Development of Multimodal User Interfaces to Internet for Common People

Debasis Samanta[1], Soumalya Ghosh[2], Somnath Dey[3], Sayan Sarcar[4],
Manoj Kumar Sharma[5], Pradipta Kumar Saha[6] and Santa Maiti[7]
School of Information Technology, Indian Institute of Technology Kharagpur, India – 721302
[1]*debasis.samanta.iitkgp@gmail.com,* [2]*soumalyaghosh@gmail.com,* [3]*somnath_dey2003@yahoo.co.in,*
[4]*mailtosayan@gmail.com,* [5]*manojsharma.net@gmail.com,* [6]*itsmepradipta@gmail.com,* [7]*santamaiti@gmail.com*

***Abstract*—In recent years, there is a rapid advancement in Information and Communication Technology (ICT). However, the explosive growth of ICT and its many applications in education, health, agriculture etc. are confined to a limited number of privileged people, who live in digital pockets. We term these underprivileged people as common people, who are digitally illiterate as well as language illiterate. According to the UNESCO report, population of such people in the globe is 64% and in developing countries like in India, China etc. it is around 76%. It is therefore an urgent need to include this large set of population in the ICT revolution. This issue has been addressed in this paper and we propose to develop a solution in ICT for common people. We propose a multimodal interface mechanism to Internet so that common people can interact to Internet with their mother languages, speech and icons. The proposed approach has been tested with people in India and results are presented.**

***Index Terms*—Information and communication technology, human computer interaction, Internet, multimodal interaction, information retrieval**

I. Introduction

People retrieve information from Internet by giving search keywords to search engine. The search engine is invoked through a web browser which then presents information returned by the search engine to navigate large information repository of World Wide Web. This large repository is mainly maintained in English. A global scenario of web contents in different languages in the Internet is shown in Fig. 1. From Fig. 1 it is clearly evident that major contents in WWW is in English. This shows a clear limitation to the 64% population of the globe[1], who are unfamiliar to English and hence unable to avail the Internet advantages. The issue is more critical in developing countries like India, China, and Pakistan etc. where 76% people in these countries[2] are English illiterate and hence deprived from the advantages of Internet facilities.

To facilitate Internet access in users' language, presently most of the web browsers support retrieving information in many users' languages. These browsers usually support UTF-8, which is variable length character encoding for Unicode to represent and manipulate text in different websites like Wikipedia etc. Nevertheless, this attempt is unable to bring complete information and meaning in English web pages to English illiterates. It is also observed that these interfaces are

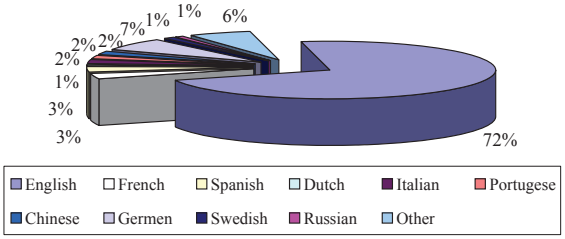


Fig. 1. Web contents in major languages in Internet

not useful to the people who are having very little knowledge about web browsing. The major problem with the existing web browsers is that in order to access web pages in users' mother language, user has to enter keywords in English but phonetically equivalent to users' mother languages [1]. Indeed this is a limitation for the users who are unfamiliar to English. In some recent efforts, virtual keyboards [2] in users' languages have been proposed such as Microsoft's on-screen virtual keyboard [2], Google's Indic keyboard [2], ILeap [2], WIVIK [3] etc. addressing the limitation. However, these approaches are yet to be matured and hardly a strong contender to the English hardware keyboard for text entry mechanism. Further, note that these approaches are limited to Unicode based search only. This means search is confined to the web pages in users' mother languages. In addition to this, there is a large population such as farmer, shopkeepers, driver, rickshaw pullers etc. who can not read/write even in their own mother language. For these people, the conventional text based information even in mother languages is not useful. For them we need some alternate approach of interaction(s). Also note that population of such illiterate people is not negligible compared to that of literate community. Figure 2 and 3 show the percentage of population in the world and India belonging to the category of underprivileged people so far the illiteracy is concerned.

Figure 2 shows the population of literate people in the world[3], country wise. It shows that there are many countries where literacy is below 50% of the population of these countries[4]. The situation is worse in countries like India where

[1]http://unesdoc.unesco.org/images/0016/001628/162808e.pdf
[2]http://theviewspaper.net/illiteracy-in-india
[3]http://en.wikipedia.org/wiki/Literacy
[4]http://www.unesco.org/en/literacy/resources/

978-1-4673-4369-5/12/$31.00 ©2012 IEEE

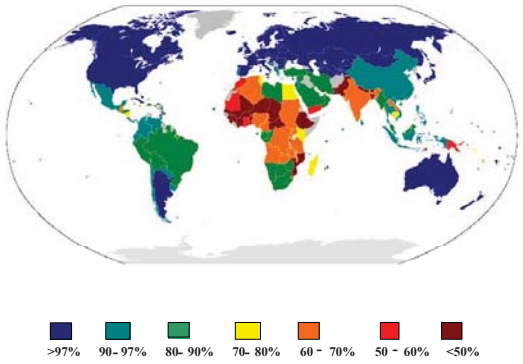


Fig. 2. Population of literacy in the world

around one billion people (64.8% of the total population) of the country are language illiterate. More contrast, if we take into account the computer illiteracy. We find that (Fig. 3) only 6.15% of the total population can use computer or know about computer applications[5].

This observation motivates us to develop a multimodal user interface so that common people irrespective of their backgrounds can access the Internet. In this work, we have properly defined users as common people and then classify them in three categories. For each category of users, we propose a mode of interaction such as text, speech and icon. We also propose required processing tasks and technology involved in each mode of interaction. The proposed framework has been developed for Indian users and our findings are summarized.

The rest of the paper is organized as follows. In Section II, we classify the users as common people and precisely state the objectives of the work. The proposed approach is discussed in Section III. Section IV discusses the case study in the context of two Indian languages, namely Hindi and Bengali. The results of user testing are discussed in Section V. Finally, Section VI concludes the paper.

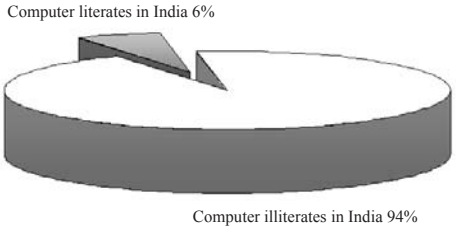


Fig. 3. Population of computer literacy in India

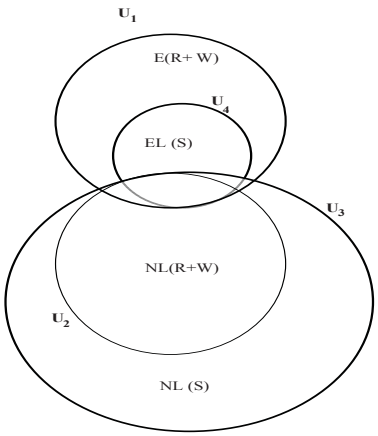
II. Scope and Objectives

We can classify population of a region on the world according to their languages compatibility. In this classification, we treat English as the main language (as it is the international language). Any other language can be treated as the mother language of users. In other words, the proposed classification is based on these two types of language compatibility of the users. The proposed classification is shown in Fig. 4(a).

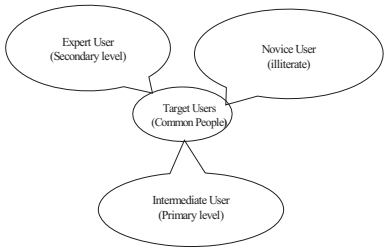
There is a group of users U_1, who are comfortable with English language so far reading and writing capabilities are concerned. There are some users who can not read or write, but can speak or understand English language (as shown U_4 in Fig. 4(a)). We categorize users' another group U_2, who are literate in their mother languages, that is, they are well conversant with respect to reading and writing skills in their mother languages. We also classify another group of users as U_3, who can not read or write but communicate their thoughts in mother languages. With respect to this classification, we may note that existing interfaces to Internet are able to cater to the needs of user group U_1 and to some extent user group U_2. There is no interface available to the best of our knowledge with which a user in group U_3 can interact with the Internet.

Further, if we consider the education levels of user in group U_2 and U_3, we again sub categorize them as shown in Fig. 4(b). Expert users are the users who are good in their mother languages. Examples of such users are teachers and students in secondary level and above, office employees, educated business persons, housewives etc. All these users have at least secondary level education and if they know English then it is poor.

Next, we categorize some users as intermediate users. These users may vary among all people as in expert users but their education is at primary level. Such a group of users are



(a) Users according to language compatibility



(b) Users according to level of literacy

Fig. 4. Classification of target users

[5]http://wiki.answers.com/Q/Computer_literacy_rate_in_india

English illiterate and have very poor level of reading/writing capabilities, if any, in their mother languages. However, they are good in communication in their mother languages. Finally, we categorize the novice users, who are either having no education or education, if any, then it is below the primary level and they are very poor to convey their messages to others. In other words, these users also not capable to convey their thought by speaking but can understand if someone speaks with them in their mother languages. Examples of such kind of users are petty vendors, rickshaw pullers, auto drivers, porters, domestic helpers etc.

We use the classification scheme to classify the common people in three categories as shown in Fig. 4(b). Now, we propose the interaction mode suitable for these target users. For expert users, text-based interaction is preferable. Those users like to enter keywords and retrieve information in their mother languages. For the intermittent users, interaction should be supported with speech so that they can utter their keywords and then retrieved information can be spoken out by the system. For the novice users, some pictorial mode of interactions such as gesture, icons etc. are useful to enter their search words and retrieved information also to be spoken out by the interface. In other words, three modes of interactions, namely text, speech and icon are useful to meet the requirement of interface so that any common people can interact with the Internet.

The above mentioned identification of target users and their requirement implies a scope of developing an efficient and user friendly interface embedding the different modes of interactions. The interface would takes care both language literacy and computer illiteracy while accessing information from web repository irrespective of whether the content is in English or in their mother languages.

The three objectives to bridge the digital divide are mentioned below.

- To develop user friendly text based interface for the users who are unfamiliar with English language but are well capable to read and write in their mother languages.
- To develop speech based interface for the people, who are not at all conversant with English language and also has lack of reading and writing power in their mother language.
- To provide a support to user, who cannot read/write and comfortably speak even in their mother language.

III. Proposed Methodology

In this section, we discuss our proposed approach to develop a user interface to Internet for the people as classified in earlier section. An architecture of the interface is shown in Fig. 5 As shown in Fig. 5, a user can interact to the Internet with three modes; text, speech and icon. To process each mode, a sequence of tasks is to be followed. It may be noted that the module "Search Engine" is not a part of the development. Rather, any existing search engine can be plugged into the proposed architecture. The processing of each mode of interaction and technology requirement is discussed in details in the following sub sections.

A. Text Based Interface

In the text based interface, users would be able to enter search queries in their mother languages. In order to compose the search keywords, we propose to build a virtual keyboard completely compatible to user's mother language [4]. Virtual keyboard would be supported with several rate enhancement strategies, namely word prediction [3], [5], word correction [6], [7], adaptation and personalization [4], [8] etc. to support the user. Further, the virtual keyboard would be designed with optimum arrangement of keys so that minimum movements of mouse pointer or finger are required to enter keywords. The module 1 labeled "Virtual Keyboard" in Fig. 5 which is represented to the task of search keywords entries.

Once a keyword is generated by a user, the keyword is translated into English. This is because the user's keyboard is in user's mother language through which we would like to fire the search engine to access English content in web [1]. To do this, we propose a language translator to translate keywords in users' mother languages to English keywords. The module 2 labeled "mother Language to English Language Translation" is planned for this. However, a user may be interested for Unicode based search, which is, searching for the web pages maintained in users' mother languages. In that case, users' query words would be directly parsed to the search engine in Unicode. No translation is required in this case.

Irrespective of the case, the module 2 generates strings which we called "query string" in English or User's mother language. A query string may consists of irrelevant component such as prepositions determiners (a, an, and the) etc. which are irrelevant as a search keywords. The module 3, namely "Keyword Extractor" will extract the appropriate keyword effective to search only. Further, in order to reduce the number of pages retrieved, query expansion [9], user personalization [10] etc. can be applied in this task. So, the "Keyword Extractor" module will refine the query strings. With the refined query string getting from "Keyword Extractor" module, the "Search Engine" would be invoked.

On successful invocation of search engine, it would result a number of snippets [11] or web pages. "Content Retriever", the module number 7 in Fig. 5, takes care of the snippets or content. The "content retriever" will maintain the structure and appearances of the pages and returns the tag-set in the XML/HTML page. The target is then would be translated into users' mother languages. The entire web page content would be translated in users' mother languages. We propose a language translator to translate English to users' mother languages. The module 10 in Fig. 5 is proposed to carry out this task. The output produced by the module 10 would be painted into web browsers of users' computers in their mother languages.

B. Speech Based Interface

Users would be able to enter search queries by speaking in their mother languages to the speech based interface. For creating the search keywords in text format from speech input, we propose to build a speech recognizer module, as

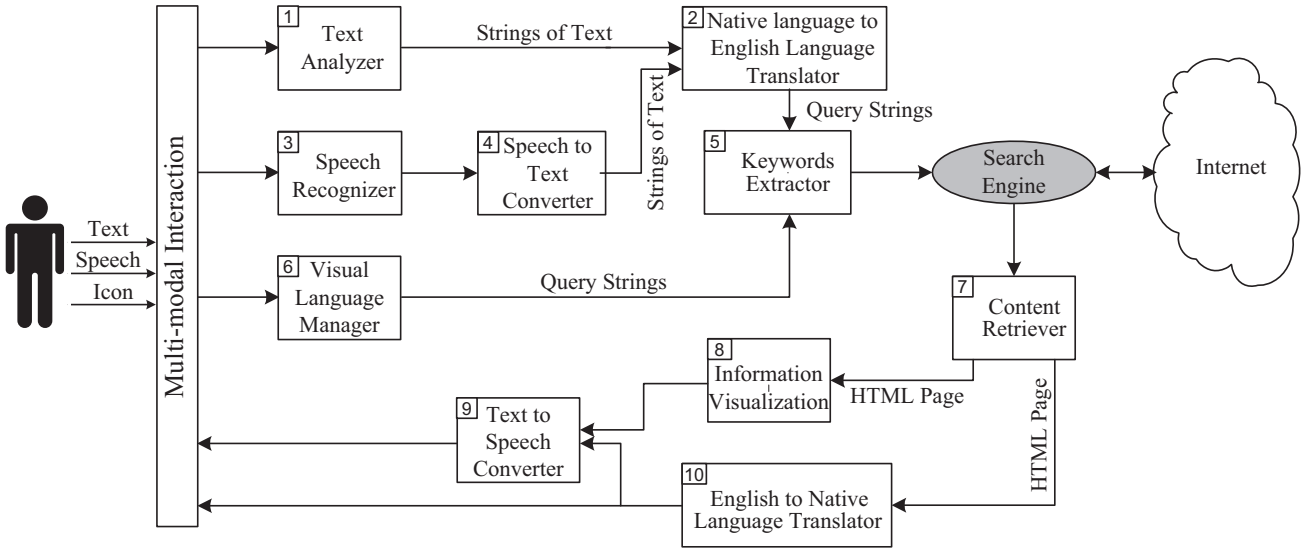


Fig. 5. Architecture of the Multi-modal user interface

suggested in Figure 5 as module 3, attuned to user's mother language. The recognizer module along with module 4 in Fig. 5, namely speech to text converter would achieve the goal by implementing sub modules like capturing the speech content spoken by user, feature extraction [12] of these audio signals, transformation of speech into feature vectors followed by the process of recognizing what was actually spoken with the help of Hidden Markov Model (HMM) [13]. The user query in text then would be formed and get inputted to translation module 2 of Fig. 5. Later tasks would be same as the text-based interaction. Further, the English to mother language translated text getting from module 10 would again be converted to speech for understanding of the intermittent user. To serve the purpose, text to speech synthesis module, pointed as module 9 in Fig. 5, would be developed. This module would comprise of several tasks like addition of letter-to-sound and syllabification rules [14], selection of text corpus to be recorded [15], recording and labeling of speech database [16], extraction of pitch markers, building of units' database etc. The output produced by the module 8 would be played from sound output device of users' computers in their mother languages.

C. Icon Based Interface

The third mode of interaction is the iconic mode [17]. A set of icons[6] would be displayed on the user's interface. Users' have to select appropriate icon in order to specify the keyword for searching. Once the keywords are selected, it would be transformed to query string. The module 6 labeled as "Visual Language Manager" would accomplish this task. Visual language also plays the role of managing a large set of icons, using icon publication to enhance the rate of icon selection, reducing (visual) search time etc. Further, an icon may be ambiguous in terms of meaning representation. An ontology mapping database in users' languages has to be maintained so that an icon can imply right meaning in the context of users' intended queries. The Visual Language would have all these implementations. The "Visual Language Manager" produces the right query string and go to search engine. The content retrieved from the search engine in the form of web pages or snippets will further get inputted to "Information Visualization" module 8. This module summarizes the web page contents and the summarized text portion will either send to speech synthesis module to get final output in speech mode or directly be mapped to visual template consists of minimal set of icons which can represent the semantics/meaning of the text.

IV. Case Study

Towards implementing the above mentioned proposed methodology, we have developed one multimodal user interface system called Bharati which supports common people in India. The Bharati web site can be accessed from http://www.nid.iitkgp.ernet.in/Bharati/. In Bharati, we have considered two Indian languages, namely Hindi and Bengali. Hindi is the official language in India and most of the people in northern, central and western part of India with 45.66% of total population speak in this language. Bengali language, on the other hand, is a majority spoken language in eastern part of India (27.78% population of India)[7] and neighboring country Bangladesh (100%). Bharati system considers text, speech and icon in these two Indian languages only. The three different modes of interfaces are briefly described in below subsections.

A. Text Interface

To support the users in query generation task for accessing search engine, keyboard would become invaluable part of the

[6]http://en.wikipedia.org/wiki/Wikipedia:Icons/Icons_for_Wikipedia

[7]http://censusindia.gov.in/Data_Products/Library/Provisional_Population_Total_link/PDF_Links/chapter5.pdf

entire text interface module. According to definition, virtual keyboard is one type of soft keyboard which replaces hardware keyboard on a computing device. It means that whenever a user is trying to type desired keyword in the search textbox on mother language, a keyboard displaying on the screen would help him of exempting from the task of key tapping in mother language or English mapped hardware QWERTY keyboard, and also providing the facility of single mouse click for entering a character. Virtual keyboards are being developed in two Indian languages which are our main concern, namely Hindi and Bengali. Unlike standardized English language virtual keyboard, the developed keyboard layouts consist of several extra language oriented features which need to be handled.

To enhance the typing speed and minimizing the error, prediction mechanism has been incorporated on keyboard in both Hindi and Bengali language [7]. The keyboard gives the freedom to user to choice suitable search engine where they are intended to search and also search types like Unicode or translation based searches. Tool strip status bar has also given to provide the help tips whenever user is typing in the search textbox. Figure 6 and 7(a) shows the pictorial representation of the developed virtual keyboard in Hindi language.

After the intended keyword has been entered in mother language, it would be given to Hindi/Bengali to English language translation engine to translate into corresponding English language. Google machine translation engine is used for Hindi to English translation task. But for Bengali to English translation task, we developed dictionary based intermediate English word or phrase generation engine. For that purpose, we have created dictionary (Bengali to English) of 60000 unique words. Then the translated English keywords are extracted and send to any of the search engine (Google, Bing or Yahoo) selected by user's own choice. In Unicode based search, only existing mother language pages will be provided rather applying web translation mechanism. User given search keywords would be sent to the search engine in mother language and if pages exist, would be displayed as a result. In this mechanism user would get only small number of exiting mother language pages.

Fig. 6. Virtual keyboard for Hindi language

(a) Keyword entry through virtual keyboard

(b) Snippets rendered in Hindi

(c) Translated webpage in Hindi

Fig. 7. Text interface in Hindi

On the other hand, translated search keyword would provide translation mechanism on the returned snippets. In the process, our system would extract tag and text portion of web pages by maintaining index table and then sends the text portion to translation engine for translation. Here, for English to Hindi and Bengali language translation, Google translation service[8]

[8]http://translate.google.com/translate_t

and Anubadok machine translation engine[9] had been used, respectively. After getting back the translated text in mother language, original English text would be replaced by translated mother language texts and painted to the browser maintaining the proper index value. Further, relative web addresses will be changed to corresponding absolute web addresses and users will get the proper translated page into their mother languages. One can understand the text based interface pictorially in a more clarified manner. Figure 7(b) and 7(c) show the translated form of resulting pages containing snippets and the required information corresponding to search keyword pictorially.

B. Speech Interface

The main two modules required in the speech based interface design are "Automatic Speech Recognition" (ASR) and "Text to Speech Synthesis" (TTS). In the present context, there is a requirement that user will not enter the keyword by typing keys in mother language keyboard placed in the web page but only tell the system through a microphone in corresponding mother language. System should catch the voice, properly convert the voice into text and then place it in the search textbox. This can be done by speech recognition engine. After keyword has been entered into the search textbox, the text interface will work as its own and produce output painted in the webpage in user's mother language. As the output mode of interaction is also speech, there is a requirement that the user will select some portion of the webpage (In case of selection of snippet link, only heading of that page will be read out) and output voice should be played from machine at user end.

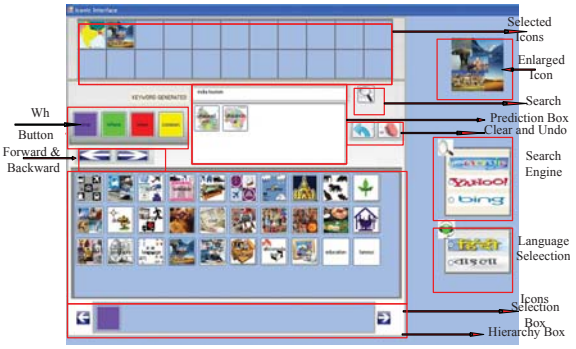
Currently, Festival text to speech system[10] is working for both Hindi and Bengali mother language. The configuration file has been changed for obtaining the support in Hindi and Bengali mother language. Currently, the speech dictionary containing utterance of 1.5 lakh unique words (1 lakh of Noun and 0.5 lakh of other parts of speech). It has been configured and hosted to a server and it provides service to any client computer by sending speech file (.wav format) resulted from running a text file which has been sent from client's end. Another system, called Sphinx, would be used for performing Speech Recognition task (Speech to proper text) also both in Hindi and Bengali language.

C. Icon Interface

The proposed design in iconic interface has major two modules, one is natural language generation with flavor of query expansion module selecting set of given icons and another is representing the main theme of the returned web page with a set of icons. Basically, all the modules working for text based interface are running in background and front interface would be decorated with set of easily understandable basic icons [18]. As the interface consists of icons independent of language, initially the iconic interface of Hindi or Bengali language is very much similar to English or any foreign language interfaces. First of all, a user would shows his/her intention in searching for some topics by clicking an icon. Then the interface tries to properly predict the next set of icons depending on the context and user's intention [19]. The interface automatically extracts the keywords from those clicked icons and generates sentences or phrases in mother language (Hindi/Bengali) in background. Figure 8(a) clarifies the scenario more clearly. Then the entire sentence or phrase is being fed to the one of the popular search engine (Google, Yahoo, and Bing) by user choice which would give the result as snippet pages as viewed in Fig. 8(b). Finally, returned English snippets or web pages are being entered to information visualization module which summarizes the web page contents and the summarized text portion will either send to speech synthesis module to get final output in speech mode or directly be mapped to visual template consists of minimal set of icons which can represent the semantics/meaning of the text.

V. User Testing

In every user-centric design, the general people play a significant role in evaluating the product which follows the design. So, here our target users are mostly digitally illiterate people who had already crossed the basic barrier of literacy but can not properly read and write in their mother language (Hindi and Bengali) also. For user selection task, we visited nearby villages, local markets, schools, domestic help centers etc. Experiment has been done with three types of users



(a) Icon selection to construct a search query

(b) Web page returned with set of icons

Fig. 8. Icon interface

[9]http://bengalinux.sourceforge.net/cgi-bin/anubadok/index.pl
[10]http://www.cstr.ed.ac.uk/projects/festival

(according to level of literacy) of different occupation and $60-70$ people in average belonging to each user type. The interface has been tested by educated persons like office staff, college students, business person and also by porters, rickshaw pullers, drivers, domestic helpers in both urban and rural areas. People from rural areas who do not know mother languages in terms of reading and writing like farmers, rice mill workers have also evaluated the interface. The detail information of categories of user and their basic skills are described in Table I.

The interface was given to each user in their own places. The developers and the experts initially trained the user for a while. The total testing phase had been executed by following different steps described below.

- A user detail form has been given to each and every user to give detail user profile containing occupation, place of living and interest area in which he intend to gather knowledge by accessing Internet.
- After filling the form, user is allowed to access the system. One instructor always helped him if he is illiterate or literate but not so familiar with computer.
- After a fixed time (30 minutes), the testing phase had been completed. A feedback form has been given to user to make comment on performance of system with respect to his proper satisfactory level.

Using the interface, many people specially illiterate people found useful information of their own query formed in any of the three modes like rail porters have accessed the interface and found status of their insurance card, rickshaw pullers have gathered useful information of current price of tyres and other parts of a rickshaw etc. Similarly, the semi−literate people like farmers and housewives were become happy when one knew the price of local rice grain in West Bengal and other came across list of hindi channels with timing of mega TV serials. Te office clerks booked the railway ticket in both Hindi and Bengali language using the interface. A number of important user parameters have been taken from generated log file entry. A few snapshots taken at the testing phase where shopkeeper and domestic helper are accessing the interface to find out price of tea and the recipe of a food item respectively are given in Fig 9(a) and 9(b).

In the testing phase, we have tried to cover the people from different parts of socio-economic structure in the state of West Bengal in India. As a result, the level of intention and also satisfaction of them after getting the results are quite different from each other. Though the testing phase was 30 minutes long, the result does not reflect the scenario of whole community. We need to do the testing in much broader perspective with more number of participating users.

After completion of the testing phase of each user, we have collected a feedback from each user and profile like education, occupation etc. The users, by filling the feedback form, actually evaluated the user interface made by us with respect to six evaluation matrices namely user-friendliness, usability, error proneness, applicability and reliability. By consolidating the feedback, we found that users certified on user friendliness and usability properties of the interface as

(a) Shopkeeper is checking for price of wheat

(b) Domestic help is checking for recipe of food

Fig. 9. User testing with the developed interface

quite impressive. The interface helps the users in forming the proper spelled query. The users of all categories have more or less satisfied by using the interface. They have faced less number of problems, found more things than expectation to be searched. The reliability of the interface had become average because in remote villages, internet connectivity felt in most of the cases. After gathering the feedback, we have plotted the result in a chart depicting the system performance which is given in Fig. 10. Larger value for user friendliness, usability, applicability, and reliability as well as smaller value of error proneness reflects good performance of the interface. The graph signifies that user testing report is satisfactory.

VI. Conclusion and Future Work

A multimodal interface to Internet, which is new of its kind has been proposed in this paper. The proposed interface able to overcome the digital divide as well as language barrier, which are the main challenges for the growth of ICT. Our experiments in the context of common people of India reveals that the proposed approach is adequately cater to the need of the users and very much user friendly. The proposed approaches is generic, in the sense that, it can be extended to any languages in the world without any major modification. Development cost only involve in language translation, speech

TABLE I
USER DESCRIPTION

User Profile	Occupation	Education	Reading/Writing in English	Reading/Writing in Hindi/Bengali	Computer Knowledge	No. of Users
User 1	Business Person	Post Graduate	Average	Good	Expert	15
	Office Clerk	Graduate	Good	Good	Intermediate	20
	School Students	Class 10-12	Poor	Good	Intermediate	25
User 2	Shopkeepers	Class 6-10	Very poor	Average	Novice	18
	School Children	Class 10-6	Very poor	Average	Novice	12
	Housewives	Secondary	Very Poor	Average	Novice	15
	Farmers	Class 10-6	Very Poor	Poor	Novice	23
User 3	Rickshaw pullers	Class 5-illiterate	Nil	Nil	Novice	18
	Porters	Illiterate	Nil	Nil	Novice	24
	Farmers	Class 5-illiterate	Nil	Very poor	Novice	8
	Domestic helpers	Illiterate	Nil	Nil	Novice	11

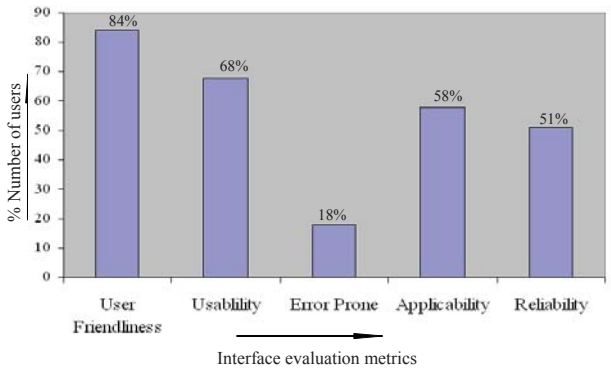


Fig. 10. Results of users' feedback

recognition and text to speech synthesis.

Also, in future, we will extend the multi-modal system from large screen computing device to small screen hand held devices. Hardware issues have to be taken care of properly in this regard.

Acknowledgement

The work is carried out under a project sponsored by Department of Information Technology, Government of India. We express our sincere thanks to all participants for their painstaking efforts in the evaluation of the interface.

REFERENCES

[1] M. K. Sharma, P. k. Saha, S. Sarcar, S. Ghosh, and D. Samanta, "Accessing Dynamic Web Page in Users Language," in *Students' Technology Symposium (TechSym)*. IEEE, 2011, pp. 35–38.

[2] S. Sarcar, S. Ghosh, P. K. Saha, and D. Samanta, "Virtual Keyboard Design: State of the Arts and Research Issues," in *Students' Technology Symposium (TechSym)*. Kharagpur, India: IEEE, 2010, pp. 289–299.

[3] M. K. Sharma, S. Dey, P. K. Saha, and D. Samanta, "Parameters Effecting the Predictive Virtual Keyboard," in *Students' Technology Symposium (TechSym)*. IEEE, 2010, pp. 268–275.

[4] S. Ghosh, S. Sarcar, S. Sarcar, and D. Samanta, "Designing an Efficient Virtual Keyboard for Text Composition in Bengali," in *3rd India HCI Conference*. ACM, 2011, pp. 90–93.

[5] H. M. Horstmann and S. P. Levine, "The Effectiveness of Word Prediction," in *Proceedings of 14th RESNA Conference*, 1991, pp. 100–102.

[6] K. Kukich, "Techniques for Automatically Correcting Words in Text," *ACM Computing Surveys*, vol. 24, no. 4, pp. 377–439, 1992.

[7] M. K. Sharma, "Word Prediction System with Virtual Keyboard for Text Entry in Hindi," Master's thesis, Indian Institute of Technology Kharagpur, Kharagpur, India, January 2012.

[8] H. Petrie, G. Weber, and W. Fisher, "Personalization, Interaction, and Navigation in Rich Multimedia Documents for Print-disabled Users," *IBM Systems Journal*, vol. 44, no. 3, pp. 629–635, 2005.

[9] Y. Qiu and H. Frei, "Concept Based Query Expansion," in *Proceedings of 16th ACM International Conference on Research and Development in Information Retrieval*. Pittsburgh, USA: ACM Press, 1993.

[10] J. Himberg, J. Häkkilä, P. Kangasi, and J. Mäntyjärvi, "On-line Personalization of a Touch Screen based Keyboard," in *Proceedings of the 8th international conference on Intelligent User Interfaces*, ser. IUI '03. New York, NY, USA: ACM, 2003, pp. 77–84.

[11] M. Thelwall, "Extracting Accurate and Complete Results from Search Engines: Case Study Windows Live," *Journal of the American Society for Information Science and Technology*, vol. 59, no. 1, pp. 38–50, 2008.

[12] I. Mierswa and K. Morik, "Automatic Feature Extraction for Classifying Audio Data," *Machine Learning Journal*, vol. 58, pp. 127–149, 2005.

[13] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–285, Feb. 1989.

[14] S.P.Kishore and A. W. Black, "Unit Size in Unit Selection Speech Synthesis," in *Proceedings of Eighth European Conference on Speech Communication and Technology*, Geneva, Switzerland, 2003.

[15] J. Matousek, J. Psutka, and J. Kruta, "Design of Speech Corpus for Text-to-Speech Synthesis," in *Proceedings of Seventh European Conference on Speech Communication and Technology*, Alborg, Denmark, 2001, pp. 2047–2050.

[16] R. Rosenfeld, X. Zhu, A. Toth, S. Shriver, K. Lenzo, and A. W. Black, "Towards a Universal Speech Interface," in *Proceedings of the International Conference on Spoken Language Processing*, Beijing, China, 2000.

[17] J. Whiteside, S. Jones, P. S. Levy, and D. Wixon, "User Performance with Command, Menu, and Iconic Interfaces," *SIGCHI Bulletin*, vol. 16, no. 4, pp. 185–191, 1985.

[18] S. Maiti, S. Dey, and D. Samanta, "Development of Iconic Interface to Retrieve Information from Internet," in *Students' Technology Symposium (TechSym)*. IEEE, 2010, pp. 276–281.

[19] S. Maiti, D. Samanta, S. R. Das, and M. Sarma, "Language Independent Icon-Based Interface for Accessing Internet," in *1st International Conference on Advances in Computing and Communications*. Springer, 2011, pp. 172–182.