Sambad – Computer Interfaces for Non-literates

Sagun Dhakhwa, Patrick A.V. Hall, Ganesh Bahadur Ghimire,
Prakash Manandhar, and Ishwor Thapa

Madan Puraskar Pustakalaya, PO Box 42, PatanDhoka, Lalitpur, Nepal
rite2sagun@gmail.com, p.a.v.hall@btinternet.com,
gghimire@gmail.com, jsaaymi@gmail.com, and rowhsi@yahoo.com

Abstract. Much of the world's knowledge is captured in writing, and shared through writing, and as such is inaccessible to the one eighth of the world's population who are illiterate. We are developing a software system for the use of this population based on speech and images without written text. We have evaluated basic interaction devices and simple interface metaphors to arrive at the design of an overall interface that is attractive to and usable by illiterate people. We report our usability experiments, and describe our system.

Keywords: HCI, illiterate, speech.

1 Introduction

Of the world population around 6 billion, almost 800 million adults are illiterate and about 100 million children are out of school [19]. While many have followed literacy programs, illiteracy is remarkably persistent, being coupled to poverty and traditional ways of surviving. Access to knowledge and the sharing of their indigenous knowledge is limited by their inability to read and write and the lack of suitable technology.

One function of writing is, clearly, to record speech, so can't we support illiterate people directly by providing technology for speech which can do everything we take for granted for writing? A move from writing to speech is a very bold step; writing has been an integral part of our society for the past 2,000 to 3,000 years. Walter Ong [15] has analyzed the distinct differences between oral and literate cultures, identifying the key contribution of writing as the ability to record thoughts and ideas reflect upon them and communicate them to others and pass them on to future generations.

All too often scholars have claimed that writing is the very foundation of our civilization such that those who cannot read and write, for whatever reason, would be regarded as "uncivilized". Roy Harris, in his book Rethinking Writing (2000), begins with: "When Boswell objected to Johnson's calling the Chinese 'barbarians', Johnson replied curtly: 'Sir, they have not an alphabet." [6 p1] and goes on to catalogue a great number of European scholars from the 17th century to the present day who equated writing, and with it literacy, with being 'civilised'.

Roy Harris [6], along with Gunter Kress and Theo Van Leeuwen [10] identify a new multimodal literacy as emerging in which the role of writing is diminished, the importance of speech is elevated, and pictures and images become central.

J. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2007, LNCS 4550, pp. 721–730, 2007.

The computer was first introduced into Nepal more than three and half decade ago, a second generation IBM 1401 to expedite the analysis of national census data. Over the years, computers have evolved from an isolated machines confined within limited organizations to an essential gadget in consumer's home, but still accessibility is limited to the literate mass. Localization, e-governance, e-medicine, and rural telecentres have helped further distribution of the technology in rural areas. But access to and use of these technologies remains extremely uneven. This disparity is a reflection of deeper socio-economic inequalities both between and within societies. Computers and the Internet are often considered effective instruments to empower people, reduce poverty and improve the lives of the people; it has just deepened already existing inequalities and divisions. Providing support in ICT for illiterates and oral communicators can be one of the most effective ways to alleviate this inequality.

On the Sambad project we aim to support illiterate people directly, as part of this move towards the new multimodal literacy. We believe if support could be given for the non literate people through appropriate interfaces, they can benefit from computers. We are developing a robust Multimodal Web Authoring and Browsing Tool (MWAB) using which non-literate users will be able to share their knowledge via their web pages using speech instead of text. We are designing computer interfaces taking into account the cultural needs of illiterates. Illiterate users may not be very sophisticated in their experience of technology, and need simple interfaces. We are proceeding cautiously, and will continue to do so, developing a prototype increment and then evaluating it. Ultimately we want functionality which:

- identifies and authenticates the user
- enables composition and editing of multimedia documents of speech and pictures
- enables storage and retrieval of collections of multimedia documents
- communicates documents to others

2 Information Needs Assessment

For assessment of the information and requirement analysis for the Sambad, we surveyed telecentres in rural areas across Nepal. We conducted semi-structured interviews with 453 local people out of which, 277 were illiterates and 176 literates. We asked all the respondents about their familiarity on computer, its usefulness and applications. 70% of illiterate respondents were from rural areas of which 67% were female. The literate participants were primarily from telecentre management, telecentre users, and computer entrepreneurs. We observed both, the rural lives in two villages of west Nepal and the services provided by 45 institutions, including cybercafés, community telecentres, FM-radio stations, and NGOs/INGOs. We also conducted 14 separate focus group discussions, 4 with the non-literate village dwellers, 3 with media forum people, 3 with ICT entrepreneurs, and 4 with development practitioners and local political leaders in order to find what kind of information and applications could be useful for these rural illiterate communities.

Though many illiterates believed that computers were not meant for them as they were unable to align with the technology, it was evident from our survey that there is a strong thirst in all illiterates for knowledge, pictorial manuals or audio books on pest controlling, farming, bee-keeping, primary health, and so on. From this experience

with illiterates, we learnt that many illiterates do at present act as indigenous, oral knowledge base for their fellow villagers on issues like curing diseases, good farming practices, etc. Preservation of this traditional knowledge, particularly from indigenous people, is seldom found. Once lost, this oral based knowledge cannot be retrieved [8].

In addition, people in villages are accustomed to indigenous communication in the forms of festivals, story telling, folk dance and songs, etc.; all of them portraying some message on social values, education, politics thereby supporting oral and cultural continuity. We felt the necessity to produce software that addresses this existing oral practice.

We also found that they can understand image and audio irrespective of their literacy level when we demonstrated our text free interfaces. The rural community is in need of information on basic health, agriculture, education, governance etc, in a form they can understand. In response to address the community needs, we are building a system that in general can generate, edit, store, and share the contents as speech in their own language and voice for accessing and contributing the information for both literates and non-literates on the same platform.

Nepal has 138th rank in the human development index out of 177 countries. The human poverty index is 38.1 [7]. The rural economy is based on agriculture mainly. Education and Information facilities are limited and basic health facilities are very poor in rural areas. There are around 28000 primary and secondary schools across the country, and only 1259 government employed doctors and 89 public hospitals in Nepal [3]. Most of the illiterates dwell in these rural villages. Information on very basic things that affect the livelihood of the villagers like agriculture, basic health, education and commerce can have a greater significance to the rural communities.

On our visits to different telecentres of mid western and far western part of Nepal we found that ICT facilities had now penetrated to rural communities, over the years with the initiatives by telecentres movement in Nepal. Till date 240 rural telecentres have been established in Nepal. However, the information that is available through these centers is mostly text, posted on a website or in print. We witnessed an active role of literate intermediaries during our field visits and interviews, and came across many non-literates entering telecentres to send and receive emails and chat with their children abroad. We found one such case in a telecentre in Nepalgunj (west Nepal), where a woman, who could not read and write, came into the telecentre and text chatted with her daughter abroad with the operator there as intermediary. For us it is an irony that she happens to be a regular visitor there and she always waits for the operator's spare time. We doubt whether the intermediary could express her exact feelings to her daughter in words.

3 Text-Free Technology

As a lead into this project in mid-2005, Roger Tucker focused on the composition of speech messages, believing that these should be as easily produced and edited as are written messages. Work in this area goes back to the 1980s on the Etherphone project (Zellweger et al [22]), part of the wider Cedar project within Xerox in California. This work was later picked up by Roger Tucker and colleagues at Hewlett Packard when looking for speech tools for personal digital assistants (PDAs), calling this 'speech as

data' (Tucker et al [17]). The idea in both these systems is to display speech in a word-like form shown in Fig. 1, a sequence of speech 'chunks' separated by non-speech 'silences'. In Tucker's design study the ideas were reviewed not by illiterate people but by another group of writing-disabled people, dyslexics. This proved very useful.

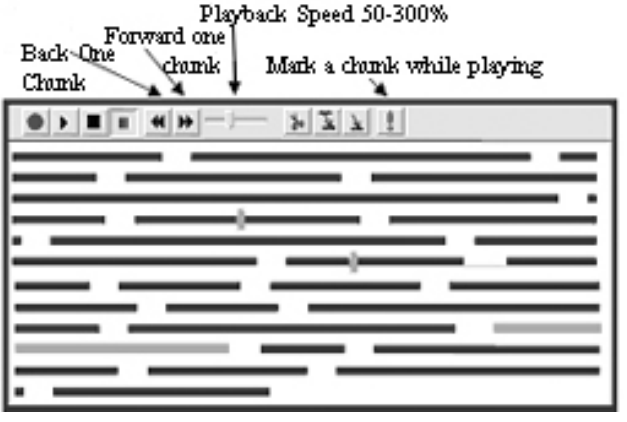

Fig. 1. Taken from Tucker's Design Report [16] for Sambad

Software to edit speech as illustrated in Fig. 1 had been implemented within Hewlett Packard but is not available to us, so on Sambad we have been developing our own version of this, using publicly available algorithms for it, and building on developments in speech processing since that early work, adding some innovations of our own.

3.1 Text Free User Interfaces for Non-literates

The normal interfaces to computers, such as Microsoft Windows, and the one shown in Fig. 1, are relatively complex, with their metaphors of file paths, menus and command lines. They are targeted mainly at literate users and encompass the metaphors of office environments. It is evident that illiterates are equally able to use modern technologies, for instance televisions, radios, telephones, and motor vehicles, if provided appropriate training and interface [18].

However computer interfaces may include text and unfamiliar metaphors. We believed that using understandable pictures instead of text could be one of the approaches to reduce such complexities. We knew from Medhi et al [13] that this should be the case, but wanted to find out what kind of interactions illiterate people felt most comfortable with.

We started by assessing basic interaction devices – keyboard and mouse, and also touch screen. For this trial we developed a simple text free audio recorder, as a prototype, with pictorial icons and buttons with audio captions. From other studies we know that pictures and audio can be used to guide illiterate users from one interface to other [11]. We designed widgets and UI components using metaphors that we believed would be readily understood by illiterate and technically inexperienced people (see Fig. 2 and 3).

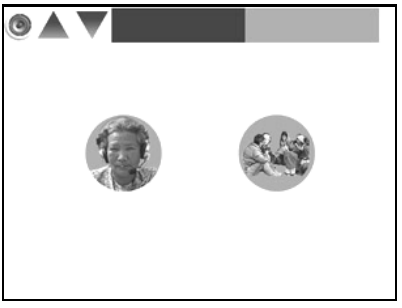

Fig. 2. Screen to select play/record

Fig. 3. Leaf and Rupee notes for rating

The Godavari Experiment

We undertook a usability trial at the Godavari Marble Works just south of Kathmandu. For this and other usability assessments we followed the combined approach involving observation, performance measure, open interviews and focus group methods [2], [14].

The main objective of the experiment was to assess basic interaction devices – keyboard and mouse, and also touch screen – to find which interaction device would be the most appropriate. Underlying this lay our overarching objective, to find out what form of text-free user interfaces were appropriate for non-literates.

We established a usability laboratory in the marble works, and arranged a number of usability assessment trials over several days, with subjects shown in Table 1.

Table 1. Subjects for Godavari experiment

Subjects		**Age**	
Sex	**N**	**Average**	**Std. Dev**
Male	13	30.92	11.21
female	23	29.96	9.06
Total	**36**	**30.31**	**9.74**

Each trial was attended by 4 or 5 people. The subjects were initially interviewed to record basic facts about them, in particular their level of literacy. They were then shown a short video on the computer about the importance of education to get them used to the computer and surroundings, and then introduced to using the computer and the interface shown in Fig. 2. This system asked them to do a number of tasks using the interface device being evaluated, before they went on to listen to pre-recorded stories or songs by other non-literate people or record their own story or song. During this we recorded all their keystrokes for later analysis, videoed them with two cameras, and observed them taking notes of all significant behaviour.

The number of subjects was really very small, so though differences were found between different categories of subject and between different devices, none were statistically significant excepting for one. Table 1 summarises the timing for interaction devices, showing that the mean time for the touchscreen is more than two standard deviations less than the times for the other devices. From this we concluded that the touch screen was significantly easier to learn and to use. We could have predicted this, but here it was in the experimental evidence.

Table 2. Mean Time Consumed by participants using different input devices

Used input device	Mean	N	Std. Deviation
Mouse	8.9982	11	2.7455
Touchscreen	4.3638	8	1.4876
Keyboard	8.5125	4	2.4151
Custom Keyboard	8.4750	4	1.1857
Total	7.4756	27	2.9342

During the experiment we also observed the performance of each participant on their attitude, ability to use and understanding of the software interface. On each screen we graded them Excellent, Good and Poor. From these grades we arrived at an overall grade shown in the bar chart in Fig. 4, confirming the conclusion above.

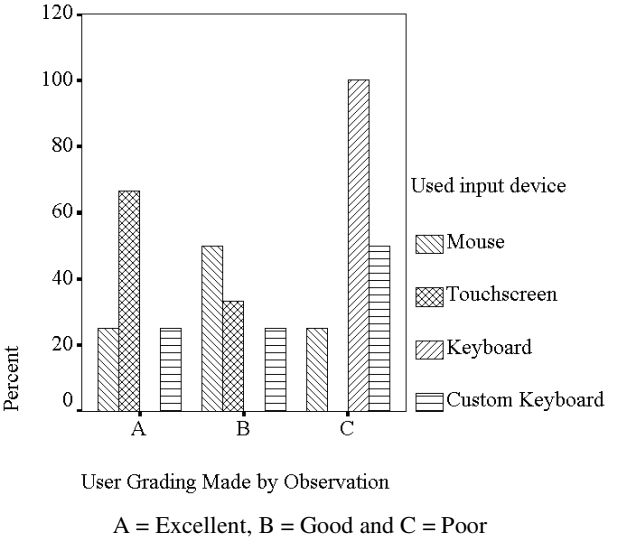

Fig. 4. User performance on different input devices

We found that the text free interface was readily understood by everybody, though some found the spoken instructions too long and stopped attending to these. We have demonstrated this system to a range of people in the context of explaining what we did. Some of the people who had seen this simple record and playback system have seen its potential for recording oral histories and indigenous knowledge.

3.2 Appropriate Metaphors

Nielsen's second user interface design heuristic is to "speak the user's language", and in particular use appropriate metaphors [14], a point also taken up by Alan Blackwell [1], with Evers [5] emphasizing the user's cultural background.

A metaphor enables people to quickly learn to use a system by mapping the real world objects onto software objects [12]. Expectations about how a metaphor object

in the software would behave will depend on the cultural use of the object in the real world, its affordance [4]. In a country like Nepal with diverse ethnicity and culture, the choice of metaphor takes on a locally cross-cultural dimension. We used knowledge from the surveys and our own knowledge of life in Nepal to select metaphors.

In our second prototype we incorporated several metaphors like "village", "cupboard" and "exit door" etc. all closely connected with daily activities in Nepal. These metaphors were assisted by audio captions. We also used image based login system, pictures and photographs are not new in any culture today; though there is some evidence that sequences are difficult for people to memorize [9]. We had explained the analogy of pass images to "Secret Spell" during usability trials which the users accepted well.

Fig. 5. The Village Metaphor

Fig. 6. The Cupboard Metaphor

The village metaphor (Fig. 5) was adopted from a CD produced by UNICEF, Hyderabad but also incorporated into a system in Indian villages [9]. This classifies the information available, in effect the root directory, the top level filing cabinet and folder structure, with sub-directories home, school, health-post, market, etc., clicked on to select. Evers [5] used a picture of a university campus in an equivalent manner in her researches.

The cupboard is the most familiar storage device used by people in Nepal, familiar even for the storage of paper files, as anybody who has visited a government office will have experienced. We presented our users with a cupboard UI component to hold their files with drag and drop to move them, in fact two cupboards, one private and one public (see Fig. 6). We included simple file previewing like audio playing with speed adjustment and volume control and an image and album preview. The combination of village and cupboard metaphors gives a fixed three level hierarchy.

Bungamati Experiment

We undertook usability evaluations of these metaphors at a Telecentre located just south of Kathmandu. We had developed a prototype having a login system, village and cupboard metaphor, and photo/album viewer. We had 13 female participants with age varying from 23 to 51 who had recently joined the CLC (Community Learning

Center) program for illiterates at the Bungamati Telecentre. These participants were functionally illiterate though some of them could read and write a few letters.

From our first experiment we had concluded that the touch screen was the most appropriate user interaction device for the illiterates; however in our second experiment we only used the mouse because the participants had already seen a mouse before and six of them had some experiences of using mouse during literacy classes. However, the Sambad software is enabled for touch screens.

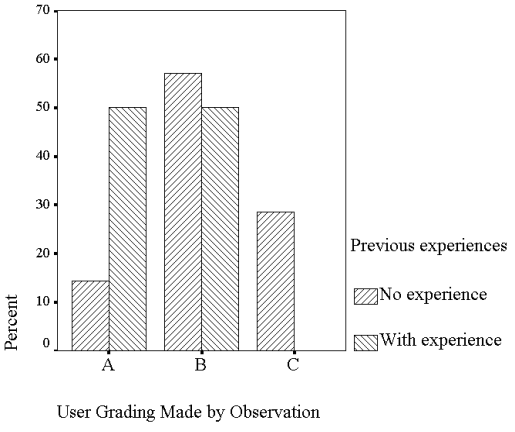

Fig. 7. Performance Grading based on previous experience

We assigned fixed tasks to individuals and two of us observed them, making notes. We found that 4 of them had difficulties about how to control the mouse and use the widgets with the mouse. One of the participants found it difficult to comprehend audio captions in Nepali while the rest of them had no such problems; we concluded that this particular subject had difficulty because Nepali is her second language. We also found that the participants with prior experience in using the mouse felt at ease using Sambad UI widgets. The first time computer users required an introduction to the computer and its peripherals, like using of mouse, etc before using the software, and then the more time to understand and use the software (see Fig. 6). Subjects without prior experience of computers were graded C with only a single A, while those who had prior experience all got A or B. The participants who got C had major difficulties in using the slider component as they didn't have full mouse control.

Since the majority of the participants could use the metaphors without much difficulty we can conclude that the metaphors we developed were usable. We found that the participants felt that our metaphors were interesting and entertaining, because the UI was colorful and simple; and it allowed users to view images and recorded speech. One participant was very enthusiastic with the interface because she had heard from her daughter that it takes at least 2 months for literate people to learn computers, but she was already using the computer using the Sambad interface. She said, "Wow! It's damn easy to use; I will take this home and use it in my daughter's computer."

3.3 A Note on Current Technology

User interfaces without text rely on audio technologies. We reviewed Voice XML [21] but found this impractical for Sambad. We use extended Java Swing components with audio encoded in SPEEX format. SPEEX is an open source and patent free speech compression library [20]. We found that 8KHz 16 bit audio file compressed by the SPEEX algorithm is of acceptable quality for voice instructions.

We have developed a TTS system for Nepali using the Festival and Festvox systems. We intend to incorporate Nepali TTS into Sambad system for all texts. We have produced a basic speech corpus, with letter (and syllable) to sound rules, plus a 6,000 word lexicon of syllabification, and a lexicon of exception pronunciations.

We are working in Nepali, though the technology we are developing, apart from TTS, is independent of any specific language. The user interface is localizable in terms of audio captions and culture specific icons, pictures and logos.

4 Conclusion

We envisage a situation where access to computers for illiterate people will be possible, so that they can undertake all the basic functions that literate people take for granted. The interface to the computer will be different, visual with speech and little or no writing. It will also be less cluttered, with most of the complexity of normal operating systems hidden. We envisage that the computers that will be used will be in publicly accessible community telecentres or knowledge centres. However we also recognize that for somebody who is illiterate such places may seem forbidding, and that there will need to be some persuasive reason for entering.

In future we intend to extend our system to be used by visually impaired and physically disabled people so that they can also benefit from computers and the Internet. We also want to extend the system to support literacy education.

We might even contemplate a future such as that suggested by Harris (2000): "It is by no means out of the question that in some parts of the world literacy will simply be bypassed by the advent of new communications technologies which make it unnecessary to teach the skills of reading and writing to the whole population" [6 p12]. Perhaps Sambad, as we develop it, will be that new technology.

Acknowledgments. We would like to thank the Leverhulme Trust for funding this project, and Madan Puraskar Pustakalaya for hosting the work.

References

1. Blackwell, A.F.: Metaphor in Diagrams, University of Cambridge (1998)
2. Barnum, C.: Usablity testing and research. Longman, New York (2002)
3. Central Bureau of Statistics, Government of Nepal (2006) http://www.cbs.gov.np/Nepal%20in%20figure/nepal%20in%20figures%202006.pdf
4. Norman, D.: The Design of Everyday Things. Currency; Reissue edition (1990)
5. Evers, V.: Cross-cultural understanding of metaphors in interface design. Cultural Attitudes towards Technology and Communication, London (1998)

6. Harris, R.: Rethinking Writing. Continuum International Publishing Group, Academic Press (2005)
7. Human Development Report. UNDP (2006) http://hdr.undp.org/hdr2006/statistics/countries/country_fact_sheets/cty_fs_NPL.html
8. Indigenous Peoples of Nepal and Traditional Knowledge. In: International Workshop on Traditional Knowledge Panama City (September 21-23, 2005)
9. Nielsen, J.: Designing for the cultural "other" - A global perspective on ICT and illiteracy (2007) http://ideas.repec.org/p/hhs/cbsinf/2005_011.html
10. Kress, G., Van Leeuwen, T.: Multimodal discourse: The modes and media of contemporary communication. Hodder Arnold (2001)
11. Goetze, M., Strothotte, T.: An Approach to Help Functionally Illiterate People with Graphical Reading Aids. Smart Graphics Symposium, UK (2001)
12. Stefan, M., Seetharam, D.: WeatherTank: Interface for Non-literate Communities and Ambient Visualization Tool (2006) http://alumni.media.mit.edu/ deva/projects.shtml
13. Medhi, I., Sagar, A., Toyama, K.: Text-Free User Interfaces for Illiterate and Semi-Literate Users. In: Paper presented at the International Conference on Information and Communications Technologies and Development, Berkeley, CA (May 25-26, 2006)
14. Nielsen, J., Jakob, N.: Usability Engineering. Academic Press, San Diego (1993)
15. Ong, W.J.: Orality and Literacy: The Technologizing of the World. Routledge (1982)
16. Tucker, R.: An Audio-Visual Webpage Editor for Limited Literacy Users. Report commissions by Pat Hall prior to starting (2005)
17. Tucker, R., Robinson, A.J, Christie, J., Seymour, C.: Recognition-Compatible Speech Compression for Stored Speech. In: Proceedings of ESCA ETRW Workshop on Accessing Information in Spoken Audio, pp. 69–72 (April 1999)
18. Understanding Non-Literacy as a Barrier to Mobile Phone Communication (2006) http://research.nokia.com/bluesky/non-literacy-001-2005/index.html
19. UNESCO Institute for Statistics (2006) http://www.uis.unesco.org/ev.php?URL_ID=6401&URL_DO=DO_TOPIC&URL_SECTION=201
20. Valin, J-M.: The Speex Codec Manual v. 1.0.4. (2006) http://www.speex.org/
21. Voice Extensible Markup Language (VoiceXML) Version 2.0. World Wide Web Consortium (2006) http://www.w3.org/TR/voicexml20/
22. Zellweger, P.T., Terry, D.B., Swinehart, D.C.: An overview of the Etherphone System and its Applications. In: Proc 2nd IEEE Conf on Computer Workstations, pp160–168 (1988)