

# Learning phrase-level vocabulary in second language using pictures/gestures and voice

**Lavanya Prahallad**  
International Institute of  
Information Technology,  
Hyderabad  
*lavanya.prahallad@research.iiit.ac.in*

**Prathyusha Danda**  
International Institute of  
Information Technology,  
Hyderabad  
*danda.prathyusha@research.iiit.ac.in*

**Radhika Mamidi**  
International Institute  
of Information  
Technology,  
Hyderabad  
*radhika.mamidi@iiit.ac.in*

## Abstract

Earlier works have explored that a foreign language can be learnt from pictures than merely using native language word translations. These efforts on language learning using pictures were limited to colors, animals, birds, vehicles and common noun categories. In this paper, we present our research that learning of verbs is also possible by using right set of pictures and gestures, and we show that this is effective in second language learning. We report the acquisition of words in tourist domain in a second language by working with subjects who are between 14 and 48 years of age. From pre-learning and post-learning evaluations, we show that acquisition of vocabulary like nouns and verbs in a new language is better with the fusion of pictures and voice. We also show the subjects are able to generalize their learning towards phrase-level vocabulary without any additional training and efforts.

## 1. Introduction

Recalling the famous saying “A picture is worth thousand words”, we explore whether a good enough vocabulary acquisition is possible using pictures as medium to learn an Indian language as a second language. India is a wide subcontinent broken down into states, where in each state has its own language; also many states

or villages do not speak English. It is necessary for an Indian to learn another Indian language as second language. Typically a language deals with objects/persons and actions. We consider the fact that the actions exist for practical purposes are limited. Signs (gestures/pictures) used to express these actions appear same throughout the world, except the languages call the same action with different names depending on the spatial, cultural regions. For example, the sign or a picture depicting the action of eating would be same, however, has different names in different languages. Thus pictures supplement learning, helps to understand the words in a more meaningful method. This will indeed enable learners to learn, understand, and utilize that vocabulary in all aspects of their lives. With the recent advances made in education and technology language learning has been easier when voice is added with the picture and the learner hears the word in the second language and learns it by repeating any number of times. This combination enhances once language learning significantly.

Many researchers have explored this area of vocabulary acquisition and demonstrated that using pictures is better than just textual format in vocabulary acquisition while learning a foreign/second language. Experiments conducted by Carpenter and Olson (2011) shows that Swahili words were learned better from pictures than from translations. It appears, therefore, that pictures can facilitate learning of

foreign language vocabulary. Taking the idea from the Dominance, Gentner and Boroditsky (2001) work, we exploit that words can be learnt by using pictures because of their concept-to-word mapping. Cognitive dominance prevails when concepts are simply named by language as in the case of nouns. Though the inspiration for this work is from earlier research, our approach differs in the following ways:

- This study is done for a set of Indian languages and can however easily be extended to other languages of the world.
- We show that even verbs can be learnt using right set of pictures in addition to nouns.
- Our study shows that the subjects can learn the phrase-level vocabulary as a generalization of their training on nouns and verbs.
- Use of voice with pictures aids in learning the second language and can relate that sound to the picture.

In our study we explore teaching a good enough way of learning verbs in Indian languages taken as second language. The medium is chosen to be pictures/signs, which are universal throughout the world; hence no first language influence is seen on the acquisition. This idea coupled with the voice has helped the learner to pronounce the word equivalent to a native language speaker. Even the illiterate folks can be benefitted with this research as objects he knows and sees everyday can be related to the audio being spoken out. This research is language independent and can be useful to learn any language as second language with the appropriate dataset. The idea is exploited from Chomsky's universal grammar that this theory suggests that linguistic ability manifests itself without being taught and that there are properties that all natural human languages share, the pictures of the actions will

immediately get related to a person's mental picture making him understand what that picture means, now with the help of voice he learns what the action is called in that second language.

## **2. Nouns and Verbs in second language learning**

Nouns and Verbs form a pivotal part of learning a vocabulary in any language. Noun is an entity that denotes objects and Verb tells us about the action in a situation. There are several kinds of verbs. Though nouns and verbs exist universally in all languages of the world, the structure and the position of the verb and the object in a sentence varies. Unlike English, Indian languages have SOV word order where verb appears at the end of the sentence. The verbs in Indian languages are inflected gender and number and also for tense, aspect modality (TAM) where as in English they appear separately. In our work, we attempt to show that nouns and verbs can be learnt in a second language by using just pictures and there by one can achieve good enough vocabulary in a chosen domain.

## **3. Dataset**

We have chosen 'tourist' domain to show that a related vocabulary in the domain can be acquired in a second language. As a pilot for our experiments and to present in this paper we have chosen the following Indian languages as second language: Kannada and Tamil. Any tourist or a visitor can benefit with this idea and can learn a good enough vocabulary when he/she visits a state where that particular language is spoken. The medium of learning is pictures, which prove to be more advantageous as a learner need not use first language to learn.

We carefully chose a set of pictures of nouns and a set of pictures of verbs that belong to tourist domain. Subjects of age 14-48 years participated in this experiment; a webpage is

created with just pictures as shown in the Figure 1. These images are collected from Internet. When a learner clicks on the picture the corresponding audio of the picture is played in the second language. Learner can hear to this audio any number of times until he gets the word and learns the word in the second language.

#### 4. Experimental Procedure

The experimental procedure consists of three phases: 1. Learning phase 2. Post-learning 3. Check for retention. Three experiments are conducted for this study dedicated for nouns, verbs and small phrases in the combination of already learnt nouns and verbs. Participants were informed at the beginning of the experiment that they would be learning Kannada/Tamil/Telugu words by viewing the picture and hearing the sound. All participants were instructed to try their best to learn the words in second language and that later their memory for these items would be tested. Once they complete the learning and proceed for test, participants cannot go back to the lesson. Participants would know their performance at the end of the test.

##### 4.1 Experiment 1 – Nouns

A set of 25 images of Nouns is collected and shown to the learner. The images denote the names of the objects like transport, food, and professionals all are related to the tourist domain. Learner clicks on the each image to listen to the corresponding audio indeed to learn the word in the second language.

**Learning phase:** The web application was shown to the 10 subjects of age group (14-48years) and they were asked to go through the application if images of nouns for 20 minutes. Learners could hear the word by clicking on the image any number of times till they were

comfortable in remembering the word. (Shown in Figure 1)



Figure1: Demonstration of learning phase of experiment-one with nouns related to tourist domain using picture/gestures. Learner clicks on the word to hear the voice in second language

**Post Learning:** After 20 minutes, each person was given a test to evaluate the learning. Participants were then given a test with set of 10 questions in which they were asked to recall the word in the second language and chose the correct answer, given the picture and the four options. The performance of the test will be shown to the participant at the end of the test. We refer to this as error in retention. (Shown in Figure 2)



Figure2: Testing phase of nouns related to tourist domain using pictures. Learner has to listen to all four options of audio files and

*selects the voice option that is a match for the given picture*

**Retention after 2 days:** This post-learning phase was conducted again after 2 days to check the error in retention.

#### 4.2 Experiment 2 – Verbs

In this experiment, a set of 25 images of verbs is collected. The images denote the actions that are carried by individuals as a tourist. These images are selected as such it is performing an action. The experimental procedure is same as explained in section 4.1 with learning phase, post learning phase and check for retention is done for verbs. Figure 3 below shows the set of images of verbs shown to user in learning phase.



*Figure 3: Demonstration of learning phase of experiment-two with verbs related to tourist domain using picture/gestures. Learner clicks on the word to hear the voice in second language*

Figure 4 shows an evaluation test of post learning for verbs.



*Figure 4: Testing phase of verbs related to tourist domain using pictures. Learner has to listen to all four options of audio files and selects the voice option that is a match for the given picture*

#### 4.3 Experiment 3 – Small Phrases

This is the last experiment in the study. Please note that this experiment does not have a training test. A test is conducted by creating a small parse in combination of noun and verb. Learner uses their knowledge of already learnt nouns and verbs Figure 5 shows the test in small phrase-level.

This section which is an evaluation to test the memory of the learner on how much he/she remembers the experiment 1 and 2. A set of 10 combinations of verbs and nouns are displayed for the learner, where he/she is required to listen to the four options of the audio and select the right audio combination of noun and verb.

**Retention after 2 days:** This process was conducted again after 2 days to check the error in retention.



Figure5: Demonstration of testing phase of phrase-like sentences related to tourist domain using pictures. Learner has to listen to all four options of audio files and selects the voice that is a match for the given picture. There is no training for this phase. Learner has to create the right sentences with previous knowledge of already learnt nouns and verbs in second language.

## 5. Results

Table 1 and Table 2 show the results of the retention of nouns and verbs including small phrases immediately after learning-retention after 2 days. The average shows that performance of the participant is better in learning nouns and the retention percentage is also good when compared to verbs. It can be observed from the Table 1 and Table 2 that the retention after 2 days is also fairly well to indicate that learning through pictures helps in remembering the vocabulary learnt.

Participant number	Nouns	Verbs	Small Phrases
1	10-10	9-9	8-9
2	10-10	9-9	10-10
3	10-10	9-9	10-10
4	10-10	8-10	10-10
5	10-10	4-5	7-6
6	10-10	10-10	10-10

7	10-10	10-9	9-10
8	10-10	9-10	9-9
9	10-10	8-7	9-8
10	10-10	10-10	10-10
<b>Average</b>	<b>10-10</b>	<b>8.3-8.8</b>	<b>9.2-9.2</b>

Table1: Retention results of Kannada Nouns, Verbs and small phrases. Each entry indicates immediate retention-retention after 2 days.

Participant number	Nouns	Verbs	Small Phrases
1	10-10	8-9	9-9
2	10-10	10-10	8-10
3	8-7	7-6	9-7
4	10-10	10-10	10-10
5	10-10	10-10	10-10
6	8-7	8-6	9-7
7	10-10	9-8	9-8
8	10-10	9-8	9-8
9	8-7	7-8	8-9
10	10-10	9-9	10-9
<b>Average</b>	<b>9.4-9.1</b>	<b>8.3-8.4</b>	<b>9.1-8.7</b>

Table2: Retention results of Tamil Nouns, Verbs and small phrases. Each entry indicates immediate retention-retention after 2 days.

### 5.1 Observations

It is observed that a learner clicks 2 or more times to hear to the word in a second language and memorize the word and relate the picture to that sound. Also learners are able to relate to the language of same type. Though Tamil, Telugu, Kannada belongs to same language family of Dravidian there is more similarity between the structure, inflections between Kannada and Telugu. For learners whose first language is Telugu are able to perform better in Kannada but they are more time in understanding the accent, words, phones of Tamil. Learners of Telugu as first language feel that Tamil takes more time in learning even through pictures.

Another observation shows that nouns are easier to learn than verbs supporting previously done research. Another crucial observation is that the performance of the learner decreases or learners are taking more time in memorizing ambiguous words or similar sounding words. For example in Kannada “kudi – Drink”, hidi – “catch”.

For words, which are multi syllable or complex, learners find difficult to memorize it, they take more time in hearing these words again and again. Example in Kannada is “kaleduhoyitu – lost and karemadu – calling”.

The performance of the learners while taking the test in constructing the small phrases is good compared to their performance in verbs as they are able to guess with their world knowledge and can relate the pictures of nouns with verbs. This shows that learning nouns and memorizing them is playing a crucial factor in constructing the small sentences. It’s a time taking factor in learning verbs but they are able to memorize it and use it better as pictures are involved.

## **6. Conclusion**

We conclude that our current study provides pivotal insight into the importance of pictures and gestures as educational tools in second language vocabulary learning. The novelty of this work is introducing a foreign word by means of a picture and no other medium of instruction. Learners are able to construct the small phrases by learning nouns and verbs and the score is around 9.2 in Kannada and 8.7 in Tamil. Our experiments explained in the section 4 are carried on 10 subjects of age group (14-48years). However, we plan to extend the scope of this experiment by extending to more languages as second language learning and conduct the experiments on more subjects of different age groups. This would also help us to analyze the learning capabilities of second

language among the different age groups from younger to older.

This could be further extended in learning advanced features of second language such as grammar structure such as inflections, postpositions, tense aspect modality and sentence comprehension.

This application helps to learn a good enough vocabulary acquisition tailored to tourist domain. Traveling nationally and internationally has exponentially increased in today’s world for various reasons like business or vacation. Learning a language of the visited places is always helpful to interact with localities, therefore as the part of phase one of this project; tourism domain is chosen as we see great potential of second language learning using pictures in this domain. However we plan to extend our project to different domains as medical, supermarkets etc.

Our idea of vocabulary learning using pictures can be easily ported to mobile phones. Moreover, the use of mobile phones and platforms could be effectively employed in acquiring a second language using contextual learning. This application is more helpful to tourists as adults are often busy and do not have the time to go through books which give an insight to a language that is spoken in other states if he/she visits a different place to cope up with transport, directions, food etc. This helps learning environment to be mobile and anytime anywhere.

## **7. Future Work**

Our study for this paper paved a way to work further in this area of language learning using pictures and gestures. Important points to be addressed in future research include exploring more on the issue of showing expression, emotion, a short clip of an action captured as a video for verbs that are cognitive related to

reduce ambiguity and bring clarity of the intended verb by showing action. The current limitation of our work is that we are not addressing to make them learn the inflections, grammar and the postpositions, which are pivotal in Indian languages; these issues need to be carefully worked out on to how to teach them using pictures. Also the issue of accuracy in recalling the pictures by the participants when checked after a week or a month should be taken care. This particular problem gave us an idea contextual learning by developing a mobile application on a smart phone where the learner can always use the app location and language based. We think that contextual learning using pictures and gestures (learning the words on the spot and using them with the native speakers) helps in learning the language better.

In keeping the view of learning a second language is a global requirement; we intend to extend our research for learning a second language that belongs to different language family with addition to add more languages to the list of same language family. Our work will be based on the idea that nouns and verbs of similar kind exist in all languages of the world; in this case pictures help to identify the words even they are of different language family. Voice coupled with the image helps the learner to build on their knowledge of language by making connections and comparisons between the home language and the target language.

## 8. Acknowledgements

We would like to thank all the participants for their participation in user studies.

## 9. References

Shana K. Carpenter and Kellie M. Olson. 2011. *Are Pictures Good for Learning New Vocabulary in a Foreign Language? Only If You Think They Are Not*, Journal of Experimental Psychology. Learning, Memory, and Cognition.

Dedre Gentner and Lera Boroditsky. *Early Acquisition of Nouns and Verbs Evidence from Navajo*.

K.G. Vijaykrishnan. 2007. *The "Noun Advantage" in English as a Second Language: A Study of the Natural Partitions Hypothesis'*. Dissertation in English and Foreign Languages University, Hyderabad, English Language Education.

Noam Chomsky. *Universal Grammar*. Retrieved from [http://www.rit.edu/cla/philosophy/quine/universal\\_grammar.html](http://www.rit.edu/cla/philosophy/quine/universal_grammar.html) and [http://en.wikipedia.org/wiki/Universal\\_grammar](http://en.wikipedia.org/wiki/Universal_grammar)

Handley Z. 2012. *Text-to-Speech Synthesis in Computer-Assisted Language Learning*. The Encyclopedia of Applied Linguistics.

*Verbs in English Grammar*. Retrieved from web <https://www.tesol-direct.com/guide-to-english-grammar/verbs>

Liz Stinson. 2014. *A Brilliant Way to Learn Chinese Through Pictures*. Retrieved from <http://www.wired.com/2014/03/chineasy/>

Timothy G. Collins. 2005. *English Class on the Air: Mobile Language Learning with Cell Phones*. Proceedings of the Fifth IEEE International Conference on Advanced Learning Technologies (ICALT'05).

Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A. Landay. 2011. *MicroMandarin: mobile language learning in context*. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). ACM, New York, NY, USA, 3169-3178.