Mobile Application for Learning Languages for the Unlettered

Gayathri G, Mohana N, Radhika Pal, Hema A Murthy

Indian Institute of Technology, Madras, India
`gayathri13990@gmail.com, hema@cse.iitm.ac.in`

Abstract

Mobile based technologies have become ubiquitous, and various applications from games to readers are mobile oriented. In this paper, we propose development of speech based language learning app. Conventionally language learning tools start with words, followed by sentences. The fundamental assumption is that the person is literate. In the Indian context, the literacy levels are as low as 65%. In addition, each Indian language has its own scripts. The objective of this work is to develop a mobile app that starts from the script to teach the language to a person who can speak the language but is unlettered. Since the focus is on the unlettered, writing should be easy. A script centric approach is used to learn a language. The application starts with teaching a simple letter of the alphabet, followed by another letter that can be obtained by simple modification to the previously learned letter, followed by words using the letters that are learned, followed by sentences using the learned words. At every step, a text-to-speech system is used which articulates the letters, and words. The learning app is based on a book called Tamil Karpom (P Nannan). The ideas from the book are adapted for learning Hindi.

1. Introduction

The development of mobile and internet technologies have resulted in low cost smart phones. This plays an important role in the adoption of these technologies in learning and educational applications. With the availability of plenty of learning apps, most users choose the apps based on two factors - easy to use and readily fulfil the purpose. In this paper, two language learning mobile apps are presented that accomplish these two requirements. These applications are developed mainly for learners who have verbal knowledge in the language but lack writing skills or is totally new to the language.

Most language learning tools for Indian languages have the tradition of teaching the letters of the alphabet in the order prescribed by lexicographic order through simple words similar to "A for Apple" in English. This kind of introduction to the letters of the alphabet does not take into account the ease of writing. Unlike these applications, an alternate approach is used our applications to rearrange the letter in the increasing order of writing difficulty as inspired by Dr. Nannan's work on Tamil learning [1]. The apps also support text to speech synthesis which enables the leaner to listen to what is written on the screen.

2. Learning methodology

The simplest of all letters is the first one to start learning. For example in Tamil, ⊏ (ta) is the first letter taught as opposed to அ (ah) in the conventional approach. The next letter for learning is ⊔ (pa) which can be obtained by appending a vertical line at the end of ⊏ (ta). ⊔ (pa) can be modified to write ம (ma). The consonant parts of respective letters can be written with a dot on the top – ட் (it), ப் (ip) and ம் (im). Now combinations of these simple letters are used to form words like படம் (padam), பட்டம்(pattam), மடம்(madam) and so on. An analogy in English would be to start with I (single vertical line) followed by L (appending a horizontal line to I) followed by E. Simple words like ILL, LIE, EEL enable the learners to identify the letters and practice the same. It is worth to note that the pronunciation of ⊏ (ta) in படம் (padam) and மடம் (madam) is voiced as opposed to pronunciation in (pattam) which is unvoiced. These differences are learned using the audio output.

One of the important benefits of these apps is that this app is learner friendly to any age group and independent of gender. From the past experience of one of the authors - as part of the literacy movement of 1990-1991 in Tamil Nadu, it was observed that men found writing a big challenge, while women found it a lot easier, owing to the tradition of making "kolam" patterns at the entrance of their homes every morning [1]. Since these apps start with simple letter, they provide a perfect platform for learning without losing enthusiasm. The method of teaching the unlettered enables people to pick up the language quickly. As the app also keeps track of the progress in learning, it could enable neo-literates to get hooked on to the same. The only limitation of this approach is that the newbies cannot understand the meaning of the words nevertheless it is not in the scope of language learning. However to aid this, words for which pictures are possible to express them are given figures. Figures 3a and 3b show the learning order of letters of the alphabet in Tamil and Hindi respectively. The increasing order of writing difficulty can be observed in the figures. The letters of the alphabet are drawn using the HTML canvas in slow motion so that the learners can comprehend the strokes involved in writing a particular letter. Figures 1 and 2 show the screenshot of four different chapters of Tamil learning app. Each subfigures (Figures 1a - 1d and Figures 2a - 2d) show different stages of teaching specific chapters.

---

[1]Kolam patterns are symmetric line diagrams that involve curves and line [2].
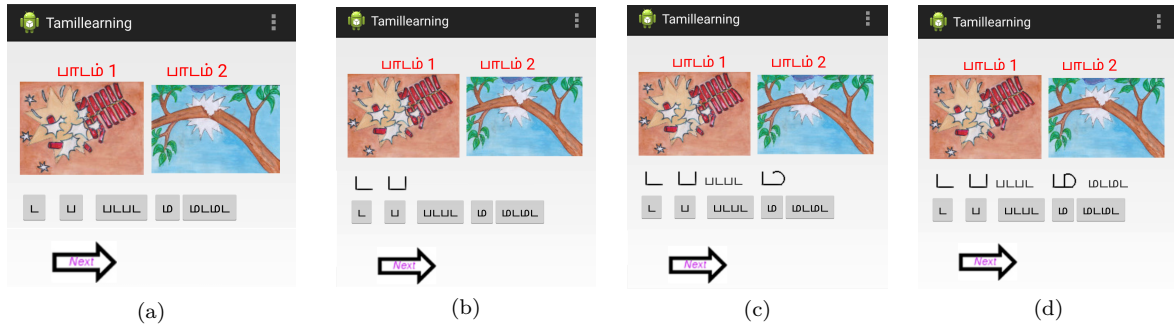
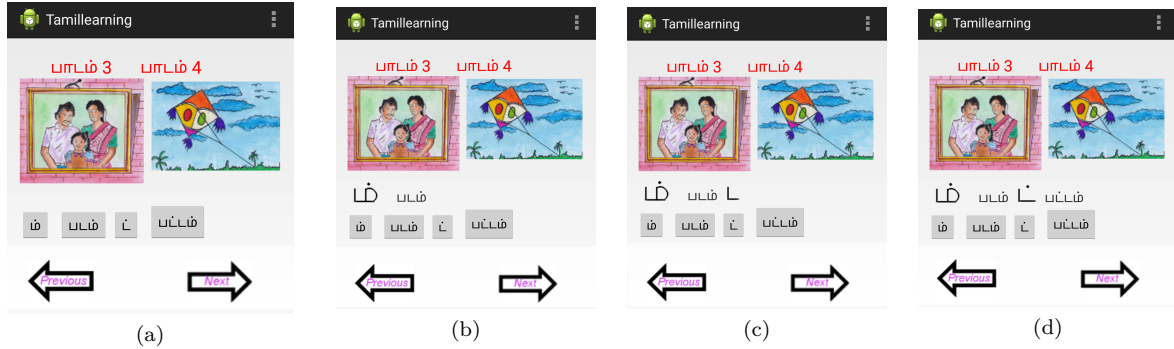Figure 1: *Images of different stages of learning Tamil characters (chapters 1 and 2)*

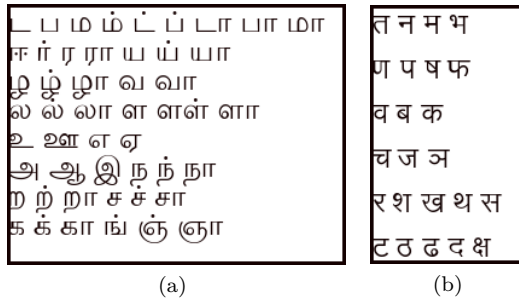Figure 2: *Images of different stages of learning Tamil characters (chapters 3 and 4)*

Figure 3: *Learning order of letters in (a) Tamil and (b) Hindi*

3. IndicTTS for speech synthesis

The apps use audio support to teach the pronunciation of letters and words. Audio of the words/letters that are currently on the screen are synthesised on the fly and can be played upon user request. To synthesis voice, HTS-STRAIGHT voices [3] are used. The voice synthesis system is trained using IndicTTS dataset [4]. The corresponding text is generated using a unified parser for Indian languages [5].

4. Conclusion

Languages learning apps for Tamil and Hindi are developed. An approach in which the letters of the alphabet are learnt in the increasing order of writing difficulty is employed for teaching the alphabet. The apps can be downloaded from https://www.iitm.ac.in/donlab/tts/androidapp.php.

6. References

[1] P. Nannan, "Tamil karpom." [Online]. Available: http://www.tamilvu.org/ta/stream-html-basic-tmbalesson-index-275501

[2] "Arts and crafts traditional customs and practices." [Online]. Available: http://www.indian-heritage.org/alangaram/kolams/kolams.htm

[3] H. Kawahara, I. Masuda-Katsuse, and A. de Cheveigne, "Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds," *Speech Communication*, vol. 27, no. 3–4, pp. 187 – 207, 1999.

[4] A. Baby, A. L. Thomas, N. L. Nishanthi, and T. Consortium, "Resources for indian languages," in *CBBLR – Community-Based Building of Language Resources*. Brno, Czech Republic: Tribun EU, Sep 2016, pp. 37–43.

[5] A. Baby, N. L. Nishanthi, A. L. Thomas, and H. A. Murthy, "A unified parser for developing Indian language text to speech synthesizers," in *International Conference on Text, Speech and Dialogue*, Sep 2016, pp. 514–521.