# Syllable based noun recognition for grounded videos

Prashant Jalan
Department of Computer Science & Engineering
Indian Institute of Technology Kanpur
Homepage: www.prashantjalan.com

April 3, 2013

# Contents

**Acknowledgements**

I would like to thank Dr. Amitabha Mukerjee under whose supervision I did this work. If not for his guidance, motivation and support, this work would not have been possible. I am thankful that he gave me an opportunity to work under him.

I am thankful to Mr. Diwakar Chauhan who was always there to help me understand the difficult concepts and motivate me throughout the project.

I acknowledge the excellent word of Nikhil Joshi in 'unsupervised language learning for complex 3D videos'.

I owe to my parents and my sister for their moral support and encouragement.

Finally, I thank IIT Kanpur for giving me this platform to pursue my interests and do this research work.

**Abstract**

We aim to make the computers learn a new language without any previous knowledge about the language. In this work, we have used a semantic syllabic approach and also a word level analysis to acquire basic linguistic units particularly, noun based on the Langacker [] theory of learning language. Based on a 2D video and co-occurring raw text, we demonstrate how this cognitively inspired model segments the world to obtain a meaning space, and combines words into hierarchical patterns for a linguistic pattern space. We try to recognize nouns in the English language and the Hindi language based on some narrations taken from different subjects using different association measures such as the mutual information, relative frequency, conditional probability and dominance weighted joint probability.

# Chapter 1

# Introduction

## 1.1 Language learning framework

The problem of language acquisition has been of great interest to many disciplines including Linguistics, Psychology, Philosophy, Neurobiology, Cognitive science and Computer Science. From Panini [] to Chomsky [] to Tomasello, there have been many attempts to formalize the theory of language. The debate is mostly two-sided. Chomsky [] argues for the innateness of language based on the argument (known as "poverty of stimulus") that the child acquiring language has access to only positive examples (grammatical sentences), and very little corrective feedback. Thus, the Chomskyan framework focuses on the syntax of a language and is largely sceptical about semantics. So, learning a language from his viewpoint is learning a "generative syntax" for that language. Langacker [], alternatively has given a central role to semantics in his language learning model. Langacker [] considers grammar as conceptualization and formalizes it as a bipolar symbolic unit interconnecting the phonological pole (linguistic representation) and the semantic pole (conceptual representation). In the view of cognitive grammar, language is entrenched in the usage and linguistic representations get their meanings because of their usage with some conceptual entity. The idea is analogous to a child's way of learning. When a child is born, he knows nothing about a language. He doesn't know anything about the noun, verb, preposition or the syntax or the word boundaries. But as he continuously hears description, slowly after many instances of a particular object or an action been referred to by a particular word, the child begins to recognize the word and associate it with the object or action.

## 1.2 Acquiring the linguistic units

Subsection text here.

## 1.3 Summary of Results

Subsubsection text here.

# Chapter 2

# Corpus Analysis

While learning the linguistic units, we did not consider the most common and frequent words used in English and Hindi. For English, we took the most commonly used words from a previously done analysis using British English Corpus, American English Corpus and recorded talks and speech []. For Hindi, we use Hindi unicode corpus, Center For Indian Language Technology, IIT Bombay []. We perform both syllabic and word analysis in the Hindi corpus to discover the most frequent words and top k-grams in Hindi.

## 2.1 Most frequent words in Hindi

## 2.2 Top k-grams in Hindi

# Chapter 3

# Conclusion & Future Work

*temp* Given the object categories discovered and visual saliency of these objects over the time, we demonstrate the ability of our system to learn nouns like in Hindi and red, blue, triangle, big, small in English. We confirm the success in learning words by analysing the strength of associations with increasing number of narrations. Discovering lal and neela from narrations describing the triangles as lal and neela and chota and bada from the narrations which describe the triangles as big or small, both in English and Hindi, confirms the success of our model. We argue that the consistent dominance of association strength of label with a visual category over the other labels is desirable and can be taken as a confirmation of the word learning. The success in learning appropriate labels even without knowing word-boundaries shows that the knowledge of word boundaries may not be a prerequisite for early word-learning. Getting the same results at a word level analysis illustrates the correctness of the association measures we have used.

# Bibliography

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.