Morphological Analyzer for Sambalpuri Odia Dialect Inflected Verbal Forms

Dhabal Prasad Sethi
Lecturer, Department of Computer Science &Engineering,
Government College of Engineering, Keonjhar, Odisha, India

Abstract: Morphological analyzer plays an important role in development of natural language software. In India and foreign countries many research work have done about morphology. They are based on standard language. But developing morphological analyzer for dialectal language is a challenging task. A non-standardized form of varieties of language spoken by group of people is called dialect .Sambalpuri language is a dialect of Odia language but their grammar, pronunciations is different from standard language. In this article I have developed the morphological analyzer for Sambalpuri Odia dialect inflected verb forms. Since research work have done at different standard languages. There should be required to develop morphological analyzer and generator for dialectal language.

Keywords: morphology, dialect, morphological analysis, morphological analyzer, stemmer, suffix stripping

---

I.   INTRODUCTION

Many people would have the experience of sitting in a public place and making conversion with other people or making phone call also. Those people, who are communicating in their own dialect, feel comfortable rather than standard language. If anyone goes research, he will find it is true that everyone speaks the language with his home district people/regional friends which are slightly different from their standard language. We can say that they are using the language which is non- standardized form of varieties of language called vernacular dialect. Vernacular language is used by local or native people for common communication which differs with official language of a state. The vernacular dialect is identified based on regional geographical background, social status, personal traits and where the people are living, what are their occupations, where they go to school and who are their friends. It concludes that we could not avoid language differences in a society which is composed of varieties of social group called dialect. There are two types of dialects e.g. a)standard dialect (supported by institutions) b)non-standard dialect(not supported by institution).The standard dialect means govt. recognized prescribed forms of a language that are published in grammar books, dictionaries and text books which are correct to spoken and written. A non-standard dialect, like a standard dialect has a complete vocabulary, grammar and syntax, but is not supported by govt. /institutions. In France and Spain, language academy have been established by their government and these academies are responsible for determining which words are to be consider as standard word and which are not. But in United States there have not the language academy but the people are speaking the correct English. To compare the characteristics of language and dialect, we may say that dialect is a small circle inside a big circle of a language. There are many similarities and dissimilarities between dialects and language.

Morphological analyzer plays an important role in development of natural language software for any language. Morphological analyzer takes the word as input and it gives/provides the output as its root word, suffix along with other information. There have been developed the morphological analyzer for many Indian and foreign languages. They are based on standard language .But little attempt have taken for developing morphological analyzer for dialectal level of Indian language. So it gives a chance to develop morphological analyzer for Odia dialect. Applications of morphological analyzer are grammar checker, parser, machine translation, spell checker, and question answering system, Speech Recognition and others.

This paper is structured as follows: Section II describes its related work. Section III explains Indian languages and morphology. A describes Odia language and morphology. B explains Odia dialect and morphology. B1 explains Sambalpuri Odia dialect. B2 explains Sambalpuri Odia dialect inflected verbal forms. C explains comparison of Sambalpuri dialect and standard Odia languages. D describes suffix stripping algorithm. E explains experiment and result. F describes conclusion.

II.   RELATED WORK

Little research work has done for dialectal level of morphological analyzer and generator in India .It is a good platform for researchers to go research about dialectal languages of India. Morphological analyzer and disambiguation for dialectal Arabic by Nizar Habash,Ryan Roth, Owen Rambow, Ramy Eskander and Nadi Tomesh[2] from center for computational learning systems, Columbia University. They have followed the MADA approach to develop Arabic morphological analyzer and disambiguation. Except that the morphological analyzer and generator are developed for Indian and foreign language that are standardized form.

III.   INDIAN LANGUAGES AND MORPHOLOGY

Indian languages are highly agglutinative [1]. There are 1652 different languages in India according to census 1961.Some of the languages are accepted our constitution where as others are accepted as dialect. The Indian constitution has recognized 22 languages. But there are nearly have 33 different languages, 200dialects approximately. The Indian languages are belong to four language families e.g. Indo-European, Dravidian, Austro-Asiatic (Austric) and Sino-Tibetan. Majority of Indian population belongs to two language families e.g. Indo-European and Dravidian. The Indo-European family is spoken in northern and central regions whereas Dravidian family is spoken in southern region of India. Since linguistic is the scientific study of a language, the various branches of linguistic analysis are morphology, phonology, pragmatics, syntax, sociolinguistics etc.

Unlike English and other foreign languages, the Indian languages are characterized by rich system of inflections (BIBHAKTI), derivation and compound formation. The numbers of words are derived from the root word by some specific orthographic rules in Indian language [1]. Morphological analysis is formulated by Indian linguistic Panini, who formulates 3,959 rules of Sanskrit morphology. Morpheme is the smallest meaningful unit in the grammar of a language. Example; the word "boys" consist of two morphemes e.g.' boy' and 's' a plural marker on noun. Different definitions of morphologies are:

Morphology deals with study of words, their internal structure and rules by which words are formed. It refers to identification of a word stem from a full form word. Morphology [1] means the word formation, analysis and generation. Morphological analysis is the process of analyzing the morphology. Or it is the process of analyzing word into their root word and associated morphemes. Morphological analyzers [1] are programs to provide morphological analysis of a language. Or morphological analyzer is a program for analyzing the morphology of an input word. It works recognize words, lexicon and algorithm to identify stem in given input word and identify the suffix. When it takes an input word, it provides related grammatical information such as lexical category, gender, number, person and tense. Morphology is categories into two classes a) Inflectional morphology b) Derivational morphology.

Inflectional morphology is the study of the words which when inflected does not change the meaning of that word and the formation of new word come from the existing stem. **Derivational morphology** is the study of the words which when inflected changes the meaning of that words and the new word formation is derived from the existing stem.

A.ODIA LANGUAGE AND MORPHOLOGY

Like Indian languages Odia language is also characterized by rich system of inflections, derivations and compound formation. The morphological analyzer for Odia language is developed by Sanghamitra Mohanty, Prabhat Kumar Santi, and K.P Das Adhikari from PG Dept of computer sc. and Application, Utkal University, Odisha. They have mentioned three types [2] of morphologies are present in Odia language e.g. 1) Pronoun morphology 2) Inflectional morphology 3) Derivational morphology and they described as: Pronoun morphology [2] is the study of grammatical classification of pronoun. Example, TUMBHEMANE (you) indicates that it is personal pronoun, 2nd person and plural number. The inflectional morphology [2] is the combination of root word with grammatical morphemes, usually resulting in a word of same class as original stem and filling some syntactic function like agreement. In other words, morpheme, which conveys 'number', 'person, 'inflection'', 'KARAKA' etc is called inflectional morphology. example MANE for making plural form(number),third (person),subject(karaka) and 1st inflection etc ,in case of nominal word. Similarly, from the inflectional morphemes like UTHANTI, we obtain information as present tense, 1st person and singular number in case of verbal word. The derivational morphology [2] is the combination of a word stem (root word) with grammatical morphemes, usually resulting in a word of different class. Example a verbal word HASHA (LAUGH) can take the derivational suffix AEIBA to produce a nominal word as HASEIBA (making laugh).

B.ODIA DIALECT AND MORPHOLOGY

 Dialectal Odia (DO) means the day to day use of spoken words of Odisha people. The proper/standard Odia is mainly spoken in urban areas. But, if you go to the village area you must find most of the spoken words which are little variation from standard Odia. So the variation of language is we called dialect. The standard Odia orthography is based on various dialect of Odia. The word orthography means correct to write. The dialectal words are lexically, phonologically, morphologically different from standard Odia. Some dialectal words are little different from standard Odia and some dialectal words are similar to standard Odia. Since dialectal Odia words have no standard orthography, it may be written in different ways to reflect written rules e.g. phonology, etymologically.

In Odia language there are seven dialects are there. They are Baleswari Odia, Ganjami Odia, Sambalpuri Odia, Midnapuri Odia, Singhabhumi Odia, Desiya Odia, and Bhatri Odia. Like standard Odia, dialectal Odia is also morphologically rich system of inflection, derivation and compound formation. The number of words is also inflecting, derivate from stem word. Therefore dialectal Odia languages have pronoun morphology, inflectional morphology and derivational morphology. Since morphological analyzers are developed for standard languages, there should be developing morphological analyzer for dialectal language.

B1.SAMBALPURI ODIA DIALECT

Sambalpuri is a beautiful dialect of Odia language. It is spoken in the districts of Sambalpur, Balangir, Baragarh, Sonepur, Kalahandi, Sundergarh, Boudh, Deogarh, Nuapada and Jharsuguda. Sambalpuri language is also called as Kosali language. There are many debates going on about Sambalpuri dialect. Sambalpuri Odia people are demanded to be become a separate state named KOSALA. They had made strike number of times for that. There have not much

research about their accent and phonetic of Sambalpuri language. So Sambalpuri language is called as dialect of Odia. From linguistically there are two speech forms in Odisha e.g. Odia/Kataki Odia and Sambalpuri/Kosali. Odia/Kataki Odia is spoken in coastal districts and Sambalpuri/Kosali Odia is spoken in north-western districts of Odisha. Sambalpuri was considered to as a dialect of Odia, but it has its own grammar and pronunciations.

B2.SAMBALPURI ODIA DIALECT INFLECTED VERBAL FORMS

From grammatical point of view, Inflection is the process of modify the word which gives different grammatical category e.g. tense, mood, aspect, person, gender, number, voice and case. Inflection can affect verb, noun, adjective, pronoun words. The inflected form of verb is called conjugation whereas the inflected form of noun, pronoun, and adjective are called as declension. The conjugation gives tense, mood, voice or aspect where declension gives number, case, and gender information. Nouns [20] are inflected in plural forms, verbs are inflected in various tenses and adjectives are inflected in comparative/superlative.

Sambalpuri Odia is morphologically rich language. A number of morphemes added with verbal root to form different grammatical function. The term verb means the word which gives information about object or person. Or the word which gives information about work is called verb. Verb is two types a) finite verb b) infinite verbs. Finite verbs: Those verbs have subject/noun and usually works as main verb in a sentence/clause are called finite verb. Finite verbs are called conjugated verb which gives person, number, tense, aspect and voice. Examples of finite verbs in English are present simple: I go, present continuous: I am going, present perfect: I have gone perfect progressive: I have been going. Whereas those verbs which cannot work as main verb and has no subject, tense, number are called non-finite verb. Examples of non-finite verbs are: he loves climbing the hill. In the sentence climbing is non-finite verbs because it is used as noun and do not work as a main verb. In this paper I discussed the finite verbal forms of Sambalpuri inflected verbs. In finite verbal forms verbal root, agreement (Agr) is compulsory while aspect (Asp), auxiliaries (Aux), Modal are optional.

Agreement morphemes mean marking person, number. **Person** means first person, second person, third person. **Number** means singular or plural. The verb form that indicates state is complete; repetition of an action or duration is called **Aspect.** Aspects are two types e.g. perfect (perfective) and another is progressive. These two aspects are combined to form perfect progressive. Perfect aspect describes events happened in past but link with present. Progressive aspect takes place a limited amount of time period. The progressive aspect is made up using-ing form in English language. **Tense** means the action of time of verb. In modern English tense is categories into two types' e.g. present tense and past tense but future tense are called as future time.

In Sambalpuri Odia dialect inflected verbal forms different categories of morphemes are:

1) Aspect (Asp) morphemes: –u for progressive, **-**i for perfect.
2) Auxiliary(Aux)morphemes: –chh/achh for present, **-**th for past and future
3) Tense morphemes: –es/s**-** for present simple, –el/l**-** for past simple **-**il-for past progressive, past perfect ,past perfect progressive, **-**em/eb or **-**m/b for future simple **-**im/ib**-** for future progressive, future perfect, future perfect progressive.

From the analysis of tense morpheme, we find there is no specific/common tense marker in Sambalpuri Odia dialect .But in Standard Odia there is specific tense marker e.g. –il **-**for past ,**-**ib**-** for future .

Here some examples

<div align="center">

Muin-khaa-es-i

I khaa-Asp-Aux- $Tense_{present}$-$Agr_{1st\ sg}$

I eat.

Muin-khaa-u-chh-e

I khaa-Asp_{prog}-Aux_{chh}- Tense-$Agr_{1st\ sg}$

I am eating.

Muin khaa-i-chh-e

I khaa-$Asp_{perfect}$-Aux_{chh} -Tense-$Agr_{1st\ sg}$

I have eaten.

Muin khaa-iasu-achh-e

I khaa-$Asp_{perfect\ progressive}$-$Aux_{achh}$-Tense-$Agr_{1st\ sg}$

I have been eating.

Muin khaa-el-i

I khaa-Asp- Aux- $Tense_{past}$ - $Agr_{1st\ sg}$

I ate.

Muin khaa-u-th-il-i

I khaa-$Asp_{progressive}$ -Aux_{th} - $Tense_{past}$ -$Agr_{1st\ sg}$

I was eating.

Muin khaa-i-th-il-i

I khaa-$Asp_{perfect}$ - Aux_{th} -$Tense_{past}$ -$Agr_{1st\ sg}$

I had eaten.

Muin khaa-iasu-th-il-i

I khaa-$Asp_{perfect\ progressive}$ Aux_{th}-$Tense_{past}$-$Agr_{1st\ sg}$

I had been eating.

</div>

Muin khaa-em-i

I khaa-Asp- Aux -Tense$_{future\ time}$ -Agr$_{1st\ sg}$

I shall eat.

Muin khaa-u-th-im-i

I khaa-Asp$_{progressive}$- Aux$_{th}$ -Tense$_{future\ time}$ -Agr$_{1st\ sg}$

I shall be eating.

Muin khaa-i-th-im-i

I khaa-Asp$_{perfect}$ -Aux$_{th}$ - Tense$_{future\ time}$ -Agr$_{1st\ sg}$

I shall have eaten.

Muin kha-iasu-th-im-i

I khaa-Asp$_{perfect\ progressive}$-Aux$_{th}$-Tense$_{future\ time}$-Agr$_{1st\ sg}$

$_I$I shall have been eating.

IV.   COMPARISION OF SAMBALPURI ODIA DIALECT AND STANDARD ODIA LANGUAGE

Due to some phonetic, syntactic, morphological difference with standard Odia, now-a-days some Odia TV channels are telecasting Sambalpuri dialectal news. Because dialectal people feel comfortable in their own language. In the table1 verbal forms of Sambalpuri words are mentioned. If we compare with standard Odia verbal words present simple, past simple, future simple forms of Sambalpuri dialect are different in pronunciation, syntax and other forms are little difference. For noun and pronoun, words are also different in pronunciation and syntax.

V.   SUFFIX STRIPPING ALGORITHM

The suffix stripping algorithm is a method used in morphological analyzer to remove the suffix and provide root/stem word along with other information by comparing the all possible suffixes with inflected word and using morpheme sequencing rules. Steps of suffix stripping algorithm for Sambalpuri odia dialect inflected verbal words are

Step1: Enter the dialectal inflected verbal word to be analyzed

Step2: If the inflected word is matched with one of the possible suffixes then

Step3: Remove the suffix and get the root/stem and provide other information like tense, aspect, number, and person.

VI.   EXPERIMENT AND RESULT

To develop the morphological analyzer data collection is important. To collect data I have applied three steps: Firstly orally interviewed with 40number of students from western districts like Sambalpur, Baragarh, Balangir, and Jharsuguda. Secondly whatever verbal inflection they are pronounced I told them to write. 3rdly I have found some research paper about Sambalpuri dialect and the inflected verbal form from blog sport in internet. I find whatever students written that is matched with collecting data from internet. In the table1 shows tense chart of Sambalpuri dialect. Again verbal suffix are clearly mentioned in table2.My work is to develop morphological analyzer of Sambalpuri dialect. I used the inflected verb as input when the inflected verb is matched with inflected suffix using Odia Unicode, then automatically remove the suffix and get the root word and the same time provide other information. I have used java program, Utkal fonts and KeyMan software for Odia and it gives maximum output correct.

Diagram1 shows odisha map which contains dialectal districts.

Present tense		Present simple	Present progressive	Present perfect	Present perfect progressive
1ˢᵗ person ପ୍ରଥମ ପୁରୁଷ	singular(ଏକବଚନ)	ମୁଁ ଖାଏସି	ମୁଁ ଖାଉଛେଁ	ମୁଁ ଖାଇଛେଁ	ମୁଁ ଖାଇଥାଅଛେଁ
	plural (ବହୁବଚନ)	ଆମେ ଖାଏସୁଁ	ଆମେ ଖାଉଛୁଁ	ଆମେ ଖାଇଛୁଁ	ଆମେ ଖାଇଥାଯୁଛୁଁ
2ⁿᵈ person ଦ୍ୱିତୀୟ ପୁରୁଷ	singular(ଏକବଚନ)	ତୁ ଖା / ତୁମେ ଖାଅ	ତୁ ଖାଉଛୁ / ତୁମେ ଖାଉଛ	ତୁ ଖାଇଛୁ / ତୁମେ ଖାଇଲନ	ତୁ ଖାଇଥାଯୁଛୁ / ତୁମେ ଖାଇଥାଯୁଛ
	plural(ବହୁବଚନ)	ତୁମେମାନେ ଖାଅ	ତୁମେମାନେ ଖାଉଛ	ତୁମେମାନେ ଖାଇଲନ	ତୁମେମାନେ ଖାଇଥାଯୁଛ
3ʳᵈ person ତୃତୀୟ ପୁରୁଷ	singular(ଏକବଚନ)	ସେ/ହେ ଖାଏସି/ଖାଏସନ	ସେ/ହେ ଖାଉଛେ/ଖାଉଛନ୍	ସେ/ହେ ଖାଇଛେ	ସେ/ହେ ଖାଇଥାଯୁଛେ
	plural(ବହୁବଚନ)	ହେମାନେ ଖାଏସନ୍	ହେମାନେ ଖାଉଛିନ୍	ହେମାନେ ଖାଇଛିନ୍	ହେମାନେ ଖାଇଥାଯୁଛନ୍

Paste tense		Past simple	Past progressive	Past perfect	Past perfect progressive
1ˢᵗ person ପ୍ରଥମ ପୁରୁଷ	singular(ଏକବଚନ)	ମୁଁ ଖାଏଲି	ମୁଁ ଖାଉଥିଲ	ମୁଁ ଖାଇଥିଲି	ମୁଁ ଖାଇଥାଯୁଥିଲି
	plural(ବହୁବଚନ)	ଆମେ ଖାଏଲୁ	ଆମେ ଖାଉଥିଲୁ	ଆମେ ଖାଇଥିଲୁ	ଆମେ ଖାଇଥାଯୁଥିଲୁ
2ⁿᵈ person ଦ୍ୱିତୀୟ ପୁରୁଷ	singular(ଏକବଚନ)	ତୁ ଖାଏଲୁ / ତୁମେ ଖାଏଲ	ତୁ ଖାଉଥିଲୁ / ତୁମେ ଖାଉଥିଲ	ତୁ ଖାଇଥିଲୁ / ତୁମେ ଖାଇଥିଲ	ତୁ ଖାଇଥାଯୁଥିଲୁ / ତୁମେ ଖାଇଥାଯୁଥିଲ
	plural(ବହୁବଚନ)	ତୁମେମାନେ ଖାଏଲ	ତୁମେମାନେ ଖାଉଥିଲ	ତୁମେମାନେ ଖାଇଥିଲ	ତୁମେମାନେ ଖାଇଥାଯୁଥିଲ
3ʳᵈ person ତୃତୀୟ ପୁରୁଷ	singular(ଏକବଚନ)	ହେ/ସେ ଖାଏଲା	ହେ ଖାଉଥିଲା	ହେ ଖାଇଥିଲା	ହେ/ସେ ଖାଇଥାଯୁଥିଲା
	plural(ବହୁବଚନ)	ହେମାନେ ଖାଏଲେ	ହେମାନେ ଖାଉଥିଲେ	ହେମାନେ ଖାଇଥିଲେ	ହେମାନେ ଖାଇଥାଯୁଥିଲେ

Future time		Future Time1	Future Time2	Future Time3	Future Time4
1ˢᵗ person ପ୍ରଥମ ପୁରୁଷ	singular(ଏକବଚନ)	ମୁଁ ଖାଏମି/ଖାଇମି	ମୁଁ ଖାଉଥିମି	ମୁଁ ଖାଇଥିମି	ମୁଁ ଖାଇଥାଯୁଥିମି
	plural(ବହୁବଚନ)	ଆମେ ଖାଏମା/ଖାଏମୁଁ	ଆମେ ଖାଉଥିମୁଁ	ଆମେ ଖାଇଥିମୁଁ	ଆମେ ଖାଇଥାଯୁଥିମୁ
2ⁿᵈ person ଦ୍ୱିତୀୟ ପୁରୁଷ	singular(ଏକବଚନ)	ତୁ ଖାଏବୁ / ତମେ /ତୁମେ ଖାଏବ	ତୁ ଖାଉଥିବୁ / ତମେ/ତୁମେ ଖାଉଥିବ	ତୁ ଖାଇଥିବୁ / ତମେ /ତୁମେ ଖାଇଥିବ	ତୁ ଖାଇଥାଯୁଥିବୁ / ତମେ /ତୁମେ ଖାଇଥାଯୁଥିବ
	plural(ବହୁବଚନ)	ତୁମେମାନେ ଖାଏବ	ତୁମେମାନେ ଖାଉଥିବ	ତୁମେମାନେ ଖାଇଥିବ	ତୁମେମାନେ ଖାଇଥାଯୁଥିବ
3ʳᵈ person ତୃତୀୟ ପୁରୁଷ	singular(ଏକବଚନ)	ସେ/ହେ ଖାଏବା	ସେ/ହେ ଖାଉଥିବା	ସେ/ହେ ଖାଇଥିବା	ସେ/ହେ ଖାଇଥାଯୁଥିବା
	plural(ବହୁବଚନ)	ହେମାନେ/ସେମାନେ ଖାଏବେ	ହେମାନେ/ସେମାନେ ଖାଉଥିବେ	ହେମାନେ/ସେମାନେ ଖାଇଥିବେ	ହେମାନେ/ସେମାନେ ଖାଇଥାଯୁଥିବେ

Table1 shows the verbal tense chart of Sambalpuri dialect which contains tense, aspect, Agreement e.g. number, person. The verbal inflected suffixes in table1 are again specified clearly in table2.

Tense	person	Singular suffix	Plural suffix
Present	1ˢᵗ person	ଏସି,ଉଛେଁ,ଇଛେଁ,ଇଥାଯୁଅଛେଁ	ଏସୁଁ,ଉଛୁଁ,ଇଛୁଁ,ଇଥାଯୁଛୁଁ
	2ⁿᵈ person	ଉଛୁ,ଉଛ,ଇଛୁ,ଏଇଲନ,ଇଥାଯୁଛୁ, ଇଥାଯୁଛ	ଅ,ଉଛ,ଏଇଲନ,ଇଥାଯୁଛ
	3ʳᵈ person	ଏସି,ଏସନ,ଉଛେ,ଉଛନ,ଇଛେ,ଇଥାଯୁଛେ	ଏସନ,ଉଛନ,ଇଛନ,ଇଥାଯୁଛନ
past	1ˢᵗ person	ଏଲି,ଉଥିଲି,ଇଥିଲି,ଇଥାଯୁଥିଲି	ଏଲୁ,ଉଥିଲୁ,ଇଥିଲୁ,ଇଥାଯୁଥିଲୁ
	2ⁿᵈ person	ଏଲ,ଏଲୁ,ଉଥିଲ,ଉଥିଲୁ,ଇଥିଲ,ଇଥିଲୁ, ,ଇଥାଯୁଥିଲୁ,ଇଥାଯୁଥିଲ	ଏଲ,ଉଥିଲ,ଇଥିଲ,ଇଥାଯୁଥିଲ
	3ʳᵈ person	ଏଲା,ଉଥିଲା,ଇଥିଲା,ଇଥାଯୁଥିଲା	ଏଲେ,ଇଥିଲେ,ଉଥିଲେ,ଇଥାଯୁଥିଲେ
future	1ˢᵗ person	ଏମି,ଇମି,ଉଥିମି,ଇଥିମି,ଇଥାଯୁଥିମି	ଏମା,ଉଥିମୁଁ,ଏମୁଁ,ଇଥିମୁଁ,ଇଥାଯୁଥିମୁ
	2ⁿᵈ person	ଏବୁ,ଉଥିବୁ,ଇଥିବୁ,ଏବ,ଉଥିବ,ଇଥିବ,ଇଥାଯୁଥିବ,ଇଥାଯୁଥିବୁ	ଏବ,ଉଥିବ,ଇଥିବ,ଇଥାଯୁଥିବ
	3ʳᵈ person	ଏବା,ଉଥିବା,ଇଥିବା,ଇଥାଯୁଥିବା	ଏବେ,ଉଥିବେ,ଇଥିବେ,ଇଥାଯୁଥିବେ

Table2 contains the Sambalpuri odia dialect verbal inflected suffixes.
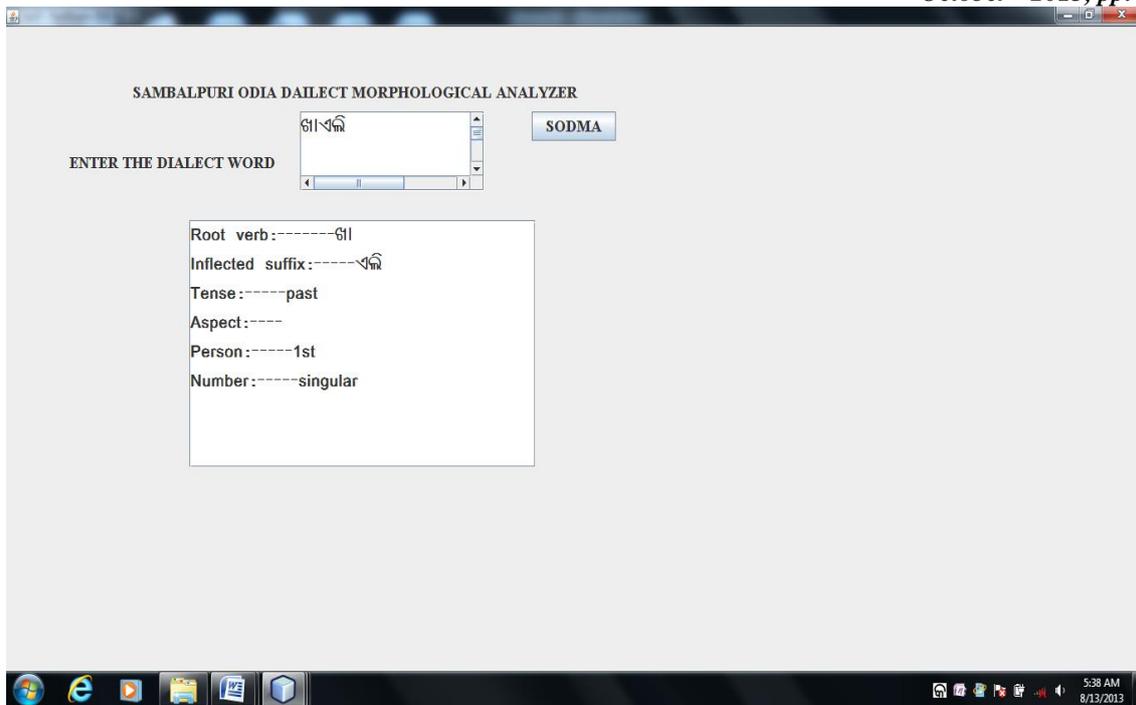
VII.   CONCLUSION

In this paper I explained the dialectal verbal inflected forms of Sambalpuri Odia. I have developed the Sambalpuri Odia dialect morphological analyzer which will gives root word, inflected suffix and other information like standard language morphological analyzer. Therefore the morphological analyzer for different dialects of Indian languages should be developed which will help the researchers to develop new software of dialectal language.

REFERENCES

[1]   Sanskrit.jnu.ac.in/rstudent/mphil/muktanand/chapter1

[2]   Morphological analyzer and disambiguation for dialectal Arabic by Nizar Habash,Ryan Roth,Owen Rambow,Ramy Eskander and Nadi Tomesh from center for computational learning systems,Columbia University.preceeding of NAACL-HLT2013,page426-432,Atlanta,Georgia,9-14june2013@2013Association for computational linguistics

[3]   Analysis and design of oriya morphological analyzer:some test with orient by sanghamitra mohanty,pabitra Kumar Santi,K.P das adhikari,PG Dept of computer science and application,Utkal university,bbsr

[4]   Politics of Sambalpuri or Kosali as a dialect of oriya in Orissa by tuna biswal,mphil at language of india volume10:11november 2010

[5]   http://kosalidentity.blogsport.in/2011/07/peculiarities-of-sambalpuri language-in.html

[6]   Finite state morphological processing of oriya verbal forms by kalyani R. sabadi,resource center for Indian language technology solutions,department of management studies,Indian institute of science,banglore,india.

[7]   A suffix stripping algorithm for odia stemmer by Sampa Chaupatnaik, Sohag sunder Nanada, Sanghamitra mohanty,PG dept of computer science and application, Utkal university, international journal of computational linguistics and natural language processing.

[8]   Developing morphological analyzers for four Indian languages using a rule based affix stripping approach by Mona Parakh, Rajesha N, Linguistic Data Consortium for Indian Languages, CIIL, Mysore, Language in India www.languageinindia.com 1 1 : 5 M a y 2 0 1 1 Special Volume: Problems of Parsing in Indian Languages

[9]   Suffix Stripping Based NER in Assamese forLocation Names by Padmaja Sharma, Utpal Sharma, Jugal Kalita

[10]   A suffix stripping based morph analyzer for Malayalam language by Rajeev R,R1, Rajendran N2 and Elizabeth Sherly.

[11]   Telemore: morphological generator for Telgu nouns and verbs by madhavi Ganapathiraju and lori Levin,language technologies institute, carneie mellon university,Pittsburgh,PA15213,USA

[12]   Development of hindi lemmatizer by snigdha paul,nisheeth josi,Iti mathur,Apaji Institute ,banasthal university,rajastan,india.

[13]   Morphology:Indian languages and European languages by shweta vikram,computer science,Banasthali University,international journal of scientific research publications,volume3,issue6,june2013

[14]   Computational morphology and natural language parsing for Indian languages:a literature survey by Antony p J,research scholar,dr k.p soman,computational engineering and networking(CEN),research center,amrita vishwa vidyapeetham university,Coimbatore india.

[15]   Verbal Inflection in Hindi: A Distributed Morphology Approach by Smrti Singh, Vaijayanthi Sarma Department of Humanities and Social Sciences, IIT Bombay, Powai, Mumbai-400076, India, 25th Pacific Asia Conference on Language, Information and Computation, pages 283–292

[16]   Morphological Analysis of Inflecting Compound Words in Bangla Sajib Dasgupta, Naira Khan, Asif Iqbal Sarkar, Dewan Shahriar Hossain Paveland Mumit Khan BRAC University, Dhaka, Bangladesh

[17]   Morphological Analyzer for Affix Stacking Languages:A Case Study of Marathi Raj Dabre1 Archana Amberkar1 Pushpak Bhattacharyya1 Indian Institute of Technology Bombay, Mumbai-400076, India Proceedings of COLING 2012: Posters, pages 225–234,COLING 2012, Mumbai, December 2012

[18]   Morphological Analysis of Verbs in Nepali Santa Bdr. Basnet, Shailesh Bdr. Pandey

[19]   Research Report on Bangla Verb and noun Morphological Analysis Md. Zahurul Islam Center for Research on Bangla Language Processing, BRAC University

[20]   www.esl.fls.edu

SAMBALPURI ODIA DAILECT MORPHOLOGICAL ANALYZER

ENTER THE DIALECT WORD

SODMA

Root verb:------ଗା
Inflected suffix:----- ଲ୍ଲି
Tense:-----past
Aspect:----
Person:-----1st
Number:-----singular

(Snapshot of Sambalpuri odia dialect morphological analyzer)

Author Profile:

Dhabal Prasad Sethi is currently working as a lecturer in Computer Science & Engineering at Government College of Engineering, Keonjhar, Odisha. He has completed his BE CSE from Bhadrak Institute of Engineering & Technology, Bhadrak, Odisha in 2006.He has completed his ME CSE with Specialization in Knowledge Engineering from Post Graduate Department of Computer Science & Application, Utkal University, Odisha in 2011.His research area of interest is Natural Language Processing and Information Retrieval. Now he has been researching about dialectal languages of Odia.