IST-Africa 2013 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2013
ISBN: 978-1-905824-39-7

Correlation Between Rapid Learnability and User Preference in IVR Systems for Developing Regions

T.J. (Jama) NDWE[1], Etienne BARNARD[2], Thato FOKO[3]
[1]*Rhodes University, Department of Computer Science, Centre of Excellence in Distributed Multimedia, P.O. Box 94, Grahamstown, 6140, South Africa*
Tel: +27 46 603 7478, Fax: +27 46 636 1915, Email: j.ndwe@ru.ac.za
[2]*North-West University, Multilingual Speech Technology Group*
P.O. Box 1174, Vanderbijlpark, 1900, South Africa
Tel: +27 16 910 3111, Email: Etienne.barnard@nwu.ac.za
[3]*Meraka Institute-CSIR, ICT in Education Group, P.O. Box 395, Pretoria, 0001, South Africa*
Tel: +27 12 841 4560, Fax: +27 12 841 4720, Email: tfoko@csir.co.za

Abstract: Access to information and communication is one of the most important needs in any population group. It is generally challenging for people in the developing world to access information because the tools and the technologies used to access information are prohibitively expensive and also require training prior to operating such tools [13]. This restrains those who are either poor, illiterate, or without computer skills from accessing information. We carried out two case studies of contrasting Interactive Voice Response (IVR) systems. The research compared the users' choice of interaction modality between Dual-Tone Multi-Frequency (DTMF) and speech-enabled IVR modalities and correlated the results with learnability of the different modalities in the milieu of the two systems. The targeted users are oral users of Southern Africa with diverse literacy levels but nevertheless numerically literate and accustomed to the telephone.

Keywords: Learnability, IVR systems, Oral Users, Developing countries, DTMF, speech-enabled IVR, Information access..

1. Introduction

According to the International Standards Organization (ISO) the definition of usability is: *The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use* [8].

Additionally the ISO standard describes the implied usability characteristics of effectiveness, efficiency, and satisfaction in Table 1.

Table 1: Definitions of Usability Characteristics [8]

USABILITY CHARACTERISTIC	DEFINITON
Effectiveness	Accuracy and completeness □ith □hich users achieve specified goals
□fficiency	□esources e□pended in relation to the accuracy and completeness □ith □hich users achieve goals
Satisfaction	□reedom from discomfort, and positive attitudes to□ards the use of the product

The ISO 9241-11 standard definition of usability is becoming the main reference of usability [9] and the standard presents a contextually oriented view of usability [5]. The contextual orientation and popularity of the ISO 9241-11 standard motivated the researchers to adopt the ISO definition as the main definition that will guide the way that usability is engaged and measured in this research study. However, other viewpoints of usability will also be integrated into the current study as they are not in contradiction with the ISO definition but rather extend on it. Nielsen [14] and [18] expand on the ISO definition by including learnability as a characteristic of usability in addition to ISO's effectiveness, efficiency and satisfaction. Learnability is defined as change of usability over time as the user becomes more acquainted with the system [18] and it is one of the main pillars of usability engineering which emphasizes that the user should be able to rapidly begin working with the system [7]; [14]. Learnability of a product can be measured by comparing the quality of use for users over time [2], and in the case of this study learnability has been achieved within a short space of time and not over continuous use of the technology over a long period and this is referred to as rapid learnability. Shackel [18] also includes flexibility as a characteristic of usability which is defined as how the system will change and adapt to the change in the context of use within the users [18], which implicates changes that occur over an extended period of time. In that context, flexibility is not applicable to this study as we are mainly concerned with usability for the first round of development iteration of the intended products whereas flexibility arises over a longer period of time in further iterations.

In comparing the two modalities of DTMF and the speech-enabled input that uses automatic speech recognition (ASR), the researchers had to design two prototypes for each of the two case studies and, in order to compare the two modalities even-handedly, we had to base the designs for both the DTMF and ASR on the same structure and the same fundamental design. The OpenPhone application involves accessing information about HIV/AIDS illness that is pandemic in the region of Southern Africa and has brought about much damaging consequences in all walks of life and either infects or affects everybody in the region. The BGR application is in strong contrast to the OpenPhone as it is a passionate and fun application that allows soccer fans to access soccer results of recently played games. Soccer is the favourite sport in the region and its status as the favourite sport has recently been enhanced by the World Cup 2010 soccer games which were held in the region. The OpenPhone case study was conducted in 2008 in Botswana and the BGR study was done in 2010 in South Africa. Before the design of the systems pre-design studies were carried out for both systems and it was established that the intended users of the OpenPhone system wanted the system to be in their indigenous language whereas the BGR system's intended users preferred their system to address them in English. The researchers are not aware of any published research studies that compare the two modalities and correlate the results with learnability of the modalities in the milieu of two disparate applications in oral communities.

2. Previous Studies

There have been a number of studies of IVR systems whereby the users were required to compare ASR and DTMF modalities. The distinction between this research study and the other studies is that the majority of those studies have been conducted in developed countries [3]; [6]; [10]; [11]. The intended users in this study differ in terms of culture which entails differences in the users' mental models. The geographical location of the other studies is also important due to the fact that this study is conducted in Africa, where oral rather than literary orientation prevails. Africa is a place where news travels from person to person by spoken word [1]. This is an important differentiation as orality affects various aspects that are of importance in the interaction between the intended users and the proposed technology including how people think, communicate, and learn [20].

The results from the aforementioned previous studies that were conducted in the developed world have shown various results with the preference of DTMF input for linear task completion and ASR for non-linear tasks [11]. Other studies, conducted by Harris Interactive and commissioned by Nuance, have shown that speech-enabled IVR systems are preferred by users over the DTMF system signifying ease-of-use, convenience, and accessibility as the key benefits in using the speech-enabled system [15]. Delogu et al. [3] found preference of DTMF over speech-enabled input designed for an isolated word interface. Both case study applications in this research use isolated word input and the tasks that are to be completed by the users are all linear tasks.

There have been a few studies within oral communities in developing countries for designing information access via speech technology but with differing cultural and user contexts, and also differing geographical locations from this study. Sherwani et al. [19] investigated the design of a speech-based access to health information by low-literate users in Pakistan and [21] established the preference of speech-enabled input compared to DTMF in a healthcare system that was intended for low literacy users. However, these studies inevitably involved limited user populations and task domains – for example, [21] did not involve stigmatized and privacy issues (the healthcare system was aimed at neonatal healthcare). Also, the aforementioned studies in the developing world did not involve a juxtaposing application that is aimed at entertainment for people in developing countries.

In the Avaaj Otalo (AO) IVR application for agricultural information access that was intended for farmers in rural India, Patel et al. [17] did a comparison between DTMF and speech-enabled modality and found user preference of DTMF over speech-enabled modality but that study focused on a longer term user experience as their experiment was conducted over a seven month period unlike this current study which focuses on the users' first encounter with the IVR systems.

3. Usability Experiments

Given that learnability is a characteristic of usability the researchers performed usability test experiments by engaging prospective users of the two different systems. Learnability can be measured as a function of effectiveness of the system as defined in Table 1. In usability experiments, and consequently in this study, effectiveness is measured as the ratio of successfully completed tasks in relation to the total number of tasks. Since this study is concerned with the correlation between users' choice of interaction modality and the modality's learnability, we will only focus on the characteristic of efficiency and the subjective users' choice of technology.

For the OpenPhone system, the researchers resolved that the imperfections and limitations of automatic speech recognition (ASR) technology would obscure the comparison of the two systems since the user utterances would not be guaranteed to be correctly recognized by the system. This is due to the fact that the language used in the OpenPhone application, Setswana, is not as extensively researched language such as English in terms of speech technology. The likelihood of recognition errors is more pronounced in less prominent languages such as the OpenPhone users' native language of Setswana. This would compromise the accuracy and integrity of the experimentation. For this reason, a wizard-of-oz (WOZ) prototype was used instead of a fully functional speech-enabled system. In the WOZ prototype, the speech recognition errors were evaded since a human being recognized the user utterances instead of a true ASR engine. The wizard, who is obscure to the user, listens to the user commands and responds by selecting the appropriate responses to be played by the system in the same way that a fully functional system would respond. The user is not aware that it is the wizard who is controlling the outputs that the user hears as the user thinks that they are interacting with the system directly without human intervention.

In contrast to the OpenPhone WOZ system, the BGR system used a fully functional ASR system where the users used verbal commands to interact and control the system which were automatically recognized by the system. These verbal commands are chosen from a given list in the system's menu. The BGR was a high-fidelity prototype that was connected to the Public Switched Telephone Network (PSTN) so that users could call from anywhere in the world.

In both applications, the DTMF modality allows the use of the telephone keypad as the sole input device and the modality utilizes a menu dialogue strategy. The DTMF system output presents a menu that instructs the user to press a particular number on the phone keypad that corresponds to a particular effect. For example:

System: To learn about hygiene and cleanliness press 2, to learn about nutrition press 3,…

The user then reacts by pressing whichever number that corresponds to the task that they want to carry out.

For the OpenPhone application, a demonstration video that illustrated how an IVR system works within the context of health information access was shown to each participant before they could engage in the tasks. As the intended users were not acquainted to the use of an IVR system, the purpose of the video was to demonstrate the abilities of such a system in a visual manner, and to build an expectation from the caregivers of what to anticipate out of the experience with the system. Demonstration videos have been found beneficial in clarifying the use of interactive information access services particularly for low-literacy users as in the case of the OpenPhone case study [12]; [19].

There was no need for a demonstration or extensive explanation for the BGR application because during the pre-design studies the BGR prospective users indicated that they were accustomed with the use of IVR systems for information access, mostly during interaction with their cellphone service providers. The BGR participants were briefed on the purpose of the study and the tasks they were being requested to perform. The participants were also informed about the operation of the system – e.g., that the information on the system was indexed by the different days that games are played and that information, therefore, was to be primarily accessed according to the day that a specific match was played. Since the BGR was a live application that was accessible over the PSTN, it was possible for remote participants to participate in the experiment unlike OpenPhone which was only tested within a laboratory environment. All the prototypes were instrumented to log participants' interaction activities including task navigation paths, task completion times, timeouts, barge-in activities, and others.

For both applications there were two test stations (two PSTN communication lines in BGR) for the experiments, one for the DTMF modality session and the other for the speech-enabled ASR/WOZ modality session. The participants were asked to participate in both the DTMF and ASR/WOZ modalities. In total there were four sets of tasks (A, B, C, and D) that could be done by the participants and out of the four each participant was required to do only two sets. Each set had two tasks within it (task 1 and task 2) and all the sets were different from each other in terms of the combination of tasks that were performed. This made a total of four tasks that were performed by each participant (two tasks in each set of two sets). The first set was performed during the first session and the second set was done in the second session. Half of the participants were asked to start with DTMF as their 1st session and the other half were asked to start with ASR/WOZ as their 1st session.

In order to distribute the task set execution evenly amongst the participants, the task sets were executed in a rotating manner. For example, the first participant did the 1st session in DTMF using set A tasks and the same participant was then asked to do the 2nd session in ASR/WOZ using set B tasks. The second participant was asked to do 1st session in ASR /WOZ using set B tasks and the same participant was asked to do 2nd session in DTMF using set C tasks. The third participant was asked to do 1st session in DTMF using set C tasks and the same participant was then asked to do the 2nd session in ASR/WOZ using set D tasks. The fourth participant was asked to do 1st session in ASR/WOZ using set D tasks and the same participant was then asked to do 2nd session in DTMF using set A tasks. The fifth participant followed the same procedure as the first participant, and the sixth followed the same procedure as the second, the seventh as the third and the rotation carried on in that manner. All the tasks were independent of each other in the sense that no task depended on the accomplishment of any other task.

The above experimental procedure of half-and-half crisscrossing of task execution combined with the rotation of set tasks was done in order to make sure that:

- The same number of tasks was performed for each of the modalities.
- Half of the participants started with DTMF modality and the other half started with the ASR/WOZ modality.
- There was an even distribution of tasks performed between the 4 possible task sets.

In each set, for both applications, the first of the two tasks was designed to be easier than the second task and the participants started with the easier task and then progressed to the slightly more difficult task. The 2nd tasks were more challenging and engaging in the sense that they needed the participants to pay more attention and have more patience as the task performance required the participants to access information that was mentioned later in the system dialogue and deeper in the dialogue structure. Figure 1 is a graphical representation of the task execution as described above.

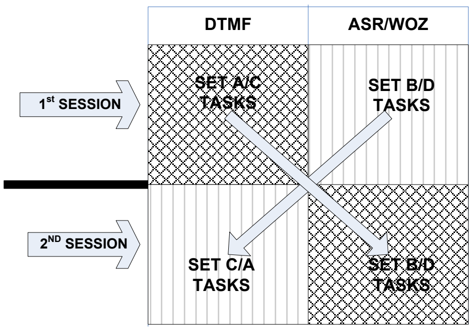

Figure 1: Graphical Presentation of Planned Task Execution

In-between the sessions, questions about the participants' experiences on the 1st session were asked and questions about comparing the two modalities were not asked since they had interacted with only one modality by that time. After completing both sessions, participants were asked a few demographic questions which requested their age, level of education, gender, and home language. They were also asked for their subjective evaluation of the systems and their choice of interaction modality between the DTMF and the speech-enabled ASR/WOZ modalities.

For both case studies the tests did not always go according to the plan as demonstrated in Figure 1 mainly due to some technical errors in the prototypes. This reduced the initial number of recruited participants on both systems. Out of the 33 OpenPhone participants, only 22 participants experienced both prototypes (both modalities) and only those participants' results are employed for the direct contrast between the two modalities. The OpenPhone participants were recruited from the pool of clients at the clinic where they usually obtain medication and consulting services for the illness.

For the BGR system there were 27 participants and these are the participants who are used in the comparison of the two modalities as they performed tasks in both modalities. There were nine BGR participants who performed their tasks in the presence of the researchers and 18 performed their tasks remotely. The participants covered speakers of all the nine indigenous languages of South Africa and came from all the nine provinces of the country. It was important for some of the BGR participants to execute their tasks remotely as the application needed to cover participants from both urban and rural areas in all nine provinces of South Africa.

4. Experimental Results

The following subsections discuss the objective and the subjective results that were obtained from the experiments.

4.1 OpenPhone Objective and Subjective Results

Out of the 22 participants for the OpenPhone application there were 21 participants who did both task 1 and task 2 of the DTMF prototype because one user's results were disqualified due to a technical error in the DTMF prototype. As stated earlier, this research focuses on the usability characteristic of effectiveness in order to explore learnability effects and therefore the study relies on the measurement of the success rate.
A summary of results that compare success rates between the first and second tasks of the DTMF modality in OpenPhone are presented in Table 2:

Table2: Results for Users Who Did Both DTMF Tasks

	DTMF Task 1	DTMF Task 2	Total
Correct	4	11	15
Incorrect	17	10	27
Average success rate	19%	52.4%	35.7%

The results in Table 2 are for 21 users who did both DTMF tasks 1 and 2 in the proper manner, i.e., within one session, consecutively without repeating any tasks, and in the proper order of doing task 1 first and then task 2. The 21 users did not necessarily perform their WOZ tasks in the above mentioned proper manner.

By using Chi-square test for significant differences in the proportion of correct answers between task 1 and task 2 in DTMF of OpenPhone:

Chi-squared value = 5.0815, p-value = 0.0242, which means that there is a statistically significant difference in the user performance for the two DTMF tasks since $p < 0.05$.

A summary of results that compare success rates between the first and the second tasks of the WOZ modality in OpenPhone are presented in the following Table 3:

Table 3: Results for Users Who Did Both WOZ Tasks

	WOZ Task 1	WOZ Task 2	Total
Correct	8	9	17
Incorrect	12	11	23
Average success rate	40%	45%	42.5%

The above results, in Table 3 are for 20 users who did WOZ tasks 1 and 2 in the proper manner. The 20 WOZ users did not necessarily perform their DTMF tasks in the above mentioned proper manner.

By using Chi-square test for the two WOZ tasks:

Chi-squared value = 0.1023, p-value = 0.7491, which means that there is no statistically significant difference in the user performance for the two WOZ tasks since $p > 0.05$.

In order to verify the above results, McNemar's test for significant differences was used instead of the chi-square test and the same results were obtained indicating a statistically significant difference between task 1 and task 2 in DTMF modality and no difference in the WOZ modality.

It is important to recall that tasks 1 (the 1st tasks) within different sessions and different modalities were different tasks but they were equivalent in their level of difficulty and the same applies for tasks 2. Also, tasks 2 (the 2nd tasks) were more difficult than tasks 1 in both sessions because they needed the participants to navigate information that was deeper in the dialogue and mentioned later in the system prompts. The difference in the level of difficulty between the 1st and 2nd tasks could have been made to be the requirement for the participants to make certain inference about the given information. The researchers did not want to test the participants' abilities or inabilities to make inferences from the given information because the researchers were only interested in the participants' ability to navigate and use an IVR system to obtain information on this first iteration.

Immediately after completing each session the participants were asked subjective questions about their experiences with the prototype that they had just used. The subjective questions enquired about factors such as ability to understand the content, pace of the conversation, perceived system response time, and whether their expectations were met or not. After having completed both sessions they were also asked which system they preferred between the two prototypes and why. Out of the 22 participants who interacted with both prototypes, 13 preferred DTMF, four preferred the WOZ system, and five were equally happy with both systems which produced a 59.1% preference of DTMF over 18.2% for WOZ, with 22.7% of users who were undecided. Two people out of the four who preferred the speech-enabled system (WOZ) said that they actually like both but they were choosing the speech-enabled system because they think that it would be a better system for elderly users who might have difficulty using the DTMF system.

The substantial majority of 13 participants who chose the DTMF system had well-defined reasons including perceived faster speed. The most common reason was the ease of use and the fact that the DTMF was easy to follow because, "it is impossible to get lost as the system (DTMF) just tells you what to do." Another participant remarked that in the WOZ system she had to pay more attention because if a command is missed or misunderstood or misinterpreted, then that can be problematic but the DTMF is more "straightforward."

4.2 BGR Objective and Subjective Results

BGR users exhibited high success rates in both modalities when compared to OpenPhone users. Results of the DTMF tasks that compare success rates of the DTMF modality are presented below:

Table 4: Results for Users Who Did Both DTMF Tasks

	DTMF Task 1	DTMF Task 2	Total
Correct	26	27	53
Incorrect	1	0	1
Average success rate	96.3%	100%	98.1%

The results in Table 4 are for 27 users who did both tasks 1 and 2 in DTMF modality in the proper manner.

Results of the ASR tasks that compare the success rates of the ASR modality in BGR are presented in Table 5 below:

Table 5: Results for Users Who Did Both ASR Tasks

	ASR Task 1	ASR Task 2	Total
Correct	23	27	50
Incorrect	4	0	4
Average success rate	85.2%	100%	92.6%

The results in Table 5 are for 27 users who did both tasks 1 and 2 in ASR modality in the proper manner. There was no statistically significant difference found in using chi-square test for comparing differences in the proportion of correct answers between task 1 and task 2in both DTMF and ASR modalities of the BGR system and the same results were obtained when using Fisher's exact test.

The BGR subjective questions were conducted in the same manner as in the OpenPhone system by asking similar questions after each session and then asking the participants to compare the two modalities after having performed tasks in both modalities. In total, out of the entire set of 27 participants who participated in the BGR tests, 23 preferred ASR and four preferred DTMF, which produced an 85.2% preference of ASR over 14.8% for DTMF. The BGR users were all decisive in their preferences and none of them preferred both modalities as in the case of OpenPhone. These users considered the ASR to be easier to use and they felt that they had more control when giving verbal commands than pressing numbers in the DTMF modality. The two users who preferred the DTMF mentioned that their preference was based on the fact that they are more used to DTMF than ASR as it is the only modality that they have used in IVR systems.

5. Analyses

In OpenPhone, the DTMF modality was preferred significantly over the WOZ modality notwithstanding the fact that the two modalities did not differ significantly in the overall success rates achieved. The researchers had imagined that, due to the intended users' oral tradition, the speech-enabled WOZ system would be preferred by a large margin since the participants use their local language to interact with it. The researchers had believed that the single modality of interaction of the speech recognition substitute system (WOZ) would be preferable because information is exchanged to and from the system using only speech. With the DTMF system users employ two modes of interaction with the system, i.e., listening to speech and pressing corresponding buttons. It was interesting to observe that even people who experienced more success with the WOZ still chose the DTMF as their preferred modality of interaction. In BGR the overall success rate was not significantly different but the users preferred the ASR with an even greater extent than the users who preferred DTMF in the OpenPhone.

In both systems there was some learning effect on both modalities whereby the 2nd tasks in both modalities were performed with equal or more success than the 1st tasks, notwithstanding the fact that the 2nd tasks were more difficult than the 1st tasks.

6. Conclusions

In the OpenPhone system the users preferred DTMF over WOZ modality and the success rates showed a higher learnability in DTMF than in WOZ. In BGR the users preferred ASR and there was also higher learnability in ASR than in DTMF. Patel et al. [17] established preference of touchtone (DTMF) over a speech input voice user interface, and stated that, "participants using the DTMF interface demonstrated a significantly greater performance improvement between the first and third task" [16]. By using the two case studies in this paper and also [16] and then deploying data triangulation we can conclude that there is a positive correlation between oral users' choice of technology and learnability of the technology. Data triangulation entails gathering data through several sampling strategies, so that slices of data at different times and social situations, as well as on a variety of people, are gathered [4]. In view of the three aforementioned case studies (the OpenPhone, the BGR and [16] - a case study that was also conducted in a developing region for rural farmers), through data triangulation we can conclude that users of IVR systems in developing regions prefer the technology that exhibits more rapid learnability.

References

[1] An analysis of African oral traditions, (2010). [Online] Available: http://www.helium.com/items/1008065-an-analysis-of-african-oral-traditions [Accessed 22 November 2010].

[2] Bevan, N., and Macleod, M., Usability measurement in context. Behaviour and Information Technology, 13, pp.132-145 (1994).

[3] Delogu, C., Di Carlo, A., Rotundi, P., and Sartori, D. (1998). A comparison between DTMF and ASR IVR services through objective and subjective evaluation. In *Proceedings of IVTTA'98* (Sydney, Australia, November 30–December 4, 1998), pp. 145–150.

[4] Denzin, N. K. (1970). The Research Act in Sociology. Aldine, Chicago, USA.

[5] Folmer, E., and Bosch, J., (2004). Architecting for usability: a survey. *The Journal of Systems and Software*, 70 (1-2) 2004, pp 61–78.

[6] Foster, J.C., McInnes, F.R., Jack, M.A., Love, S., Dutton, R.T., Nairn, I.A., and White, L.S., (1998). An experimental evaluation of preferences for data entry method in automated telephone services. In *Behaviour & Information Technology*, Volume 17, Issue 2, pp. 82-92.

[7] Holzinger, A. 2005. Usability engineering methods for software developers. Commun. ACM 48, 1 (Jan. 2005), pp. 71-74.

[8] ISO 9241-11. Ergonomic Requirements for Office Work with Visual Display Terminals - Part 11: Guidance on Usability, 1998.

[9] Jokela, T., Iivari, N., Matero, J., and Karukka, M., (2003). The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11. In *Proceedings of the Latin American Conference on Human-Computer interaction*. Rio de Janeiro, Brazil, August 17 - 20, 2003. CLIHC '03, vol. 46. ACM, New York, NY, pp. 53-60.

[10] Karis, D., (1997). Speech recognition systems: Performance, preference, and design. In *Proceedings of 16th International Symposium on Human Factors in Telecommunications*, May 12–16, 1997, Oslo, Norway, pp. 65–72.

[11] Lee, K. M. and Lai, J., (2006). Speech Versus Touch: A Comparative Study of the Use of Speech and DTMF Keypad for Navigation. International Journal of Human-Computer Interaction, 19 (3), pp. 343-360.

[12] Medhi, I. and Toyama, K., (2007). Full-Context Videos for First-Time, Non-Literate PC Users. In *IEEE/ACM International Conference on Information and Communication Technologies and Development*, Bangalore, India.

[13] Ndwe, T.J., Dlodlo, M.E., Mashao, D.J., (2007). Usability engineering of an interactive voice response system in a diverse-cultured and multilingual setting. *International Conference on Systems, Computing Sciences and Software Engineering (SCSS 2007).* 3-12 December 2007, pp. 554-559.

[14] Nielsen, J., (1993). Usability Engineering. AP Professional, Boston, MA, USA.

[15] Nuance Communications Inc., (June, 2003). Speech Satisfies a Broad Consumer Market. [Online]. Available: http://investor.nuance.com/ireye/ir_site.zhtml?ticker=nuan&script=410&layout=6&item_id=425519 [Accessed: 28 July 2003].

[16] Patel, N., Agarwal, S., Rajput, N., Nanavati, A., Dave, P. and Parikh, T.S., (2009). A Comparative Study of Speech and Dialed Input Voice Interfaces in Rural India. In *Proceedings of CHI 2009*, pp. 51-54.

[17] Patel, N., Chittamuru, D., Jain, A., Dave, P., and Parikh, T.S., (2010). Avaaj Otalo — A Field Study of an Interactive Voice Forum for Small Farmers in Rural India. In *Proceedings of CHI 2010*.

[18] Shackel, B., (1991). Usability—context, framework, definition, design and evaluation. In *Human Factors For informatics Usability*, B. Shackel and S. J. Richardson, Eds. Cambridge University Press, New York, NY, pp. 21-37.

[19] Sherwani, J., Ali, N., Mirza, S., Fatma, A., Memom, Y., Karim, M., Tongia, R., and Rosenfeld, R., (2007). HealthLine: Speech-based access to health information by low-literate users. In *IEEE/ACM International Conference on Information and Communication Technologies and Development*, Bangalore, India.

[20] Sherwani, J., Ali, N., Rosé, C., and Rosenfeld, R., (2009a). Orality-Grounded HCID: Understanding the Oral User. *Information Technologies & International Development, 5*(4), pp. 37-49.

[21] Sherwani, J., Palijo, S., Mirza, S., Ahmed, T., Ali, N., and Rosenfeld, R., (2009b). Speech vs. touch-tone: Telephony interfaces for information access by low literate users. In *Proceedings of the IEEE International Conference on ICTD*, pp. 447–457.