

Creation of Multilingual Speech Resources: Academic & Technical issues



Title: Dictionary Optimization for Large
Vocabulary Automatic Speech Recognition

Authors Names: R.Ravindra Kumar, K.G. Sulochana,
Pradeep Balachandran

Affiliation: Language Technology Section C-DAC,
Thiruvananthapuram

www.cdactvm.in

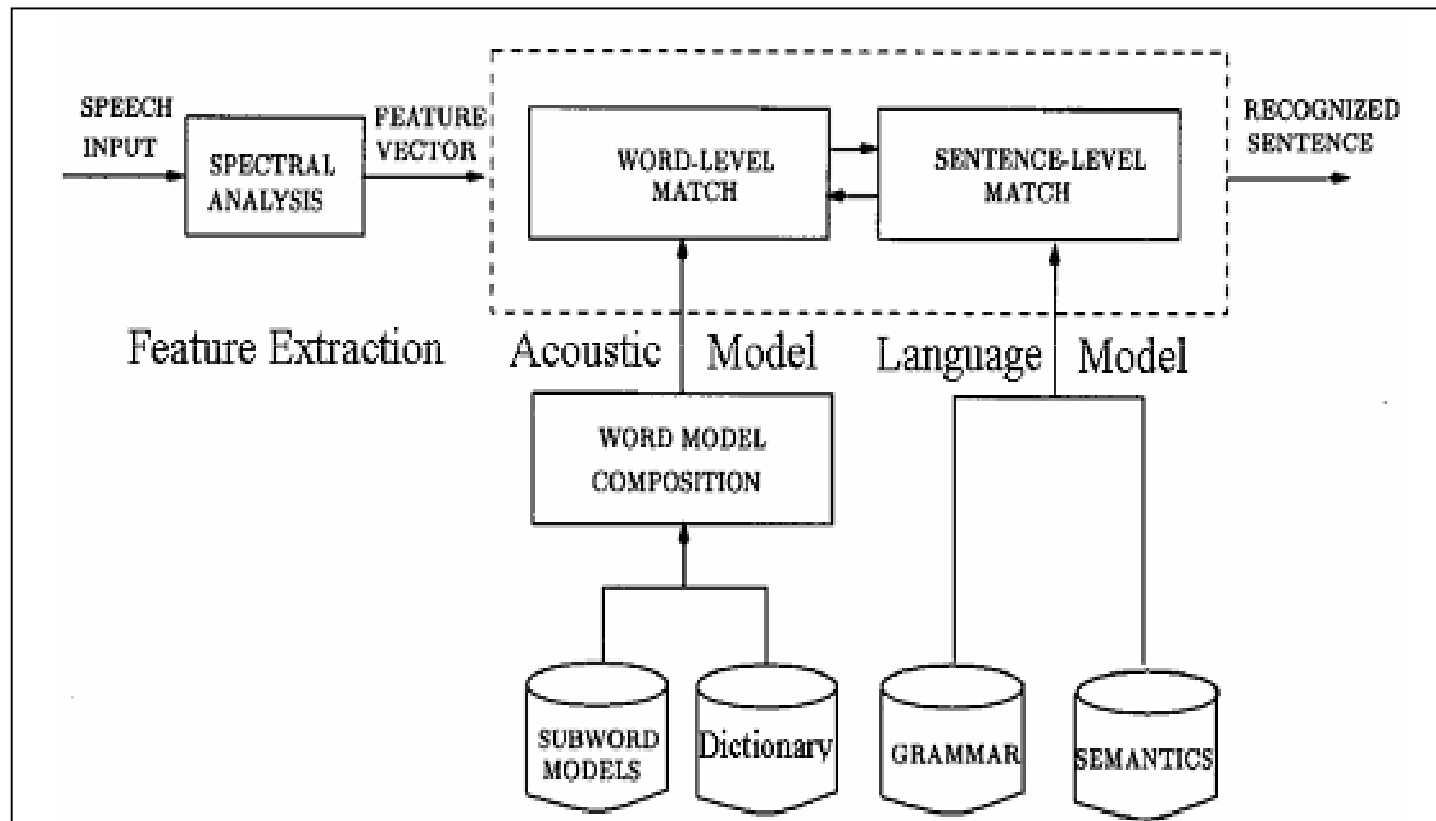
Contents



- Automatic Speech Recognition System: An overview
- Sphinx4 Speech Recognition System
- System Constraints for large Vocabulary ASR
- Need for Dictionary and Grammar Optimization
- Proposed Method
- Performance Evaluation
- Conclusion

www.cdactvm.in

Automatic Speech Recognition System: An overview



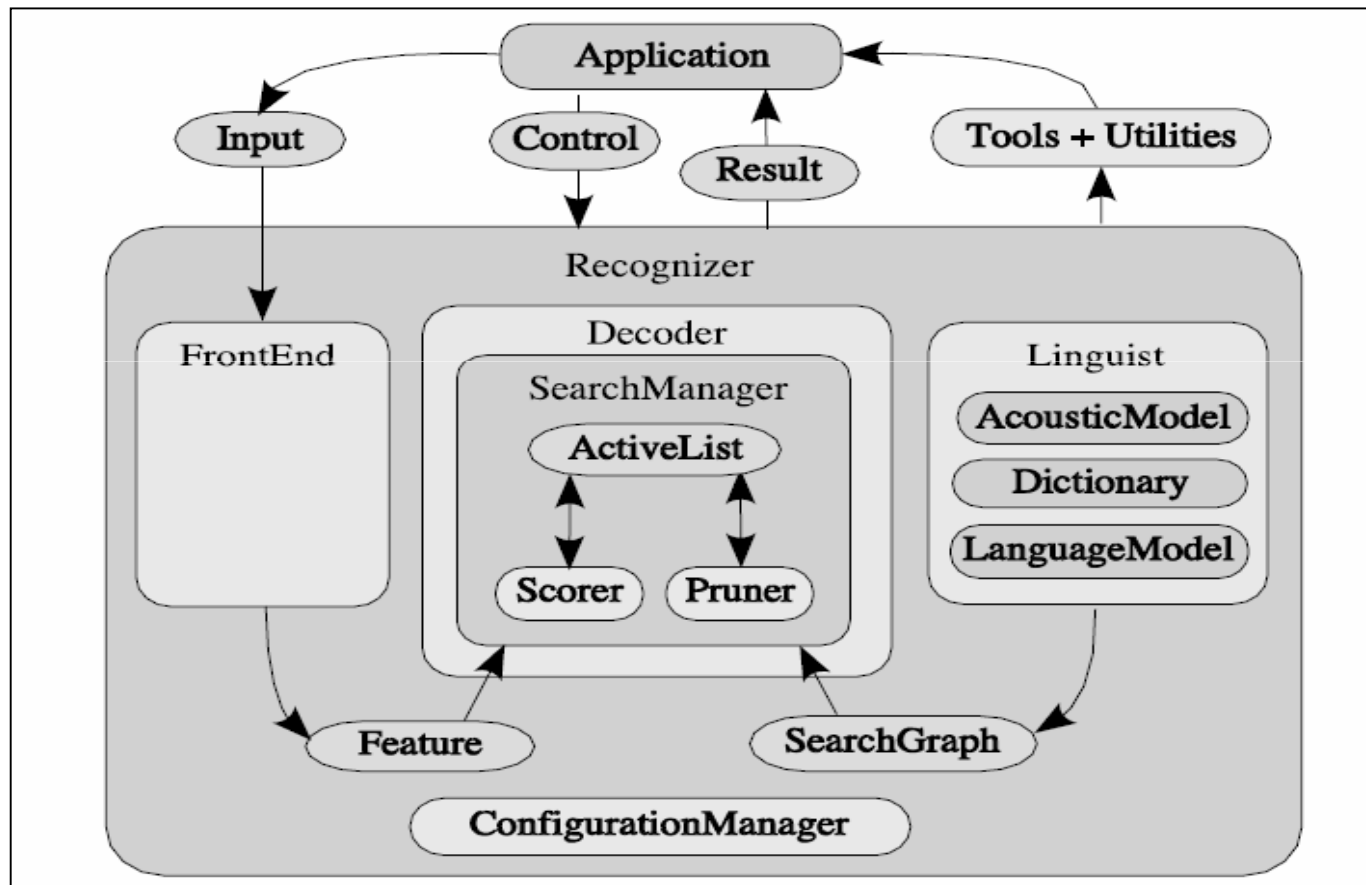
Automatic Speech Recognition System



- **Dictionary:** Contains pronunciations for words to be recognized
- **Grammar:** Defines the syntax of the words to be recognized

www.cdactvm.in

Sphinx 4 Speech Recognition System: Architecture



Sphinx 4 Speech Recognition System



Dictionary Design

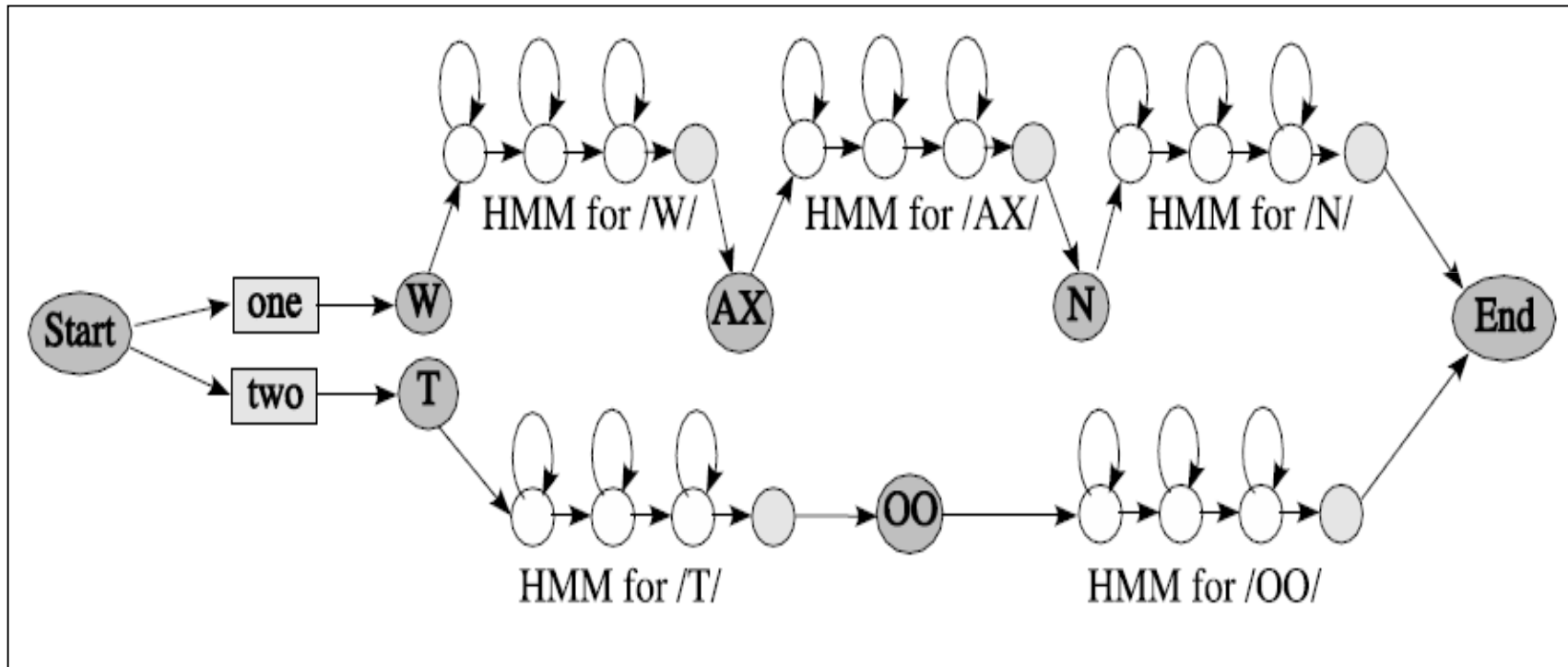
```
aadya aa vbd d y a
ks'he'tranat'a   clk k s'h e' clt t r' a nl a clt' t' a
sam'ghat'ita    s a ng' clk kh a clt' t' i clt t a
vikram'         v i clk k r' a m
kat'utta        clk k a clt' t' u clt tt a
vimukhata       v i m u clk kh a clt t a
```

www.cdactvm.in

Pronunciation can be defined using a standard transliteration scheme

Sphinx 4 Speech Recognition System

Grammar Design



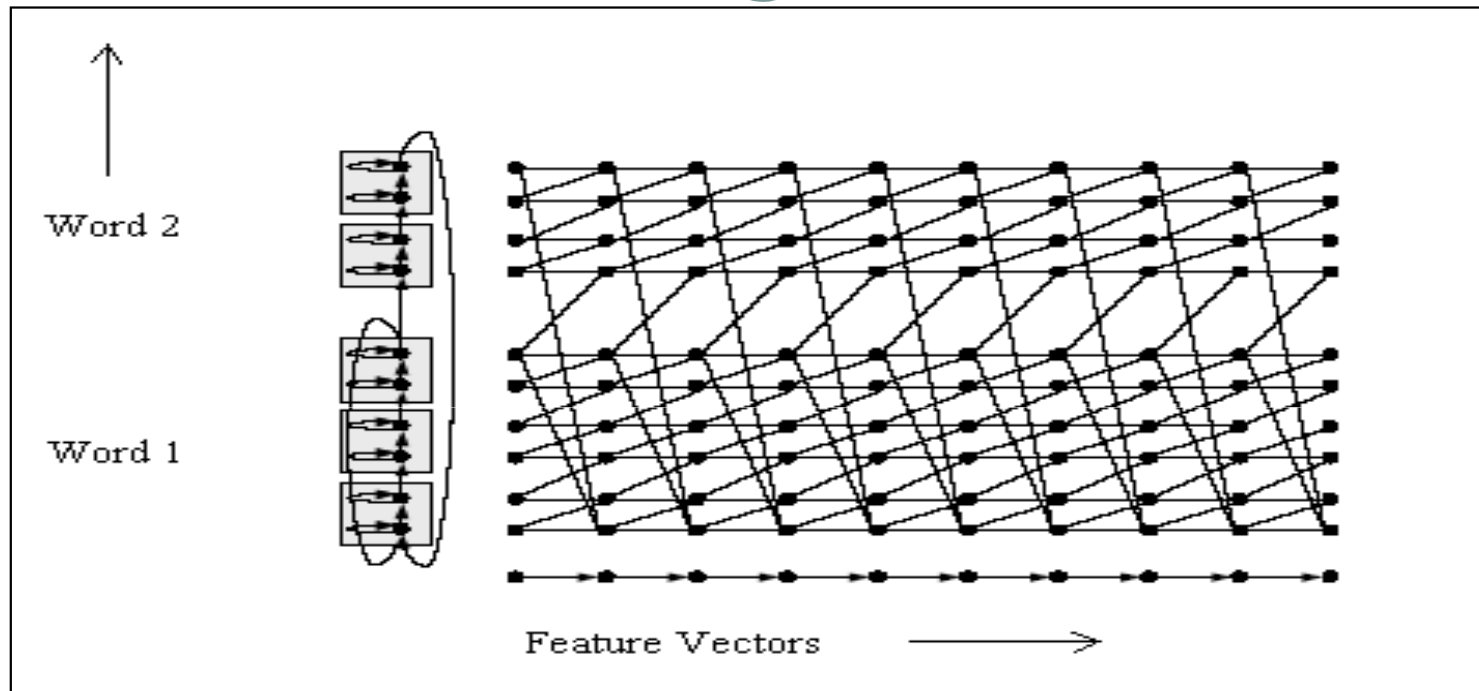
System Constraints for large Vocabulary ASR



- Grammar makes use of Dictionary information to construct a data structure called SearchGraph primary data structure used during the decoding process
- As the vocabulary size increases the size of the Dictionary as well as the Grammar increases and hence the size of the SearchGraph
- SearchGraph is loaded on to the memory during system initialization
- Hence the need to optimize the memory usage while constructing the SearchGraph

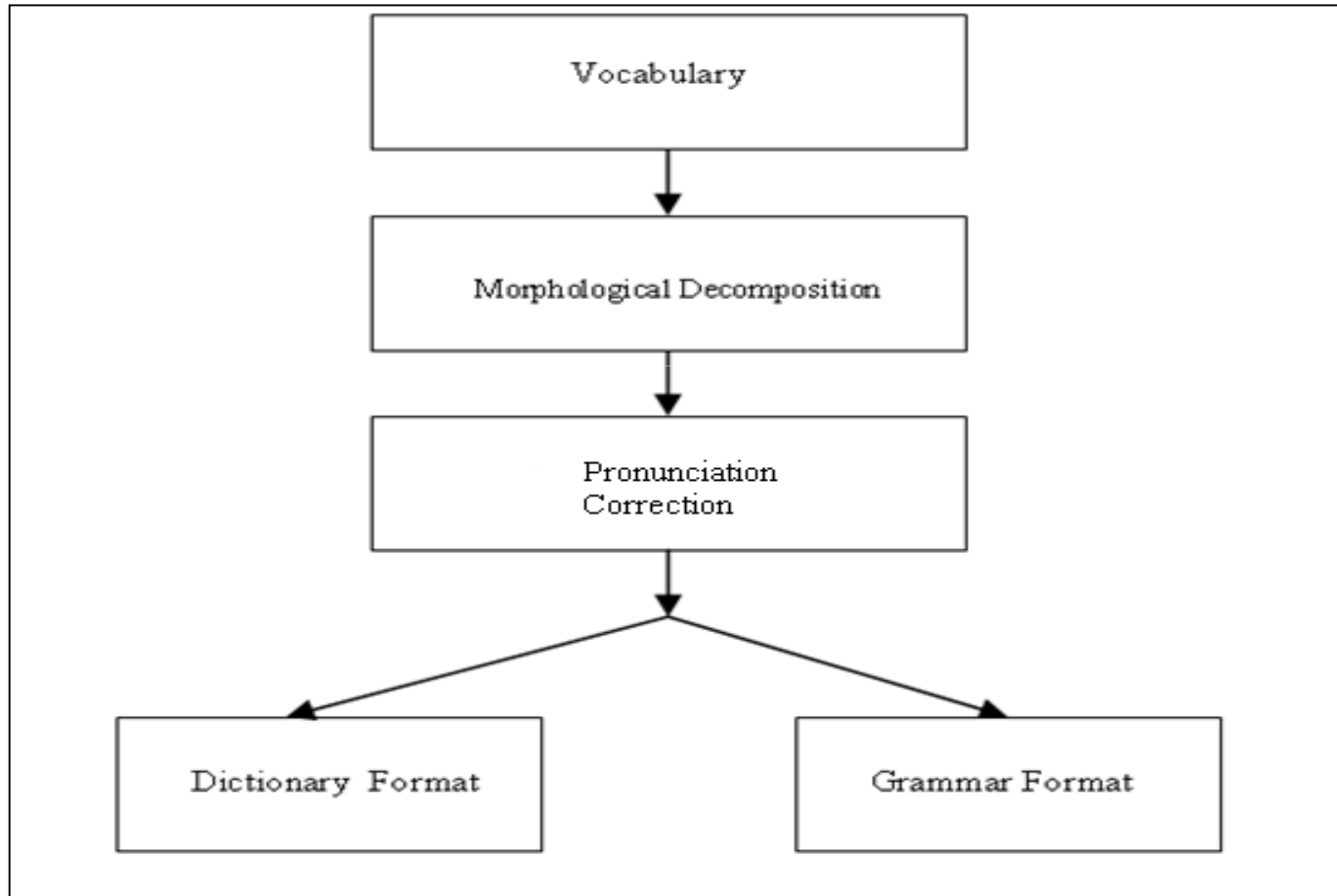
www.cdactvm.in

Need for Dictionary and Grammar Optimization



- Recognition is essentially a search problem
- Search is carried out through a SearchGraph ,the primary data structure used during the decoding process

Proposed Method

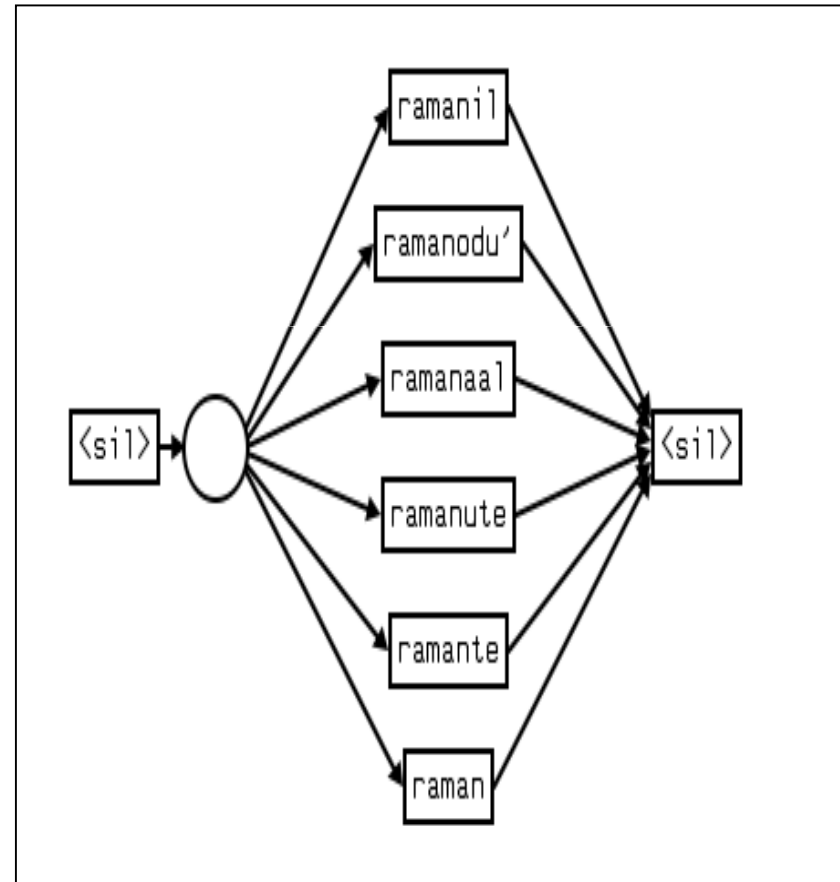


Performance Evaluation

Dictionary (Default)

r a m a n i l	r a m a n i l
r a m a n o d u '	r a m a n o d u '
r a m a n a a l	r a m a n a a l
r a m a n u t e	r a m a n u t e
r a m a n t e	r a m a n t e
r a m a n	r a m a n

Grammar (Default)

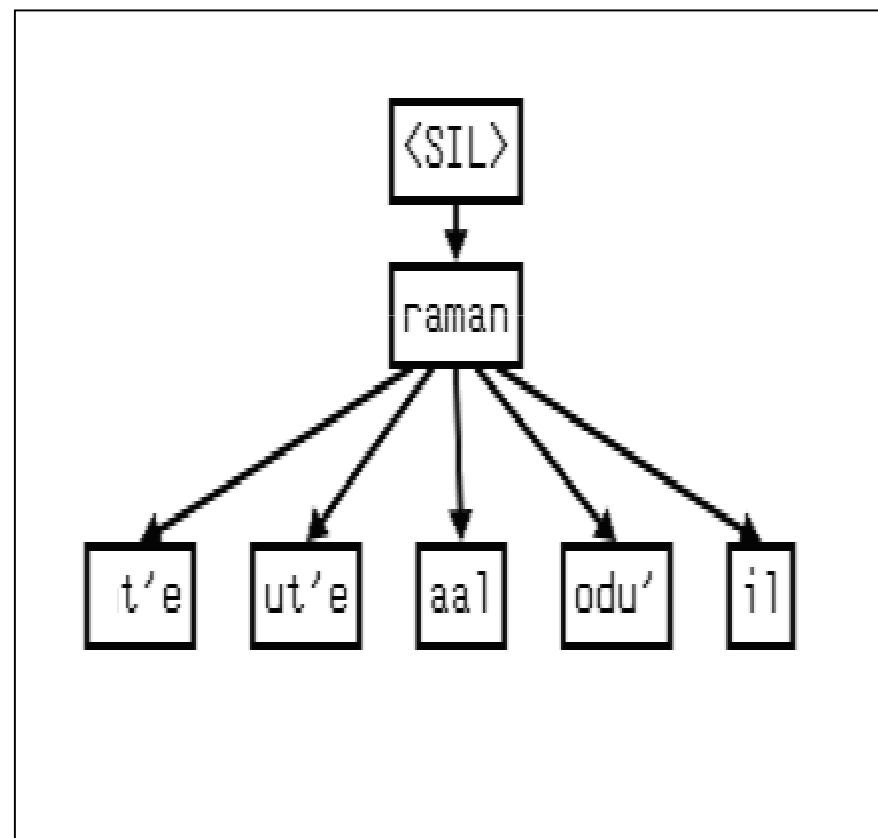


Performance Evaluation

Dictionary (proposed)

r a m a n	r a m a n
i l	i l
o d u'	o d u'
a a l	n a a l
u t e	u t e
t e	t e

Grammar (proposed)



Experimental Verification



- A 1000 word random vocabulary taken as input
- Morphological decomposition performed to obtain root and suffix units
- 24 unique roots and 18 unique inflectional suffixes obtained

www.cdactvm.in

Experimental Verification



How Dictionary got optimized ?

- able to replace 432($18 * 24$) out of 1000 words with just 42 words in the Dictionary.
- able to compress the 432 morphologically rich vocabulary size by almost a factor of 10.
- The total dictionary vocabulary size is reduced from 1000 to 610

www.cdactvm.in

Experimental Verification

How Grammar got optimized ?

- With the tree representation we are able to share the common units to all the 432 words leading to better memory utilization
- We measure the memory efficiency in terms of the Grammar Load time

	Minimum Time	Maximum Time	Average Time
Default method	0.1090s	0.1090s	0.1090s
Proposed method	0.0630s	0.0630s	0.0630s

Conclusion



- Proposed a new method for Dictionary and Grammar representation for large vocabulary automatic Malayalam Speech Recognition systems.
- We show that the proposed method based on the morphological decomposition of words has better efficiency in terms of system memory utilization.
- The proposed method holds promise in the case of other languages of similar agglutinative nature.



Thank you for your time