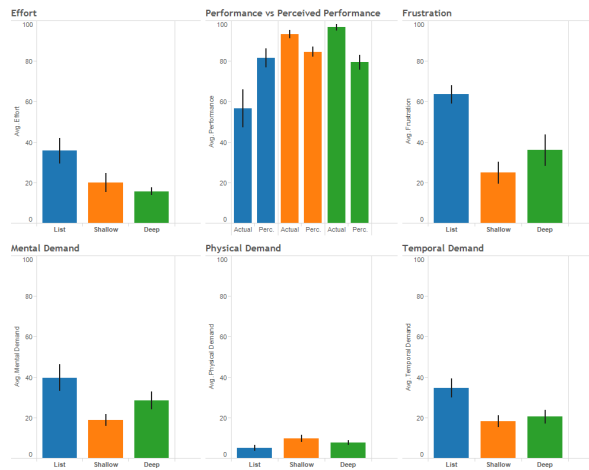
Figure 4. User self assessment of workload during the experiment

provide subjective measures on a scale of 0 to 100 to ~~gauge~~ mark their effort, frustration, and perceived performance during the task. They were also asked to provide the degree of mental, physical, and temporal demand while completing the task. The results of this self-reported feedback are displayed ~~on a per experiment basis~~ in Figure 4.

~~User self assessment of workload during the experiment.~~

On average, users found the list structure to require the most effort~~, be the~~ . They also found it most mentally and temporally demanding, and ~~to be the most~~ frustrating. The shallow and deep hierarchies were comparably less physically and temporally demanding than a list. However, the deep hierarchy was on average more mentally demanding for ~~subjects~~ participants. This mirrors the findings ~~of Medhi et al., who found~~ [15] that successful navigation of the deep hierarchy required greater abstraction and reasoning capabilities than the other structures.

Also shown in Figure 4 is perceived performance compared with actual performance. After completing the task, ~~subjects~~ participants were asked to indicate how well they felt they performed on a scale of 0 to 100. ~~Subjects using the list felt that they had performed better than they in fact did, rating themselves on average as~~ Participants using rated themselves as high performing (average score of 81.5 out of 100~~while in fact only achieving a~~). This is much higher than their actual performance score (57.7 ~~% success rate~~ out of 100). Users of the hierarchical ~~structure~~ structures saw the opposite effect~~, rating~~ ; they rated themselves lower than they actually performed. Subjects with the shallow hierarchy claimed an average of 84.5% ~~correct~~ correctness rate but achieved 93.3%, while ~~subjects~~ participants using the deep hierarchy claimed an average of 79.3% ~~correct~~ correctness rate but achieved 96.7%~~correct~~.

ANALYSIS AND RECOMMENDATION

~~TO DO What are the salient observations? What did you find out - Try to rationally explain the findings how is it different from the hierarchical comparisons done for graphical interfaces for US population? Why? (Read~~

~~literature survey, some people have done the same for graphical interfaces where test subjects were university students) tell problems in our hierarchy - ambiguity in Rings, Shirt, Pants tell participants frustration, confusion while navigating hierarchies tell how two participants while navigating deep hierarchy chose a wrong path but then figured it out and selected the right option eventually.~~

This work compared three design structures to organize information in an IVR system ~~. Our goal was to make recommendations on the preferred structure when creating an IVR~~ with the aim to increase information absorption and decrease information seek time. Looking at correctness and length of call measures, it is clear that the list structure is not a good choice. ~~It generated the worst correctness by a wide margin, and subjects experienced a wide range of call times. These factors indicate that users of the list do not obtain a consistent user experience, and indeed these results were borne out using the feedback survey~~The correctness rate for the tasks performed on the list structure was worse, and the participants also spent much more wide time distribution while navigating the list. Users of the linear list structure expended the ~~the~~ most effort, were ~~the~~ most frustrated, and experienced ~~the~~ highest mental and temporal demands. Under some circumstances, correctness might be put at a premium over usability. In these cases, one might argue that the list structure could still be the best choice for certain situations. If the IVR system was providing important information concerning a disease outbreak, for instance, it could be most important that the correct ~~decision was made~~ information is fetched rather than that users enjoyed using the system. The list structure also fails in this scenario, however, as users were dramatically less likely to use the correct ~~targets~~ target items using the list.

The shallow and deep hierarchies ~~fared better . Performance using the deep hierarchy was slightly better than that of the shallow hierarchy, at 96.7% and 93.3%, respectively. The deep hierarchy was more frustrating and mentally demanding than the shallow hierarchy, although it required less effort (Figure 4). The two were rated as being comparably physically and mentally demanding.~~

performed better than the list. Part of the high performance in the hierarchical structures may be attributed to the fact that users receive ~~feedback during navigation~~instant feedback while navigating the structure. For example, two users of the list structure did not mentally parse the call correctly. Consequently, they associated numbers with the incorrect item. These two users obtained ~~a 0%~~ zero success rate. For instance, if the system read ~~"For pants, press one. For shirt, press two", the subjects~~ *"For pants, press one. For shirt, press two",* the participants bracketed the recording incorrectly and heard it as ~~"press one for shirt". This was true even though they had completed the demonstration task and though by this parsing of the call there would be no way to select the first item. Using the hierachies, however , if subjects selected clothing but heard~~ *"Press one for shirt".* They incorrectly deciphered the grouping despite completing the demonstration task. In hierarchical structures, however

while listening to a prompt *"For clothing Press one, For jewelry Press two"*, if a participant wants to press a digit for *clothing* but hear a list of jewelry items, they would be more likely to realize ~~they were making the~~ that they are making an incorrect selection. Users of the list did not have this opportunity, as their first selection was a terminal item. Indeed two users of the hierarchies selected the wrong list and returned to the previous level by pressing zero, indicating that incorrect choices can be corrected by the user in a hierarchical structure.

Additional sources of error can be attributed to disagreements with hierarchy structure, confusion over items, and frustration. During the *find Pants* task, several users of the deep hierarchy expressed concern that the item could be in any of the clothing categories, as men, girls, and women all wear pants. During the *find Blouse* task, two female users selected shirt. Later in the experiment, when hearing blouse as an explicit option, they remarked that they realized they had made a mistake. Finally, one user of the list structure became frustrated and simply entered a ~~guess in~~ random number while performing the last task rather than ~~listen to the items~~ listening to the list items carefully again. This ~~indicates accentuate~~ the importance of ~~usability in creating an effective IVR system~~ designing usable IVR interfaces.

Between the two non-list structures, the shallow hierarchy is recommended over the deep hierarchy. The two gave comparable accuracy scores. The shallow structure, however, required an average of 10 seconds less per task, and was found relatively less frustrating and less mentally demanding by users.

Sam's recommendations

~~Lists have a wide distribution of completion times and are frustrating to users Hierarchies provide the opportunity for performance feedback while completing the task Real world hierarchies will have classification ambiguity.Shallow hierarchies limit the impact of hierarchy decisions and keep the choices intuitive for a broader group of users Shallow hierarchies outperform the other two categories in frustration and performance Use a shallow hierarchy~~

Aditya's recommendations ~~List of size 9 is better than deep or shallow of~~ Although using a hierarchical structure resulted in higher accuracy and lower seek time than a linear list for forty household items, it might be better to use list if your dataset is small (i.e. it contains less than 10 items). The results for time completion rate for tasks (see Figure 3) indicates that the seek time to find the information is much lower than the hierarchical structures till task 2. Task 2 required people to listen to up to 9 items ~~because time taken to traverse a list of this size is less than navigating hierarchical structure~~ in a linear list. Moreover, some users also feel confused while pressing multiple digits to provide input. The correctness in the list can be improved by recording more clear prompts. Thus, we would recommend using a linear list over hierarchical structures if the dataset has less than 10 items. Within deep and shallow hierarchies, recommend using the list of items less than 9 items at the last level. We

would also recommend using a list of less than 10 items while navigating the last level of the hierarchical trees.

FUTURE WORK

We will recruit visually impaired people, and low-literate people from India to participate in our study. We will work with a local organization in Madhya Pradesh to get access to the marginalized communities. We will measure the recruited participants on three variables:

- Literacy level (low, medium/high)

- Cognitive skills (low, medium/high)

- Familiarity to IVR usage (limited, medium/high)

Previous research has indicated strong correlation between literacy level and cognitive skills [24, 15]. Thus, we will group participants in 4 groups on the basis of the level of cognitive skills and familiarity with IVR:

- People with low cognitive skills and limited IVR experience

- People with low cognitive skills and medium/high IVR experience

- People with medium/high cognitive skills and limited IVR experience

- People with medium/high cognitive skills and medium/high IVR experience

The fact that some users disagreed with the choices made in the deep hierarchy highlights an issue that is not discussed in this paper: hierarchy quality. Our hierarchies were chosen from ~~Medhi et al.~~ the prior work [15, 18] and retained for the sake of comparability~~to prior work~~. The appropriateness of the hierarchy is an important element of usability, and has been addressed in other work partly through the notion of information scent ~~[?]~~ [23]. Ideally all hierarchies used in a comparison study would pass an objective measure of quality, however, valid conclusions can still be drawn with imperfect structures. The list, for example, was outperformed by both hierarchical structures even if they were not optimal, suggesting that a flat list is not an appropriate IVR interface for 40 items.

An additional question is how well ~~the conclusions from this work and the recommendation~~ our conclusion to use shallow hierarchies over deep ~~and list hierarchies~~ hierarchy and linear list will scale to ~~sets~~ datasets comprising more than 40 items. With very large datasets, the distinction between shallow and deep hierarchies might begin to blur. ~~For instance, if a collection of hundreds of items led to a deep hierarchy height of 11 and a shallow hierarchy height of 11, it is possible that the two would be indistinguishable to users.~~

~~We will also measure (**TO DO**): **Add Sam's point here**~~

~~Impact of hierarchy appropriateness on usability Items that are more easily categorizable (e.g. animal species broken by taxa) Greater than 40 items--when does a shallow hierarchy become insufficient? Is shallow preferable to deep simply~~

~~because there were fewer questions asked? What if shallow was a single hop but had 15 intermediate questions? Would it still be found preferable? Perhaps users don't like having to stay alert and make decision in the hierarchy What is the ideal number of items in a list at any given point? E.g. some users seemed confused by having to press two buttons on the keypad What is the ideal number of questions in intermediary levels? E.g. perhaps a deep binary tree would out perform our deep tree with a variable branching factor Compare different shallow and deep hierarchies with proportional branching factor and depth of the tree. In our experiment, the ratio of depth was 2: 4 for shallow and deep. It would be great proportions like 4: 8~~ We would run experiments to measure performance of various sets of deep and ~~3:6 to understand which is better among deep and shallow.~~ shallow hierarchical structures.

CONCLUSION
~~TO DO~~

IVR systems are a valuable ~~tool to~~ information and communication tool for a large number of visually impaired and low-income, low-literate users in the developing world. Making these systems effective and user-friendly in terms of information absorption and search time are key components of their usability. Although this work has focused on educated graduate students in the ~~developing~~ developed world, it has laid the groundwork for future studies that will be ~~undertaken~~ executed with the target population in India.

This study is motivated by real world problems and is directly comparable to existing work. We find that, for a set of 40 items, a shallow hierarchy outperforms a linear list and a more finely detailed hierarchy. In order to reach populations that remain out of reach of developments in graphical user interfaces, it is important that researchers focus on problems such as these that leverage existing ~~technology that is relevant to problems faced by millions of people every day outside of the developed, sighted world.~~

MISCELLANEOUS TO-DOS
~~GitHub page for the project Upload paper and other artifacts in the repository.~~ technologies to collect data and provide information to marginalized communities in the developing world.

REFERENCES
1. NASA TLX: Task Load Index.
   `http://humansystems.arc.nasa.gov/groups/tlx/`.

2. Agarwal, S. K., Kumar, A., Nanavati, A. A., and Rajput, N. User-Generated Content Creation and Dissemination in Rural Areas. *Information Technologies and International Development 6*, 2 (2010), 21–37.

3. Allen, R. Cognitive Factors in the Use of Menus and Trees: An Experiment. *Journal on Selected Areas in Communications SAC 1*, 2 (1983), 333–336.

4. Chakraborty, D., Medhi, I., Cutrell, E., and Thies, W. Man versus machine: evaluating IVR versus a live operator for phone surveys in India. In *DEV* (New York, New York, USA, Jan. 2013).

5. Chaudry, B. M., Connelly, K. H., Siek, K. a., and Welch, J. L. Mobile interface design for low-literacy populations. *Proceedings of the 2nd ACM SIGHIT symposium on International health informatics - IHI '12* (2012), 91.

6. Commarford, P. M., Lewis, J. R., Smither, J. A.-A., and Gentzler, M. D. A Comparison of Broad Versus Deep Auditory Menu Structures. *Human Factors: The Journal of the Human Factors and Ergonomics Society 50*, 1 (Feb. 2008), 77–89.

7. Cuendet, S., Medhi, I., Bali, K., and Cutrell, E. VideoKheti: Making video content accessible to low-literate and novice users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), 2833–2842.

8. Ghosh, K., Parikh, T. S., and Chavan, A. L. Design considerations for a financial management system for rural, semi-literate users. In *CHI '03 extended abstracts on Human factors in computer systems - CHI '03*, ACM Press (New York, New York, USA, 2003), 824.

9. Grisedale, S., Graves, M., and Grunsteidl, A. Designing a Graphical User Interface for Healthcare Workers in Rural India. In *Proceedings of the 1997 ACM SIGCHI Conference on Human Factors in Computing Systems* (1997), 471–478.

10. Halstead-Nussloch, R. The design of phone-based interfaces for consumers. *CHI '89 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (1989), 347–352.

11. Huguenard, B., Lerch, F., and Junker, B. Working-memory failure in phone-based interaction. *ACM Transactions on Computer-Human Interaction 4*, 2 (June 1997), 67–102.

12. Jeon, M., Walker, B. N., and Srivastava, A. Spindex (Speech Index) Enhances Menus on Touch Screen Devices with Tapping, Wheeling, and Flicking. *ACM Transactions on Computer-Human Interaction 19*, 2 (July 2012), 1–27.

13. Kummamuru, K., Jujjuru, A., and Duggirala, M. Generating facets for phone-based navigation of structured data. *Proceedings of the 21st ACM international conference on Information and knowledge management* (2012), 1283.

14. McCarthy, T., Pal, J., Marballi, T., and Cutrell, E. An analysis of screen reader use in India. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12*, ACM Press (New York, New York, USA, 2012), 149.

15. Medhi, I., Lakshmanan, M., Toyama, K., and Cutrell, E. Some evidence for the impact of limited education on hierarchical user interface navigation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, ACM Press (New York, New York, USA, 2013), 2813.

16. Medhi, I., Patnaik, S., Brunskill, E., Gautama, S. N., Thies, W., and Toyama, K. Designing mobile interfaces for novice and low-literacy users. *ACM Transactions on Computer-Human Interaction 18*, 1 (Apr. 2011), 1–28.

17. Medhi, I., Sagar, A., and Toyama, K. Text-Free User Interfaces for Illiterate and Semiliterate Users. *Information Technologies and International Development 4*, 1 (2008), 37–50.

18. Medhi, I., Toyama, K., Joshi, A., Athavankar, U., and Cutrell, E. A comparison of list vs. hierarchical UIs on mobile phones for non-literate users. In *Human-Computer Interaction INTERACT 2013 in Lecture Notes in Computer Science* (2013).

19. Meeker, M., and Wu, L. 2013 Internet Trends. **https: //www.kpcb.com/insights/2013-internet-trends**.

20. MSDN. Best Practices in Designing Speech User Interfaces. **http://msdn.microsoft.com/en-us/ library/ms994646.aspx**.

21. Mudliar, P., Donner, J., and Thies, W. Swara : A Voice Forum for Citizen Journalism in Rural India. *Information Technologies and International Development 9*, 2 (2013), 65–79.

22. Ndwe, T., Dlodlo, M., and Mashao, D. Usability engineering of an interactive voice response system in a diverse-cultured and multilingual setting. *Innovative Techniques in Instruction Technology, E-learning, E-assessment, and Education* (2008), 554–559.

23. Pirolli, P., Card, S. K., and Wege, M. M. V. D. The effects of information scent on visual search in the hyperbolic tree browser. *ACM Transactions on Computer-Human Interaction 10*, 1 (Mar. 2012), 20–53.

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

25. Suhm, B. Ivr Usability Engineering Using Guidelines And Analyses Of End-to-End Calls. In *Signals and Communication Technology*. 2008, ch. Human Fact.

26. Vashistha, A., and Thies, W. IVR Junction. **www.ivrjunction.org**.

27. Virzi, R. A., and Huitema, J. S. Telephone-Based Menus: Evidence that Broader is Better than Deeper. In *Proceedings of the Human Factors and Ergonomics Society* (1997), 315–319.

28. WHO. Factsheet on Visual impairment and blindness, 2013. **http: //www.who.int/mediacentre/factsheets/fs282/en/**.

29. Yalla, P., and Walker, B. Advanced auditory menus: design and evaluation of auditory scroll bars. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility* (2008), 105–112.

30. Zhao, S., and Dragicevic, P. Earpod: eyes-free menu selection using touch input and reactive audio feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2007).