

Vimprint: Exploring Alternative Learning through Low-end Mobiles

Sheetal K. Agarwal¹, Jyoti Grover², Anupam Jain², Arun Kumar²

¹ IBM Research – India, 8th floor, G2, Manyata Embassy Business Park, Bengaluru, India

² IBM Research-India, Plot 4, Block C, Vasant Kunj, Phase II, New Delhi, India

{sheetaga, jyogrove, anupamjain, kkarun}@in.ibm.com

Abstract. The Internet today provides a plethora of applications to assist anyone wanting to learn a new subject, language or a concept. Resources available include dictionaries, translation engines, downloadable e-books, tutorials, online courses etc. The rapid proliferation of smart phones has further provided richer visual applications that assist the user in learning on the go. However, all these applications are dependent upon the availability of Internet and/or an expensive computing device such as a smartphone or a computer. This puts them out of reach for a large section of society that consists of underprivileged people (economically or literacy-wise) and who probably need such tools the most. Also, many learning applications are pull-based and depend on the user's motivation to keep coming back for more. We present Vimprint - a system that offers an alternative mode of learning through telephony voice applications over any touchtone phone with a push-based interaction module. We present its design in the context of vocabulary building. Application of Vimprint system in the field is a work-in-progress and we present the results of a preliminary study conducted to assess its effectiveness.

Keywords: Social Computing; Developing regions; Voice Applications; Education; Diversity; Inclusiveness.

1 Introduction

The Internet has revolutionized the way learning can happen today. Online resources form a critical part of the learning process for any concept. Interactive web and smartphone applications engage the learner making it a fun experience. However, barely 35% of the world's population has access to Internet [1]. This makes all the rich online resources out of bounds for a vast majority of its potential users. Even for those who do have access to the Internet, illiteracy or English illiteracy acts as another major deterrent as most learning tools are offered in English.

In India, ability to converse in English not only improves chances of employment, it also elevates one's social image [4]. To address this demand from users in addition to classroom coaching and school based courses, books are published in local languages to teach English, cable TV operators provide simple English lessons [19] daily among other popular tools for language learning. While books are not interactive, cable TV based learning programs are broadcast in nature and do not offer an interactive experience. Classroom coaching is still the best option available but is expensive for many and skilled teachers are hard to find.

Recent research has attempted to introduce ICT in education in developing countries with the aim of providing a good supplementary channel to classroom teaching. Educational games [8, 5] to encourage math skills, improving vocabulary, learning shapes and colors etc., have been created and tested in primary schools in developing countries and have been found to be beneficial. While these games have been shown to be effective, they require a mobile, either a smart phone or a feature phone, or a computer where the game can be downloaded. Even though the mobile penetration has improved phenomenally, mobiles are often a shared resource in a family and its primary use is to make and receive phone calls. Users are barely able to use basic mobile phone features such as address books and the Short Messaging Service (SMS).

Among various technology solutions proposed to empower the underprivileged [3], telephony voice based systems have been particularly successful. Such applications break the barrier of illiteracy and make technology accessible in local language even to those who cannot read or write. Examples of scenarios where such telephony voice applications have been deployed include educational services in rural India [12], game based learning [18], voice based community knowledgebase [6], information dissemination and online discussion for farmers [9], healthcare [11] and job match-making [16]. In the context of learning, BBCs Janala¹ is a popular service in Bangladesh where users can access 3 minute English tutorials for 5 Taka (0.01 USD). The pervasive access to phones along with low call costs has driven the success of this service.

However, most of the work done for this user segment, in learning space, focuses primarily on delivering the content. A pull model is adopted relying on the learner's own interest and enthusiasm to revisit the content for retention. In the absence of a push mechanism, the less motivated learners tend to drop off thus defeating the purpose. In a classroom setup, the push based interaction is somehow enforced by the teacher through prompt repetitions in class or through assignments and tests. In this paper, we present a telephony voice based learning tool - *Vimprint*, focusing on the design of its vocabulary building module. Vimprint adopts a push based interaction method to increase retention of words accessed. It employs an adapted version of PimSleur's Graduated-Interval Recall method [10]. We believe that an active push

¹ <http://www.bbcjanala.com/>

based approach would simulate the aspect of a teacher that ensures retention of concepts taught, and hence result into more effective learning applications.

2 Related Work

A few SMS/MMS based systems have been employed to teach a language [15, 13]. However, SMS based learning is an impractical option for many who are not literate enough to read and understand SMS content. For others, lack of spoken content leaves out the important elements of language teaching such as pronunciation, accent, etc. MMS on the other hand, does not work on most low-end phones which dominate in the developing world.. Telephony applications have been tried in the context of teaching English through multi-lingual story playback [2, 14] but they are pull based and depend upon learner's own motivation to pick up the concepts taught. Kumar et al [8] use application installed on the phone and mention about the possibility of indicated teaching over telephony voice as an explorable alternative.

3 VIMPRINT: Vocabulary Building

Vimprint is a learning system accessible through voice interaction over any touch tone mobile phone (including a landline). In this paper, we specifically focus on its vocabulary building module that aims to promote word literacy by enabling users to lookup words they have heard from others or read somewhere. On placing an ordinary phone call to the application, user is prompted to speak the word he wants to lookup. Vimprint retrieves and plays a list of words sounding similar to the user's utterance.

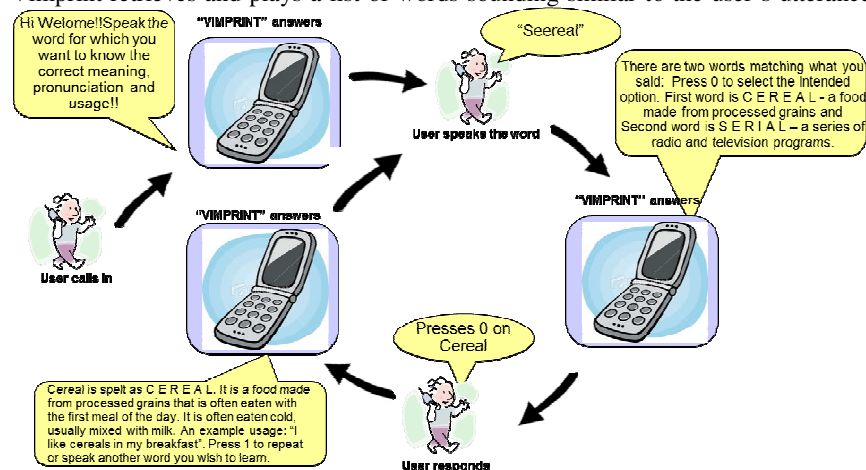


Figure 1: User Interaction with Vimprint System for learning new words

For example, if the user says ``cereal'' the API will return ``serial'' and ``cereal''. Along with every word it also plays its meaning to enable the user to identify the intended word. Figure 1 shows the interaction of a user with the Vimprint system for learning a new word.

Once the user identifies and selects the intended one, it's pronunciation, it's spelling and it's usage in a sample sentence are played. Vimprint focuses on enabling users to retain the words by following Pim Sleur's Graduated-Interval Recall method [5] for learning.

Graduated-Interval Recall

Typically, a number of techniques are used to improve the stickiness of a word in memory. These include using sample sentences from learners' day-to-day scenarios, synonyms or antonyms, repeating the word several times etc. Pim Sleur method is a popular method for learning a foreign language. The method follows the principle of recalling words at progressively increasing intervals to increase the possibility of moving the word from user's short term memory to long term memory. This is also known as the spacing effect. Pim Sleur's 1967 memory schedule is as follows: 5s 25s 2min 10 min 1hr 5hrs 1 day and so on. It follows the equation:

$$T_1, T_2 = 5 \times T_1, T_3 = 5 \times T_2, \dots T_n = 5 \times T_{(n-1)}$$

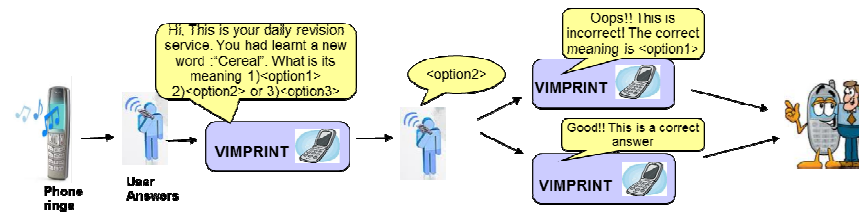


Figure 2: User Interaction for Revision calls

In Vimprint, word recall is facilitated by proactively placing outbound calls to users to revise words previously seen. A user profile is created for each caller into the application. The user's preferred time for receiving revision calls; words accessed by the user, and actual calling times are maintained in a database. Since receiving a call requires a user's immediate attention, we adapted the Pim Sleur's method to take into account user's preferred time slots for receiving reminder calls. Figure 2 depicts the user interaction for revision calls. At each recall, a multiple choice quiz is rendered where user is asked to select the correct meaning of the word from multiple choices. While this method has been used to teach languages through books and audio courses, none of them are push-based in nature. The figure shows revision for a single word only, but a single revision call is generated on a scheduled day for all the words that need to be revised for that user.

4 VIMPRINT - Implementation

The Vimprint application is a Java based voice application platform that runs on an application server and generates Voice XML (VXML)² at runtime. These VXML pages are then rendered by a Voice Platform (Cisco voice gateway as shown in Figure 3). We use the Nuance Automatic Speech Recognition (ASR) Engine (i.e. Nuance Recognizer v9) for word recognition.

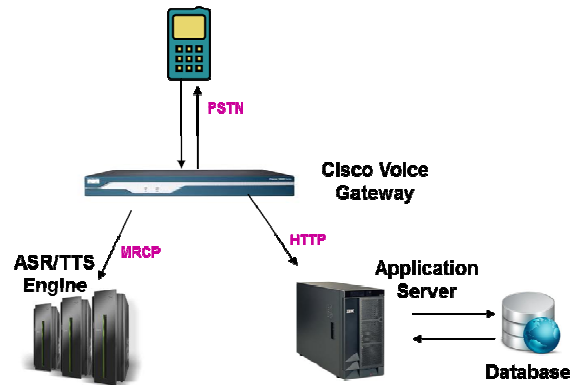


Figure 3: Vimprint Infrastructure Setup

For each word the user speaks, we use the Nuance's API to return a list of similar sounding words and allow the user to select the word he wants to access. The selected word is then fed into a Vocabulary lookup module which fetches and plays back the details for that word. At the same time, that word is also scheduled for revision calls as per the schedule described above (refer **Error! Reference source not found.**).

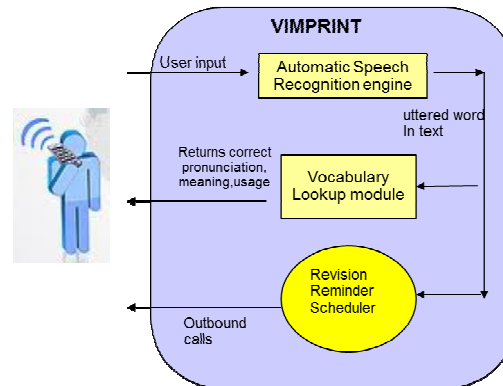


Figure 4: Vimprint System Design

² <http://www.w3.org/TR/voicexml20/>

Our voice application platform that forms the core of Vimprint system has been tested to scale to 200 simultaneous calls on an 8-core cpu, 4 GB RAM Intel machine running Linux and hosting Vimprint system on Apache Tomcat application server and MySQL database server. Nuance was hosted separately on a similarly configured machine.

5 Preliminary User Study

Deployment and evaluation of Vimprint's Vocabulary learning module is a work-in-progress. Before going all out and recruiting users from the actual end segment who are often too busy in earning a livelihood, we conducted a preliminary user study with users outside the target segment. The aim was to test the effectiveness of the technique rather than fine tune the user interface for the actual target user group.

The preliminary study was conducted with 9 subjects over a period of 3 days. Since this is a preliminary test, we recruited users from our social network who are not native English speakers but have undergone formal education in English. Given their educational background, users were exposed to 10 pre-selected English words having good difficulty level. None of the users were familiar with these words nor were these words used in their day to day conversations. Of the 10 words, revision reminders were set for only 5 words. Instead of scheduling reminders for them, the researchers simulated the scheduler by calling the subjects and requesting them to call into the system at the intervals of 1 hr, 5 hrs and 24 hrs from the time they accessed the system the first time. During each revision call, the users were quizzed by the system for meanings of words they had learnt. They were prompted to select the correct option for the word meaning from multiple choices. After the last revision call, the users were quizzed on all the 10 words – 5 for which revision calls were scheduled and 5 for which no revision call were done. Words which were revised were recalled with 93.3% accuracy while the words without revision were recalled with 55.5 % accuracy.

These preliminary results are encouraging and suggest that the classroom effect of regular prodding by teacher on a subject can probably be simulated through a push based mechanism over the phone. This motivates further investigation. Given the limiting nature of this study, an extended user study is being planned to obtain statistically significant results, with our target user group i.e. non-English speaking users with little access to web based learning tools. Although an incoming revision call can be intrusive, but if the learner is comfortable with it, such push mechanism can instill discipline and ensure continuity of learning which is missing from traditional web based learning tools that rely on pull model. Similar model of structured pull cum push based learning is now also being explored and deployed in the web world through the concept of massive open online courses (MOOC) [17].

6 Conclusion

This paper presented Vimprint, a system for exploring telephony voice applications as an alternative means for enabling effective learning and memorization for underprivileged in developing countries where numerous people have no or limited access to internet. It employs a push based interaction over outbound calls in addition to initial pull from the learners, to assist users in memorizing and to keep them disciplined and engaged in the learning process. The paper presented design and implementation of its Vocabulary building module along with the results of a preliminary study that encourages us to explore further with underprivileged users. While we tested the system for learning English, Vimprint itself is language agnostic and can be used to offer vocabulary building for other languages supported by Nuance Speech recognition engine.

Going forward, we are extending this application with additional learning modules and conduct a full scale study towards establishing telephony voice applications as a potential vehicle for structured learning for the underprivileged. For that purpose, we have tied up with an English learning school in the area who take everyday classes for students who may not have had the luxury to study in expensive English medium schools. With this user set, we intend to compare how Vimprint's vocabulary building module fares against classroom based vocabulary building methods.

References

1. Stats, I. W. World Internet Users and Population Stats. <http://www.internetworldstats.com/stats.htm>.
2. Learn English - enableM and OnMobile. <http://www.enablem.com/learn-english-mobilephone.aspx>.
3. Boyera, S. et al. Mobile Web for Social Development Roadmap W3C IG Note. <http://www.w3.org/TR/mw4d-roadmap/> (Nov 2009).
4. Kachru, B. Indian English: A sociolinguistic profile of a transplanted language. In *Studies in Language Learning*, 1, no. 2, 1976
5. Kam, M., Mathur, A., Kumar, A., and Canny, J. Designing digital games for rural children: a study of traditional village games in india. In Proceedings of the 27th international conference on Human factors in computing systems, ACM (2009), 31–40.
6. Kotkar, P., Thies, W., and Amarasinghe, S. An Audio Wiki for Publishing User-Generated Content in the Developing World. In HCI for Community and International Development (Workshop at CHI 2008) (Florence, Italy, April 2008).
7. Kumar, A., Agarwal, S., and Manwani, P. The Spoken Web Application Framework: User Generated Content and Service Creation through Low-end Mobiles. In Proc. of W4A (2010).
8. Kumar, A., Reddy, P., Tewari, A., Agrawal, R., and Kam, M. Improving literacy in developing countries using speech recognition-supported games on mobile devices. In Proc. of CHI (2012).

9. Patel, N., Chittamuru, D., Jain, A., Dave, P., and Parikh, T. S. Avaaj Otalo - A Field Study of an Interactive Voice Forum for Small Farmers in Rural India. In Proceedings of ACM Conference on Human Factors in Computing Systems (CHI) (2010).
10. Pimsleur, P. A memory schedule. *The Modern Language Journal* 51, 2 (1967), 73–75.
11. Padman R., Li X., Damir J., Kumar, A., Khera, P., Kumar, A., Agarwal, S., and Nanavati, A. Engineering a Spoken Web enabled Health Information System (SW-HIS) for Diabetes Management. In Proceedings of Mayo Clinic Conference on System Engineering and Operations Research in Health Care (2011).
12. Ruohonen, M., Turunen, M., Mahajan, G., Linna, J., Kumar, V., and Das, H. Mobile Phones and Voice-Based Educational Services in Rural India: PROJECT RURALVOICE. In Proceedings of Open and Social Technologies for Networked Learning (OST) (2012).
13. Saran M., S. G. Supporting foreign language vocabulary learning through multimedia messages via mobile phones. In H. U. *Journal of Education* (2010) 38,252-256.
14. English Seekho - Tata Indicom. <http://www.tataindicom.com/tata-zone-mobilelearning.aspx>.
15. Thornton, P., and Houser, C. Learning on the Move: Vocabulary Study via Email and Mobile Phone SMS. In Proc. of World Conference on Educational Multimedia, Hypermedia and Telecommunications EDMEDIA. 2001.
16. White J., Duggirala M., Kummamuru K., and Srivastava S., Designing a voice-based employment exchange for rural India, In Proceedings of the Fifth International Conference on Information and Communication Technologies and Development, 2012.
17. Carey, K, Into the Future With MOOC's. *Chronicle of Higher Education*. <http://chronicle.com/article/Into-the-Future-With-MOOCs/134080/> Retrieved March 20, 2013.
18. Larson M, Rajput N, Singh A, Srivastava S, "I want to be Sachin Tendulkar! A Spoken English Cricket Game for Rural Students," International Conference on Computer Supported Cooperative Work and Social Computing (CSCW) 2013, Feb 2013, TX
19. Tata Sky Active English, <http://www.tatasky.com/education/active-english.html>