European Journal of Computer Science and Information Technology
Vol.2, No.1, pp.18-29, March 2014
Published by European Centre for Research Training and Development UK (www.ea-journals.org)

SPEECH USER INTERFACE FOR LOW LITERACY USERS OF ICT SERVICES

Edim Azom Emmanuel
Department of Maths/Stats., and Computer Science
University of Calabar, Nigeria.

Muyingi Hippolyte N.
Department of software Engineering
The Polytechnic, Namibia.

ABSTRACT: *Speech user interface has the potentials to accommodate different users irrespective of the level of education. This type of user interface can provide interactions that meet the capabilities of low literacy users in underserved rural communities. One of the challenges facing the development of Information and Communication Technology in rural underserved communities is the inability of the users to interact and communicate with the services provided effectively. The task of designing user interfaces to accommodate low literacy users is still a big challenge. We conducted a study with the aim of designing a customized speech user interface to enable low literacy users in a rural community to interact and have access to a mobile commerce service. We applied different ethnographic research methods in the process of data collection, design and evaluation of the user interface. The user evaluation results show that the users were able to successfully interact with the application. The users feel that the speech user interface is easy to understand and use, and the interface meets their capabilities.*

KEYWORDS: Voice Interaction, Mobile Commerce, Mobile Phone, ICT, Low Literacy.

INTRODUCTION

Information and Communication Technology (ICT) is a viable medium for enhancing the socio-economic demands of low literacy people living in underserved communities. ICT services have been used in different regions to contribute to the socio-economic wellbeing of the people. Significant successes have been recorded in low literacy communities where ICT services have been applied to give added value to the inhabitants (Medhi et al., 2006; Sherwani et al., 2007). With the advent of new technologies, interactions between people, communities and cultures have been altered to a more flexible form (Huang and Deng, 2008). The proliferation of mobile devices has brought changes on the socio-cultural life of the people. The way people interact with one another and the environment has changed (Sato and Chen, 2008). However, the medium for interacting with ICT services and technologies is an important consideration if the users, especially the low literacy users are to have full advantage of the services provided. The user interfaces that provide the means to communication the services should be able to cater for the social and cultural requirements of the users (Boyera, 2007).

In order to improve usability and user experience for ICT services, the user interface should be intuitive, localized to meet the natural experience of the users, and easy to use. This is because in different regions and cultures, people hold different perception about technology and how they use technology (Evers and Day, 1997). Knowledge of users' culture, values, natural experience within their environment and perception will improve users' interaction with ICT services.
Rural communities in developing regions have literacy challenges and interaction with new technologies is also difficult. The users often find the language of presentation difficult to understand (Edim and Muyingi 2010a). Speech user interface has the potential for accommodating different user groups even the illiterate and semi-illiterate users. If the language of interaction is localized to meet the cultural experience of the users, the interactions will be easy and the usability of the interface will be improved.

The results of several efforts to make information accessible to low literacy users in underdeveloped regions are encouraging. But much is still desired because in so many underserved communities, the inhabitants still find it difficult to have access to information. The user interface and the interaction technique may not have been presented to accommodate these groups of users. Graphical user interface with interaction elements that depict the socio-cultural environment of the target users have yielded significant successes (Medhi et al., 2007; Edim and Muyingi, 2010a). Speech user interface offer a more natural form of interaction that is intuitive if properly adapted to the socio-cultural environment of the target users, especially in rural communities.

This study was conducted to design a speech user interface with enabling interactions that will make information accessible to low literacy users in Dwesa rural community. We have designed a localized speech user interface with dual tone and speech input modalities to enable low literacy users in the community to have access to a mobile commerce web service. This will enable local online shop-owners to interact with the application and attend to their customers' queries. This study was carried out with the involvement of the local community members as well as other literate persons who share the same culture with the community members. The highlights the potentials of speech user interface as a good medium for making information accessible to users with low level of literacy.

LITERATURE REVIEW

The use of ICTs for poverty alleviation in underserved communities by government and non-governmental agencies is gaining support (Donner, 2004; Dalvit et al, 2007). Efforts to bridge the digital gab have seen the proliferation of the mobile technology to the rural communities. Establishment of ICT centers in rural communities is also receiving much attention such as the Siyakhula living lab (Dalvit et al., 2007; Edim and Muyingi 2010c). All these efforts are geared towards making information accessible to low literacy people living in rural communities. The problem is still with the medium of interaction provided for these groups of users to access the desired information.

In order to develop ICT services in low literacy environment, several researches have been conducted and have highlighted how ICT can be harnessed to improve life in rural communities.

Few examples are worth mentioning here. A healthcare service and user interface for the provision of healthcare information to rural community health workers is presented in Sherwani et al. (2009). The application is to assist the health workers to get healthcare information and take care of patients within the rural community. Also, a financial management service for self-help groups in rural India is discussed in Gosh et al. (2003). The user interface enables the group members to record and manage savings and lending transactions within the group. Gosh et al. identified lack of literacy skills as a major challenge, but the users were able to remember and enter numbers and perform transactions on the interface. The HealthLine (Sherwani et al., 2007) is a speech user interface that was designed to help semi-literate rural health workers to have access to health information on the web.

With the rapid changes in the technology landscape, mobile technologies have become viable platforms to make ICT services available to low literacy communities. Mobile phones are found everywhere, but the nature of the mobile device makes user interactions difficult. More so the literacy challenge among rural users. The mobile user interface will accommodate few number of graphics symbols and as such limit the number of interaction elements in a graphical user interface. For low literacy users, this may reduce their understanding of the user interface and usability. Speech user interface can surpass the limitations of the graphical user interface and improve access to information (Patel et al., 2009). In low literacy environment, speech user interface will improve understanding and usability if designed to meet the capacity and cultural preferences of the target users.

Low literacy ICT users in rural environment want to interact with user interface elements that depict their culture and values, and interactions that meet their cultural preferences (Edim and Muyingi, 2010a). User interfaces with interaction elements that are not adapted to their cultural preferences have been identified as one of the causes of usability problems. Most services provided by the mobile communication service remain unusable as a result of lack of understanding of the user interfaces among low literacy users in rural communities.

Mobile device enabled ICT service is a developing market in the developing world (Boyera, 2007; Donner, 2004). In order to help users to adopt this emerging market for development, the services provided and the medium of interaction must be able to address the social, cultural and literacy factors. And so speech user interfaces present simple and intuitive interactions that meet the socio-cultural form of interaction in every culture. Speech interactions are very simple and easy to adapt to the social and cultural experience of the users. Mobile devices support speech interactions and personalization. A large population of user in low literacy communities can have access to information with speech user interfaces irrespective of the educational level attained. Speech interface is easy to use and also easy to localize for different cultural group. The users do not require much experience to be able to use the interface. Interacting with technology using the users' cultural language will improve the interaction and easy access to information. The level of literacy of the target users will not be a factor.

ICTs are now considered as important instruments for social and economic development of any region especially regions without function social amenities. These new technologies create positive influence on communities where they have been applied. The mobile phone has

contributed to the socio-economic life of people in developing countries (id21, 2007). This technology has also changed the behavior and communication pattern between people (Huang and Deng, 2008). Apart from the above, ICT brings added value that cannot be overlooked in any community. The challenge still lies in the usability of ICT services even though people in different communities are rapidly embracing these technologies. Usability problems among different regions and user groups can be reduced if the user interfaces designed for interactions meets the capabilities of the users. User interface adaptation to the users' culture and preferences will improve the usability, user experience and build natural sense of interaction as they continue to use the system (Huang and Deng, 2008; Shen et al., 2006).

METHODOLOGY

There is much attention focused towards the socio-economic development and sustainability in low literacy rural communities through the development and implementation of ICT products and services. The acceptance and usefulness of these products and services in these communities is very important if this ICT goal is to be realized. The design of user interfaces for ICT products and services for rural users' needs proper adaptation of the design approach and techniques that will enhance effective participation of the target users. The rural environment presents a different setting and makes user accessibility and participation difficult. Direct involvement of low literacy users through research methods that will motivate the rural users for effective participation will yield the required results (Medhi et al., 2009). This study was conducted with adequate participation of the Dwesa community members to design a user interface that is accessible, easy to understand and use by low literacy users.

In order to apply appropriate and practical ways to enable local community members to effectively take part in the data collection and design process, different field studies were conducted. These include on-site visits and facilitated focus groups to understand the users, the problems and micro-businesses within the community. Formal and informal interviews were also carried out for data collection. Scenarios representing user tasks, interactions and activities were designed and presented to the target users through facilitated focus groups.

We also design prototype and carried out focus groups formative evaluation for data collection and improvement on the prototype application and user interface. Cultural adaptation of the speech user interface through interviews, focus groups discussions and consultation of computer experts knowledgeable in the language and culture of the target users was also conducted. We also made recording of Isixhosa language clicks as interface interaction elements (e.g. metaphors). The nature and pronunciation of the language clicks makes it very difficult to implement directly with code. The verification and validation of the speech metaphors and language clicks through focus groups with community members was also conducted.

The speech user interface was first evaluated though a pilot test. In the pilot test, few participants were drawn from the different user groups in the community (school learners, crafters, educators, farmers and others). Based on the experience we gathered through the pilot test, modifications were made, few errors were identified and corrected. The complete evaluation of the speech user interface was carried out to assess users' capabilities to interact and understand the

interface. The evaluation took place through focus groups in the community. In each user group, participants interacted with the interface one after the other. At the end of each group session, a questionnaire was completed by each participant.

Focus groups participation in the course of this study provided valuable data and also made the process less costly. The process helped the participants because group interaction built confident and makes the participants to be free during the interactions. The criteria used in the grouping of the participants include the literacy level, occupation and group size. Gender was not used as a condition because it was difficult to have enough participants to make up a group based on gender.

RESULT AND DISCUSSION

During this study, we maintain neutrality in the process of data collection and interpretation and did not allow our own beliefs, interest, expectations and knowledge of the research and participants to affect our judgment. We adopted open-attitude in the understanding and explanation of the data. In most cases, more than one data collection method was applied in order to avoid being restricted to a single research viewpoint. The study outcome shows attempt to design usable speech interface for low literacy users by adapting the user interface to the capabilities and cultural experience of the users. The user interface was evaluated with the participants in the community. The users perceived the interface as easy and simple to interact with.

Table 1 presents the demographic information of the participants during this study. 18.33% of the participants are within the age bracket of 14 – 19 years. This is one group of the school learners in the community. 7.5% make up the age bracket of 20 – 25 years. This group is also part of the school learners. The school learners are students in the grade 11 and 12th classes (26.6%). The older adults from the age of 25 years and above made up the larger population of participants. Among these adults include the educators (21.6%), and the remaining population of crafters, farmers, local clerks, clerical staff of schools and others made up 51.8%. Table 1 represents the demographic information of the community members who participated in one process or the other during the study. The processes include the user study, prototype design and evaluation.

Table 1. The Demography of the study participants.

	Demographic Information	Frequency (person)	Frequency (%)
Age	14 – 19	22	18.33
	20 – 25	9	7.5
	26 – 30	16	13.34
	31 – 35	13	10.83
	36 – 40	7	5.83
	41 – 45	19	15.83
	46 – 50	18	15
	51 and above	16	13.34

TASKS SUCCESS RATE

The participants were given two tasks to perform to determine the level they will be able to interact with the user interface. The participants include school learners (n = 19), crafters (n = 12), clerical staff (n = 9) and educators (n = 15). The selected tasks include add an item and update the attributes the product and to access and check the status of an online shop. Figure 1 shows the rate of tasks completion success and failure. The result shows that 76% of the participants were able to perform the two tasks successfully to the end. 24% of the participants were unable to complete either one or the two tasks. Those who successfully completed the tasks did so with or without making mistakes. Those who could not complete the tasks gave reasons (e.g." this process was new to me", "the system did not hear my input" and "the system repeated I did not hear you"). We also noticed that poor pronunciation also contributed to the failure of some of the participants to complete the tasks.

The result indicated a high level of tasks completion success. The study found out that the use of the local language of the participants served as a motivation during the interaction. It was much faster for the users to enter data into the system using DTMF than voice input. Localization of the interactions makes it very easy for the users to understand the interface. If the users continue to use the interface, there will be improvements and some of the problems encountered with the voice input will no longer be experienced. The users will be able to improve the use of the user interface to the ICT service.

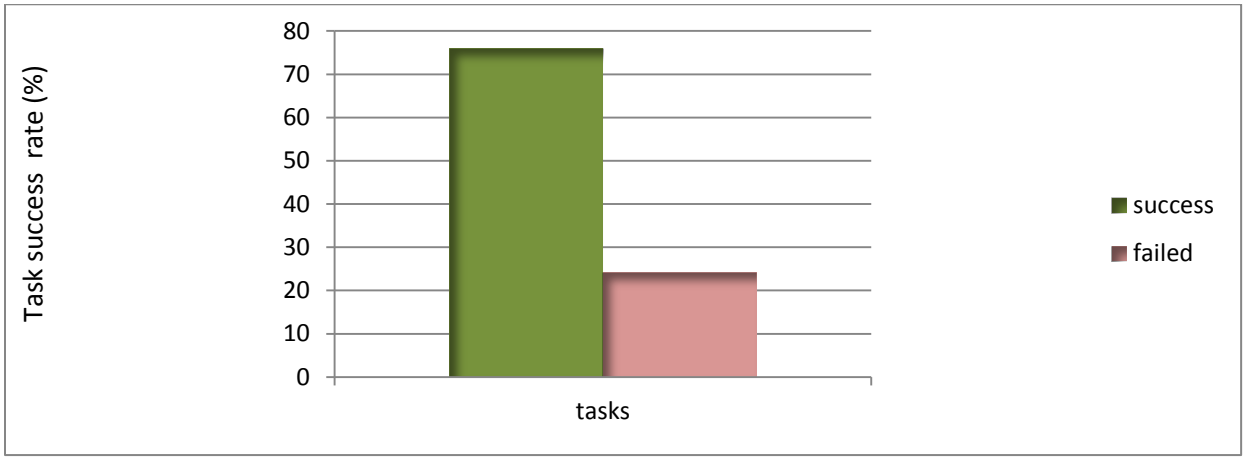

Figure 1 Tasks performance success rate

USER INTERFACE LEARNABILITY

A test was also conducted to assess how easy it was for the users to quickly learn and understand the speech interface. The participants were selected from the different user groups in the community. They include school learners (n = 15), educators (n = 13), clerical staff (n = 10), and crafters (n = 9). Two different tasks were presented to each participant to perform at three different times. The test was conducted for two days with each group. We took notes and recorded the time taken by each participant to complete a task and when they had difficulties and the nature of the problem. The results are presented in Figure 2 and 3 illustrating the learning curves for tasks 1 and 2 respectively.

The results indicate that as the participants performed the tasks repeatedly, they became more familiar with the interactions and completed the tasks faster than the previous time taken to complete the same tasks. That is, a downward trend in the time for each subsequent trial of the task. During the performance of the first task, an average time of 56.23 seconds (SD = 10.52) was used to complete the task. 50.50 seconds (SD = 7.60) and 47.21 seconds (SD = 5.24) was recorded for the second and third trials respectively. For the second task (Figure 3), participants showed improvement during interactions. During the first trial, the users completed the task at 48.50 seconds (SD = 5.21). The second and third trials recorded 43.35 seconds (SD = 4.20) and 42.10 seconds (SD = 4.00) respectively. There was significant improvement with the task performance. Better accuracy and speed were achieved by the participants.

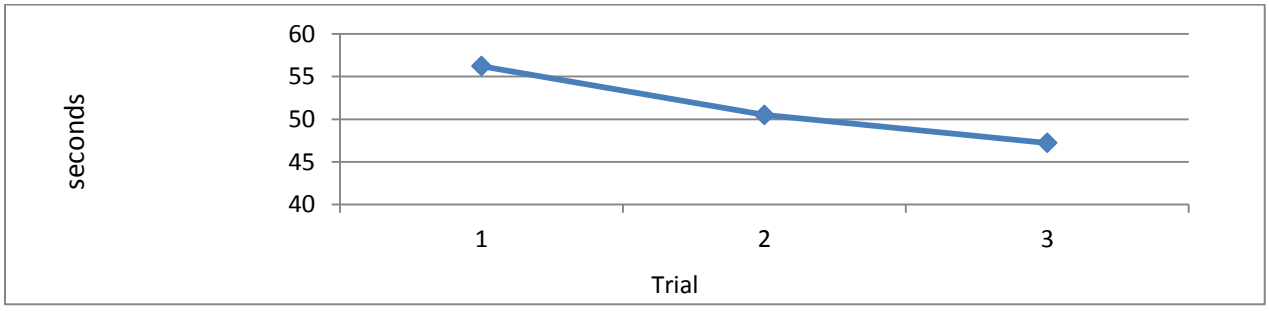

Figure 2 Learning curve for the first task

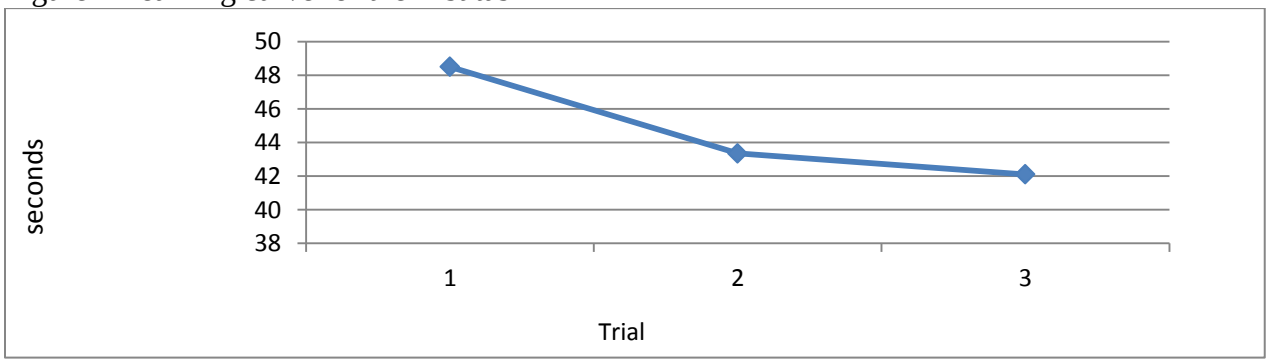

Figure 3 learning curve for second task

The above results show that learning had taken place during subsequent trials of the tasks. The users were able to learn and understand the user interface as the continued to interact with it. The study found that the users became familiar with the interactions during subsequent trials and were able to complete the tasks with less amount of time. This is an indication that the users will be able to use the interface and further improve the level of understanding and interaction. The number of errors encountered reduced during the second and third trials this also contributed to the users' improvement on the performance time. It also points out that the speech user interface is within the users' capabilities.

USER INTERFACE GENERALIZABILITY

The speech user interface provides two modalities of inputs (DTMF and voice) to the user when carrying out any given task. It is an attempt to make the user interface flexible and allow the

users to choose the mode of interaction that they prefer and are more capable to use. We tested the users' ability to extend their experience of interaction with the user interface from one interaction modality to the other. This was done by enabling the users to use one mode of input to complete a task and then later perform the same task using the other input mode. The participants include school learners in grade 11 (n = 12), school learners in grade 12 (n = 12), crafters (n = 9), clerical staff (n = 11), and educators (n = 14). The participants were asked to perform two tasks with DTMF-input mode first and then later with the voice-input mode. We observed the participants, took note of the time, errors encountered and the percentage of users that successfully completed the tasks. The result is presented in Figure 4.

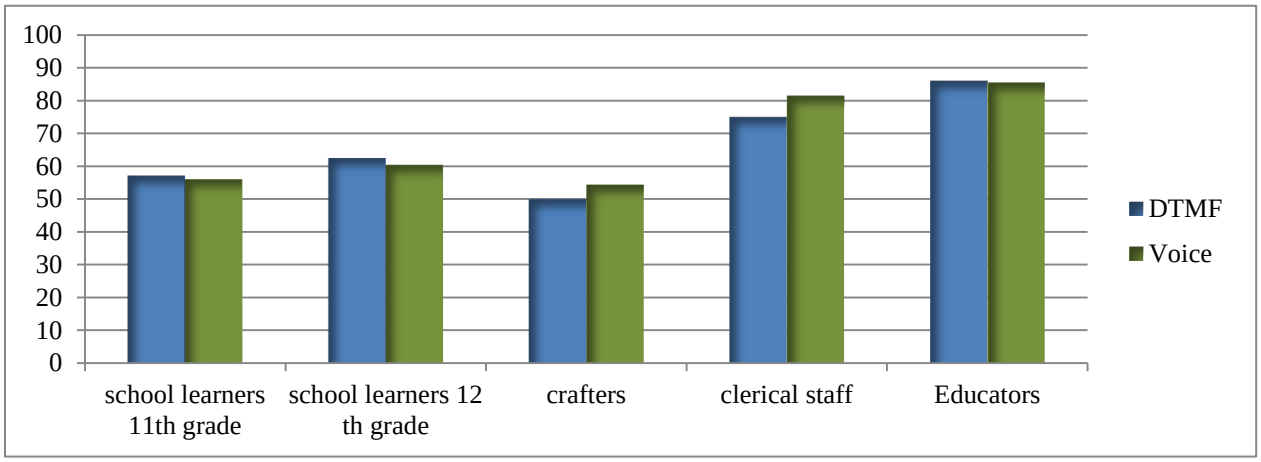

Figure 4 Tasks performance success DTMF-input and Voice-input

The result above shows the percentage of users who performed the tasks successfully and did not encountered any error as the migrated from one input mode to the other. 57.14% of school learners in grade 11 did not have any difficulty when the performed the tasks using DTMF-input and 56.12% in the same group performed successfully with the voice input without an error. Among the grade 12th school learners 62.5% of them performed the tasks successfully without an error using the DTMF input, while 60.42% in the same group used the voice input to perform the same tasks without an error. In the other groups of participants, the results were similar in trend. The Educators are the most enlightened group in the community. The success rate was 86% DTMF vs 85.5% voice input. The overall result shows consistency in interaction from one input mode to the other. This is observable on the level of success recorded for the two interaction modes. The differences in tasks performance success rate between DTMF input and voice input are minimal. This result shows that the users will be able to use any input mode on the user interface to access information.

USER PERCEPTION OF DTMF VS. VOICE INPUT

The participants after interacting with the DTMF and voice input modes were asked to give their perception of the interaction experience. The Likert-scale questionnaire contained positively rated questions (1 – strongly disagree to 5 – strongly agree). Figure 5 presents the users' perception of the input modes. The participants positively rated the simplicity of interactions with the DTMF and the voice input modes (DTMF – mean score = 3.69 and voice – mean score

= 3.31). The DTMF had a higher rating, meaning that it was easier for the participants to interact with it than the voice input. This could be so based on our observation. Participants experienced pronunciation difficulties and noise within the environment. The voice input and DTMF input interactions were generally considered not difficult to use.

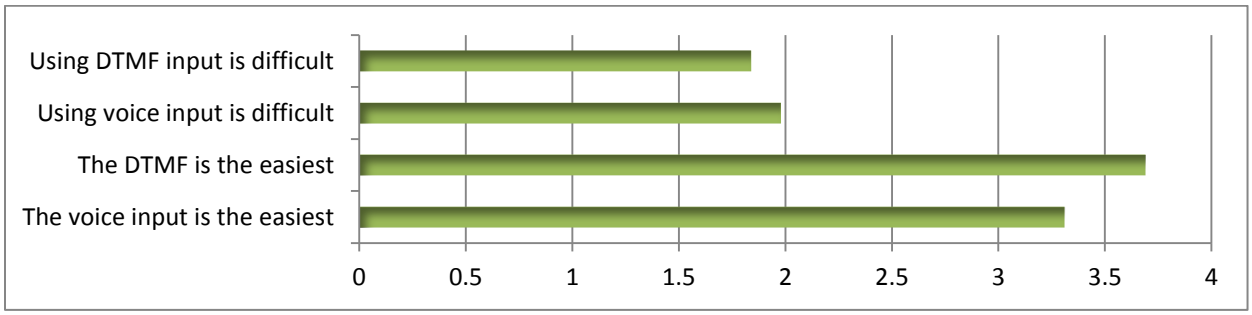

Figure 5 User Perception

The result in Figure 5 shows that the user interface is simple and the interactions are flexible enough to enable low literacy users to access information. The input modalities are within the capabilities and experience of the target users.

SPEECH USER INTERFACE EVALUATION DATA AND ANALYSIS

The participants who interacted with the speech user interface during the evaluation process were given a post-test questionnaire to complete. This was to enable the participants to make a general assessment of the user interface. A likert-scale questionnaire was used in the process. The result of 17 participants from the different user groups in the community is presented in Table 2. The users level of understanding and use of the interface was positively rated (mean score = 4.12, SD = 0.93). The deviation indicates that majority of the users agree to the ease of understanding and use of the interface. The users are also very satisfied with the performance of the user interface (mean score = 3.82, SD = 0.95). The participants feel they were able to hear and understand the interface output (mean score = 3.71, SD = 0.85) and could navigate through the interface effectively (mean score = 3.82, SD = 0.95). The users are positive that they do not require special skills to be able to use the interface because the information provided on the interface is enough to help the users to perform any given task.

The overall rating of the items on the questionnaire shows that the participants were positive that the interface is within their capabilities. The study found that continuous interaction with the speech user interface, improved the user experience, interaction speed and understanding. We also discovered that with speech interfaces, poor pronunciation will always be a problem to the users especially for novice technology users. But with continuous interaction with the interface, the users will have better experience as they continued to learn from the interface and improve their performances. The low level of literacy in the community will not hinder the inhabitants from interacting with the mobile web service.

Evaluation data and analysis (Likert-scale questions: 5 strongly agree – 1 strongly disagree) (n = 17)

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Mean	SD
1. I am satisfied with the speech interface	4	5	4	3	4	4	5	3	2	4	4	3	4	5	2	4	5	3.82	0.95
2.The menu is simple and easy to remember	5	4	4	5	2	2	3	5	4	2	3	5	3	3	5	4	3	3.65	1.11
3. The interface is easy to use	4	4	5	5	5	4	3	3	3	4	5	5	2	5	5	4	4	4.12	0.93
4. I did hear and understand the system output	3	2	3	4	4	3	4	5	3	5	4	3	4	3	4	4	5	3.71	0.85
5. The system output is difficult to understand	2	3	3	3	2	3	3	1	2	1	1	2	1	1	2	2	3	2.18	0.81
6. It is easy to use the voice input	3	4	5	5	4	3	4	4	5	3	4	4	3	2	5	5	5	4.00	0.94
7. Navigation is simple on the interface during task performance	5	3	4	3	3	4	3	4	3	4	5	5	3	3	2	5	4	3.82	0.95
8. No special skill is required to use the interface	4	3	5	5	5	5	4	4	5	4	3	5	2	2	4	4	5	4.12	0.10
9. The information on the interface is good and adequate	5	4	3	3	4	4	4	3	5	4	3	5	3	5	3	4	2	3.77	0.88

CONCLUSION

The user evaluation shows that proper design of user interfaces and interactions for low literacy users in rural communities will make it possible to deliver information effectively to underserved people. A common strategy for ICT development in underserved communities is to provide services and information based on their lifestyle, culture and motivations. These groups of users are willing to receive and interact with services that meet their present requirements. The issue of adaptation of the user interface to suit the local culture, interaction experience and environment of the users was a major consideration. It is believed that this will enhance user experience and the adoption and usability of the interface. Flexibility of the interactions will also bring about task performance consistency among low literacy users.

The results show that the speech interface is easy to use and the users will be able to use it to access information. The users were able to successfully perform interactions with the interface to complete the tasks. They were able to send inputs using DTMF and voice commands and also understand the output from the system. This is remarkable for low literacy users in underserved rural community in Africa.

In the design of usable interface for low literacy users, participatory challenges are common. The use of research methods that conform to the socio-cultural environment will aid the participation of the community members. In this study, we have used the method of facilitated focus groups extensively to encourage community members' participation. Methods that did not meet their capabilities were tried and abandoned because we could not interact and communicate effectively and the participants could not participate with ease.

Further research will be carried out to explore the potentials of other types of user interfaces and interaction techniques in low literacy rural communities.

REFERENCES

Boyera, S. (2007) The mobile web to bridge the digital divide? In *IST-Africa Conference 2007*, 09-11, May, Maputo, Mozambique.

Dalvit, L., Thinyane M., Muyingi H. and Terzoli A. (2007) The Deployment of an e-Commerce Platform and Related Projects in a Rural Area in South Africa, *International journal of Computing and ICT Research*, 1, 1, 9-18.

Donner J. (2004) Micro-entrepreneurs and Mobiles: An Exploration of the Uses of Mobile Phones by Small Business Owners in Rwanda**.** Centre for Global Health and Economic Development. The Earth Institute at Columbia University. New York. *Information Technologies and International Development*, 2, 1, 1-21.

Edim, A.E. and Muyingi, H.N. (2010a) A Cultural User Interface for a Rural Community in South Africa, *Proceedings of SATNAC-2010*, Stellenbosch, South Africa, 231-236.

Edim, A.E. and Muyingi, H.N. (2010c) A User Interface For Micro-Entrepreneurs in a Rural Community, *Electronic Journal of Information Systems in Developing Countries*, 3, 2, 1-19.

Evers, V. and Day, D. (1997) The Role of Culture in Interface Acceptance In: *IFIP TC13 International Confonference on HCI INTERACT'97,* Chapman & Hall, London.

Ghosh, K., Parikh, T. and Chavan, A. (2003) Design Considerations for a Financial Management System for Rural, Semi-literate Users, *ACM Conference on Computer-Human Interaction-2003.*

Huang, K.H. and Deng, Y.S. (2008) Social Interaction Design in Cultural Context: A Case Study of a Traditional Social Activity, *International Journal of Design*, 2, 2, 81-96.

id21 insights (2007) Mobile phones and development. Available: www.id21.org/insights/insights69/ insights69.pdf. 2007 (Accessed: 12 March, 2009).

Medhi, I., Sagar, A. and Toyama, K. (2006) Text-Free User Interfaces for Illiterate and Semi-Literate Users. In *Proceedings International Conference on Information and Communications Technologies and Development*, 72-82.

Medhi, I., Gautama S.N.N. and Toyama, K. (2009) A Comparison of Mobile Money-Transfer UIs for Non-Literate and Semi-Literate Users, *Proceedings of CHI 2009*, ACM Press, Massachusetts, 1741-1750.

Patel, N., Agarwal, S., Rajput, N., Kumar, A., Nanavati, A., Dave, P. and Parikh, T. S. (2008) Experiences Designing a Voice Interface for Rural India. In *proceedings of the IEEE Workshop on Spoken Language Technology for Development (SLT)*, 21-25.

Sato, K. and Chen, K. (2008) Cultural Aspects of Interaction Design, *International Journal of Design*, 2, 2.

Shen, S. Woolley, M. and Prior, S. (2006) Towards Culture-centered Design, *Interacting with Computers*, 18, 820-852.

Sherwani, J., Ali, N. Mirza, S., Fatma, A., Memon, Y., Karim, M., Tongia, R. and Rosenfeld, R., (2007) HealthLine: Speech -based Access to Health Information by Low-literate Users. In *Proceedings International conference of Information & Communication Technologies and*

Development, Bangalore, India. Available: http://ww.cs.cmu.edu/~jsherwan/pubs/ictd07.pdf (Accessed: 20th October, 2008)

Sherwani, j., Palijo, S., Mirza, S., Ahmed, T., Ali, N., and Rosenfeld, R. (2009) Speech vs. Touch-tone: Telephony interfaces for Information Access by low literate users, *Proceedings IEEE/ACM 3rd Int. Conf. on ICTD*, Doha, Qatar, 447-457.