



CHARACTERISTICS OF INDIAN  
LANGUAGES

BY  
MADHAVI VARALWAR  
NIXON PATEL

**Some of the features of Indian languages and the scripts used to express them are :**

- **PHONEME SET**
- **DIFFERENT GRAPHEME'S**
- **REDUPLICATION**

# PHONEME SET

Akshara forms the fundamental linguistic unit in Indian languages .

An akshara consists of 0, 1, 2, or 3 consonants and a vowel.

The typical forms of Akshara are V, CV, CCV and CCCV, thus have a generalized form of C\*V

Words are made up of one or more aksharas.

Akshara can be pronounced independently as the languages are completely syllabic.

Aksharas with more than one consonants are called samyuktaksharas or combo-characters

# DIFFERENT GRAPHEMES

- Each language uses different scripts consisting of dissimilar graphemes for printing.
- There are 10-12 major scripts in India.
- Some have a head-line or shirorekha that persists for a whole word. Others have non-touching graphemes .
- அ -Tamil
- అ - Telugu
- अ - Hindi
- - Malayalam
- आज -aaj -today
- పగలు -pagalu - morning

# REDUPLICATION

- All languages employ reduplicated form
- Extensive use of reduplication is a particular characteristic of Indian Languages .
- An overview of the different kinds of reduplicative expressions found in the subcontinent
  - 1.ONOMATOPOEIC
  - 2.EXPRESSIVE
  - 3.PAIRED WORDS
  - 4.ECHO

# ONOMATOPOEIC

- Onomatopoeic expressions and expressive share the general property that neither of the two halves of the expression are independently meaningful
- Onomatopoeic expressions may or may not be reduplicated in structure, and the repetition may be exact or there may be segmental differences between the two halves, or a non-reduplicated case.
- kiiccukiiccu-ηηυ — ‘chirping’ (TAMIL)
- taaraa zuvva — ‘like a star’ (TELUGU)
- sura-sura batti — ‘sparklers’ (HINDI)

# EXPRESSIVE

- This involves terms that are not sounding symbolic in the strict sense.
- Onomatopoeias and expressives are largely limited to positive utterances.
- Examples of expressive from Hindi, Telugu and from Bengali
- Chi~ chi~ - 'dirty, filthy' (Telugu)
- chip chip - 'sticky' (Hindi)
- Rama rama - 'expressing disgust' (Telugu)
- pil pil - 'sense of being overcrowded'

(Bengali)

# PAIRED WORDS

- These involve the juxtaposition of two lexical items and are variously referred to in the literature as ‘synonymic compounds’, ‘synonymic repetition’, ‘semantic reduplication’ and ‘redundant compounds’.
- Both parts belong to the same semantic field, and may stand in several possible relations to one another.
- kuti makkal - ‘subjects children’ i.e. ‘citizens’  
(Tamil)
- taay takappan - ‘mother father’ i.e. ‘parents’  
(Tamil)
- pilla jalla - ‘children’ (Telugu)
- dhan daulat - ‘wealth wealth’ i.e. ‘wealth, riches’ (Hindi)



# ECHO

- Echo expressions, form yet another category of reduplicated forms.
- In terms of their semantic structure, they fit between the onomatopoeias and expressives, in which neither part is independently meaningful
- Echo expressions with idiosyncratic phonology, and possible Semantics, seem to be a common feature of the Indian subcontinent such formations are highly productive

- uppu cappu - 'taste' (Telugu)
- aisa vaisa - 'this way that way' (Hindi)
- idhar udhar - 'here or there' (Hindi)
- ikkada akkada - 'here and there' (Telugu)
- kaappi kiippi - 'coffee and other beverages'

(Tamil)

- This is certainly a marginal phenomenon, but represents a potential development from the Echo expressions, with the two halves beginning to diverge.
- Echo expression merely intensifies or emphasizes the meaning of the base word.
- The echo words are said to 'encode the speaker's affective state.'

# Previously proposed extensions are

- 1. Syllable Element <syllable>
- 2. Loan-Word Element <alien>
- 3 Dialect Element <dialect>

# PROPOSED EXTENSIONS IN THIS WORKSHOP ARE

## 1. Echo element <echo>

Echo is in widespread use in the spoken language, at all levels of society.

```
<?xml version="1.0"?>
```

```
<speak version="1.0" xml:lang="tel-in">
```

```
<emphasis>Welcome</emphasis> to the echo sample tag.
```

```
  <s> neku leda <echo>illu gillu </echo></s>
```

```
</speak>
```

## 2.EXPRESSIVES :

- `<?xml version="1.0"?>`
- `<speak version="1.0" xml:lang="hindi">`
- 
- `<s> <say as> dhan dualat </say as> pa ne ke liye <say as>`  
`din raat </say as> kaam karna hai </s>`
- 
- `</speak>`

### 3. ONOMOTOPEIA

- `<? xml version="1.0"?>`
- `<speak version="1.0" xml:lang="tel-in">`
- `<p> pillalu andaru deepawali rojuna <say-as> </say-as>sura-sura bati`
- `<say-as> taara zuvva <say-as </say-as> kaalcharu</p>`
- `</speak>`

# CONCLUSIONS

- In this paper, we discussed the position of Bhrigus Inc. Hyd in building the TTS systems for Indian languages
- We described the nature of Indian Languages Keeping in view of the issues with the characteristics of these languages, we discussed the importance of **Paired words** and **Echo elements** and their behavior to improve the quality of TTS system in the context of Indian languages and other similar languages in the Asia-Pacific.

THANK YOU