

Role of Syllabification for Intonation Model of Assamese Language

Parismita Sarma¹, Shikhar Kr. Sarma²

Dept of IT, Gauhati University, Assam^{1,2}

Abstract: Analysis of speech is important for identification of emotional state of person which has different parameters like pitch, formant frequencies, MFCC etc. These parameters are used for extraction of features from speech. For unit selection based text-to-speech system, syllabification acts as a backbone. Based on different structures of different languages syllabification rules also vary. The main aim of this study is to examine and analyze the syllabification rules for Assamese language. Here we are designing the syllabic structure to form the corresponding syllables of Assamese words which later can be integrated into a text-to-speech system.

Keywords: MFCC, formant frequency, pitch.

I. INTRODUCTION

Syllabification is actually phonological building block of word formation. Syllables generally forms from a central vowel (a nucleus) and two margins at beginning and end may be consonant. Syllable can affect rhythm, prosody and stress pattern of any sentence.

There are different types of syllabic structures for syllabification process of any language. A single syllable word is called monosyllabic, with two syllables is called disyllabic etc. The syllables are separated by a hyphen and use a period when transcribing in IPA chart. Due to many reasons, like insufficient relation between letters and sounds syllabification is mostly based on morphological rules, not on phonetic rules.

Intonation Annotation System

A broad study of pitch for speech signal is important to get naturalness of speech synthesizer. Intonation model provide us this facility. It helps in phonetic study of speech signal. Generally speaking intonation can be defined as combination of acoustic parameters like tonal and tonal features. Fundamental frequency is the primary factor, based on which intonation model can be formed. Based on f0 (fundamental frequency) different types of intonation annotation model can be formed[3]. The term prosody is also used to mean dynamic and temporal features of speech signal. The term intonation is used interchangeably with the term prosody, which involves on their description temporal and dynamic features. Any intonation event which convey facts to the listener about structure, carry some practical information. On the other hand intonation model integrates data about dynamic features. This model allows us to study relationship between....

- (1) An annotation system which includes dynamical contours.
- (2) The precise alignment between tones and contours with the phonetic chain.

Syllabification issues

Syllabification is a procedure that identifies different syllables in a word or sentence. Syllable is the smallest speech sounds which are distinguishable from one another.

Basically syllabifications are done manually yet some algorithm can be designed to identify syllables for a particular language. It is seen one or two algorithm is not enough for syllabification; as it depends on the language. On the other hand syllabification rules are the basic backbone for any task related to text-to-speech conversion system [2]

Methodology for Syllabification

Methodology for our study is

1. Study Assamese language syllabic structure from literature.
2. Observed variation of dialect.
3. Previous work study.

Syllable structure

We have studied the syllabification process of the Bodo language and the Assamese language. For Bodo language it is already there , we studied for Assamese language based on which the following syllable structure of Assamese phonemes of vowels (as V) and consonants (as C) can be listed[1]-

1. V
2. VC
3. CV
4. VV
5. CVC
6. CVV
7. CCVC
8. CCV
9. VCC
10. VCVC
11. VVC

Syllabification rules

The syllabification rules we can group as follows:- [1] Generally it is observed that vowel gives the boundary of syllable in a word. When there are two vowels (not together and not at first) then after first vowel syllable boundary we get. So the rules can be listed as follows:

1. Single vowel and CV structured words have syllable at end. Eg. পৰি, মহী, তাই

- ii. For a word having CVCC*V* structure, syllable boundary is marked after first vowel. i.e. CV/C, CV/CVV, CV/CV, CV/CVC eg. গইনা, etc.
- iii. When a word starts with a vowel but following a consonant then after that again vowel then a consonant , it means having VCV* structure, then syllable boundary is marked after SECOND vowel. Because it is seen that only first vowel cannot give a syllable boundary. i.e. VCV/, VCV/C, VCV/CV, VCV/CVC eg. অকাই পকাই
- iv. When there are two conjugative vowels that is CVV* structure, syllable boundary is marked after SECOND vowel. i.e. CVV/, CVV/C, CVV/CV, CVV/CVC, CVV/V eg. etc.
- v. For a word having VCC* structure, then syllable boundary is marked after second CONSONANT. i.e. VCC/VCV, VCC/V, VCC/VC etc.
- vi. If there is only one vowel that is at the end like CCV structure no syllabic boundary will be there. After completion of the vowel only we get a syllable.
- vii. For a word having CCVC* structure, then syllable boundary is marked after first vowel. i.e. CCV/C, CCV/CV, CCV/CVC, CCV/CVV etc.
- viii. For a word is having CVCV structure, then syllable boundary is marked after first vowel. I.e. CV/CV, CV/CVV etc.
- ix. For a word having CVVCX structure, then CHECK X

- A) If x=vowel (V) i.e. CVVCV, boundary will found after second vowel
- B) If x=consonant(C) i.e. CVVCC, boundary will found between the two vowels.

Apart from these set of rules there are still some obligations which we have not defined like if a word consists of two vowels consecutively then the syllable structure will be (VV).

II. EXPERIMENT

As our work is based on Assamese language, an algorithm can be designed for syllabification of Assamese language. It is seen that location of a syllable plays an important rule, as they exhibit various behaviour in the form of articulation features and pronunciation whether it appears at beginning, at middle or end. In this work we are examining different patterns of syllables (eg. Mono syllabic, disyllabic, tri syllabic etc.). Syllables are considered as the set of smallest speech sound in a language that distinguishes one word from another. The key objective of this study is to set the 45 syllabification rules for Assamese language using an algorithm which syllabify a word. In this research work the syllabification is actually achieved by looking at different patterns of syllables means that we can experiment with monosyllabic, disyllabic, tri syllabic and polysyllabic words. Thus, the purpose of this work is to formulate an algorithm for dividing Assamese words into syllables.

Analysis of Intonational Model

In this paper we are analyzing, how the same sentences can be expressed in different tones to convey different meanings. We have considered normal tone, anger and

surprise tones e.g. the sentence “**কি কৈছা**”(Ki koisa) can be expressed in all of the three tones(normal, anger and question) even if it is the same sentence. If conveyed with a higher pitch and varied energy it will be sounded as a surprise or anger in the emotion compared to a normal voice. And according to the study performed we observed that the emotional difference in the voice is measured by fundamental frequency, duration, rhythm and different aspects of voice quality. In our work we divided the male and female wav files into distinctively 2500 wav files. The same text was given to both the speakers but the emotional tone e.g. The tone in Anger, Normal and Surprise etc. were comparatively different in both the cases. In terms of energy between the male and female voice a huge difference can be observed especially in case of anger mood. In respect to its own sound such as the male voice shows a significant change in its energy while speaking the same sentence in all the three modes – anger, surprise, normal. We find a huge difference in anger and surprise tone in comparison to a normal tone when we considered the pitch. In the figure we are showing energy, pitch and normal speech wave form for a male speaker with anger, normal and surprise tone. Figure 1 shows it for male and Figure 2 shows the same for female.

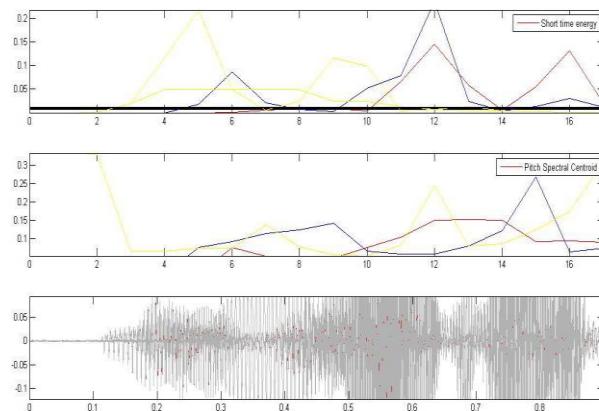


Fig1: Histogram representation of energy, pitch and the recorded text of a male speaker with anger, normal and surprise tone of an already recorded speech signal.

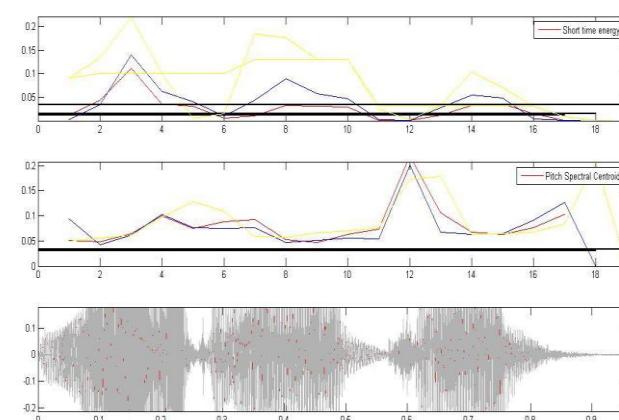


Fig.2 Histogram representation of energy, pitch and the recorded text of a female speaker with anger, normal and surprise tone.

III. RESULTS AND DISCUSSION

The objective of this work is to give a very basic introduction to the work for developing of intonation/annotation model for Assamese Language with respect to MLDS data modelling. The model is developed in natural speaking mood. Naturalness and other quality assessment will be the main consideration.

This paper discusses the importance of intonation model for Assamese language. Intonation model shows the variation in a sentence when the mood of the speaker is different. This is important while in case of designing a TTS system since every sentence has a meaning of its own also every single (common) sentence has a different meaning when it is spoken in different mood or so called emotions. Hence, it is noticeable that for the TTS to produce the sound or speech with almost exactly the same meaning in which the text is given i.e. it should be able to detect the exact attitude of the speaker while conveying the output speech from the input text.

IV. CONCLUSION

For synthesizing speech we need to have the root level syllables which further with the help of tools like Festival will be able to implement. The syllables will must help in letter to sound (LTS) rule or Grapheme to Phoneme (G2P) conversion. So this is our humble try to get an efficient algorithm for syllabification of Assamese language such that it can further help the whole speech synthesis process.

REFERENCES

- [1] Labo Kr. Thakuria, Prof. P.H. Talukdar, "Automatic Syllabification Rules for Assamese Language".
- [2] Jyotismita Talukdar, Chandan Sarma, Prof.P.H Talukdar, "Automatic Syllabification Rules for Bodo Language".
- [3] Peri Bhaskararao,"Salient phonetic features of Indian languages in speech technology".