

Online Picture Book Learning of Multi-lingual Words through Images, Text and Audio

Bhaskara Rao N.

Dept. Of CSE, Dayananda Sagar College of Engineering Bangalore

Email: bhaskararao.nadahalli@gmail.com

Abstract—A web enabled Java Script Picture Book Learning software is developed for identifying words of different objects in multiple languages. Images of perceptible objects from different categories are presented to the user. The associated matching words in selected languages are displayed for the perusal of the user. Also the corresponding pronunciations are provided in the audio format for listening experience. In this software, cross-language transliterations are also provided to learn new words easily. The transliteration process uses a few elementary techniques of Natural Language Processing.

Keywords—Indian languages, Natural Language Processing, transliteration,

I. INTRODUCTION

Vision is a direct and simple method of recognizing a physical object. The objects are represented by words for the purpose of identification and communication. Both spoken words and written words are used to identify and to describe the physical objects. But different languages use different words to represent the same object. Therefore a beginner aspiring to learn a new language has to learn the names of various objects in that new language. Even though dictionaries are universally used in learning new languages, specific dictionaries in the native language, say of a tribal or a minority community, may not be available. Also for young children who have just started learning the language, use of a dictionary is not easy. To take care of these situations, we are introducing the words through the pictures (images) of the corresponding objects. Here the input is a picture (image) and the entry level user (say a tribal or an early learner) can easily access the matching word on the screen. In this way our software is basically an **Image to Text** converter. The software also provides the corresponding audio pronunciations. The children and other beginners learn a new language quickly and efficiently through the medium of pictures and audio. In fact the learning of alphabets, digits, common colours etc. become very easy. We mainly concentrate on Indian languages.

A. Existing related works

Several online picture based web sites exist for early learners in English and other non-Indian Languages [1, 2, 3]. There are a few sites for Indian language learning [4, 5]. There is some resemblance between our work and that of reference [4]. In [4], user has to select a word, then, the corresponding image appears. But in our approach, the input is an image, and our method is highly interactive and is user friendly. In our work the output format is designed to be very useful compared to that in [4]. Our work is scalable and can be fine-tuned for different requirements.

II. WORKING OF IMAGE TO TEXT CONVERTER

The **Image to text** converter is an on line utility. Here the user need not type the particular word in any language. He has to select the image (picture) of the object by the mouse. Then the software displays the corresponding textual words in the chosen languages on the screen in the specified text boxes. The corresponding transliterated texts are also displayed properly as shown in Fig. 1. Even though English and kannada are shown in Fig. 1, other language pairs can be chosen. Also more than two languages can be accommodated.

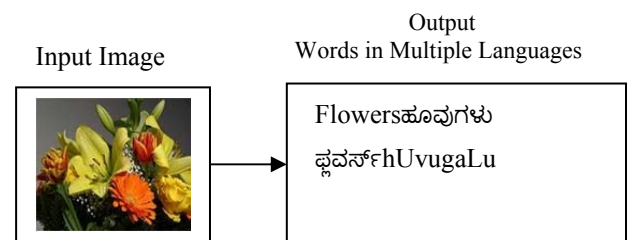


Fig. 1. Basic working of **Image to Text** converter

In our scheme, input is the images of different objects. These images are obtained from various sources.

A. Images

Images of the objects are the basic building blocks of this software. They have to be carefully acquired and stored properly.

1) *Acquisition of Images*: Some of them are created from graphic packages like MS Paint, PC paintbrush, Adobe Photoshop etc. Many of the images are obtained through internet and open sources. Several others are acquired through the digital cameras. Then these images are stored in jpeg format so as to keep the image file sizes low.

2) *Hierarchical Arrangement*: The images are classified into different categories so that accessing a specific image belonging to a category becomes easy. Some of the examples of categories used are,

1. Fruits.
2. Vegetables.
3. Flowers.
4. Domestic animals.
5. Vehicals.

Categories are also represented by images so that the user can simply click on the specific category image to select it. For example, Images of different types of fruits are stored within the Fruits category (group). Thus, to access say banana, select Fruits group and then select banana. In this way the selection of an image becomes easy. Thus, the image data is arranged in the form a tree as shown in Fig. 2. The user selects the category (group) first, then the desired image within that category is selected.

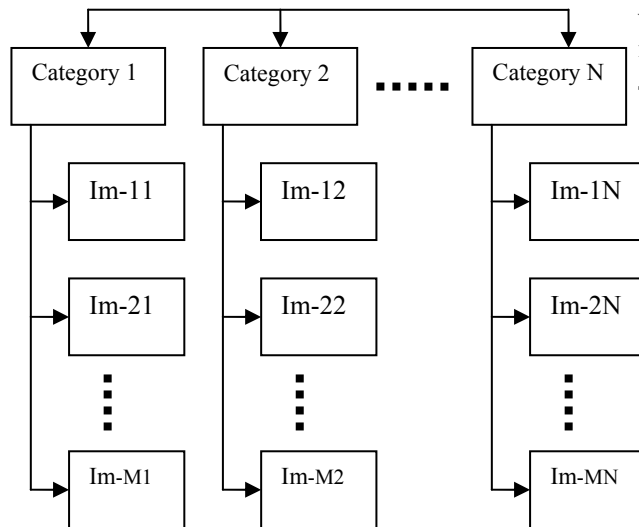


Fig. 2. Hierarchical Arrangement of Images

B. Linguistic Text data

Linguistic word for each object (whose image is already generated and stored) is derived by the knowledge of the object and by knowing the matching word for that image in that language. The dictionaries of the relevant languages are used to find out the matching words in the respective languages. **Our Image to Text** converter is basically a look up table.

1) *Look up table formation of linguistic words:* The images and the matching words are stored in a look up table as shown in Table 1.


Table 1. Arrangement of words matching the images

Image →	Im-1 Col-1	Im-2 Col-2	----	Im-M Col-M
Word in Language -1 (Row-1) →	Word-11	Word-12	----	Word-1M
Word in Language -2 (Row-2) →	Word-21	Word-22	----	Word-2M
-----	-----	-----	----	-----
Matching Word in Language -L (Row-L) →	Word-L1	Word-L2	----	Word-LM

For a given image, Im-M in column M, the matching words in different languages are stored in the corresponding rows of column M. Thus the word in column-M and row-L gives the matching word for image M in language L. In other words, the column and row indices of a word identify its image and its language. Therefore, once the image M is selected, the matching words in different languages are obtained from rows of column-M.

For example, the words for banana fruit are stored as shown in Table 2. Let the image represent Col-J. The matching words in English and different Indian languages are shown in the successive rows of that column.

Table 2. Storage of Words matching banana image

	Col-J	-----
Image →		
Word in Lang-1	Banana	
Word in Lang-2	kEla	
Word in Lang-3	bALehaNnu	
Word in Lang-4	vALaippaLam	
Word in Lang-5	araTipaMDu	
Word in Lang-5	vAlappalam	

2) *Use of common English language script:* In our scheme, all the words are stored in the English language script so that the data can be stored and manipulated as strings in our programming language. The English language script is relatively easy to enter while creating the table.

3) *Audio Interface:* The correct pronunciation of each word in its native language is recorded and then stored as an audio file. The audio pronunciation file corresponding to image Im-J and with word in language-I, is stored with column index J and row index I, so that it can be correctly accessed to match the image and the word in the specified language. (We could not find proper text to speech converters for Indian languages.)

III. IMPLEMENTATION IN HTML AND JAVASCRIPT

Our **Image to Text** converter is implemented in html and Javascript so that it is web enabled. The user can run the software using a suitable browser.

A. Display and Selection of a specific image

The images are displayed as backgrounds for click buttons which are placed strategically on the web page. In a specific category, let the number of distinct images stored be M. (Each

image represents an object.) These images are represented by M click buttons arranged in a rectangular array as shown in Fig. 3. Each click button carries a specific image as its background. In Fig. 3, each image is essentially a button. Each button is indexed and the Jth button contains the Jth image for J = 1 to M. Indices from 1 to 9 is shown in Fig. 3. By clicking on that image button, that image is selected and the corresponding column index number is activated. This in turn selects the words corresponding to that column index and these words are displayed in the appropriate text boxes provided.



Fig. 3. Buttons with background Images

B. Display of the word of the selected Image

Once the image is selected, its column index J is fixed. Then the user selects the language in which to get the word. This sets the row index say to I. See Table 1. Then the word selected is Word-IJ. This word which is stored in English script should be converted from English script to the script of the selected language. This operation is **Transliteration**.

Transliteration: This process converts the script from English to the chosen (target) Indian language. The process uses the standard transliteration rules with reference to the target language. Here, detection and conversion of basic phonemes are done using a few elementary techniques of Natural Language Processing of the Basically the conversion takes place from English characters (mainly alphabets) of the word to the corresponding 16 bit Unicodes for the target language as shown in Fig. 4. Valid phonemes are detected using the *regular expression matching*[6] and then, they are appropriately replaced by the Indian language characters. Designated Unicode blocks are used for

different Indian languages. The string replacement Javascript function used is, `str.replace(regexexpression, targetstr);` The java script function, `String.fromCharCode(Unicode value);` converts the Unicode value into its equivalent Indian language character in the specified font. Technical references are given in [7].

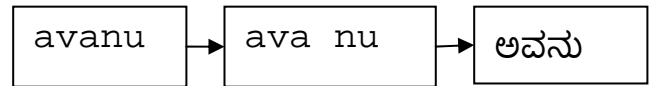


Fig. 4. Transliteration by phoneme replacement

The range of Unicode blocks for some prominent Indian languages are given in Table .3. These values are used in the transliteration process. Languages are selected using a drop down menu. When a different language is selected, the Unicode block (range) needed will be correspondingly changed to its present value.

Table 3. Unicode range for Indian languages

Index (x)	Language x	Unicode range in Hex	Unicode range in decimal
0	Devanagari (Hindi)	0900–097F	2304–2431
1	Bengali	0980–09FF	2432–2559
2	Gurumukhi	0A00–0A7F	2560–2687
3	Gujarathi	0A80–0AFF	2688–2815
4	Oriya	0B00–0B7F	2816–2943
5	Tamil	0B80–0BFF	2944–3071
6	Telugu	0C00–0C7F	3072–3199
7	Kannada	0C80–0CFF	3200–3327
8	Malayalam	0D00–0C7F	3328–3455

The target languages are selected using drop down menus which are crated using the <select> and <option> tags. The selection process returns the index position x, of the selected language. In Table 3, x varies from 0 to 8.

From Unicode Table 3, we see that the Unicode of the starting character of Indian language whose index is x, is given by,

$$\text{Starting Unicode of language } x = 2304 + 128 * x \quad (1)$$

Therefore the transliteration algorithm applied for Devanagari (whose index x =0) equally holds good for other Indian languages listed in Table 3, by shifting the offset of generated unicodes by 128*x. some additional corrections are needed to Tamil because of shortages of some alphabets in that language.

C. Pronunciation Audios

The pronunciation audio of each word is recorded and stored in the *mp4* format. The audio files are also indexed like words. Thus the audio file corresponding to Word-IJ is stored as Af-IJ. This file is stored at row-I and column-J of the AF array. Therefore, once J and I are selected, the matching audio file is accesses and played.

To play the audio file, our software uses the `<audio id="name_audio" ></audio>`.

When the user selects an image, the corresponding words as well as their audio files are selected. The user can listen to the pronunciation when he clicks the corresponding "PLAY_SOUND" button.

IV. WEB PAGE LAYOUT AND OPERATIONS

The **Image to Text** converter operates in two levels. In the first level we select the categories. In the next level, we select the objects of that category. Initially the category page is opened in the browser using the 'url' of the **Image to Text** converter. The screen shot of the category page for a trial example is shown in Fig. 5. (The category items are incomplete). There are 12 categories in the page of Fig. 5. The user selects a category by mouse left clicking a chosen image. (in this case 'birds' category is selected). Then the objects within that category are displayed in the *new tab* as shown in Fig. 6.

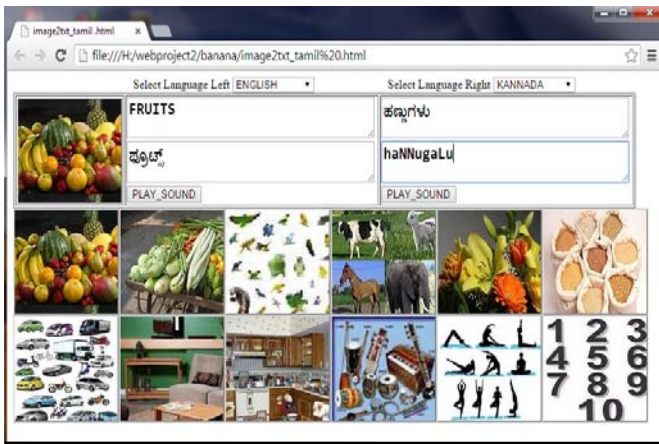


Fig. 5. Category page (Main page)

The user can select any desired image by clicking on (inside) that image. The matching words are displayed in the designated text areas.



Fig. 6. Images of objects within 'birds' category

In Figs. 5 and 6, only two languages are active. The web pages can be easily redesigned for more languages.

V. EXTENSION AND FUTURE SCOPE

The basic software can be extended to cover all categories of objects from every knowledge stream like history, geography, physics, chemistry, wildlife etc. We can also provide more number of hierarchical levels, so that users can access objects depth wise. The data base of objects and words used in the software can be classified according to the learning levels of the users like young learners (say 5-8 age group), medium level groups (9-12 age group), senior school group and so on. Also the software can be specially built for tribal people, rural population, migration workers etc.

The principles used in this software can be modified for generating quizzes, puzzles and learning games. Thus the learning process can be evaluated and it can be made a funny affair..

VI. CONCLUSION

A very useful Multi-media software is presented that provides easy learning of new words in new languages. The introduction to different objects is through images which makes the operation of the software very user friendly. Since the **Image to Text** converter software is web enabled, it can be easily distributed and deployed extensively over a wide region. The techniques used here can be extended to other non-Indian languages like Arabic, Chinese, Russian etc.

REFERENCES

- [1] A picture is worth a thousand words: applying image-based learning to online course design. www.cirtl.net/node/7533.
- [2] Starfall: Learn to Read with Phonics, Learn Mathematics. www.starfall.com/.
- [3] Enchanted learning home page. www.enchantedlearning.com.
- [4] Learn Indian Languages - LanguageReef - Your one stop solution for all your needs. www.languagereef.com/.
- [5] Preschool Hindi for Kids www.akhlesh.com/.
- [6] Java Script Reg Exp Object -W3 Schools. www.w3schools.com/js/js_regexp.asp.
- [7] Antony P J and Dr.Soman K P. "Machine Transliteration for Indian Languages: A Literature Survey," International Journal of Scientific & Engineering Research, Volume 2, Issue 12, December-2011.