

## Problem Statement:

How many words do you know in your native language?

Write a short report with an answer to this question. I don't want a generic answer. Make your answer specific to you and your language. Your answer should have some empirical basis such as sampling with a dictionary or analysing your writing. You should also specify how you interpreted "how many", "word", and what "know" means.

## Literature Survey:

Kannada is my native language and is a Dravidian language spoken predominantly by the people of Karnataka, India. Kannada has over 45 million native speakers and 8<sup>th</sup> most spoken language in India. Kannada is a morphologically rich language. There are 49 letters in the alphabet. Kannada literature comprises of great literature works like the Sri Ramayana Darshanam by Kuvempu which received the "Jnanpith award".

## Empirical Analysis:

Aim: To calculate the number of familiar words in Kannada language.

The following approaches are adapted to check the knowledge on words in the native language:

- 1) **An empirical analysis conducted on the 49 alphabets of the Kannada language with word length of 3 is as follows:**

Kannada Alphabets	Number of 3 letter words	Knowledge basis of words of length 3
ಅ,ಆ,ಇ,ಈ,ಉ,ಊ,ಋ	15, 20, 7, 2, 5, 1, 0	10,17, 6, 2, 4, 1, 0
ಎ,ಏ,ಐ,ಒ,ಓ,ಔ,ಅಂ,ಅಃ	8, 5, 3, 7, 4, 4, 3, 0	7, 4, 2, 7, 4, 3, 2, 0
ಕ ಖ ಗ ಘ ಙ	47, 4, 23, 3, 0	40, 2, 15, 2, 0
ಚ ಛ ಜ ಝ ಞ	12, 1, 8, 0, 0	10,1, 6, 0, 0
ಟ ಠ ಡ ಢ ಣ	8, 0, 1, 0, 0	5, 0, 1, 0, 0
ತ ಥ ದ ಧ ನ	8, 0, 16, 3, 34	7, 0, 14, 2,30
ಪ ಫ ಬ ಭ ಮ	18, 2, 11, 3, 16	14, 2, 9, 2, 10
ಯ ರ ಲ ವ	3, 9, 3, 7	2, 7, 3, 5
ಶ ಷ ಸ ಹ ಳ	5, 4, 12, 8, 0	3, 1, 10, 6, 0

## **Conclusion:**

Iterating through the above table of small sample set leading to the calculation of being familiar with 278 words out of 354 words of length 3 with a percentage of **78.53** awareness.

- 2) **Sampling with a dictionary:**

A careful sampling of the dictionary named "A Kannada-English dictionary by author Kittel, F. 1832-1903", it is observed that the dictionary comprises 1770 pages of words. Each page comprises of an average of 50 words. Analysis of the number of words familiar from the dictionary out of 20 visited pages is on an average 16 per page.

## **Conclusion:**

Extrapolation as the result of empirical analysis leads to a familiarization of **28,336 words**.

(16 words per page\*1770 number of pages)

**Interpretation definition: The word**

The corpora thus considered here include various transformations of the Kannada vocabulary word as a unique word added to the known word list.

**Knowledge Base:**

The percentage of awareness of words starting with the alphabets of Kannada is greatly influenced by the following factors:

- Communication: The language is being acquired from parents and teachers who taught me the language during childhood during their interaction with me.
- Media/Entertainment: My vocabulary increased in Kannada by watching television series and movies time to time.
- Education: A formal education of 12 years with Kannada as first language.

**References:**

1] <https://catalog.hathitrust.org/Record/012154563>