AUTOMATIC ANALYSIS OF ODIA INFLECTIONAL MORPHOLOGY: A RULE BASED APPROACH

Dissertation submitted to Jawaharlal Nehru University
in partial fulfillment of the requirements
for the award of the Degree of

MASTER OF PHILOSOPHY

Submitted by

SONALI MAHANTA

Under the supervision of

Prof. GIRISH NATH JHA

CENTRE FOR LINGUISTICS
SCHOOL OF LANGUAGE, LITERATURE & CULTURE STUDIES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI, 110067, INDIA

2016

DECLARATION

Dated 22nd July 2016

This dissertation entitled "Automatic Analysis of Odia Inflectional Morphology: A Rule Based Approach" submitted by me for the award of the degree of Master of Philosophy, is an original work and has not been submitted so far in part or in full, for any other degree or diploma of any University or Institute.

Ms.Sonali Mahanta
M.phil Student
Centre for Linguistics
SLL&CS
JNU

Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University
New Delhi-110067, India

Dated: 22nd July 2016

CERTIFICATE

This dissertation entitled "**Automatic Analysis of Odia Inflectional Morphology: A Rule Based Approach**" submitted by **Ms. Sonali Mahanta** to the Centre for Linguistics, School of Language, Literature and Culture Studies, Jawaharlal Nehru University, New Delhi, for the award of the degree of **Master of Philosophy** is an original work and has not been submitted so far in part or in full, for any other degree or diploma of any University or Institution.

This may be placed before the examiners for evaluation for the award of the degree of Master of Philosophy.

(PROF. PKS PANDEY)
CHAIRPERSON

Chairperson
Centre for Linguistics
School of Language, Literature & Culture Studies
Jawaharlal Nehru University, New Delhi - 110067

(PROF. GIRISH NATH JHA)
SUPERVISOR

SUPERVISOR
Dr. Girish Nath Jha
Professor
Center for Linguistics
School of Language Literature & Culture Studies
Jawaharlal Nehru University, New Delhi - 110067

Acknowledgement

I would like to extend my gratitude to *Prof. Girish Nath Jha,* my supervisor and guide for his constant help and encouragement in completing my research work. I am a novice in the field of Computational linguistics but the mentorship of Prof. Jha has not only helped me to smoothly surpass the hurdles but also has made the journey of research interesting. I want to express my sincere thankfulness to *Prof. Pramod Pandey* for providing a substantial knowledge of phonetics, Phonology and generative phonology which is of enormous importance for constructing a sound linguistic background. I am indebted to *Prof. Franson Manjali* for showing the relativity of philosophy with linguistics. I would like to thank *Prof. Pradeep kumar Das* for introducing me to the field of typological classification of languages. My heartiest appreciation and thankfulness goes to *Prof. Ayesha Kidwai* for her precious lectures on generative syntax. I would like to thank *Prof. Vaishna Narang* for sharing the knowledge of Applied Linguistics, Acoustics and Pedagogical theories. I also thank *Asst. Prof. Hari Madhab Ray* for introducing me to the concept of Morphology and morphological processes which turned out to be very helpful for my research work. Most of all I would like to thank *Prof. Anvita Abbi* for her contribution in inducing an interest for Linguistics in me.

I am thankful to *Prof. Panchanan Mohanty,* Centre for Applied Linguistics and Translation Studies, University of Hyderabad for clearing my confusions related to Odia grammar while doing the research. I would like to thank *Mr. Abhishek Sharma* for helping me in writing the java code in developing the Morph Analyzer and tackling the computational problem posed by the system. I acknowledge the help of *Mr. Atul kumar Ojha* and *Dr. Abhijit Dixit* for their linguistic and computational help. I would like to thank Mr. *Pitamber Behera* for clearing my queries related to the Odia POS tagger which plays a very important role in the analysis done by the developed Morph Analyzer. Finally I would like to express my deep gratefulness to all my class mates and friends for their help and support.

CONTENTS

LIST OF ABBREVIATION USED

ABL	Ablative
ACC	Accusative
CL	Computational Linguistics
DAT	Dative
GEN	Genitive
ILCI	Indian Languages Corpora Initiative
IMPF	Imperfect
INS	Instrumental
LOC	Locative
MA	Morph Analyzer
MT	Machine Translation
NLP	Natural language processing
NOM	Nominative
PL	Plural
PNG	Person, Number, Gender
POS	Part of speech
PRF	Perfect
PST	Past
SG	Singular
TAM	Tense, Aspect, Mood
VOC	Vocative

LIST OF TRANSLITERATION KEY USED

Odia	IPA	Odia	IPA
ଅ	ɔ	ଙ	ŋ
ଆ	ɑ	ଚ	c
ଇ	i	ଛ	c^h
ଈ	i:	ଜ	ɖʒ
ଉ	u	ଝ	$ɖʒ^h$
ଊ	u:	ଞ	ɲ
ଋ	ru:	ଟ	ʈ
ଏ	e	ଠ	$ʈ^h$
ଐ	ɔi	ଡ	ɖ
ଓ	o	ଢ	$ɖ^h$
ଔ	ou	ଣ	ɳ
କ	k	ତ	t̪
ଖ	k^h	ଥ	$t̪^h$
ଗ	g	ଦ	d̪
ଘ	g^h	ଧ	$d̪^h$

ନ	n
ପ	p
ଫ	p^h
ବ	b
ଭ	b^h
ମ	m
ଯ	ʤ
ର	r
ଳ	ɭ
ହ	h
କ୍ଷ	kʃ
ଯ୍ଞ	j
ଲ	l
ଶ	ʃ
ଷ	ʃ
ସ	s

INTRODUCTION

This M.Phil dissertation is an attempt for developing an online tool for Morphological analysis of Odia inflectional Morphology. A Morphlological Analyzer is a computational program for providing grammatical information of a lexical category. The developed Morph Analyzer produces the Morphological analysis of nominal and verbal inflections of Odia language and conveys the information regarding the Morpho-syntactic properties of the words it analyses.

Aims and Objectives behind developing the system

I. So far no such online tool is available for public platform which can provide Morphological analysis with the relevant features of Odia part of speech categories. Most of the previous works have been publications only. The claimed applications in these publications have been largely missing. So this system will be helpful for the analysis of such linguistic features for nouns and verbs.

II. This dissertation has tried to enlist all the available inflectional forms of nouns and verbs. The inflectional particles are collected both from standard written and spoken forms with the help of available written grammar works of eminent Grammarians of Odia language in both the diachronic and synchronic phases.

III. This Morph Analyzer has used an inbuilt Odia POS tagger as an interface .The tagger has been developed as a part of an M.Phil dissertation (2015) submitted to Jawaharlal Nehru University, New Delhi by Mr. Pitamber Behera under the guidance of Prof. Girish Nath Jha.This tagger has 95% of accuracy rate. The use of the Tagger can significantly enhance the performance of the Morph Analyzer.

IV. One of the major motivations for working on this topic is that in the present scenario Odia may have been announced as the sixth classical language due to its marvelous literary and linguistic history but it lacks in the e-resourcefulness segment. There are many well analysed manual grammar books available but the section of a complete and apt online availability of linguistic analysis at various levels like phonetics, Morphology,

syntax and semantics is still missing. So the system developed through this M.Phil dissertation will contribute towards making Odia a more digitally enable language and a substantial premise for further NLP applications.

Chapterization:

This dissertation consists of five chapters as enlisted below:

I. The survey of the literature of Morph Analyzers

The first chapter gives an overview related to the existing approaches for developing Morph Analyzers. This chapter first discusses the various methods and techniques for developing Morph Analyzers. It also summarises the research work done in the field of computational analysis of Morphology done outside India and in India in various languages by the application of various methods and approaches. It also talks about the available online links, tools for different languages. The major focus of the literary survey is on the detail account of works done in Odia in the field of Morphological Analyzer and its current scenario.

II. Research Methodology

The second chapter discusses the methodology for the current M.phil research in developing the Morph Analyzer. This chapter accounts for the collection of the linguistic data from primary as well as secondary resources for preparing the hand crafted dictionary with the file of nominal and verbal inflections. This chapter will also discuss about the formulation of generalized rules. As this dissertation is about a rule based technique for automatic analysis of Odia, so a detail description of the rules for inflection formation has been done. This chapter also discusses the programming aspects of the system development.

III. Odia Language and Morphological Processes

The third chapter initiates with a brief description of the language and its linguistic varieties. This chapter further does a comprehensive discussion of the Morphological processes for nominal declension and verbal conjugation with ample examples. This chapter also gives an overview of how the rules operate for marking the grammatical features in the inflected speech categories of noun and verb.

IV. Odia Inflectional Morph Analyzer

This chapter is divided into two sections. The first section deals with the design, development and scope of the proposed Odia Morph Analyzer. In the series of chapters this chapter discusses the implementation aspect of the algorithm discussed in the second chapter. This chapter will also give the system description and how it functions. The second elemental composition of this chapter also includes validating the morphanalysed data and then after putting forth the issues, challenges and the process of disambiguation of the morphanalysed data.

V.Conclusion

The fifth chapter is the concluding chapter. This chapter does an overall evaluation of the result and output generated by the system. This chapter discusses its strength, weaknesses future directions and extension of the work.

CHAPTER 1

MORPH ANALYZER AND A SURVEY OF THE AVAILABLE LITERATURE

This chapter is serving as a description of Morph Analyzer as a linguistic analysis tool in computational grammar and also enlists the various approaches and techniques available for building it. Major focus of this chapter is on the analytical overview of the various works done to develop Morphological Analyzers in foreign languages and Indian languages with special emphasis on Odia language till date.

1.1 Morphological Analyzer

A Morph Analyzer provides the analysis of the Morphology of the input word with grammatical information as reflected from its affixes attached to the word. As an application of NLP this type of system is very much helpful for information retrieval and machine translation. The development of a Morph Analyzer comes after the POS tagging which tag the input data into various grammatical categories such as noun, verb, adverb, adjective and many more. The Morph Analyzer goes a step beyond and gives the dissection of the various word forms and supplies the proper categorical features. This Morph Analyzer is dealing with inflectional features only.

1.1.1 Approaches and methods for developing Morphological Analyzer

There are various methods available in computational grammar for developing a Morphological Analyzer. Few of them are enlisted below:

1.1.1.1 Two-Level Morphology Based Approach

The two level approach for developing a computational model for recognition and generation of Morphology was developed by a Finnish computer scientist named Kimmo Koskenniemi in 1983. As the name reflects the two level Morphology based approach has two levels. Out of the two levels one is lexical level and the other one is surface level. Surface level encompasses a linguistic form which represents the actual written representation or the actual utterance where as the lexical level displays the various kind of information related to lemma form of word. The lexical level foregrounds the concatenated Morphemes to form the word. The linking of Morphemes in the lexical

level is determined by a particular pattern which is again language specific. Moreover the two level morphology is about establishing a correspondence between the lexical to surface level through grammatical features.[1]

As for instance if the surface level representation of an inflected Odia word is pʰulɔguḍikɔ (flowers) then its lexical level representation will be Noun + Pl, where /pʰulɔ/ is the noun and /-guḍikɔ/ is the plural marker.

pʰulɔguḍikɔ → /pʰulɔ/+ /-guḍikɔ/
flower.N.Pl→ flower.N.Sg + Pl marker
flowers

1.1.1.2 Corpus Based Approach

As the name suggests corpus based approach is statistical based where the machine is required to be trained with a large sized corpus. An algorithm is written to train the machine. The training of machine is done through a large set of corpora and thereafter retrieving the statistical information and other features from the available corpus. The performance of the system is directly proportional to the features and size of the corpus. The more is the size of the corpus the more will be its accurate output. This approach which is based on corpus bears a trait of being less favorable as it requires a plenty of time. However this approach is getting used profitably for developing language softwares.[2]

1.1.1.3 Suffix Stripping Algorithm

Suffix Stripping Algorithm works independent of a lookup table that consists of root and their inflected forms. This process uses a stem dictionary which helps in identifying a valid stem. The composition of the dictionary has all the possible affixes that noun or verb have in a particular language, the Morph tactics rule as well as the Morphophonemic rules or *sandhi* rules. This form of Morphological analysis is an economical one. The process for obtaining the stem of the whole word is initiated with the suffix identification

[1]http://ijarcet.org/wp-content/uploads/IJARCET-VOL-3-ISSUE-3-623-625.pdf

[2] Ibid.

and then after by the application of proper *sandhi* rules (J.P.Jayan,R.R. Rajiv & S. Rajendran,2009).

1.1.1.4 Finite State Automata Based Approach

This approach is based on the technique of acceptance or rejection of a string. It uses regular expressions as its input. The acceptance of a string is dependent on the condition whether it reaches the final state of FSA or not. If it reaches the final state the string is said to be accepted else said to be rejected. The system deals with the regular expressions are powerful tools for text searching. The regular expression is a text material of a definite structure. FSA can be utilised for the recognition of the Morphological lexicon.[3]

1.1.1.5 Paradigm Based Approach

This approach revolves around the representation of a paradigm of a stem. The paradigm of a stem is provided with all the possible forms in addition to the related pattern of features. The appropriacy of this approach is more for languages with frequent inflections. The whole idea of this approach is vested on a table. The table contains the possible root forms and the generated word forms with the affixes. The paradigm approach is programmed for generating, adding, and deleting the string while morphologically analyzing an input. The Morphological behaviour is the determining factor in this process of Morphological analysis of various word classes. The subsequent step deals with the classification of the part of speech categories into specific paradigms depending on their Morphophonemic affiliations. The dominance of the paradigm approach can easily be figured out in the Indian languages.[4]

1.1.1.6 Rule based approach

Rule-based approach revolves around the substantial knowledge of linguistic rules. This approach has got a favorability factor for a language which is reach in resource as well as the less resourced one. This approach is more prevalent than the corpus based approach. As the corpus based requires an enormous amount of data. At the same time the rule

[3] Shaji.A & Sindhu, P.(2014).*Morphological Analyzer for Malayalam: A Literature Survey*

[4]Ibid.

based approach operates through a hand crafted dictionary which is formed on the basis of the available linguistic repertory of grammar of a language.[5]

1.1.1.7 DAWG (Directed Acylic Word Graph)

DAWG does have different applications, which can be efficiently used for lexicon representation along with string matching. This structure can store finite strings in a compact way; also it can take the advantage of common affixes in the string. This method has been successfully implemented for Greek language by University of Partas Greece. Non-deterministic DAWG data structure can be used for both morphological analysis and generation and if we only need one function then it is better to use deterministic DAWG with attached to the single language with a better response time. This method is effective for Indian Languages too. This approach does not utilize any morphological rules or any other special linguistic data.[6]

The above paragraphs have enlisted various kinds of approaches available for developing Morph Analyzers. The forth coming sections of this chapter are dealing with the survey of the works done in different languages for developing Morph Analyzer adopting different approaches. If we will start surveying the available literature in the current scenario then the prominence of multitude of online links can be noticed. This is definitely a substantial and successful part of computational evolution. The following section on review of literature will first discuss the works done outside India and then with in India. Subsequently the discussion will have a focus on the works done in Odia language.

1.2 Works done outside India

1.2.1 Morphological Analyzers for multiple languages

Open Xerox is an online tool which gives Morphological analysis of languages like Czech, English, French, German, Spanish, Hungarian, Italian, Polish and Russian. It has

[5] Akilan, R., & Naganathan, E.R. (2015).Morphological Analyzer for classical Tamil text: a rule based approach .

[6] Shanji,A.& Sindhu,P.(2014) .Morphological Analyzer for Malyalam:A Literature Survey

beendeveloped using finite state technology.[7]It uses the submitted text and provides the output with English glosses.

Masato Hagiwara and Satoshi Sekine(2009) at Rakuten Institute of Technology, New York has developed a Lightweight Client-Side Chinese/Japanese Morphological Analyzer. It is available online. It is based on an online Learning and employs an online learning algorithm SCW(Soft Confidence-Weighted), which enables client-side model update and domain adaptation. This has claimed of achieving a compact model size of 5MB while maintaining the state-of-the-art performance, via techniques such as feature hashing, FOBOS, and feature quantization.[8]

Majka is a free online Morph Analyzer which has databases for Czech, English, Portuguese, Catalan, Welsh, Spanish,Galician, Slovak, Polish, Swedish, German, French, Italian Asturian and Russian. This Morphological Analyzer accepts one entry (word, lemma, or string lemma tag, according the data file in use) per line on its standard input and prints the requested information on its standard output. It is free downloadable for windows and Linux platform.[9] Rakuten MA is a Morphological Analyzer which is a combination of word segmenter and POS Tagger for Chinese and Japanese.[10]

1.2.2 Morphological Analyzers for English

NLP.C# is an online parser. This captures and displays the morphological process of any English word produced through inflection or derivation. It also gives an analytical output of any word produced playfully by someone.There are various free softwares available for the English morphological analysis like:

- HFST - Helsinki Finite-State Transducer Technology - FST library, command line tools, hfst-twolc (a rule compiler for two-level rules), and several spellers and Morphological Analyzers (GPL)

[7]https://open.xerox.com/Services/fst-nlp-tools/Consume/Morphological%20Analysis-176
[8]http://anthology.aclweb.org/C/C14/C14-2009
[9]https://nlp.fi.muni.cz/polish-Morphology-Analyzer/
[10]https://github.com/rakuten-nlp/rakutenma

- FOMA - finite-state toolkit (similar to Xerox XFST), created and maintained by Måns Huldén (GPL)
- Lttoolbox -- lexical processing tools for building Morphological Analyzers/Generators with XML specification files. Includes data for English (both analysis and disambiguation). (GPL)
- PC-KIMMO - a Two-level Processor for Morphological Analysis, including KGEN, KTEXT, and Englex
- SFST - Stuttgart Finite State Transducer Tools (GPL)
- MULTEXT mMorph - (unmaintained) two-level Morphology, package includes some data for English and German, (GPL2 or later).[11]

1.2.3 Morphological Analyzer for Chinese

A Chinese Morphological Analyzer with Character-level POS Tagging has been developed by Mo Shen, Hongxiao Liu, Daisuke Kawahara, and Sadao Kurohashi(2014). This tool has demonstrated that how the introduction of a character-level part of speech information can significantly improve the performance of a Morphological Analyzer.

1.2.4 Morphological Analyzer for Arabic

Kenneth R. Beesley(1995) has developed a Finite-State Morphological Analyzer and generator for Arabic language. This system analyses and generates Arabic words represented in the standard orthography, whether fully voweled, partially voweled or unvoweled. It has followed an earlier KIMMO-style two-level Morphological system and restructured extensively on the framework of Xerox Finite-State Morphology tools.[12] This Morph Analyzer analyses and produces the root, pattern and all other affixes together with feature tags. The tags are with features like number, person, tense, aspect, mood, voice etc.

Mourad Gridach and Noureddine Chenfour (2011) from the department of Mathematics and Computer Science of Sidi Mohamed Ben Abdellah University adopted an approach

[11]http://aclweb.org/aclwiki/index.php?title=Morphology_software_for_English

[12]http://www.aclweb.org/anthology/C96-1017

based on Arabic Morphological automaton (AMAUT) for developing a Morphological Analyzer and generator for Arabic language. XMODEL language is used in this technique for developing the system.[13]

MADAMIRA is another Morph Analyzer which claims to be a fast and comprehensive tool for analysis of Arabic text. It also helps in disambiguation. This system is being built up in combination of the two used systems for Arabic processing named MADA (Habash and Rambow, 2005; Habash et al., 2009; Habash et al., 2013) and AMIRA (Diab et al., 2007).Both of the systems were very common in use. The system claimed to be more efficient and economical than the previous ones. The Java platform has been described to be more robust.[14] Araflex, Sourceforge, Aramorf are few other online tools available for Morphological analysis of Arabic words.[15]

1.2.5 Morphological Analyzer for Japanese

Kuromoji is an open source Japanese Morphological Analyzer. The system has been described with multitasks like word segmentation, Part-of-speech tagging, getting dictionary forms for inflected verbs and adjectives and extract readings for Kanji.[16]

1.2.6 Morphological Analyzer for Spanish

SMM is a Morphological Analyzer for Spanish language developed by Cerstin Mahlow(2009). The constructional basis for the Spanish Morph Analyzer draws inspiration from the formalism of Left-Associative Grammar, developed by Roland Hausser. The Malaga grammars are basically used for automatic analysis of Morphology or syntax. This claims about handling all kinds of word formation processes which includes inflection, derivation, and compounding. SMM analyses single word forms as well as large corpora.[17] Span Morph is an open-source Morphological Analyzer for Spanish.[18]

[13]https://arxiv.org/abs/1101.5494?context=cs.CL
[14]http://www.lrec-conf.org/proceedings/lrec2014/pdf/593_Paper.pdf
[15]https://www.google.co.in/?ion=1&espv=2#q=+the+list+of+arabic+Morph+Analyzer
[16]https://github.com/atilika/kuromoji
[17]http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1870-90442009000100007
[18]http://www.academia.edu/1805446/SpanMorph_an_open-source_Morphological_Analyzer_for_Spanish

1.2.7 Morphological Analyzer for Russian

Lissa Vilkki(1997) has presented a paper named as "RUSTWOL:A System for Automatic Recognition of Russian words". This paper presented RUSTWOL as a tool for automatic Russian word form recognition. The theoretical foundation of this program is the two-level model. This tool processes the written standard Russian and gives the Morphological information of Russian word forms. It deals with the computational analysis of the inflected, derived word forms and also the compounds.[19]

1.3 Morphological Analyzers for Indian languages

1.3.1 Morphological Analyzer for multiple Indian languages

IIT Kanpur in collaboration with HCU, Hyderabad has developed a Morphological Analyzer for Sanskrit, Hindi, Marathi, Kannada, Telugu and Punjabi by Akshara Bharati group. The Analyzer claims for coverage of 88% and 95% coverage for Hindi and Telugu respectively.[20] It is free downloadable and available for windows and Linux platform.

1.3.2 Hindi-Marathi-Telugu Morphological Analyzers

Language technologies research centre, IIIT Hyderabad has developed an online system for Morph analysis of three Indian languages. The three languages are Hindi, Marathi, and Telugu. This Morphological Analyzer allows choosing a language as well as font. The font or the coding in which a word will be entered provides option for ISCII, RomanReadable, Itrans, Roman WX and Shusha. It is hyperlinked to web pages with other linguistic resources for Indian languages and English dictionaries.[21]

Rajesh N. and Mona Parakh (2011) have developed a Morph Analyzer for four Indian languages viz., Odia, Bengali Assamese and Bodo. It has used rule based suffix stripping approach for analysis the inflections of the respective languages.

TDIL in collaboration with the Special centre for Sanskrit studies, JNU, New Delhi has undertaken a project for developing shallow parser tool for Indian languages.The sole purpose is enriching the Indian languages technologically. This SPT-IL includes the development of Morph Analyzers (MA), Part of speech Taggers and chunker. In the

[19]http://www2.lingsoft.fi/doc/rustwol/rustwol.txt
[20]http:// ltrc.iiit.ac.in/showfile.php?filename=onlineServices/Morph/index.htm
[21]Ibid

present context it is working for eleven languages such as Assamese, Bodo, Odia, Bangala, Kashmiri, Gujrati, Manipuri, Maithili, Konkani, Nepali and Santhali.[22]

1.3.3 Morphlogical Analyzer for Hindi

IIT Kharagpur has developed a Morphological Analyzer for Hindi at Media Lab Asia Research Library. This Morph Analyzer uses a Directed Acyclical Morphological structure and identifies the tense, aspect, modality, person, gender and number along with the root of an inflected verb form.[23]

The Hindi Morphological Analyzer developed by Vishal Goyal and Gurpreet Singh (2008) used paradigm based approach in developing the Morph Analyzer. Their Hindi Morphological Analyzer stores all the commonly used word forms of all the Hindi root words in a database. This Analyzer claims to take very less time in comparison to other Morphological Analyzer based on paradigm Morphology with a very good accuracy rate. But this approach commends only for those languages in which the number of possible inflections for a word is not very high.

IIT Bombay has also developed an online tool named as beta for Morph analysis. It provides the feature gender, number, person, case, tense, aspect and mood of word classes.[24]

A FST Based Morphological Analyzer for Hindi Language is being developed by Deepak Kumar, Manjeet Singh, and Seema Shukla(2012) in the Department of Information Technology, JSS Academy of Technical Education Noida, Uttar Pradesh. This Morphological Analyzer claims to give 97% of correct result.

A Statistical Morphological Analyzer for Hindi was developed by Deepak Kumar Malladi and Prashanth Mannem (2013) in LTRC, International Institute of Information Technology Hyderabad. Nikhil Kanuparthy, Abhilash Inumella, Dipti Misra Sharma(2012) of IIIT Hyderabad presented a paper on Hindi Derivational Morphological Analyzer. Their Hindi Derivational Analyzer is an extensional development of an already existing inflectional Analyzer developed at IIIT Hyderabad.

[22]http://sanskrit.jnu.ac.in/projects/sptools.jsp?proj=sptools
[23]http://sanskrit.jnu.ac.in/rstudents/mphil/muktanand/Chapter%201.pdf
[24]http://www.cfilt.iitb.ac.in/~ankitb/lexentry/about.php

1.3.4 Morphological Analyzer for Sanskrit

A Morph Analyzer for Sanskrit is being developed by a consortium of seven universities like University of Hyderabad; Jawaharlal Nehru University; IIIT-Hyderabad; Sanskrit Academy, Hyderabad; Poornaprajna Vidyapeetha, Bangalore; Rashtriya Sanskrit Vidyapeetha, Tirupati; JRR Sanskrit University, Jaipur. The tool is available online on the site of Indian language technology and proliferation centre. This tool processes a Sanskrit word and pours out its nominal stem or verbal stem along with the linguistic features such as lexical-category, gender, number, case, person etc.. Special centre for Sanskrit, JNU, New Delhi has developed Morph Analyzer known as Subanta. This System does the analysis of inflected nouns with the help of two relational databases of examples and rules.[25]

1.3.5 Morphological Analyzer for Bangla

IIT Kharagpur in Media Research laboratory has designed and developed a Morphological Analyzer for Bengali. The Morphological Analyzer deals with verbal inflection only. The system processes the verbal inflection and identifies the tense, aspect, modality and person of an inflected verb form (Muktanand,2006:15).

1.3.6 Morphological Analyzer for Kannada

Antony P.J., M Anand Kumar and KP Soman (2010) proposed a Kannada Morph Analyzer and generator tool using a paradigm approach. This approach used trie or prefix tree as the data structure for the storage of suffixes and root words. This tool claims about handling upto 3700 root words and around 88,000 inflected forms. RCILTS, Hyderabad has also developed a Kannada Morph Analyzer and generator by using Network and process model.[26]

1.3.7 Morph Analyzer for Telugu

G. Sai Kiranmai, K. Mallika, M. Anand Kumar, V. Dhanalakshmi and K. P. Soman (2010) haveproposed a Morphological Analyzer for Telugu. This Analyzer utilises Support vector machine. This system has done claim of an accuracy rate of 94% and

[25]http://sanskrit.jnu.ac.in/Morph/analyze.jsp
[26]http://www.iitg.ernet.in/rcilts/phaseI/rc.htm

97% in case of Telugu verbs and nouns respectively.[27]A Morphological Analyzer was also developed for Telugu by RCILTS.

1.3.8 Morphological Analyzer for Tamil

RCILTS-T has developed a Morph Analyzer in Anna University named as *Atcharam*. It has got two modules for noun and verb Analyzer based on 125 rules .It handles verbal and nominal inflections. It is also used for extraction of the root words from inflections in Tamil Search Engine and Online Tamil Dictionary. It is open for expansion and can add more rules to increase the dictionary size. It claims about providing approximately 75% result.[28]Anusaraka group of researchers under the guidance of Rajendran(2006) in Tamil University prepared a Morphological Analyzer for translating Tamil into Hindi at word level. AUKBC Morphological parser for Tamil prepared a Morphological parser for Tamil under the supervision of Rajendran.It uses the finite state machinery like PC Kimmo.[29]

M Anand Kumar, V. Dhanalakshmi, K. P.Soman, and S.Rajendran (2010) havepresented a paper to propose a novel approach to solve the Morphological Analyzer problem using machine learning methodology. The approach has adopted a sequence labeling and training by kernel methods that capture the non-linear relationships of the Morphological features. This has been done by training the data samples.

1.3.9 Morphological Analyzer for Malyalam

V. Pa Abeera, Aparna, Sa, Rekha, R. Ua, M Kumar, A., Dhanalakshmi, Va, Soman, K. Pa, and Rajendran, Sb. (2012) have proposed a Morphological Analyzer for Malayalam.It has been developed by using machine learning approach.The result shows that the system is very effective and after learning it predicts correct grammatical features even the entries which are not in the training set.

Jancy Joseph, Dr. Babu Anto(2015) has proposed an Analyzer through a paper named "Rule Based Morphological Analyzer for Malayalam Nouns: Computational Analysis of

[27]https://www.amrita.edu/publication/Morphological-Analyzer-telugu-using-support-vector-machine

[28]http://sanskrit.jnu.ac.in/rstudents/mphil/muktanand/Chapter%201.pdf

[29]https://www.researchgate.net/publication/258665032_Parsing_in_Tamil_Present_State_of_Art

Malayalam Linguistics". The system has a medium of input and output that is done in Malayalam without doing any transliteration and retransliteration process.

1.3.10 Morphological Analyzer for Punjabi

Punjabi Morphological Analyzer has been developed at Punjabi University, Patiala, India.[30]Another online link for Morphological Analyzer and generator is available on http://punjabi.aglsoft.com/punjabi/?show=Morph.

1.3.11 Morphological Analyzer for Assamese and Manipuri

RCILTS and IIT Guwahati have developed Morphological Analyzers for Assamese and Manipuri. Both the Morphological Analyzers use the technique of stemming. Stemming deals with the deletion or addition of affixes to arrive at the root words.[31]

1.4 Works done in Odia

Sanghamitra Mohanty, Prabhat Kumar Santi and K.P.Das Adikary (2004) presented a paper named as "Analysis and Design of Odia Morphological Analyzer: Some Tests with OriNet" in the Symposium on Indian Morphology, Phonology & Language Engineering in IIT Kharagpur. The paper mentioned about three type of Morphology such as pronoun Morphology, inflectional Morphology and derivational Morphology.The system designing is based on object oriented approach.[32] The architectural design of Odia Morphological Analyzer (OMA) has been described with composition of five parts such as Ori Net data base(OD) which stores the odia lexicon(root words),OMA Engine(OE) which processes the system, Morphological Parser(MP) which parses the word according to orthographic rule, Decision Tree(DT) which decides to classify the Morphemes. The system development is based on the syntactic approach of Sanskrit language.

Kalyanmalini sahoo (2004) had proposed a model based on two level processing for Odia in the symposium on Indian Morphology, Phonology & Language Engineering held at IIT Kharagpur. The proposed system was based on the finite state transducer model. This paper talked about the noun Morphology of the language. The nominal form of Odia are narrated with few major affixing categories like numeral (Nmrl), classifier (Cl/Clas),

[30]http://www.learnpunjabi.org/websitepics/Splash.jpg
[31]http://www.iitg.ernet.in/rcilts/phaseI/Morph.html
[32]http://www.ciil-ebooks.net/html/simple/total.pdf

quantifier(Quan), number marker (Nmb), qualitative affirmative marker(Q.Aff), negation marker(Neg), Case marker and postpositions (PP) etc in the paper.

Itisree Jena, Sriram Chaudhury, Himani Chaudhry, Dipti M. Sharma(2011) had proposed a Morphological Analyzer for Odia language through a paper named as "Developing Oriya Morphological Analyzer Using Lt-Toolbox". It had followed the paradigm approach. The paper talked about various paradigms of nouns, adjectives, indeclinable (avyaya) and finite verbs. The paradigms were created using an XML (extensible markup language) based Morphological dictionary from the Lt-toolbox package.[33]

Dhabal Prasad Sethi (2014) had presented a research paper called "Analyzer for sambalpuri Odia dialect inflected verbal forms". The paper dealt with the Morphological analysis of inflected verbal forms.Sambalpuri language is one of the linguistic varieties of Odia language spoken in the western part of Odisha.

Kalyani R. Shabadi proposed a model for the Morphological processing of verbal forms in Odia using a deterministic finite state automation.This paper has talked about the computational analysis of the verbal forms appear with information of lexical, Morphological and syntactic features(Sethi,2014:623).

R.C. Balabantray, M.K. Jena and S. Mohanty (2014) presented a paper entitled "ShallowMorphology based complex predicates extraction in Odia". This Paper describes about the extraction of the complex predicates from the sentences possessing the lexicon pattern {[MMM](n/adj)[NNN](v)}.

Kalyanamalini Sahoo (2016) has published an article named as "multi-Verb Constructions - Parsing with a Deterministic Finite State Automaton (DFA)". The paper has proposed about the parsing of multi-verb construction (V-V) sequences, and the case of passivation with three verb sequences.The the paper thus proposed is focusing on the discussion of co-occurrence restrictions of main verb as well as the light verb in verbal forms.

[33]http://link.springer.com/chapter/10.1007%2F978-3-642-19403-0_20

CHAPTER 2

RESEARCH METHODOLOGY

This chapter encapsulates the methodology for this research in order to create the Odia Morph Analyzer for inflectional Morphology. Since this Morph Analyzer has a theoretical underpinning of the rule based approach, so the necessary stages for creating a rulebased system has been discussed in detail through this chapter. The chapter talks about the data collection of the linguistic resources from varied sources.This chapter also programs an algorithm for recollecting the desired output from the submitted data to the system with all possible Morphosyntactic features proper to nominal declension and verbal conjugation in the context of Odia inflectional Morphology.

2.1 Role of Morph Analyzer in Computational Morphology

Morph Analyzer plays a vital role in NLP for the analysis and processing of the Morphology of input words and gives the grammatical categories of the root and the prefixes/suffixes of each word. Morphological analysis follows the stemming as the next step during processing. The analysis and generation of word forms through computational means is known as computational Morphology.[34] The tool or program that does the Morphological analysis of a language is called Morph Analyzer. Morph Analyzer computationaly processes a word to extract out the root form in addition to the other features such as gender, number, and tense of the word. Morph Analyzer includes a recognition engine, lexicon and an algorithm to find out the stem within an input word and identify the affixes. Thus the information extracted by it is made use of by the later following layers of processing.

This Morph Analyzer in the present context is dealing with the analysis of inflectional Morphology of verbs and nouns. Few sample of the analytical output provided by a Morph Analyzer for Odia in relation to inflections is as follows:

i.gʰɔrɔ-ku →gʰɔrɔ+ku

To home home.SG.LOC

[34]Jha,Girish N.(2007),Introduction to Computational Morphology, *Lecture in PGDMLT course*

ii.ʈɑ-ku → ʈɑ+ ku

to you you.SG.LOC

iii.kʰɑucʰi →kʰɑ+ u + cʰi

eating eat.PRS.PROG

2.2 Manual analysis of the behaviour of the Odia nominal and verbal inflections

Every language has its own specific ways of forming inflections. Among the various kinds of affixations, Odia majorly uses the suffixation as a process of inflection formation. Since the Morphological Analyzer is adopting a rule baesd approach, the first step is the creation of a hand crafted dictionary. The dictionary is manually created and fed in to the system. This dictionary will serve as a database for the Morphological processing of Odia inflections.In order to build a dictionary the identification of all inflectional suffixes are done in accordence with the characteristics of the inflectional lexical categories. For this research the pattern of noun and verb inflection is being noticed.Then the forms are analysed in accordence to their appearance with the root words in variuos environments. This will help in the derivation and formulation of certain generalised rules pertaining to their inflectional Morphology. The ultimate goal behind the done analysis is to patternise the word formation process. So the manual work involves both the analysisand insertion of the rules into the data files.

2.3 Rule based analysis of Morphology

System which operates on rule based approach is dictionary dependent. In the rule based approach when certain text is given as an input to the Morphological Analyzer, the system first searches, if the corresponding Morpheme is present in the database for generating the output. If the searched Morpheme is present the system produces the relevant features of the corresponding Morpheme. But if the Morpheme is missing in the dictionary, then the rule based system suffers. In a rule based system each rule is dependent on the previous rule. So for this reason if one rule fails the entire set of subsequent rules poses invalidity and might produce error in result. The inflectional particles in addition to the rules of word formation are entered into the dictionary. This

dictionary serves as a database for the processing of the lexical entries by the developed tool in the next stage of computation.

2.4 Dictionary of nominal and verbal inflections

Once the manual calculations is done for determining the pattern of occurence of the inflectional particles, they are inserted into the datafile.

The dictionary of nominal inflections have a series of seven columns. The columns are arranged in a sequence of captions like nominal Inflection, surface form ofinflection, POS (part of speech), meaning, example in words, exception and condition.

The first column has the entries of all possible inflectional markers for the existing cases in odia both in singular and plural form. The gathering and grouping of the markers for various case relationships hail from both standard spoken and written Odia. The cases entered into the dictionary are nominative, accusative, dative, ablative, genitive, instrumental, locative and the vocative particles. Each inflectional particle entered into the first column has representative marker for both singular and plural forms exhibiting various case relationships. The second column is containing the surface form of the markers. This entails regarding the actual appearance of the markers in written form.

The third column posesses the part of speech category. The POS for the inflected nominal particle is marked as N. The fourth column demonstrates the examples in words containing the respective inflectional markers in reference to the first column. The fifth column has enlisted the exceptions. The exception column has inserted all the environments where the particle resides apart from its usual occurance.

The list of verbal inflections is also handcrafted in the data file. This sheet has got eleven columns.The first column has inflectional forms of verbs,second column represents the surface representation of the inflectional form, third column has given example in words of the related verbal inflection,fourth colum has given the POS category, the fifth and sixth talks about the person (1^{st},2^{nd} and 3^{rd}), number(Sg and Pl) respectivel, seventh column talks about the tense forms (present, past and future). Eighth is about the aspects of tense like simple,continuous, perfect and habitual etc. Nineth column is about the honorificity/degree of respect pertaining to ontological properties like intimacy,

familiarity and formality for the second person singular form and honorific/non honorific for the third person singular from. The second last or the tenth column has given a list of the moods. The moods in Odia are indicative mood,subjuntive mood, imparative mood, optative mood. The last or eleventh column has included the verbal feature called as voice.

2.5 Linguistic Resources

In the word of Biber et al (1998:246) Corpus is not only an accumulation of texts but a corpus seeks to represent a language or some portion of it.[35] The corpus that is used for this purpose of study is the ILCI corpora of Odia. It is being developed by special centre of Sanskrit Studies, JNU in collaboration with Ravenshaw University and University of Hyderabad. ILCI (Indian languages corpora initiative) is a project sponsored by Technology Development for Indian Languages (TDIL) program of Ministry of Communication and Information Technology (MCIT). The major goal of ILCI project is to build annotated text corpora for major Indian languages including English.[36] The languages in the current phase include a total of 12 languages which include 8 from Indo-Aryan language family, that is Hindi, Urdu, Bangla, Odia, Punjabi, Gujrati, Marathi, Konkani and 3 from the Dravidian family such as Telugu, Tamil, Malayalam and English. The main objective of this project is to build an annotated parallel corpus in the domain of health and tourisms with Hindi as their source language. The encoding and annotation of the corpora is as per global standards which are currently being examined. The generated resource is supposed to feed in various MT and other LT projects in the country. The ILCI projects are based on sound research methodology for corpus collection, corpus encoding and Marking, corpus annotation, corpus storage, editing and search. The corpora are hoisted on the server with centralised code management systems.

2.6 Algorithm

The algorithm followed for developing the Odia Morph Analyzer is as follows:

When an text input is fed into the system, it processes the data through various levels.

[35]www2.gslt.hum.gu.se/~leifg/gslt/doc/rep021202.pdf

[36]Jha,Girish N.,(2010)The TDIL program and the Indian Language Corpora Initiative(ILCI).

i.The input data will first move into the pre-processing stage. First the system checks whether the submitted text is a word or a sentence.

ii.The pre-processing stage involves the following steps

a.Tokenisation/Word recognition

The recognizer analyses the input file to be a word or a text. If the entered data is a word then there is no need of tokenization and if the data is a sentence then the tokenisation occurs. Then the input will be assigned with appropriate tags through the inbuilt tagger.[37] Identification of the indeclinable and declinable form continues with each data fed into the system.

iii.When the input enters this stage it goes to the Morphological Analyzer. This is the main processing stage.

a. In this stage the words are identified according to their speech categories and get analysed. If a nominal inflection is found than the process of analysis proceeds according to the nominal rule base and if verbal inflection is found than the analysis happenend through the verbal rule base. The processing of other inflection can aslo be done in the manner by the system provided the rule bases are fed in to the system.The nominal output merges with the features like number,person and case relationship. The verbal output comes out with verbal properties like person, number, tense, aspect, mood, voice, honorificity/degree of respect.

b. After the system identifies the inflection as noun or verb it starts slicing the inflectional form and with each split it checks with the rule base for the suffix. Once it finds the form the processing stops else it continues. If the search is successful then the system produces the root word and suffix separately.

c.The system is with the rules fed into it. So when the system finds a word with an inflection maintained in the datafiles, it produces all the possible features related to the inflectional form of as provided in the rule base.

[37] The tagger has been developed by Mr. Pitamber Behera (2015) through his M.phil dissertation submitted to JNU New Delhi.

2.7 Rule for Inflection formation in Odia language

Odia has got both inflectional and agglutinative Morphological properties. So it has got the prevalence of both concatenation and synthetic processes for forming inflections. Concatenated forms are the simple arrangement of the morphemes whereas the second one marked with the synthesis of sounds. Concatenated forms appear same on the surface level whereas the synthetic forms undergo a change on their appearance. However for both the forms dominant rule for inflection formation in Odia is word+suffix.

The following examples will make the concept clear:

concatenated form

gʰɔrɔ-ʈʰu (from house)

Synthetic form

pɔɖʰe(reads)

/pɔɖʰ/is the root word where from the above form has its derivation.

CHAPTER 3

ODIA LANGUAGE AND MORPHOLOGICAL PROCESSES

This chapter can be divided to two parts. The first part is an introduction to the language with its distribution and its current status. The second portion of this chapter is a linguistic introduction to the language. This portion proceeds with an account of the various grammatical phenomenon of Odia language with special reference the word formation for inflections with special reference to noun and verbs.

3.1 Background of the language

According to Suniti Kumar Chatterjee -"Of the three speeches Oriya, Bengali and Assamese, Oriya has preserved a great many archaic features, in both grammar and pronunciation and it may be said without travesty of linguistic truth that Oriya is the eldest of the three sisters, when we consider the archaic character of the language" [Indian Historical Quarterly Vol-XXIII, 1947, P- 337]

The linguistic structure places Odia in the Indo-Aryan family of languages. The similarity of Odia to Assamese, Bengali and Maithili can be traced in terms of Morphological features. The phenomenon of linguistic assimilation has imported few traits from the neighbouring regional languages of the Indo-Aryan and Dravidian families, as also that of the Austric group of languages current among the tribal groups.Odia can be heard and observed with multiple varieties with a touch of regional features in diversified parts.The existing varieties of Odia are Baleswari (Balasore), Bhatri (Koraput), Laria (Sambalpur), Sambalpuri (Sambalpur and other western districts), Ganjami (Ganjam and Koraput), Chhatisgarhi (Chhatisgarh and adjoining areas of Orissa) and Medinipuri (Midnapur district of West Bengal). The language varieties have possibly got their derivationfrom accentual peculiarities.[38]

According to B.P. Mahapatra (1998:6) Odia language on the basis of its dialectical variation is being divided into two parts. One part is western Odisha and the other one is known as the coastal belt.The language spoken in Sundergarh, Sambalpur, Bolangir,

[38] Prusty, S.K. (2014).*Classical Language: Odia*.Odisha review.Classical status to odia. P.6.

Kalahandi and Boudh is known as Sambalpuri and all together taken as the representative of Western Odisha.

Odia hails majority of linguistic feature from both the Indo-Aryan (IA) group and has also got few features pertaining to the Dravidian languages (Pattanaik, 2004). This influence can be noticed from the Morphological typology of Odia language.

In the year 2014 Odia is being spaced in the classical language category. The other languages already included in the classical category areTamil, Sanskrit, Kannada, Telugu and Malayalam.[39]

A diachronic view of Odia grammatical forms will provide the evidence of the kind of metamorphosis it has under gone. However this dissertation aims at discussing the nominal declension and verbal conjugations in the modern context of both standard spoken and written form of Odia.

3.2 Morphological description of Odia language

The typological classification based on word formation process places odia language with inflectional as well as agglutinative features.

According to Sapir (1921:136-146) agglutinative languages does regular affixation and Morphemes can easily be dissected from each other. According to Hippisley (2011:515) Inflectional or fusional languages are the languages where the words modified without changing the category which inturn brings in a difference between its Morphosyntactic and lexicosemantic properties.

This current chapter elaborates the inflectional pattern of odia part of speech categories. Inflection is one of the processes of word formation which results in the formation of a new word form with new grammatical features but retains the previous part of speech category. The chapter is dedicated to the detailing of the Morphological features shown on the inflected categories after inflection. Odia part of speech has been described with eleven grammatical categories like noun, verb, adjective, adverb, pronoun, postposition, demonstrative, conjuction, particles, quantifiers and residual (Behera,2015:37). As

[39] http://www.thehindu.com/news/national/odia-gets-classical-language-status/article5709028.ece

mentioned earlier now the purview of discussion is concentrated on the inflectional pattern shown by various categories.

3.2.1 Noun

The inflectional Morphology of noun shows the features like number, person, gender, case and non-case relationships. As per as Odia language is concerned the nominal declension is with grammatical features like number and case. Let us discuss about all the nominal features and how do they affect the language. Odia Morphological process in terms of inflection for an inflected noun can be formulized as follows:

Word → Root/Stem + Suffix

Noun:

gɾɑm guɖikɔ → gɾɑm + guɖikɔ

village.N.PL → village.N. + Pl.Marker

Villages

3.2.1.1Gender in Odia

Masica (1991:219) narrated gender in the New Indo-Aryan languages as an inherent and classificatory property of one class of words (nouns) and a variable or inflectional property of others (adjective, certain verbal forms, sometimes of pronouns and adverbs and one extremely important postpositions)". But the exception can be found in languages like Odia and Bengali inwhich some nouns designating professions are classified as masculine or feminine without any syntactic implications. Gender in eastern Indo-Aryan languages has almost disappeared (Abbi,2001:26). Gender has both syntactic and morphological features in Indo-Aryan languages but Odia does not show any grammatical gender which means gender is restricted to Morphology and has got no expression syntactically.

According to Mohanty (2007:149) the archaic distinction of assigning odia with a three gender system (masculine, feminine and neuter) is no more valid on grammatical level. Masculine and Neuter gender unifies to lose the distinction between them and thus provide two gender forms such as masculine and feminine. Odia bears biological gender at word level which doesn't prevail over any other Morphosyntactic properties of a

sentence and therefore gender assignment of a noun doesn't control the agreement feature in the rest of a sentence.

For example:

i.bikɑʃ miʈʰɑ khɑe

bikɑʃ.3.MAS.SG sweet.N eat.PRS.SIM

Bikash eats sweet.

ii.rɑɖʰɑ miʈʰɑ kʰɑe

raɖʰɑ3.FEM.SG sweet.N eat.PRS.SIM

Radha eats sweet.

The above examples (i) and (ii) are clarifying that for both feminine and masculine gender the verb is retaining the same form. There is no gender agreement happening between subject and verb. So it can be concluded that Odia lacks the grammatical gender. However Gender is basically used for semantic purpose for making the distinction between male entities from female entities. Nouns designating professions are classified as masculine and feminine without any syntactic implication. In odia the masculine gendered forms are considered to be the fundamental one and by attaching affixes the feminine form of the nouns and adjectives are structured (P.Mohanty,2007,43). Few such affixes which are used for deriving the feminine forms from male counterparts are

/-ɑ /, /-i/, /-ɑɳi/,/-uɳi/,/-ni/,/-ɳi/,/-ikɑ,/-ri/

affix	**Masculine**	**Feminine**
/-ɑ /	sɔbʰjɔ	sɔbʰjɑ
/i/	puɽrɔ	puɽri
/-ɑɳi/	sɑdʰɔbɔ	sɑdʰɔbɑɳi
/-ɳi/	mɑsʈɔrɔ	mɑsʈrɑɳi
/-ikɑ /	bɑlɔkɔ	bɑlikɑ
/-ri/	neɽɑ	neɽri
/-ni/	biɖeʃi	biɖeʃini
/-ɑ /	ɖusʈɔ	ɖusʈɑ
/-uɳi/	ɔnɖʰɔ	ɔnɖʰuɳi

Example:

Root	+ suffix →	word
mɑsʈɔrɔ	+ -ɳi →	mɑsʈrɑɳi
teacher.N.M.SG	suffix	teacher.N.F.SG

3.2.1.2 Number system in Odia

Number is a grammatical feature for both nouns and pronouns. Number displays the numerical contrasts of both the word classes.[40] Odia grammatical system shows binary number system: singular /ekɔ bɔcɔnɔ/ and plural /bɔhu bɔcɔnɔ/.

3.2.1.2.1 Singular Number

Odia nominal singular form remains unmarked. But it can take particles like /-ʈi/, /-ʈɑ/.These are the classifier Morphemes in Odia.The use of such particles is done for categorising the nouns into various classes. According to Aikhenvald (2003) classifiers can be of various kinds such as numeral classifiers which appear in numerical expressions, locative classifiers within a locative expression,possessive classifiers in possessive constructions, noun classifiers within a noun phrase and verbal classifiers on a verb or a predicate. Classifier Morpheme constructions has a lot of semantic relativity to its referent. A classifier can be a word or affix which is used to accompany nouns.The main purpose is to"classify" the noun depending on the type of its referent.

According to B. Mohanty (1998:198) Odia the classifiers /-ʈi/ and /ʈɑ/ are the markers for definitiveness as well as intimacy. Classifier /-ʈi/ and /-ʈɑ/ can further take the particle /-e/ in order to form /-ʈie/ and /-ʈɑe/. Both /-ʈie/ and /-ʈɑe/ get amalgamated to from a particle /-ʈe/ by phonetic erosion. /-ʈe/ carries the feature of singularity but not specificity. /-ʈie/ and /-ʈɑe/ markers represent inspecificity and inintimacy where as /-ʈi/ and /-ʈɑ/ markers show specificity and intimacy.

/-ʈie/ and /-ʈɑe/ also indicate the singular number for the noun. It can be summarised that the change in form is also inducing a change in the semantic aspect for the referent.

[40]SUP(Nominal Inflection). P.159.Retrieved from: http://shodhganga.inflibnet.ac.in/bitstream/10603/17616/14/14_chapter%204.pdf

Examples:

i.gɔcʰɔ + - ʈie → gɔcʰɔ ʈie

tree.N.SG one.NUM

A tree

ii.pokɔ + -ʈɑe → pokɔ ʈɑe

germ.N.SG one.NUM

One germ

iii. cɑbi + -ʈi → cɑbiʈi

key.N.SG the.ART

The key

iv.nɑcɔ + -ʈɑ → nɑcɔʈɑ

dance.N.SG the.ART

The dance

The use of /-ʈi/ and /-ʈɑ/ has got an additional interpretation on semantic level leading to some ontological perspective. Both the particles bear the symbolic significance of love, affection and dislike. Both markers can be used with the singular form of the noun. However /-ʈi/ projects a certain degree of respect,love and affection towards the referent where as /-ʈɑ/ invokes some sort of dislike (Mahapatra,2007:59).

For example:

i.puɔ ʈi kʰeɭ- ucʰi

boy.N.M.SG the.ART play-PRS.PROG

The boy is playing.

ii.puɔ ʈɑ ʤʰɔrɔkɑkɑcɔ bʰɑŋgi delɑ

boy.N.M.SG the.ART window glass.N.SG break.PRS give.PST

The boy broke the window glass.

The differential contrast between /-ʈi/ and /-ʈɑ/ can clearly be grasped from their syntactic level representation./-ʈi/ is applied to a boy who is engaged in reading but /-ʈɑ/ is used for a person who has done some notorious deed. Another deviating feature between /-ʈi/ and /-ʈɑ/ is that /-ʈi/ is used with human nouns only where as /-ʈɑ/ can be used with non human nouns also.

For example:

i.kɔlɔmɔ -ʈɑ

pen.N.SG the.ART

The pen

Example of /-ʈi/ and /-ʈɑ/ with numerals:

i. pɑncɔ ʈi bɔhi

Five.NUM the.ART book.N.SG

Five books

ii. ɑʈʰɔ -ʈɑ ɑmbɔ

eight.NUM the.ART mango.N.SG

eight mangoes

In addition to /-ʈi/ and /-ʈɑ, /-ɖʒɔɳɔ/ is another particle which is used as a numeral classifier.

For example:

i. ɖui ɖʒɔɳɔ lokɔ

two.NUM CL person.N

two persons

ii. ʈini ɖʒɔɳɔ lokɔ

three.NUM CL person.N

three persons

Unlike /-ʈi/ or /-ʈɑ/,/-ɖʒɔɳɔ/ can never be used for non human subjects. The construction /kʰɑʈɑɖʒɔɳɔ/ is impossible and ungrammatical.

After going through a detail account related to the use of classifiers in Odia. It can be concluded that there is no singular number marker for odia in subjective case and particles like /-ʈi/,/-ʈɑ/ are getting used to classify the language in addition to some other forms. These markers are not only related to some semantic features like definiteness, intimacy, singularity but also come in association with various case relationships which can be seen in the section on case.

3.2.1.2.2 Plural Number

Plurals are formed in Odia through the following processes.

I.In Odia Suffixes of plural markers are added with the singular base form to get the plural form.

For example:

pɔʃu + mɑne → pɔʃumɑne

animal.N.SG + Pl.marker animal.N.PL

Animals

Odia plural morphology shows a clear distinction between the uses of different plural markers for animate and inanimate objects. Many Odia grammarians coexist on the opinion that noun do not always take any plural marker in case of living creatures and the plural form remains unmarked. The following example will make it clear.

Example:

Bɔŋɔ -re koili ɡɑuc^{h}nʈi

Forest.N.SG in.PREP cuckoo.N.PL sing.PRS.PROG

Cuckoos are singing in the forest.

II. The plural form used for non human is treated with singular number (Mohanty, 2007:159). The plural markers like /-e/, /mɑnɔ/,/mɑne/ are used for living creatures only. /-e/ and /mɑne/ can be placed alternatively to show the plural inflection. So they are said to be in free variation.[41]

For example:

i.Pilɑ + -e → pile /pilɑe

boy.N.SG Pl marker boy.N.MAS.PL

boys

ii.hɔnsɔ + -mɑnɔ → hɔnsɔmɑnɔ

swan.N.SG Pl marker swan.N.PL

swans

iii. ɖɑkʈɔrɔ + - mɑne → ɖɑkʈɔrɔmɑne

doctor.N.SG PL marker doctor.PL

doctors

[41]http://www.ciil-lisindia.net/oriya/oriya.html/

III. Few other plural markers are /-guɖi/, /-guɖɑ/. Both the plural markers can take particle /-e/,/-kɔ/ and produce few more plural marker forms such as:

-guɖi- + -e → guɖie
-guɖɑ- + -e → guɖɑe
-guɖi-+ -kɔ → guɖikɔ
-guɖɑ- +-kɔ → guɖɑkɔ
/-guɖie/,/-guɖɑe/ can occur before as well as after the noun.
For exapmle:
i.pɔisɑ + -guɖie / -guɖɑe → pɔisɑguɖie / pɔisɑguɖɑe
Penny.N.SG PL marker penny.N.PL (Pennies)

ii.-guɖie/-guɖɑe + Pɔisɑ → guɖie Pɔisɑ / guɖɑe Pɔisɑ
PL marker Penny.N.SG penny.N.PL. (pennies)

/-guɖikɔ/, /-guɖɑkɔ/ are also used for human nouns for those who are placed at a socially dejected status and position.These particles have got a semantically negative connotation.
For example:
corɔ + - guɖikɔ /-guɖɑkɔ → corɔ guɖikɔ / corɔ guɖɑkɔ
Thief N.SG PL marker thief.N.PL
Theives

/-guɖikɔ/ and /-guɖɑkɔ/ can be employed for both non-human as well as human noun but the same can't be done while using /-mɑne/. It can be sffixed as a plural marker with human nouns only. /Pɑniɑmɑne/ is an agrammatical construction.Plural markers like /-e/ and /-kɔ/ stands for definiteness and indefiniteness respectively. Definite markers like /-ʤɑkɔ/ and /-ʈɔkɔ/ can also be added to the nounbase to get the plural form(B.Mohanty,2010,198).

For example:
a. kɔʈʰɑ ʤɑkɔ
Talk.N.SG PL marker.DEF
All the talks

b. bɔhi ʈɔkɔ

book.N.SG PL marker.DEF

All the books

IV.The plural formation in Odia is done by suffixation of both free and bound Morphemes. Few examples of free Morphemes used for making plural are as follows:ɑɖi, sɔbu, sɔmɔsʈe, sɔmuhɔ, ʃreɳi, bɔrgɔ, gɔɳɔ,ʤuʈʰɔ,ɖɔlɔ, mɑlɑ,pɔɳkʈi,mɔnɖɑ,goʈʰɔ,rɑʤi

Examples:

ʤɔnɔ + sɔmuhɔ → ʤɔnɔ sɔmuhɔ

These nouns of multitude in additiion to the plural suffix are also used for plural formation.

bɔnɖʰu (friend) +bɔrgɔ (Pl. Free Morpheme) +mɑne(Pl.marker)=bɔnɖʰubɔrgɔ mɑne(friends)

For example:

sɔbu+neʈru+brundɔ+ mɑne= sɔbu neʈrubrunɖɔ mɑne(ungrammatical)

V.The plural marker sɔbu can come in both prefix and suffix position.

For example:

pilɑ sɔbu / sɔbupilɑ (All boys)

If a noun stands for a class or in generic sense than also it indicates a plural form (Mahapatra &Das,2009:37-38).

Noun:

mɑʈʰiɑ mɑʈʰiɑ ɖɔhi (vessels and vessels of curd)

Adjective:

bʰɔlɔ bʰɔlɔ puɔ (good boys)

3.2.1.3 Case in Odia

Blake defines case broadly as ‘a system of marking dependent nouns for the type of relationship they bear to their head’ (Asburry,2008:15). Case is a grammatical category

determined by the syntactic or semantic function of a noun or pronoun. In a narrow sense, case is the relationship between the constituents and the verb of the sentence.Odia uses both case and post positions. The case system in Odia includes cases like Nominative, Accusative, Instrumental, Dative, Ablative, Genitive, Locative, and the Vocative markers.Subjects in Odia can take nominative, instrumental, dative, locative, and genitive case markers. Accusative case and dative case represent direct and indirect object respectively. Oblique objects take locative, ablative case markers. Many consider vocation also to be a case but it does not have any role to play in the sentence. These vocative expressions usually came in the beginning (Mahapatra,2007: 72).

Example:

hɔri re or re hɔri (oh Hari)

3.2.1.3.1 Nominative Case

Nominative case is a direct case and marked with the subject of the sentence. Nominative case is unmarked in Odia singular form but have many markers for plural. The following enlists few of the nominative case markers in odia.

Case	Number	Inflectional marker	Example in words
NOM	Sg	Absent	-------
NOM	Pl	e	pile
		mɑnɔ	kɔʈʰɑm ɑnɔ
		m ɑne	cʰu ɑm ɑne
		gudˌ ɑ	pʰulɔgudˌ ɑ
		gudˌikɔ	kʰɑʈɑgudˌikɔ
		gudˌ ɑkɔ	dˌɔb ɑgudˌɑkɔ
,		gudˌie	kɔʈʰɑgudˌier
		gudˌɑe	micʰɔgudˌɑe
		ʈɔkɔ	bɔhi ʈɔkɔ

3.2.1.3.2 Accusative/Objective Case:

This case is marked on direct object and generally governed by a transitive verb. According to Mahapatra (2007:74) the accusative case marker /-ku/ has three other alternatives /-ki/, /-ʈe/ and null. /-ki/ comes with an object ending with sound /-i/./-ʈe/ comes with pronominal forms of 1st and 2nd person singular number.

Examples using the accusative particles:

i. kɔbi + -ki → kɔbiki
kabi.N.SG ki.ACC to kabi
To kabi

ii. mo + -ʈe → moʈe
i.PR.REF ʈe.ACC to me
To me

Few accusative case markers for both singular and plural in odia are enlisted below:

Case	Number	Inflectional marker	Example in words
ACC	Sg	ku	puɔku
		ʈʰɑku	ʈɑtʰɑku
		ki	kaɔbiki
		ʈe	moʈe
		ʈiku	kɑmɔʈiku
		ʈɑku	ʈʰɑlʲiʈɑku
		ŋku	bɑpɑŋku
ACC	Pl	ŋku	pilaɳku
		ŋki	mɔnʈriɳki
		ʈʰɑku	ʈɑtʰɑku
		mɑnɔŋku	puɔmɑnɔŋku
		mɑnɔŋkɔʈʰɑku	semɑnɔŋkɔʈʰɑku
		ʤɑkɔku	gãdʒɑkɔku
		ʤɑkɔŋku	ʤʰiɔdʒɑkɔŋku

3.2.1.3.3 Instrumental Case

This case denotes something which helps in the completion of an act.

Case	Number	Inflectional marker	Example in words
INS	Sg	e	kuʈʰare
		re	pɔnikʰire
		ɖʊara	buldɖodʒɔr ɖʊara
		ŋkɔɖʊarɑ	dʒɔnʈriŋkɔɖʊarɑ
		ɖei	paipɖei
		ŋkɔɖei	sɑŋkɔɖei
		ɖehi	rɑmɔɖehi
		ŋkɔ ɖehi	ɖebIŋkɔɖehi
		ɖehe	lokɔɖehe
INS	Pl	ŋkɔɖehe	lokɔŋkɔɖehe
		kunei	ʈɑkunei
		ŋkunei	bɑbuŋkunei
		ŋkɔdʊarɑ	bɑpɑŋkɔdʊarɑ
		mɑnɔŋkɔ ɖʊarɑ	sainikɔmanɔŋkɔ ɖʊarɑ
		mɑnɔŋkɔɖei	misʈri mɑnɔŋkɔɖei
		mɑnɔŋkɔɖehi	mɔʃɑmanɔŋkɔɖehi
		mɑnɔŋkɔɖehe	lokɔ mɑnɔŋkɔɖehe
		mɑnɔŋkunei	dʒibɔ mɑnɔŋkunei
		gudɖikɔɖʊarɑ	mɑcʰɔgudɖikɔɖʊarɑ
		gudɖɑkɔɖʊarɑ	pilɑgudɖɑkɔɖʊarɑ

3.2.1.3.4 Dative Case

The dative case is marked on an indirect object. It is marked on indirect object.This case has become practically merged with the Accusative or Objective case. Both the case share same particles for designating their cases. However there are few other markers used for Dative case in addition to the accusative case markers. Those markers are lɑgi, pɑĩ,nimɔnɻe,sɔkɑse. All the mentioned particles used for dative case stands for English meaning of *for*. These mentioned particles take the singular as well as plural markers to form the dative form of noun. For singular dative inflection they get added to /-ɳkɔ/ and for plural inflection /-ɳkɔ/, /-mɑnɔɳkɔ/,/-guɖikɔ/,/guɖɑkɔ/.

For example:

i. ɔɖʒɑɳkɔ lɑgi
Grandfather.N.SG for.DAT
For grandfather

ii. ɖʒhiɔmɑnɔɳkɔ pɑĩ
girl.N.Pl for.DAT
For girls

3.2.1.3.5 Ablative Case

This case is an indication of movement from something or also used to project the contrast (Mohanty.2015:151). /-ʈʰɑru/ is an ablative particle in Odia which is used to highlight the difference or to make comparison between persons or things.

Example:

se mo- ʈʰɑru sunɻɔrɔ
She.N.Sg i.PRO than.ABL beautiful.ADJ
She is more beautiful than me.

Case	Number	Inflectional marker	Example in words
Ablative	Sg	ru	ɔpʰisru
		ʈʰɑru	pɔɖiɑʈʰɑru
		ŋkɔʈʰɑru	ɑiŋkɔʈʰɑru

		u	gʰɔru
		ũ	sʊɔrgɔlokũ
		sũ	beɭɔsũ
		ʈʰu	gʰɔrɔʈʰu
		ʈʰɔu	kali ʈʰɔu
		ʈiru	ɖʒamaʈiru
		ʈaru	baksɔʈaru
Ablative	Pl	manɔŋkɔru	ɖeʃɔ manɔŋkɔru
		ŋkɔʈʰaru	praɳiŋkɔʈʰaru
		manɔŋkɔʈʰaru	biɖeʃi manɔŋkɔʈʰaru
		guʈakɔru	ɖʰaɖu guʈakɔru
		guʈikɔru	kʰɔɳiguʈikɔru
		manɔŋkɔʈhu	semanɔŋkɔʈhu
		manɔŋkɔʈhɔu	bagʰɔmanɔŋkɔʈhɔu

3.2.1.3.6 Locative Case:

Locative case shows the place or time of an event.According to B.Mohanty(2010:211)the main inflectional forms in Odia for locative case are /-re/, /-ʈʰare/ and /-ʈʰi/ and /-ʈʰeĩ/ are the spoken form of the/-re/ and /-ʈʰare/.Here follows few locative markers for both singular and plural form.

Case	Number	Inflectional marker	Example in words
LOC	Sg	e	gʰɔre
		re	akʰire
		ʈʰare	gʰɔrɔ ʈʰare
		ŋkɔʈʰare	bapaŋkɔʈʰare
		ʈʰi	bɔgicaʈʰi
		ŋkɔʈʰi	kakaŋkɔʈʰi
		ʈʰeĩ	pokʰri ʈʰeĩ
		ʈʰare	gʰɔrɔ ʈʰare
		ʈire	skulʈire
		ʈare	cʰɔʈare

LOC	Pl	mɑŋɔkɔre	gruhɔmɑŋɔkɔre
		mɑŋɔkʈʰɑre	semɑŋɔkʈʰɑre
		ŋkɔʈ̪ʰi	ʈ̪ɑŋkɔʈ̪ʰi
		mɑnɔŋkɔʈ̪ʰi	mɑŋɔkʈʰi
		gudɻikɔre	kɑmɔgudɻikɔre
		gudɻɑkɔre	luhɑgudɻɑkɔre

3.2.1.3.7 Genitive Case:

Genitive case expresses a grammatical relationship between nouns or pronouns among themselves by means of inflections. It typically exhibits a possessive relationship. ‘rɔ’ is the default suffix for genitive case in Odia.It has also got other plural and singular forms of inflection. Few genitive case markers are as follows:

Case	Number	Inflectional marker	Example in words
GEN	Sg	rɔ	rɔgʰurɔ
		kɔrɔ	ʈɔŋkɑkɔrɔ
		kɑrɔ	eʈʰikɑrɔ
		kɑ	kɑlikɑ
		ʈikɔrɔ	kɔʈ̪ʰɑʈikɔrɔ
		ŋkɔ	ʤeʤeŋkɔ
		ŋkɔrɔ	debiŋkɔrɔ
		ri	ʈ̪ɑri
		ŋkɔri	ʈ̪ɑŋkɔi
GEN	Pl	mɑnɔŋkɔ	sɑŋgɔmɑnɔŋkɔ
		mɑnɔŋkɔrɔ	ɑmɔmɑnɔŋkɔrɔ
		mɑnɔŋkɔri	ɑmɔmɑnɔŋkɔri
		ʈ̪ɔkɔrɔ	bɔhiʈ̪ɔkɔrɔ
		ʤɑ kɔrɔ	gʰɔrɔʤɑkɔrɔ
		gudɻikɔrɔ	gʰɔdɻigudɻikɔrɔ
		gudɻɑkɔrɔ	kɑrgudɻɑkɔrɔ
		gudɻiŋkɔrɔ	menɖʰɑgudɻiŋkɔrɔ

gudɑŋkɔrɔ menḍʰɑgudɑŋkɔrɔ

3.2.1.3.8 Vocative Case Markers

Apart from the above mentioned cases few odia grammarians also include vocative markers in the case list. However these markers are usualy placed in the outset and do not play any role in the sentence(Mahapatra,2007:72).

The vocative particles in Odia are enlisted below:

/re/ ,/he/ ,/ɑhe/, /lo/,/ɑlo/, /go/ ,/ɑgo/, /e/,/be/ ,/kire/ , /ɑre/ /hɔikie/ /hɔikire/ /ɑbe/ /hɔikilo/ /ei/,/hei/ ,/hɔire/ ,/hɔilo/ /kio/ ,/ɑgjã:/,/ɑgɳãmɑne/

Example:

ɑre	rɑmɔ ,	toɻe	bɑpɑ	kʰodʒucʰɔnʈi
ɑre.VOC	ram.N.M.SG	you.ACC	father.N.M.SG	search.PRS.PROG

Ram, father is looking for you.

3.2.2 Pronoun

Pronoun is defined as a term which is used inplace of a word or a sentence (D.Mahapatra, 2013:24). The following list shows the list of personal pronouns in Odia. Pronouns is with the following categorical division such as personal, demonstrative, interrogative, relative, reflexive, intensive, Indefinite, possessive.

3.2.2.1 Personal pronoun

Person	Number	Pronoun (Subject)
1st	Sg	mũ
	Pl	ɑme(we),ɑmbʰe, ɑme mɑne(we), ɑmbʰemɑne(we)
2nd	Sg	ʈu(-H)-You, ʈume(+H)/ ʈumbʰe(+H)-You, ɑpɔɳɔ(+H) -You
	Pl	ʈumemɑne/ ʈudʒmbʰe mɑne(You), ɑpɔɳɔmɑne(You)
3rd	Sg	se(he/she)
	Pl	semɑne(they)

As said earlier case relationship operates with both nouns and pronouns, So all the pronominal forms also inflect for case relations as that of noun. In Odia the noun form remains morphologically unaltered after suffixation for case relationships but pronominal undergoes change (Pradhan et al,1998:77).

Let us look at the 2nd person singular non honorific form(ʈu) for various case inflection

subjective/ nominative Case	:ʈu
Objective/Accusative/Dative Case	:ʈoʈe
Genitive/Possessive Case	:ʈorɔ
Locative Case	:ʈoʈʰɑre
Ablative Case	: ʈoʈʰɑru
Instrumental Case	: ʈoɖei

3.2.2.2 Other pronominal forms

As similar to the above person the case and number inflection can be found for all the persons. The list of few other pronouns in Odia are as follows:/e/,/ehɑ/, /sɔbu/, /ɑpe/, /nidʒe/, /ubʰɔjɔ/, /ʈɑhɑ/, /ehi/, /dʒɑhɑ/, /ɔnjɔ/, /ɑrɔ/, /kie/, /kɔɳɔ/, /kisɔ/, /kehi/.

3.2.3 Postposition

Postpositions are used interchangeably with the case markers. According to Abbi(2001:128). Indian languages exhibit various case relations with distinct postpositions and the same postpostion can be used to display more than one case relationship. For example Odia language used the case marker /-ku/ for both accusative and dative case.

For example:

Somi-ku (to Somi)

somi.N to (ACC/DAT)

3.2.4 Adjective

Adjectives has got semantic properties like dimension,physical property, speed, age, color, value, qualification, human propensity, similarity, taste, quantification to modify their noun head (Abbi,2001:132).

Odia Adjectives shows variation in forms for degree of comparison. The three degrees are positive, comparative and superlative degree. /-ʈ̪ɔrɔ/ and /–ʈ̪ɔmɔ/ are the variant suffixes for designating comparative and superlative degreefor the positive degree. Another pair of particles such as /-iɔs/ or /-iɑn/ and /-isʈɔ/ are used to show the comparative and superlative degree respectively.

For example:

Positive Degree	Comparative Degree	Superlative Degree
nikɔʈɔ (near)	nikɔʈɔʈ̪ɔrɔ (nearer)	nikɔʈɔʈ̪ɔmɔ (nearest)

Odia adjectives also show agreement with the biological gender of noun. But this phenomenon sustains in an optional condition. Odia adjectives does not always necessarily takes the gender of noun and the retention of the default form also subsides with the grammaticality (Pradhan et al.,1998:73).

Example:

sunḍ̪ɔrɔ bɑlʲikɑ sunḍ̪ɔri bɑlʲikɑ

Both the forms are grammatically acceptable.

3.2.5 Verb

This part will discuss about the inflection of verbs in Odia in relation to the various verbal features. Verbs are mainly devided intofinite and non finite forms. Finite verb forms display inflection whereas the non finite does not. Odia verbs show inflection for a number of the features like person, number, tense, aspect, honour or degree of respect and mood.

3.2.5.1 Person and number

Odia does have a sum total of seven persons i.e. 1st Sg, 1stPl , 2ndSg(+H), 2nd Sg(-H), 2nd Pl, 3rd Sg and 3rd Pl. Odia inflected verb forms take different particle for different persons and number. For the inflected forms refer to the list of personal pronouns in P.

3.2.5.2 Moods in Odia

Odia have four types of grammatical moods such as indicative mood, subjunctive mood, imperative mood and optative mood (Mahapatra,1999,*Pandit Nilakanthanka Bhasatatwa*)

Odia verbs get inflected for the prior mentioned moods and can replace the tense and aspect suffixes.

3.2.5.2.1 Indicative mood

The indicative mood stands for the general statements.

Example:

se	kɑli	ɡʰɔrɔ	-ku	ʤɑucʰi
she.3.F.SG	tomorrow.ADV	home.N.SG	To.PSP	go.PRS.PROG

She is going home tomorrow.

3.2.5.2.2 Subjunctive Mood

The subjunctive mood deal with the stated expressions which are contrary to fact at the time of the utterance. It basically comes with the conditional statements.Odia verbs take a different inflectional marker /–nʈɔ/ for displaying this mood.

Example:

ɑʤi	bɔrʃɑ	hoi	nɔ	ʈʰile	mũ	skul	
today.ADV	rain.N	happen.V	neg	happen.V.PST.SMP	i.1st.SG	school.N.SG	

ʤɑiʈʰɑnʈi

go.COND

If it would not have rained, I would have gone to school.

3.2.5.2.3 Imperative Mood

The imperative mood is related to sentences that make direct commands, express requests, and grant or deny permission. In odia the imperative mood can appear with second person singular number and for all the three forms like: ʈu(Intimate), ʈume(familiar), ɑpɔɳɔ(formal) (Mahapatra,2007:146). Moreover this mood can be grasped from the intonational pattern of the speaker

Example:

ɑpɔɳɔ	kɑli	ɑsi	pɑrɔnʈi
You.2.SG.HON	tomorrow.ADV	comes	can

You can come tomorrow.

3.2.5.2.4 Optative mood

This is about wishing or hoping from someone. In odia Optative mood can be used only with the third person singular and plural number (Mahapatra,2007:147).

Example:

Se kɑmɔ -ʈɑ kɔru

he.3.M.SG work.N.SG the.ART does. OPT

I am expecting that he will the work.

3.3 Aspect and Auxiliaries in Odia

Odia has got three auxiliary forms /-ʈ^{h}ɔ/ and /-c^{h}ɔ/ or /-cɔ/ and /-nɔ/. When the root word take the aspectual marker /-i/ (perfect) or /-u/ (imperfect) they ought to take auxiliary forms. All the three mentioned auxiliaries can be used only in present tense in contrast to past and future. In present tense the auxiliary forms /-ʈ^{h}ɔ/ and /-c^{h}ɔ/ are used to show indefiniteness and definiteness respectively (Mahapatra,2007,134-143).

Aspect	Person	Number	Inflectional Markers
SIM	1st	Sg	-e
		Pl	-u
	2nd	Sg	-u, -ɔ,-ɔnʈi
		Pl	-ɔnʈi
	3rd	Sg	-e, -ɔnʈi
		Pl	-ɔnʈi
PROG	1st	Sg	-uʈ^{h}ɑe
		Pl	-uʈ^{h}ɑu
	2nd	Sg	-uʈ^{h}ɑu,-uʈ^{h}ɑɔ,-uʈ^{h}ɑnʈi
		Pl	-uʈ^{h}ɑɔ,-uʈ^{h}ɑnʈi
	3rd	Sg	-uʈ^{h}ɑe,-uʈ^{h}ɑnʈi
		Pl	-uʈ^{h}ɑnʈi
PRF	1st	Sg	-iʈ^{h}ɑe
		Pl	-iʈ^{h}ɑu
	2nd	Sg	-iʈ^{h}ɑu, -iʈ^{h}ɑɔ,-iʈ^{h}ɑnʈi

	Pl	-iʈʰɑɔ,-iʈʰɑnʈi
3rd	Sg	-iʈʰɑe,-iʈʰɑnʈi
	Pl	-iʈʰɑnʈi

3.4 Inflectional forms for various tense

The exponential markers for past tense, future tense and conditional are /–il/, /-ib/, /-ɔnʈ/ respectively whereas Present tense remains unmarked.

3.4.1 Inflectional markers for Present Tense

The following list is presenting the various inflectional markers for present tense in agreement with the three persons (1st,2nd,3rd) , number (singular and plural) and in all the three aspects of present tense.

Tense	Aspect	Person	Number	Inflectional Markers
Present	SIM	1st	Sg	-e
			Pl	-u
		2nd	Sg	-u,-ɔnʈi
			Pl	-ɔ, -ɑnʈi
		3rd	Sg	-e, -ɑnʈi
			Pl	-ɑnʈi
	PROG	1st	Sg	-uɔcʰi/-ucʰi
			Pl	-uɔcʰu /-ucʰu, -ucʰe /uɔcʰe
		2nd	Sg	-uɔcʰu/-uɔcʰɔ,-ucʰɔnʈi/-ucʰnʈi
			Pl	-uɔcʰɔ /-uɔcʰɔ,-ucʰɔnʈi/-ucʰnʈi
		3rd	Sg	-uɔcʰi/ucʰi/, -ucʰɔnʈ/ucʰnʈi/
			Pl	-uɔcʰɔnʈi/ucʰɔnʈi
	PRF	1st	Sg	-iɔcʰi/icʰi
			Pl	-iɔcʰu/icʰu
		2nd	Sg	-iɔcʰu/icʰu,-icʰɔ /iɔcʰɔ,-icʰɔnʈi, -icʰi
			Pl	-icʰɔ/iɔcʰɔ,-icʰɔnʈi/icʰnʈi

3rd	Sg	-iɔcʰi/icʰi/,-iɔcʰɔnʈi/icʰɔnʈi	
	Pl	-iɔcʰɔnʈi /icʰɔnʈi	

It can be noticed from the list of inflectional marker that first person plural form (ɑme) has got two alloMorphs. The marker -uɔcʰu/-ucʰu is related to exclusiveness property and -ucʰe/-uɔcʰe is related to the semantic concept for inclusiveness. So it can be derived from the inflectional ending with /-u/ sound is about excluding others and the form ending with sound /-e/ is about inclusiveness. This phenomenon can also be found with 1st person plural form for other tenses also.

3.4.2 Inflectional markers for Past Tense:

Tense	Aspect	Person	Number	Inflectional Markers
PAST	SMP	1st	Sg	-ili
			Pl	-ilu,-ile
		2nd	Sg	-ilu,-ilɔ,-ile
			Pl	-ilɔ, -ile
		3rd	Sg	-ilɑ, -ile
			Pl	-ile
	PROG	1st	Sg	-uʈʰili
			Pl	-uʈʰilu,-uʈʰile
		2nd	Sg	-uʈʰilu,-uʈʰilɔ, -uʈʰile
			Pl	-uʈʰilɔ,-uʈʰile
		3rd	Sg	-uʈʰilɑ,-uʈʰile
			Pl	-uʈʰile
	PRF	1st	Sg	-iʈʰili
			Pl	-iʈʰilu,-iʈʰile
		2nd	Sg	-iʈʰilu,-iʈʰilɔ,-iʈʰile
			Pl	-iʈʰilɔ,-iʈʰile
		3rd	Sg	-iʈʰilɑ,-iʈʰile
			Pl	-iʈʰile

3.4.3 Inflectional markers for Future Tense:

Tense	Aspect	Person	Number	Inflectional Markers
FUT	SMP	1st	Sg	-ibi
			Pl	-ibu, -ibe
		2nd	Sg	-ibu, -ibɔ,-ibe
			Pl	-ibɔ,-ibe
		3rd	Sg	-ibɔ, -ibe
			Pl	-ibe
	PROG	1st	Sg	-uʈʰibi
			Pl	-uʈʰibu,-uʈʰibe
		2nd	Sg	-uʈʰibu, -uʈʰibɔ, -uʈʰibe
			Pl	-uʈʰibɔ, -uʈʰibe
		3rd	Sg	-uʈʰibɔ,-uʈʰibe
			Pl	-uʈʰibe
	PRF	1st	Sg	-iʈʰibi
			Pl	-iʈʰibu,-iʈʰibe
		2nd	Sg	-iʈʰibu,-iʈʰibɔ,-iʈʰibe
			Pl	-iʈʰibɔ,-iʈʰibe
		3rd	Sg	-iʈʰibɔ, -iʈʰibe
			Pl	-iʈʰibe

The Odia verbal inflection has got five slots for filling in the various verbal inflectional features. So let us see the process of formation of various verb forms for features like TAM (tense, aspect, mood) and PNG (person, number, gender) respectively.

kʰɑucʰi → kʰɑ + u + cʰ i

Eat.PRS.PROG Root word + Aspect + Tense +P.N.G agreement

From the above example it can be deciphered that the root verb form is placed in the first slot and then the aspect, tense, person, number fills in the places respectively. The formula for the concatenation of verbal inflection is as follows:

Inflected form → root word + aspect +mood+ tense + PNG agreement

As we know that Odia has only the lexical gender but not grammatical gender. The following example will make the point explicit.

For example

i. sili bʰɔʤɔn gɑucʰi.
Sili.N.FEM.SG prayer.N sing.PRS.PROG
Sili is singing prayer.

ii. bibʰuʈi bʰɔʤɔn gɑucʰi.
Bibhuti.N.MAS.SG prayer.N sing.PRS.PROG
Bibhuti is singing prayer

For both the feminine as well as masculine gender the verb (gɑ) is getting inflected with the same marker. So the deletion of the gender feature is reducing the PNG to PN agreement.

3.5 Inflectionalsuffixes for conditionals, habitual and imperatives.

3.5.1 Inflectional suffixes for conditionals

Aspect	Person	Number	Inflectional Markers
SIM	1st	Sg	-ɑnʈi
		Pl	-ɑnʈu,-ɑnʈe
	2nd	Sg	-ɑnʈu,-ɑnʈɔ,-ɑnʈe
		Pl	-ɑnʈɔ,-ɑnʈe
	3rd	Sg	-ɑnʈɑ,-ɑnʈe
		Pl	-ɑnʈe
PROG	1st	Sg	-uʈʰɑnʈi
		Pl	-uʈʰɑnʈu,-uʈʰɑnʈe
	2nd	Sg	-uʈʰɑnʈu,-uʈʰɑnʈɔ,-uʈʰɑnʈe
		Pl	-uʈʰɑnʈɔ,-uʈʰɑnʈe
	3rd	Sg	-uʈʰɑnʈɑ, -uʈʰɑnʈe
		Pl	-uʈʰɑnʈe

PRF	1st	Sg	-iʈʰanʈi
		Pl	-iʈʰanʈu,-iʈʰanʈe
	2nd	Sg	-iʈʰanʈu, -iʈʰanʈɔ, -iʈʰanʈe
		Pl	-iʈʰanʈɔ, -iʈʰanʈe
	3rd	Sg	-iʈʰanʈɑ, -iʈʰanʈe
		Pl	-iʈʰanʈe

The conditional verb morphology also follows the same pattern for verbal inflection as other tense inflections. The conditional mood Morpheme is /-ɔnʈ-/, whose alloMorphs are /-ɔnʈ-/ and /-nʈ-/.[42]

Example:

ʤɑiʈʰanʈɑ → ʤɑ+iʈʰanʈɑ
Would have gone
ʤɑ +iʈʰanʈɑ

In the above example /ʤɑ/ is the verb root and the inflectional suffix for conditional is /-iʈʰanʈɑ/, where /-i/ is the perfective aspect and /-ʈʰɑ/ is the auxiliary, /-nʈ-/ shows the conditional form. /-ɑ/ shows the morpheme agreement for person and number.[43]

3.5.2 Inflectional marker for Optative Mood

Optative mood is about expressing a wish for third person only. The inflectional markers for optative mood are:

Aspect	Person	Number	Inflectional Markers
SIM	3rd	Sg	-u,-ɔnʈu
		Pl	-ɔnʈu
PROG	3rd	Sg	-uʈʰɑu,-uʈʰanʈu
		Pl	-uʈʰanʈu
PRF	3rd	Sg	-iʈʰɑu,-iʈʰanʈu
		Pl	-iʈʰanʈu

[42]http://www.ciil-lisindia.net/oriya/oriya.html.
[43]Ibid.

For example:

nɑcɔ + iʈʰɑu→ nɑciʈʰɑu

Apart from the above enlisted inflectional markers for the optative mood construction there are two other markers which can be used are /-idʒɑu/,/-ineu/ in perfective aspect.

Example: pɔɖʰineu

pɔɖʰidʒɑu

3.5.3 Inflectional Markers for imperative mood

Even though imperative mood has been given the inflectional forms for all the three persons (1^{st}, 2^{nd} ,3^{rd}) but in reality Imperative mood (expressing order) in odia can be stated only for second person. The first person mũ kɔre(I do) or ɑme kɔru(we do) is similar to simple present tense and the third person singular form /se kɔre/ or plural form /semɑne kɔrɔnʈu/ is suggestive of a request or wish (Mahapatra:98).

In Odia the Inflectional markers for imperative mood (expressing order) can be stated only in case of second person singular number. It has got three inflectional forms.The inflections for the imperative mood (expressing order) can be done by using three types of marker. Those are enlisted below:

Inflectional form -1: -ɔ, -ɔnʈu

Inflectional form -2: -iɖe, -iɖiɔ, -idiɔnʈu

Inflectional form -3: -iʈʰɑ, -iʈʰɑɔ,-iʈʰɑnʈu

Example:

gɑ +-ɔnʈu→gɑɔnʈu

gɑ + -iʈʰɑ→gɑiʈʰɑ

3.6 Inflectional forms for passive construction

3.6.1 Inflectional markers for present, past, future

In passive construction mainly two forms are used /-ɖʒɑ/,/-ho/. Any one of the form can be used to form the passive. The following enlists the inflectional particles for all the three tense forms with the aspect, person and number.

Tense	Aspect	Person	Number	Inflectional Markers
PRS	SIM	1st /2nd/3rd	Sg/Pl	-ɑdʒɑe / -ɑhue
	PROG		Sg/Pl	-ɑdʒɑucʰi / -ɑheucʰi
	PRF		Sg/Pl	-ɑdʒɑicʰi /-ɑheicʰi
PST	SIM		Sg/Pl	-ɑgɔlɑ /-ɑhelɑ
	PROG		Sg/Pl	-ɑdʒɑuʈʰilɑ or -ɑheuʈʰilɑ
	PRF		Sg/Pl	-ɑdʒɑiʈʰilɑ /-ɑheiʈʰilɑ
FUT	SIM		Sg/Pl	-ɑdʒibɔ/-ɑhebɔ
	PROG		Sg/Pl	-ɑdʒɑuʈʰbɔ/-ɑheuʈʰibɔ
	PRF		Sg/Pl	-ɑdʒɑiʈʰbɔ/ -ɑheiʈʰib

3.6.2 Passive markers showing the present imperfect indefinite forms

Aspect	Person	Number	Inflectional Markers
SIM	1st, 2nd,3rd	Sg /Pl	-ɑhue/-ɑdʒɑe
CONT			-ɑheiʈʰɑe/-ɑdʒɑiʈʰɑe
PRF			-ɑheuʈʰɑe/-ɑdʒɑuʈʰɑe

Example:

lekʰ +-ɑheiʈʰɑe/-ɑdʒɑiʈʰɑe → lekʰɑheiʈʰɑe/ lekʰɑ ɑdʒɑiʈʰɑe (should have been written)

3.6.3 Passive markers showing the present perfect indefinite forms

Aspect	Person	Number	Inflectional Markers
SIM	1st, 2nd,3rd	Sg /Pl	-ihue/-idʒɑe
CONT			-iheiʈʰɑe/-idʒɑiʈʰɑe
PRF			-ɑheuʈʰɑe/-ɑdʒɑuʈʰɑe

Example:

kʰeɭ +-ɑheuʈʰɑe/-ɑdʒɑuʈʰɑe → kʰeɭɑheuʈʰɑe / kʰeɭɑ ɑdʒɑuʈʰɑe

3.6.4 Passive Markers for conditional

Aspect	Person	Number	Inflectional Markers
SIM	1st, 2nd, 3rd	Sg/Pl	-ɑhuɔnʈɑ or -ɑdʒanʈɑ
CONT			-ɑheiʈʰɑntɑ or-ɑdʒaiʈʰɑntɑ
PRF			-ɑheuʈʰɑntɑ/ or/-ɑdʒauʈʰɑntɑ

Example:

ɖekʰ + -ɑhuɔnʈɑ/ -ɑdʒɑnʈ → ɖekʰɑhuɔnʈɑ/ ɖekʰɑdʒnɑʈɑ (could have been seen)

3.6.5 Passive Marker for Optative mood

Passive inflectional markers for imperative mood (expressing order) in odia are

Aspect	Person	Number	Inflectional Markers
SIM	3rd	Sg/Pl	-ɑheu/-ɑdʒɑuu
CONT			-ɑheuʈʰɑu/- ɑdʒɑuʈʰɑu
PRF			-ɑheiʈʰɑu/ -ɑdʒɑiʈʰɑu

Example:

kɔr + -ɑheuʈʰɑu/- ɑdʒɑuʈʰɑu→ kɔrɑheuʈʰɑu/kɔrɑdʒɑuʈʰɑu (keep continue the work)

3.6.6 Passive marker for imperative mood

Passive inflectional markers for imperative mood (expressing order) in odia are

Inflectional form -1: -iɖiɑheu /-iɖiɑdʒɑu

Example: l ekʰ+ -iɖiɑheu/ -iɖiɑdʒɑ → lekʰiɖiɑheu / lekʰiɖiɑdʒɑu

Inflectional form -2: -ɑheuʈʰɑu /-ɑdʒɑuʈʰɑu

Example: lekʰ+ -ɑheuʈʰɑu/-ɑdʒɑuʈʰɑ→ lekʰɑheuʈʰɑu or lekʰɑdʒɑuʈʰɑu

3.7 Adverb

According to Mohanty (2010:357) adverbs are used to establish a relationship with subject, object or verb. Odia adverbs take postposition to inflect.

Example

ɑdʒ + -rɔ → ɑdʒrɔ

today.ADV -of

of today

CHAPTER 4

ODIA INFLECTIONAL MORPH ANALYZER

4.1 Design and Development of the Morph Analyzer

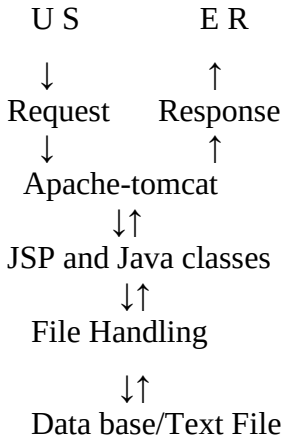
Chapter-4 is an implementation to the algorithm demonstrated in the chapter 2(Research Methodology). The tool has been developed on the theoretical basis of the rule based approach for inflectional Morphology of Odia. The development of the computational tool is modelled on the java based technology. As the system is based on the rule based approach, it contains a hand crafted dictionary where the rules are manualy written and fed into the system. The developed system takes the text information and prepares it for processing the input for identification of inflections. Then it segments the input text or word into their root forms and case suffixes. The inbuilt Odia POS Tagger is the intermediate module before the input information gets Morphanalysed. In the recognition process the tagging and tokenization happen simultaneously. Subsequently the Morphological analysis of the data goes on. The present model uses Java in the web format for the recognition and analysis of nominal and verbal inflection from the input Odia texts. This system accepts word, sentences and text in UTF-8.

4.2 Scope

The developed system analyses the text and generates its necessary features with its root words. The result produced are with grammatical features having categories like part of speech, gender, number, person, tense, aspect, mood, voice for various types of inflections.

4.3 System Description

The coding for the system is done by Abhishek Sharma and Prof. Girish Nath Jha. The system architecture has got two ends, the front end and the back end. The developed system will be available on the internet at http://sanskrit.jnu.ac.in/odia/Morph.



4.3.1 The Front End

The front end is the Graphical User Interface (GUI) of the Odia Morph Analyzer, viewed by the users. It has elemental composition of JSP (java server pages) and HTML(hyper text markup language) components. The main JSP file index.jsp allows the user to give input in UTF-8 using HTML text area component. The text box area has two buttons below to it, the submit and the reset button. After submitting the text in the textbox area the users have to click on the submit button to find out the results and if they want to reset the input text then they have to click on reset. Below the text area the result area is highlighted in red colour. The output will be displayed below the heading of result.

Here follows two figures. The first figure represents the system module and second figure is a screen shot of the front end, which is a plain text area with out any submitted data. Here follows two figures. The first one is the flow diagram of the system and the second is the screen shot of the front end.

Figure 1: Architecture of the Odia Inflectional Analyzer

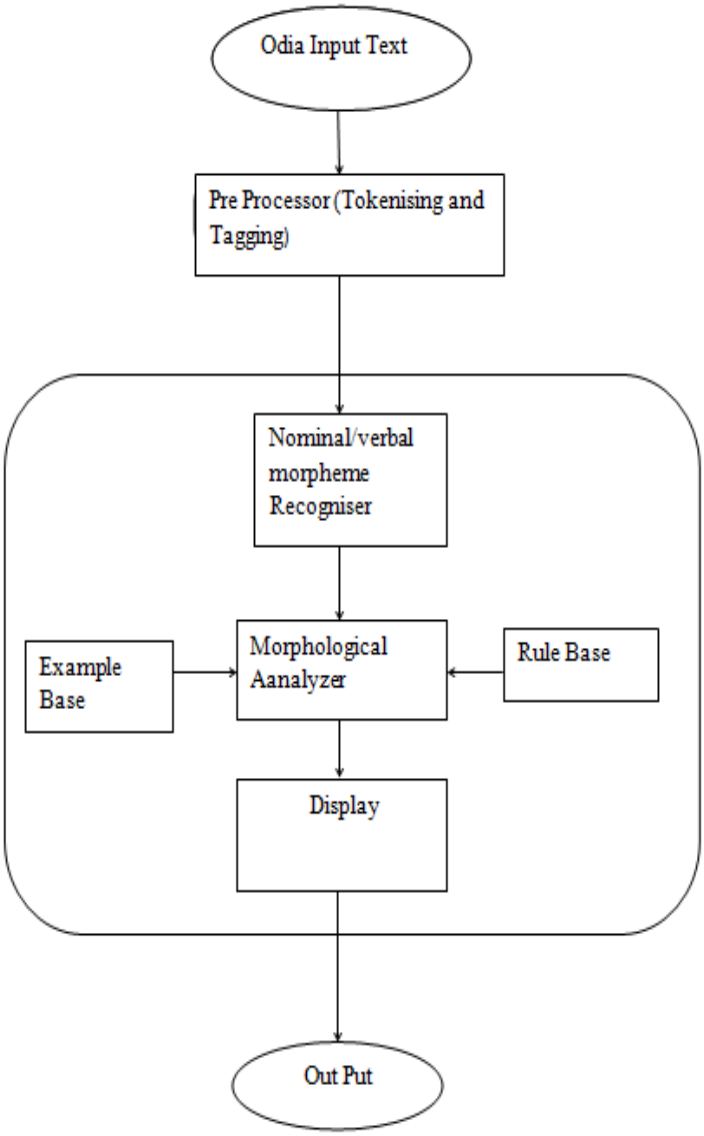
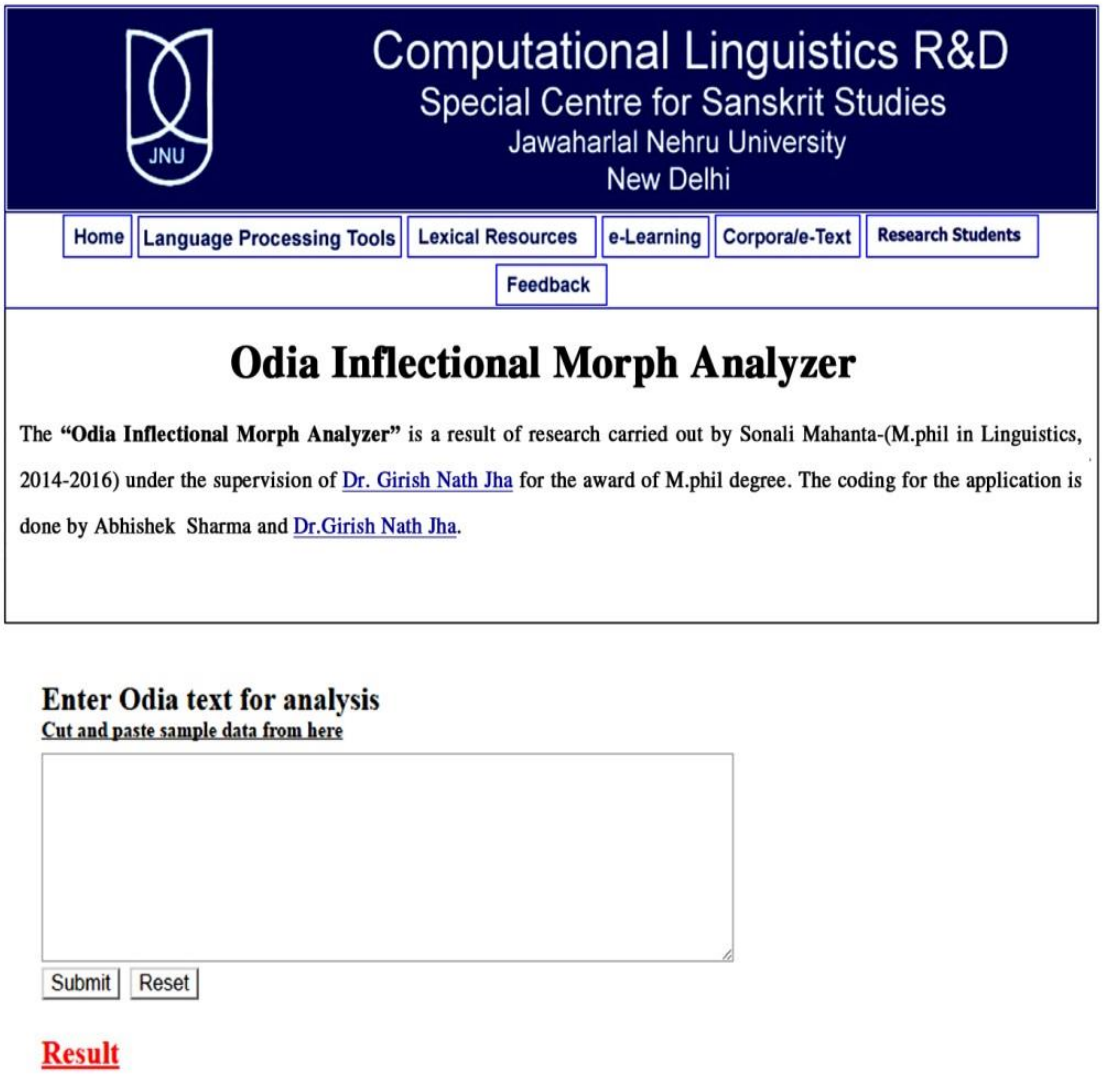


Figure -2 Front end of the Morph Analyzer



4.3.2 The Back End

The back end deals with the internal structure of the system and how the system works. The internal structure of this system consists of the textfiles and webserver.

4.3.2.1Text File

The text files serves as the database for the system. The text files are containing the inflectional rules for the nouns and verbs. The sample of the data stored in the text files is given below.

Sample of Odia Noun Morph.txt:

Nominal Inflection	**Surface form of inflection**	**POS**	**Meaning**	**Example in words**	**Exception**	**Condition**
ଏ	େ	N	Nom_Pl	ପିଲେ	(i)ଆଖିଏ) ii)ଘରେ	'ଏ' is used for(i) derivation(ii)locative case
ମାନ	ମାନ	N	Nom_Pl	ହାତୀମାନ	(i)ଭାସମାନ)ii)ବୁଦ୍ଧିମାନ)iii) କଥାରମାନ	ମାନ' is used to form meaning(i)floating(ii)male gender(iii)standard
ମାନେ	ମାନେ	N	Nom_Pl	ବାଘମାନେ	(i) କଥାମାନେନାହିଁ)ii)କଥାରମାନେବୁଝିଗଲି	In example(i) it means 'obey' and in (ii)meaning
ଗୁଡ଼ା	ଗୁଡ଼ା	N	Nom_Pl	ଫୁଲଗୁଡ଼ା	NIL	NIL
ଗୁଡ଼ି	ଗୁଡ଼ି	N	Nom_	ବହିଗୁଡ଼ି	NIL	NIL

			Pl			
ଗୁଡ଼ିକ	ଗୁଡ଼ିକ	N	Nom_Pl	ଘରଗୁଡ଼ିକ	NIL	NIL
ଗୁଡ଼ାକ	ଗୁଡ଼ାକ	N	Nom_Pl	ପିଲାଗୁଡ଼ାକ	NIL	NIL
ଗୁଡ଼ିଏ	ଗୁଡ଼ିଏ	N	Nom_Pl	ପକ୍ଷୀଗୁଡ଼ିଏ	NIL	NIL
ଗୁଡ଼ାଏ	ଗୁଡ଼ାଏ	N	Nom_Pl	କଥାଗୁଡ଼ାଏ	NIL	NIL
ଯାକ	ଯାକ	N	Nom_Pl	ପକ୍ଷୀଯାକ	NIL	NIL

Sample of Odia Verb Morph txt:

Verbal Inflection	**Surface Form of Inflection**	**Example in words**	**POS**	**Person**	**Number**	**Tense**	**Aspect**	**Honour/ Degree of respect**	**Mood**	**Voice**
ଏ	ଏ	ଖାଏ	V	1st	Sg	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଏ	େ	ଖେଳେ	V	1st	Sg	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଉ	ଉ	ଖାଉ	V	1st	Pl	PRS	SIM	NIL	INDICATIVE	ACTIVE

ଉ	ୁ	ପଢୁ	V	1st	Pl	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଉଛି	ଉଛି	ଖାଉଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛି	ୁଛି	ପଢୁଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛି	ଉଅଛି	ଖାଉଅଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛି	ୁଅଛି	ଲେଖୁଅଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛୁ	ଉଛୁ	ଖାଉଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛୁ	ୁଛ	ଲେଖୁଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛୁ	ଉଅଛୁ	ଖାଉଅଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛୁ	ୁଅଛୁ	ନାଚୁଅଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛେ	ଉଛେ	ଖାଉଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛେ	ୁଛେ	ଲେଖୁଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛେ	ଉଅଛେ	ଖାଉଅଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE

ଉଅଛେ	ୁଅଛେ	ପଢୁଅଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଇଛି	ଇଛି	ଖାଇଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଛି	ିଛି	ନାଚିଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଅଛି	ଇଅଛି	ଖାଇଅଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE

4.3.2.2 The web server

The web server acts as a container of the whole system. The Odia Morph Analyzer runs on Apache Tomcat 4.0 platform.The details for this Java based webserver follows http://java.sun.com/products/servlet/

4.3.2.2.1 Apache Tomcat 4.0

Apache Tomcat is the servlet container that is used for the Java Servlet and Java Server Pages technologies. The Java Servlet and Java Server Pages specifications are developed by Sun under the Java Community Process. Apache Tomcat is developed in an open and participatory environment and released under the Apache Software License. Apache Tomcat is intended to be a collaboration of the best-of-breed developers from around the world.[44]

4.3.2.2.2 Java Server Pages

Java Server Pages (JSP) technology provides a simplified, fast way to create dynamic web content. JSP technology enables rapid development of web-based applications that

[44]http://tomcat.apache.org/

are server and platform-independent. It is one of the most sophisticated tools available for high performance and secures web applications.[45]

4.4 The Components of the System

The developed system has three components such as POS tagger, Tokeniser-Recogniser, Morph Analyzer.

4.4.1 POS tagger

The Odia MorphAnalyzer is using an inbuilt POS tagger as an interface. Which means as soon as the raw data is being submitted into the text box area it gets processed through the POS tagger. Thereafter the by- produced data comes out with the appropriate tags.

Here follows few screen shots of the POS tagger which is used as an interface in the processing of the Odia Inflectional Morph Analyzer.

Out of the following three screen shots of the Odia POS Tagger, the first one is the front end, the second one is with the input data and the third one is with the output of the processed data.

[45] http://www.oracle.com/technetwork/javajavaee/jsp/index.html

Figure 5: Front end of the Tagger

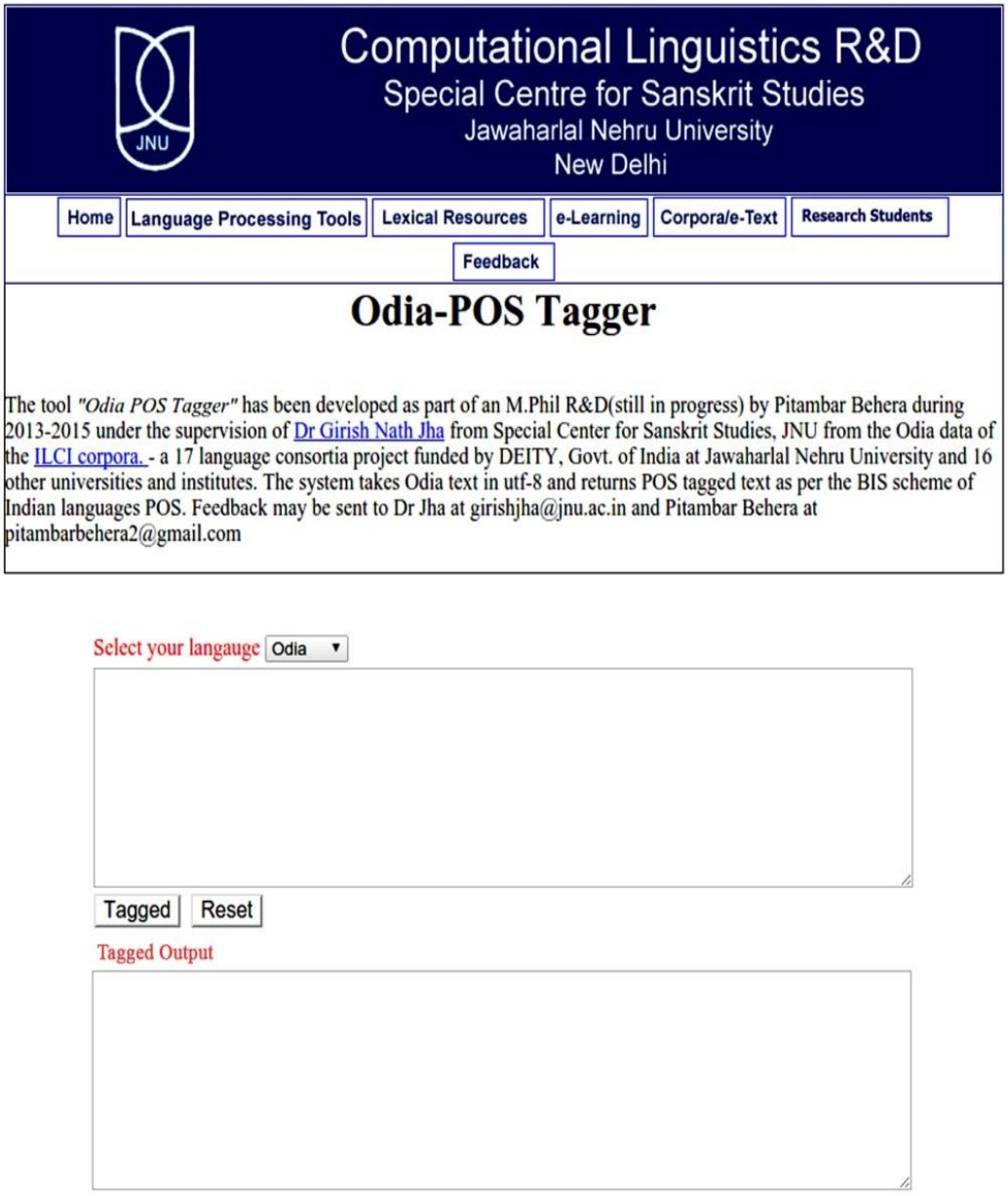


Figure 6: Tagger with raw data

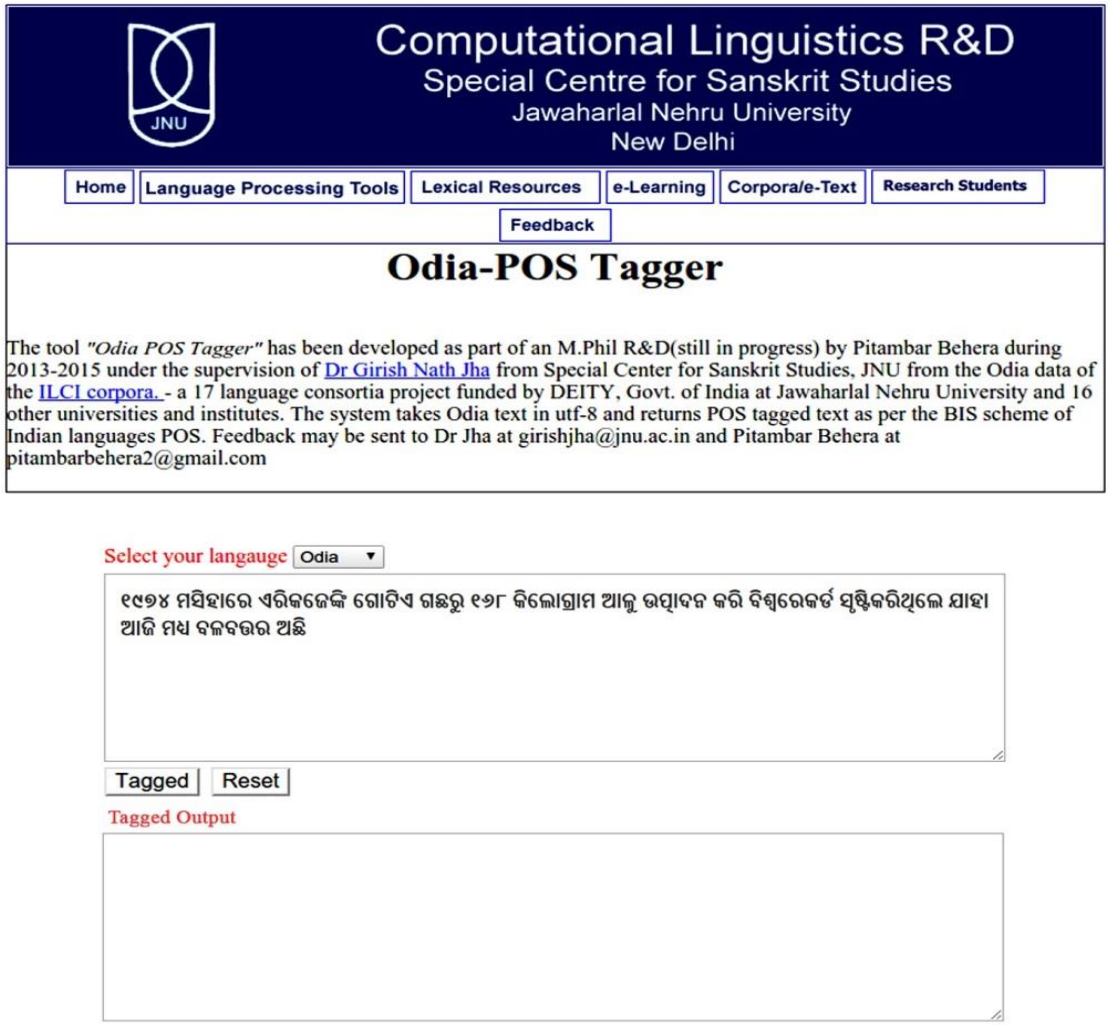
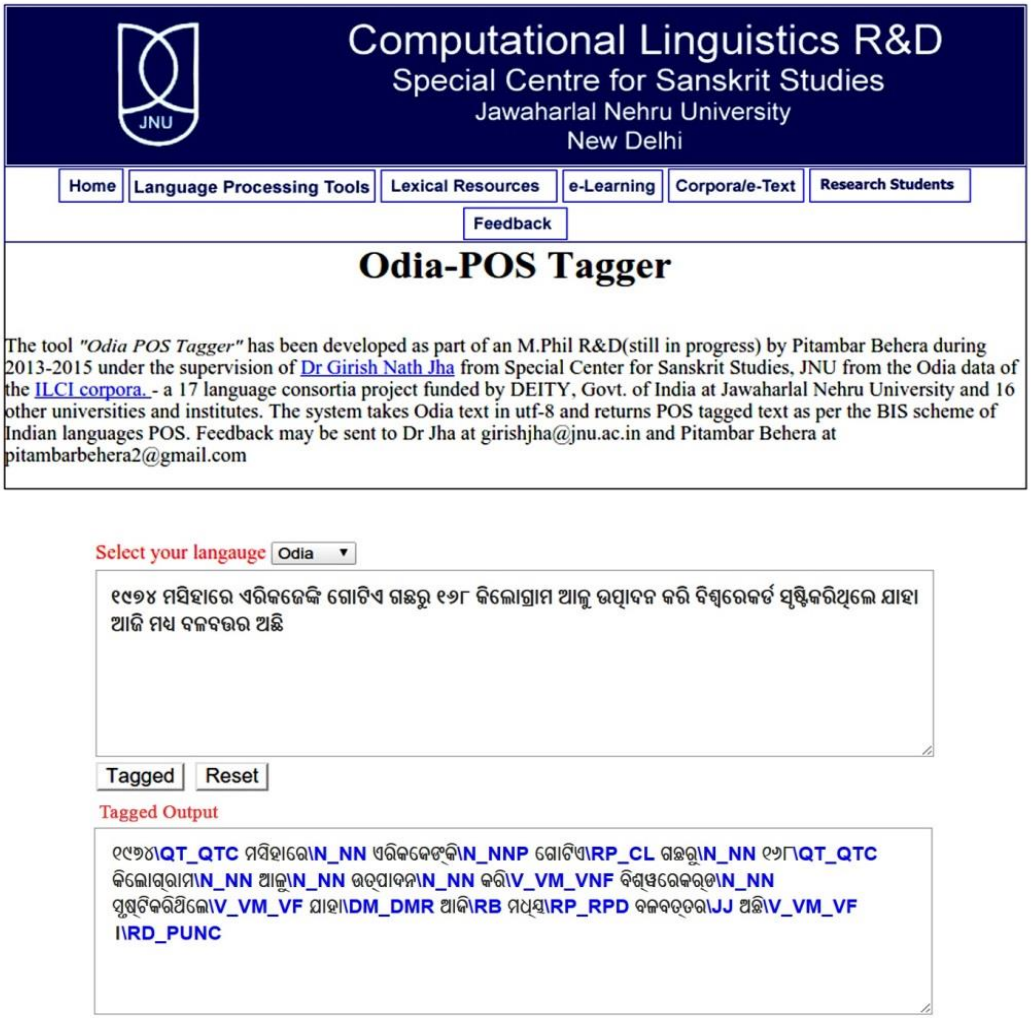


Image 7: Tagger with tagged output



4.4.2 Tokeniser-Recognizer

The subsequent phase is data tokenisation. Tokenization breaks the text into meaningful elements called tokens.Tokens may be words or phrases. The list of tokens becomes input for further processing. The recognizer then checks for the nouns, verbs and other inflectional categories simultaneously and are stored in a hash table.

Sample code for the recogniser:

```
import java.util.HashMap;
import java.util.Map;
import java.util.Map.Entry;

public class OdiaMorph {
public HashMap<String, String> MorphNounVerbAnalysis(String sentence){
HashMap<String, String> ht = new HashMap<String, String>();
try{
String DELIMITER = "\t";
if(!sentence.equals("")){
String[] words = sentence.split(" ");
for(String word : words){
if(word.contains("N_NN") || word.contains("N_NNP") || word.contains("N_NNV") ||
word.contains("N_NST") || word.contains("PSP")){
ht.put(word, "Noun");
}
else if(word.contains("V_VM") || word.contains("V_VM_VF") ||
word.contains("V_VM_VNF") || word.contains("V_VM_VINF") ||
word.contains("V_VM_VNG") || word.contains("V_VAUX")){
ht.put(word, "Verb");
}
else{
ht.put(word, "None");
```

```
}
}
}
}catch(Exception e){
    System.out.println("Exception caught : " + e.getMessage());
}finally{
}
return ht;
}}
```

4.4.3 Morph Analyzer

The third and decisive component of the system is the Morph Analyzer. Once the process of recognition is over the Morph Analyzer starts analyzing the inflectional forms.

The system searches for all the inflectional forms. As for instance the Analyzer scans a noun then it searches for whether it has inflection or not. If the inflection is found then the processing happens according to the nominal inflection rules stored in the text file and the same mechanism is also followed for the other categories having inflections.

The analysis is done with the help of the text files containing the rules. The Analyzer proceeds through the following steps. First it checks for solution in the text file. If solution is found then it proceeds to check the rule base by splitting a certain number of last characters from the inflcted forms. With each split, it checks if the base and affix combination it arrives are valid or not. The validity of the base and affix is done by checking the affix in the word. If the affix is found then it assumes the analysis to be correct and reports back. If the affix is not found then it creates another combination of base and affix. Thus the process goes on for each inflectional form. For each inflection it checks out with the total list of inflections according to their respective speech categories. Here follows two more screen shots of the Odia Inflectional Morph Analyzer. The first in the sequence is with the raw data and the second one is with the result after processing.

Figure 8: Morph Analyzer with text input

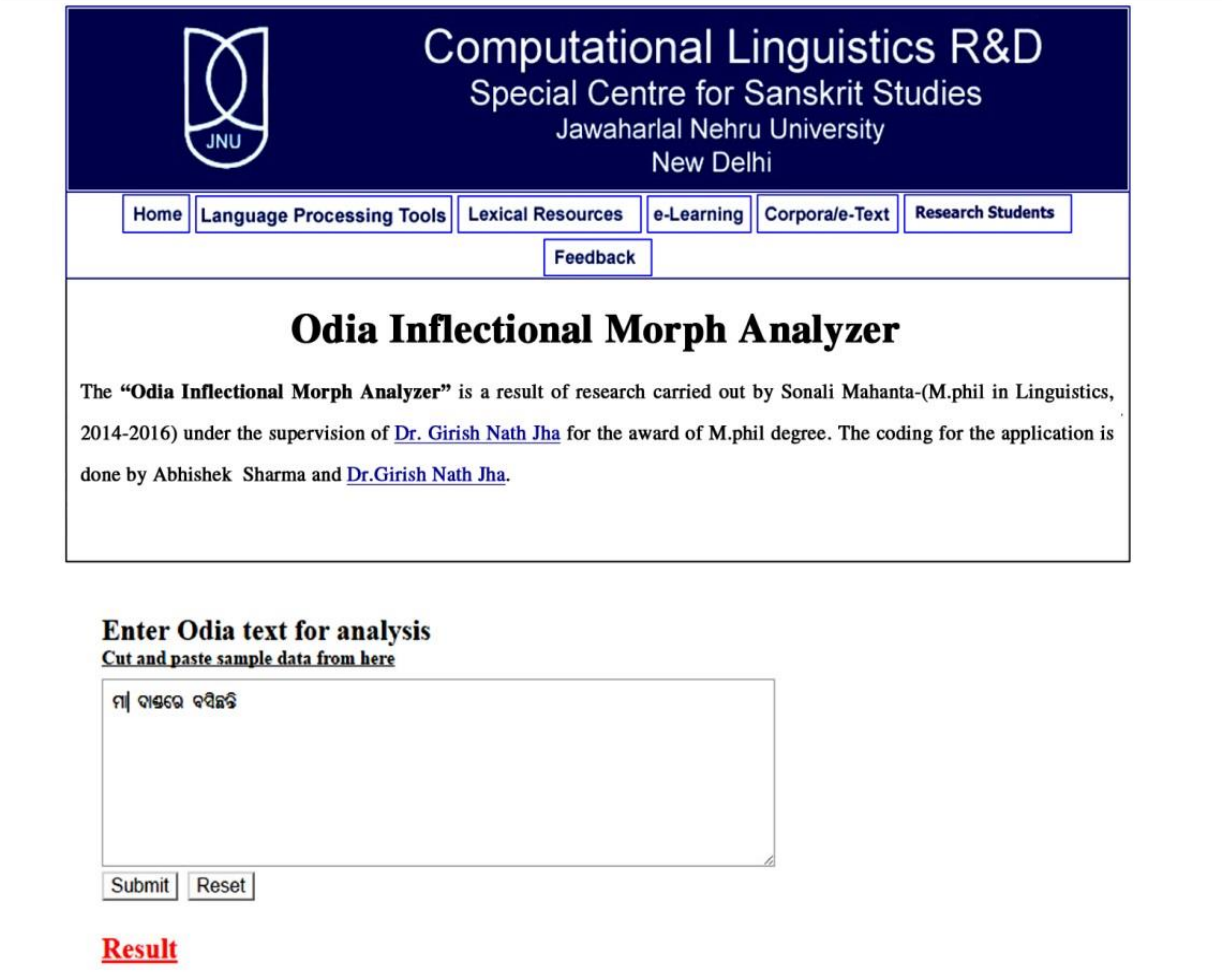


Figure 9: Morph analysed output

Cut and paste sample data from here

ମା ଦାଣ୍ଡରେ ବସିଛନ୍ତି

Submit Reset

Result

Nominal Inflection

Processed words	Root Word	Nominal Inflection	Surface form of inflection	POS	Meaning	Exception	Condition
ଦାଣ୍ଡରେ	ଦାଣ୍ଡ	ରେ	ରେ	N	Loc_Sg	(i)ଜାଣ୍ଟିରେ(ii)ଖୁରାରେ	'ରେ' stands for(i) vocative case(ii) instrumental case

Verbal Inflection

Processed words	Root Word	Verbal Inflection	Surface Form of Inflection	POS	Person	Number	Tense	Aspect	Honour/Degree of respect	Mood	Voice
ବସିଛନ୍ତି	ବସ	ଛନ୍ତି	ିଛନ୍ତି	V	2nd	Sg	PRS	PRF	Formal	INDICATIVE	ACTIVE

No Inflection

4.7 Validation of the Morph analysed data and issues

While validating the Morph analysed data it has been noticed that the system is producing correct results as well as suffering from few issues. This section first discusses about the achievements and then the issues related to the system.

4.7.1 Achievements

In the present context the system is working with the untagged data but soon the interface of the POS Tagger will be added. As already mentioned Odia language is with both agglutinative as well as inflectional features. The system is producing correct result for both concatenated as well as synthetic Morphology.

For example:

Concatenated form : biɖjɑlɔjɔʈʰɑru

The result produced on the front end is in the form of a table with the following results

Processed word → biɖjɑlɔjɔʈʰɑru (ବିଦ୍ୟାଳୟଠାରୁ)

Root word → biɖjɑlɔjɔ(ବିଦ୍ୟାଳୟ)

Nominal inflection →ʈʰɑru(ଠାରୁ)

Surface form of inflection →ʈʰɑru(ଠାରୁ)

POS→N

Meaning→ LOC_Sg

Exception Condition→NIL

Synthetic forms are present in both nouns and verbs but the dominance has been seen in verbal inflection. The following is an example of an inflected verb with synthesis of phonetic sounds.

kʰeɭuɔchi (ଖେଳୁଅଛି)

Processed word →kʰeɭuɔchi (ଖେଳୁଅଛି)

Root word →kʰeɭ (ଖେଳ)

Verbal inflection →uɔchi (ଉଅଛି)

Surface form of inflection → (ୁଅଛି)

POS→V

Person→1st

Number →Sg

Tense →PRS

Aspect →PROG

Mood →Indicative

Voice →Active

4.7.2 Issues and Challanges

The data validation by the developed system encountered with some issues. The issues are enlisted below:

I. Odia has the same accusative/dative inflectional marker/-ɳkɔ/. This marker is being used for both the singular and plural form. So the system is showing ambiguous result while dealing with the inflectional marker/-ɳkɔ/. There are a few other constructions formed in combination with /-ɳkɔ/ are also showing ambiguous result. The other inflectional forms which are formed in combination of other markers are /-ɳkɔrɔ/,/-ɳku/,/-ɳkɔʈʰi/,/-ɳkɔpaĩ/, /-ɳkɔnimɔnʈ̪e/, /-ɳkɔʈʰiki/,/-ɳkɔʈʰiku/.

II. There are a few words (noun) which do not possess any nominal inflections but the words include such endings which has the same phonetic combination equal to the inflectional form of a noun. So the system is also processing those words to produce the respective inflection markers.

For example:

The tool is taking the last letters of *gʰimiri(prickly heat)* to be inflectional marker /-ri/. /-ri/ is a genitive singular case marker in Odia.

III. /-re/ is an inflectional marker which is at once used for locative case, instrumental case and also as a vocative address particle. So the result provided by the Analyzer is producing ambiguous result. However the datafile has enlisted the exceptions and conditions for each particle if any, written against it. This will help in resolving the ambiguities emerged by such particles.

CHAPTER 5

CONCLUSION

This chapter gives an overview of the current undertaken research. This chapter can be devided into two sections. The first section is a representation of the summary of the contents covered through the previous four chapters and the second section is talking about the future perspective of the research.

5.1 Summing up the Research

The first chapter has dealt with the enlistment and description of the various approaches and techniques available for building Morph Analyzer and has also given a major focus on the various works done to develop Morph Analyzers in foreign languages and Indian languages with special emphasis on Odia till date. The second chapter has dealt with the research methodology for data collection, analysis and preparation of an algorithm for developing the Odia Inflectional Morph Analyzer. The third chapter encapsulates the Morphological phenomenon for Inflection formation with special reference to noun and verb.

The fourth chapter has depicts the computational implementation of the algorithm written in second chapter. This chapter has also given a comprehensive description of the system architecture and has discussed the output produced by the developed system.

The current chapter is the final chapter which is dealing with the limitation as well as the future perspective of the research.

5.2 Description of the developed tool

The tool developed for analysing the Odia inflectional Morphology by rule based approach is done through a hand crafted dictionary. The dictionary contains the nominal and verbal inflection forms with appropriate examples, exceptions and the conditions for the occurence of such exceptions. It has tried to incorporate almost all the possible rules for the inflection formation for nouns and verbs in Odia.

The system developed is an online system which runs on Apache Tomcat Platform. This has used the Odia POS Tagger as an interface before the Morphological analysis. It has tried to provide the inflectional analysis of all Odia noun and verb forms within the constraints of the topic. The aim behind the attempted study was to create an online tool for Morph analysis in Odia which will be freely available on the web and can run on any operating system with the sole condition that there should be a Java environment.

5.3 Limitations of the system

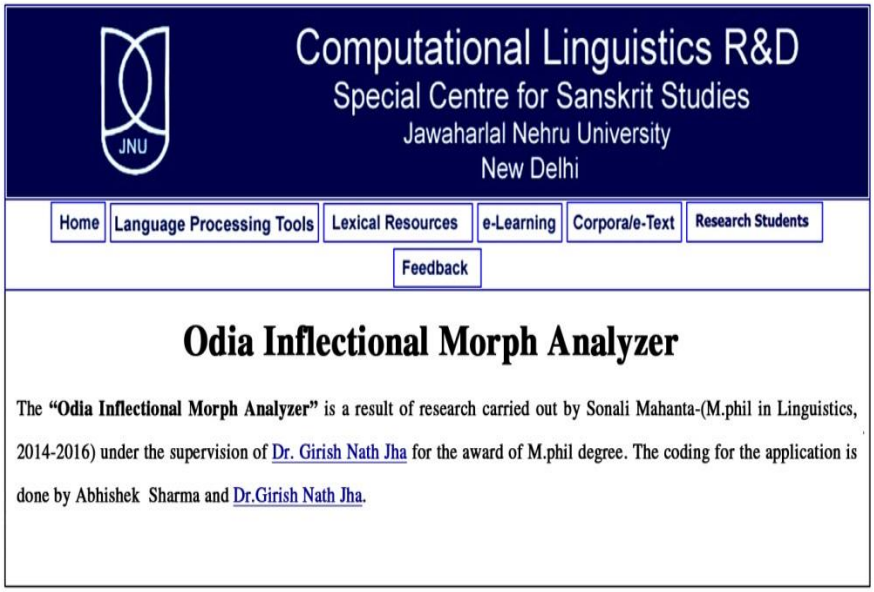
I. This system has been developed as a rule based system which works on string manipulation and dictionary. So it is eviedent that the system is dependent on the data and the rules. If the table lacks any of the rules for forming the nominal or verbal inflecion then the tool cannot generate the desired result.

II. As it is already mentioned that the POS tagger is the computational interface before the Morphological analysis of odia inflections. So if the POS tagger produces any error, the Morphological analysis of the system will also suffer accordingly.

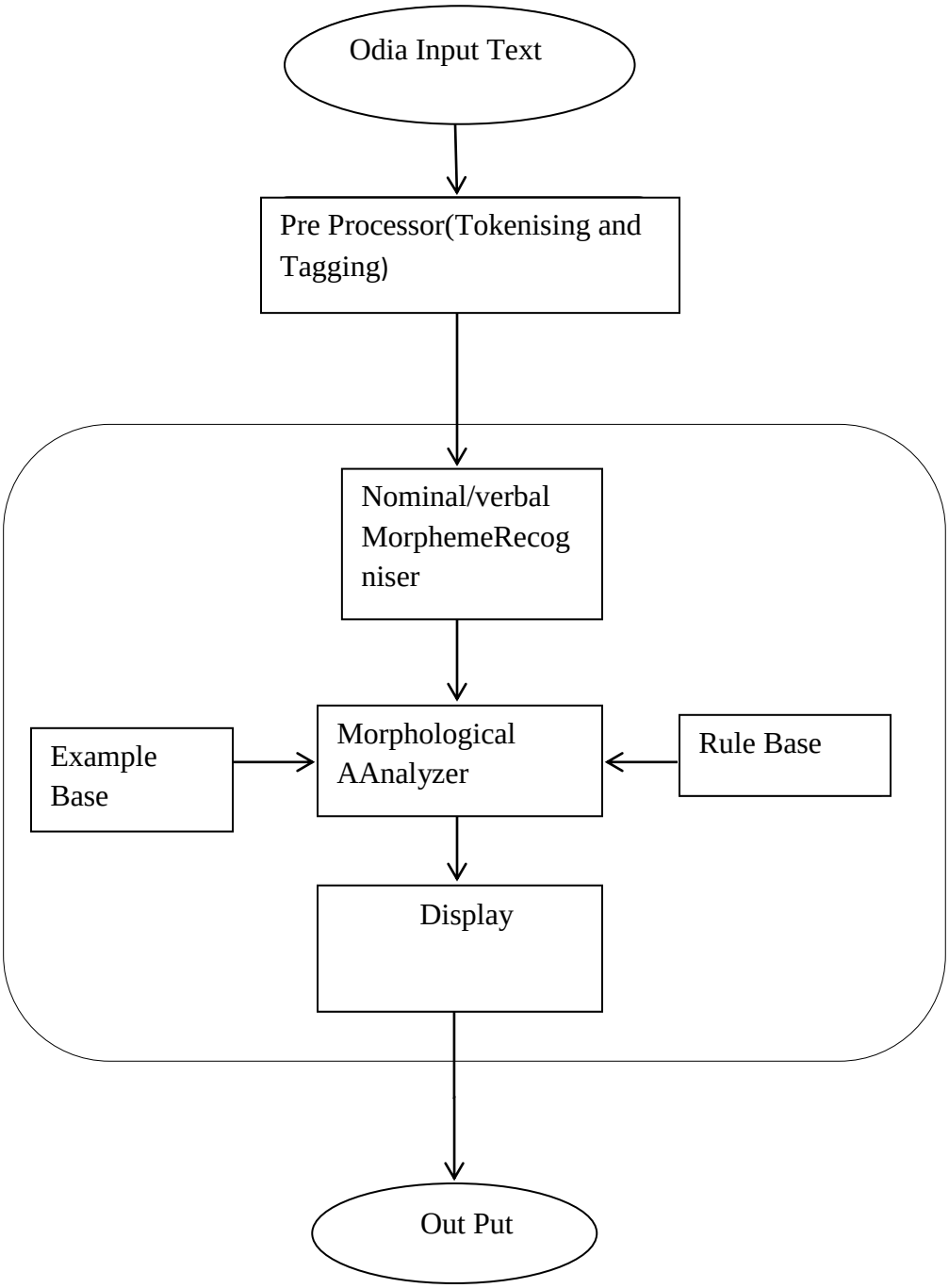
5.4 Scope for future research and development

In the present context the Odia Morph Analyzer is handling the inflectional forms of noun and verb only. However the analysis coverage of the Morph Analyzer will be extended to the other inflectional categories with in the circumference of the present research work. For further research the computational analysis can also be extended to derivational and compound Morphology of Odia. Moreover the dissertation is dealing with the inflectional forms of nouns and verbs only, so the causative derivations, complex verb predicates like compound verb and conjunct verb as well as the serial verb constructions are not in the discussion. However the passivation of the active speech counterpart is included. This Morph Analyzer will hopefully be a supplement for other NLP applications in Odia language.

Appendix I


Computational Linguistics R&D
Special Centre for Sanskrit Studies
Jawaharlal Nehru University
New Delhi
JNU
Home
Language Processing Tools
Lexical Resources
e-Learning
Corpora/e-Text
Research Students
Feedback
Odia Inflectional Morph Analyzer
The "Odia Inflectional Morph Analyzer" is a result of research carried out by Sonali Mahanta-(M.phil in Linguistics, 2014-2016) under the supervision of Dr. Girish Nath Jha for the award of M.phil degree. The coding for the application is done by Abhishek Sharma and Dr.Girish Nath Jha.
Enter Odia text for analysis
Cut and paste sample data from here
Submit
Reset
Result

Appendix II


Odia Input Text
Pre Processor(Tokenising and Tagging)
Nominal/verbal MorphemeRecog niser
Example Base
Morphological AAnalyzer
Rule Base
Display
Out Put

Appendix III

Nominal Inflection	**Surface form of inflection**	**POS**	**Meaning**	**Example in words**	**Exception**	**Condition**
ଏ	େ	N	Nom_Pl	ପିଲେ	(i)ଆଖିଏ) ii)ଘରେ	'ଏ' is used for(i) derivation(ii)locative case
ମାନ	ମାନ	N	Nom_Pl	ହାତୀମାନ	(i)ଭାସମାନ)ii)ବୁଦ୍ଧିମାନ)iii) କଥାରମାନ	ମାନ' is used to form meaning(i)floating(ii)male gender(iii)standard
ମାନେ	ମାନେ	N	Nom_Pl	ବାଘମାନେ	(i) କଥାମାନେନାହିଁ)ii)କଥାରମାନେବୁଝିଗଲି	In example(i) it means 'obey' and in (ii)meaning
ଗୁଡ଼ା	ଗୁଡ଼ା	N	Nom_Pl	ଫୁଲଗୁଡ଼ା	NIL	NIL
ଗୁଡ଼ି	ଗୁଡ଼ି	N	Nom_Pl	ବହିଗୁଡ଼ି	NIL	NIL
ଗୁଡ଼ିକ	ଗୁଡ଼ିକ	N	Nom_Pl	ଘରଗୁଡ଼ିକ	NIL	NIL
ଗୁଡ଼ାକ	ଗୁଡ଼ାକ	N	Nom_Pl	ପିଲାଗୁଡ଼ାକ	NIL	NIL
ଗୁଡ଼ିଏ	ଗୁଡ଼ିଏ	N	Nom_	ପକ୍ଷୀଗୁଡ଼ି	NIL	NIL

			Pl	ଏ		
ଗୁଡ଼ାଏ	ଗୁଡ଼ାଏ	N	Nom_Pl	କଥାଗୁଡ଼ାଏ	NIL	NIL
ତକ	ତକ	N	Nom_Pl	ବହିତକ	NIL	NIL
ଯାକ	ଯାକ	N	Nom_Pl	ପକ୍ଷୀଯାକ	NIL	NIL
ଟାଯାକ	ଟାଯାକ	N	Nom_Pl	ଗୋଠଟାଯାକ	NIL	NIL
ଟିଯାକ	ଟିଯାକ	N	Nom_Pl	ଘରଟିଯାକ	NIL	NIL
କୁ	କୁ	N	Acc_Sg	ପୁଅକୁ	NIL	NIL
କି	କି	N	Acc_Sg	କବିକି	(i)ଘଡ଼ିକିଘଡ଼ି)ii)ଶୋଇକି)iii)କି	କି' used(i)for reduplication(ii)with verb(iii)as an interrogative marker
ତେ	ତେ	N	Acc_Sg	ତୋତେ	NIL	NIL
ଟିକୁ	ଟିକୁ	N	Acc_Sg	କଥାଟିକୁ	NIL	NIL
ଟାକୁ	ଟାକୁ	N	Acc_Sg	ଘରଟାକୁ	NIL	NIL
ଙ୍କୁ	ଙ୍କୁ	N	Acc_Sg	ଦିଅଙ୍କୁ	NIL	NIL
ଠାକୁ	ଠାକୁ	N	Acc_S	ତା'ଠାକୁ	NIL	NIL

			g			
ଙ୍କଠାକୁ	ଙ୍କଠାକୁ	N	Acc_Sg	ତାଙ୍କଠାକୁ	NIL	NIL
ଠିକି	ଠିକି	N	Acc_Sg	ସେଠିକି	NIL	NIL

Appendix IV

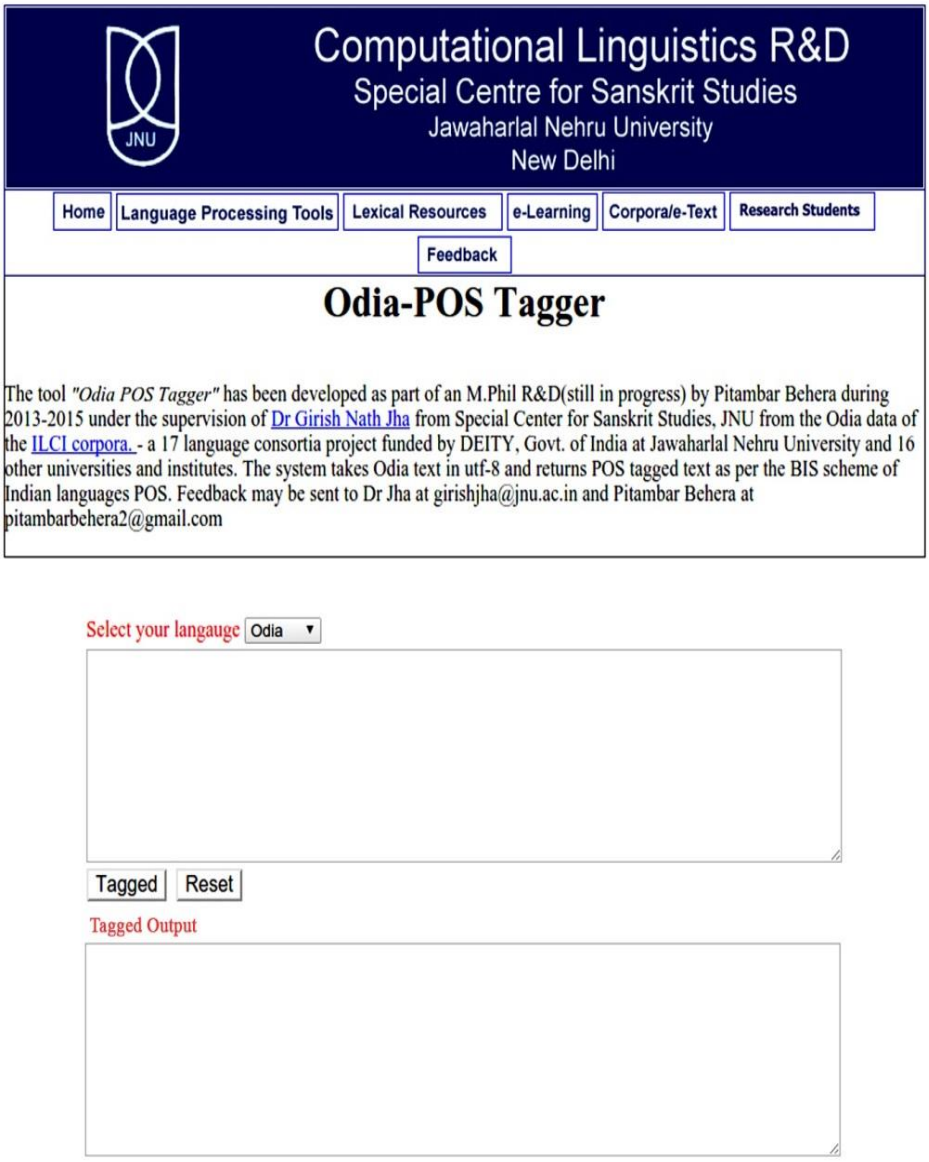
Verbal Inflection	**Surface Form of Inflection**	**Example in words**	**POS**	**Person**	**Number**	**Tense**	**Aspect**	**Honour/ Degree of respect**	**Mood**	**Voice**
ଏ	ଏ	ଖାଏ	V	1st	Sg	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଏ	େ	ଖେଳେ	V	1st	Sg	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଉ	ଉ	ଖାଉ	V	1st	Pl	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଉ	ୁ	ପଢୁ	V	1st	Pl	PRS	SIM	NIL	INDICATIVE	ACTIVE
ଉଛି	ଉଛି	ଖାଉଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛି	ୁଛି	ପଢୁଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE

ଉଅଛି	ଉଅଛି	ଖାଉଅଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛି	ୁଅଛି	ଲେଖୁଅଛି	V	1st	Sg	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛୁ	ଉଛୁ	ଖାଉଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛୁ	ୁଛ	ଲେଖୁଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛୁ	ଉଅଛୁ	ଖାଉଅଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛୁ	ୁଅଛୁ	ନାଚୁଅଛୁ	V	1st	Pl(Exclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛେ	ଉଛେ	ଖାଉଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଛେ	ୁଛେ	ଲେଖୁଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛେ	ଉଅଛେ	ଖାଉଅଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE
ଉଅଛେ	ୁଅଛେ	ପଢୁଅଛେ	V	1st	Pl(Inclusive)	PRS	PROG	NIL	INDICATIVE	ACTIVE

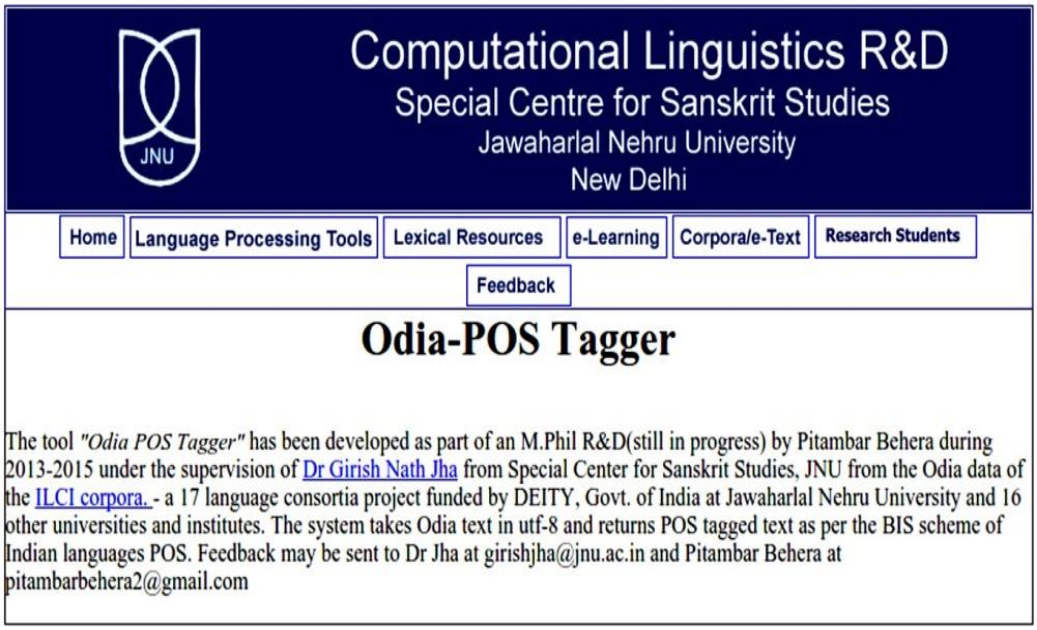
ଇଛି	ଇଛି	ଖାଇଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଛି	ିଛି	ନାଚିଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଅଛି	ଇଅଛି	ଖାଇଅଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଅଛି	ିଅଛି	ପଢିଅଛି	V	1st	Sg	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଛୁ	ଇଛୁ	ଖାଇଛୁ	V	1st	Pl(Exclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଛୁ	ିଛୁ	ପଢିଛୁ	V	1st	Pl(Exclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଅଛୁ	ଇଅଛୁ	ଖାଇଅଛୁ	V	1st	Pl(Exclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଅଛୁ	ିଅଛୁ	ପଢିଅଛୁ	V	1st	Pl(Exclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଛେ	ଇଛେ	ଖାଇଛେ	V	1st	Pl(Inclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE
ଇଛେ	ିଛେ	ପଢିଛେ	V	1st	Pl(Inclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE

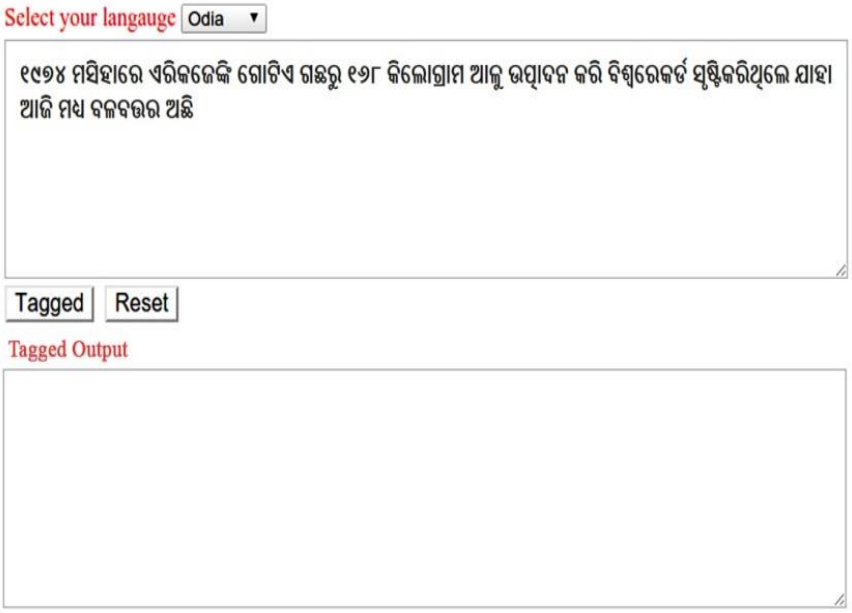
ଇଅଛେ	ଇଅଛେ	ଖାଇଅଛେ	V	1st	Pl(Inclusive)	PRS	PRF	NIL	INDICATIVE	ACTIVE

AppendixV

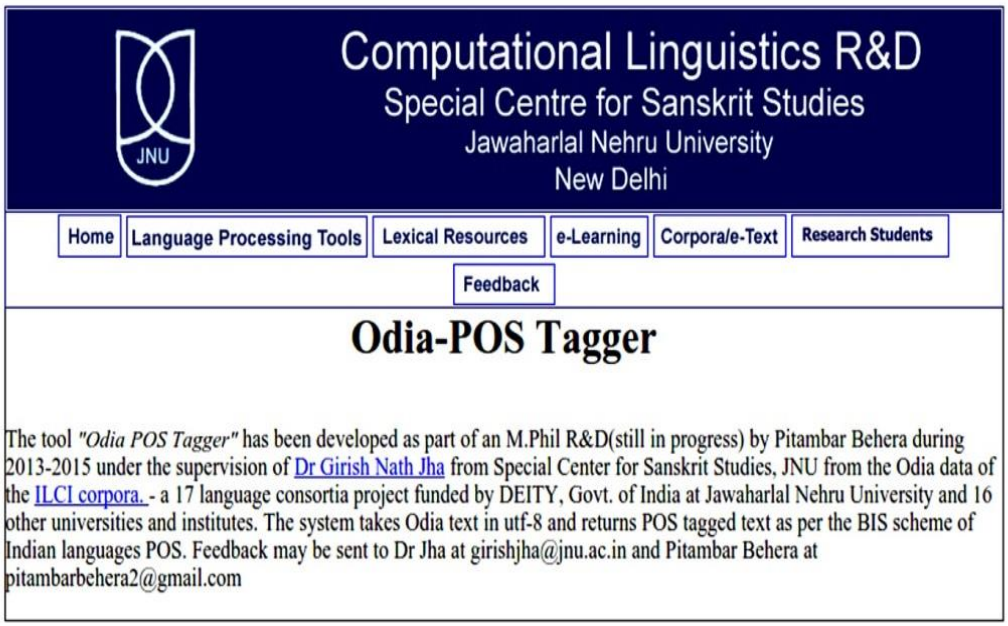


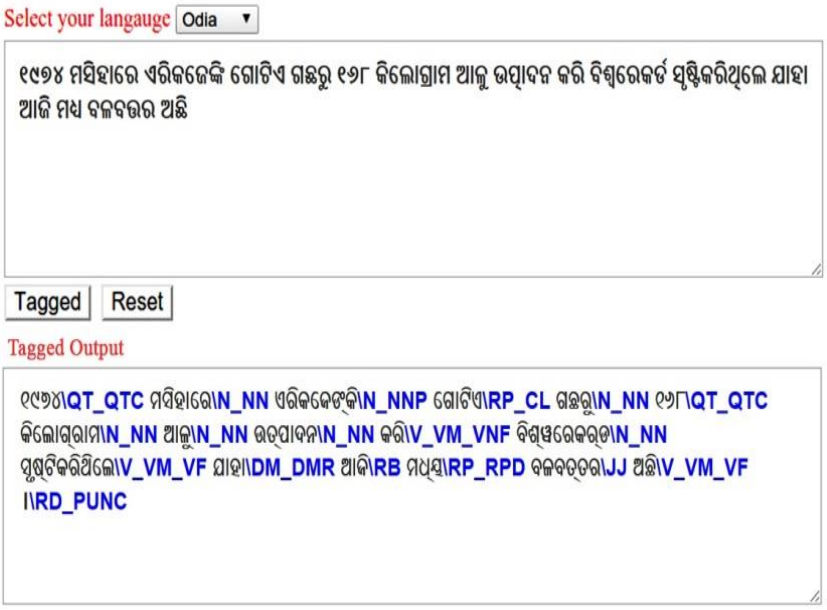
Appendix-VI


Computational Linguistics R&D
Special Centre for Sanskrit Studies
Jawaharlal Nehru University
New Delhi
Home
Language Processing Tools
Lexical Resources
e-Learning
Corpora/e-Text
Research Students
Feedback
Odia-POS Tagger
The tool "Odia POS Tagger" has been developed as part of an M.Phil R&D(still in progress) by Pitambar Behera during 2013-2015 under the supervision of Dr Girish Nath Jha from Special Center for Sanskrit Studies, JNU from the Odia data of the ILCI corpora. - a 17 language consortia project funded by DEITY, Govt. of India at Jawaharlal Nehru University and 16 other universities and institutes. The system takes Odia text in utf-8 and returns POS tagged text as per the BIS scheme of Indian languages POS. Feedback may be sent to Dr Jha at girishjha@jnu.ac.in and Pitambar Behera at pitambarbehera2@gmail.com

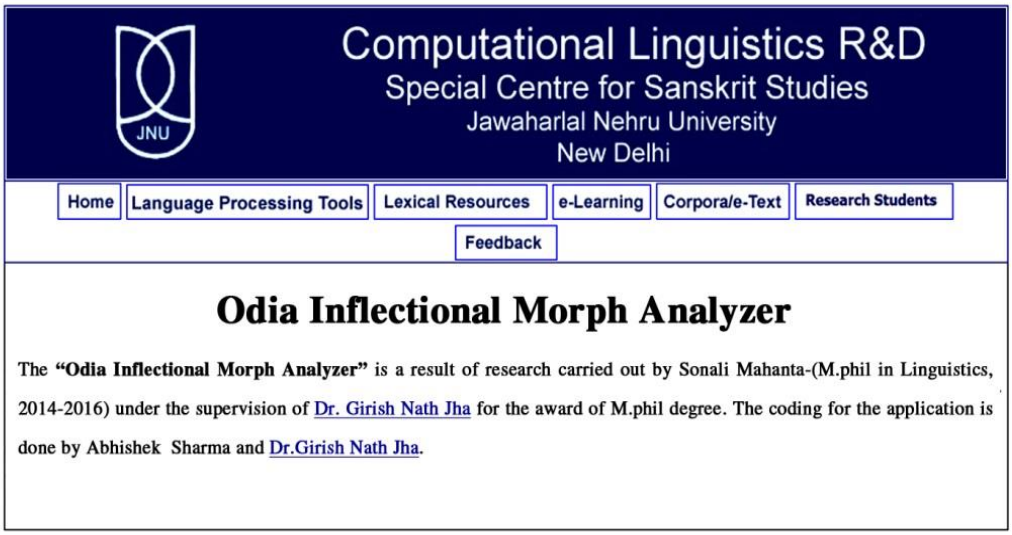

Select your langauge
Odia
୧୯୭୪ ମସିହାରେ ଏରିକଜେଙ୍କି ଗୋଟିଏ ଗଛରୁ ୧୭୮ କିଲୋଗ୍ରାମ ଆଳୁ ଉତ୍ପାଦନ କରି ବିଶ୍ୱରେକର୍ଡ ସୃଷ୍ଟିକରିଥିଲେ ଯାହା ଆଜି ମଧ୍ୟ ବଳବତ୍ତର ଅଛି
Tagged
Reset
Tagged Output

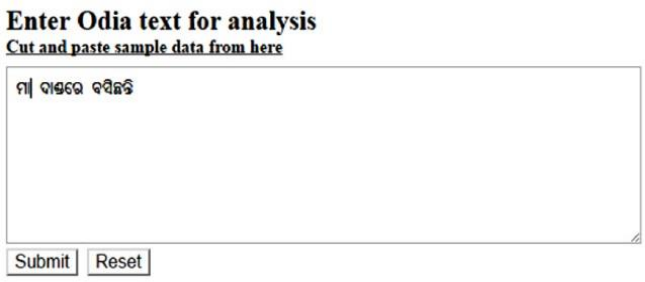
Appendix VII





Appendix VIII





Appendix IX

Cut and paste sample data from here

ମା ଦାଣ୍ଡରେ ବସିଛନ୍ତି

Submit Reset

Result

Nominal Inflection

Processed words	Root Word	Nominal Inflection	Surface form of inflection	POS	Meaning	Exception	Condition
ଦାଣ୍ଡରେ	ଦାଣ୍ଡ	ରେ	ରେ	N	Loc_Sg	(i)ଜାନୁରେ(ii)ଛୁରାରେ	'ରେ' stands for(i) vocative case(ii) instrumental case

Verbal Inflection

Processed words	Root Word	Verbal Inflection	Surface Form of Inflection	POS	Person	Number	Tense	Aspect	Honour/Degree of respect	Mood	Voice
ବସିଛନ୍ତି	ବସ	ଇଛନ୍ତି	ିଛନ୍ତି	V	2nd	Sg	PRS	PRF	Formal	INDICATIVE	ACTIVE

No Inflection

BIBLIOGRAPHY

Booksand Encyclopedias:

Abbi,A.(2001).*A manual of Linguistic Fieldwork and Structures of Indian Languages*(Vol. 17). Lincom Europa.

Anderson, John M. (1971).The grammar of case: towards a localistic theory.Cambridge University Press: Cambridge.

Bharati, A., Chaitanya, V., & Sangal, R. (2004).*Natural language Processing: A Paninian Perspective.*Prentice Hall of India Private Limited.New Delhi.

Brown,K.(2006). Encyclopedia of Language and Linguistics.National Research council of Canada:Elsevier

Comrie, B. (1976).*Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*.Cambridge: Cambridge University Press.

Fasold, R. &Connor-Linton, J. (2006).*An Introduction to Language and Linguistics*. New York: Cambridge University Press.

Mahapatra,B.P.(1999). Pandit Nilakanthanka Bhasatatwa.Cuttack:Friends Publishers. Press.New York.

Mahapatra,B.P. (2007).*Prachalita Odia Bhasara ek Byakarana[Standard Kathita o Likhita]*.Cuttack:Vidyapuri.

Mahapatra, D. (2013).*Adhunika odia byakarana*.Cuttack: Kitab Mahal.

Mahapatra,D.(2003).*Odia Bhasatatwa :bichar O bishlesana* .Cuttack:Bijayini Publication.

Mitkov,R.(2003). *The Oxford Handbook of computational Linguistics*. Oxford University Press, New York

Mishra, B.(1962). *Odia bhasara puratatwa*.Cuttack: Friends Publisher.

Mishra,G.C.(1973).*Bhasa Bigyan Paichaya.* Cuttack: Friends Publishers.

Mitkov, R. (2003).*The Oxford Handbook of Computational Lingustics.*Oxford University Press, New York.

Mohapatra,D.(1977).*Sarala Mahabharatara Bhasatatwika Alochana*.Cuttack:Friends Publishers Binod Behari.

Mohanty, B. (1998).*Odia Byakarana Bichar*.*Bhubaneswar*: Odisha Samabaya Samiti ltd.

Mohanty,B.(2008).*Bhasa Bhabana odia bhasa alochana osankalana*.Bhubaneswar.Vidya.

Mohanty, P. (2007).*Brutti e mo poshe kutumba, Odia bhasa baigyanika carccara nutana diganta (A linguistic analysis of different aspects of the Oriya language).*Bhubaneswar:Prajutsu Mohanty, Prabuddha Mohanty.

Mohanty,S.C.(2015).*Prayogatmaka Odia Grammar*.Cuttack:A.K. Mishra Publishers private limited.

Pradhan,Das,Mohanty,Sadangi,Srichandansingh.(1998).*Odia byakarana astam,nabam o dasam shreni*.cuttack.madhyamika sikhya parishad Odisha.

Sober, M. M., & Benedito, J. R. M. (Eds.). (2010). *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*.Information Science Reference.

Nanda,B.,Nanda,R.(2002).*Odia Byakarana sankalana* .Bhubaneswar:Mahabir Prakashan

Katamba,F.(1993). *Morphology*. Palgrave Macmillan

Articles and Papers:

Akilan, R., & Naganathan, E.R. (2015).*Morphological Analyzer for classical Tamil text: a rule based approach* .vol.10, No.20.ARPN Journal of Engineering and AppliedSciences. P.9326.Retrieved from http://www.arpnjournals.org/jeas/research_papers/rp_2015/jeas_1115_2876.pdf

Gridach,M.&Chenfour,N.*"Design and Realization of an Arabic Morphological Automaton: New Approach for Arabic Morphological Analysis and Generation*", International Journal of Computer Science Issues (IJCSI)

Goyal,V.,&Singh Lehal, G.(2008). *Hindi Morphological Analyzer and Generator*, P. 1156–1159. IEEE Computer Society Press.California.USA.Retrieved from: http://www.academia.edu/482680/Hindi_Morphological_Analyzer_and_Generator

Jena.I.,Chaudhury, S.,&Sharma, H. (2011) Developing*Oriya Morphological Analyzer Using Lt-Toolbox*.P.124-129.
Retrieved from:http://link.springer.com/chapter/10.1007%2F978-3-642-19403-0_20

Jayan,J.P.,Rajeev,R.R.,&Rajendran,S.(2009).*Morphological Analyzer for Malyalam-A comparision of different Approaches*.IJCSIT.Vol.2,No.2.P.157.
Retreieved from:http://www.iiitmk.ac.in/vrclc/images/Documents/IJCITMORPH.pdf

Jha,GirishN.(1995). *Proposing a computational system for Nominal Inflectional Morphology in Sanskrit*. Proceedings of national seminar on recognition of Sanskrit Shastras with a view to prepare their computational database

Jha, Girish N.(2007).Introduction to Computational Morphology.*Lecture in PGDMLT course*
Jha,GirishN.(2010).The TDIL program and the Indian Language Corpora Initiative(ILCI)

Kanuparthi, N.,Inumella, A.,&Misra Sharma,D.(2012) *Hindi Derivational Morphological Analyzer*. Proceedings of the Twelfth Meeting of the Special Interest Group on Computational Morphology and Phonology(SIGMORPHON2012),P.10–16,Montreal, Canada, June 7, 2012. Retrieved from http://www.aclweb.org/anthology/W12-2302

Kumar, D., Singh, M.,& Shukla.,S.(2012).*FST based Morphological Analyzer for Hindi Language*International Journal of Computer Science Issues (IJCSI). Retrieved from:http://arxiv.org/abs/1207.5409

M. Kumar,A., Dhanalakshmi, V., Soman, K. P.,&Rajendran, S.(2010).*A sequence labeling approach to Morphological Analyzer for Tamil language*.International Journal on Computer Science and Engineering, vol.2.
Retrieved from:https://www.amrita.edu/publication/sequence-labeling-approach-Morphological-Analyzer-tamil-languag

Mohanty, S., Santi, P.K., Das Adhikary, K.P. (2004) *Analysis and Design of Oriya Morphological Analyzer: Some Tests with OriNet*. P.118.Retrieved from: http://www.ciil-ebooks.net/html/simple/total.pdf

Nair, Latha R., and S. David Peter.(2011) *"Development of a rule based learning system for splitting compound words in Malayalam language*" IEEE Recent Advances in Intelligent Computational Systems

Parakh, M.,& Rajesh,N.(2011).*Developing Morphological Analyzer for four Indian languages using a rule based Affix stripping approach*.Linguistic data cosortium for Indianlanguages CIIL Mysore Special volume :Problems in parsing Indian languages. Retrieved from http://www.languageinindia.com/may2011/monarajesha.pdf

Patnaik, B. N. (2001). Nominative and non-nominative constructions in Oriya.
Retrieved from http://home.iitk.ac.in/~patnaik/documents/nnom.pdf

S. Kiranmai, M Kumar, A., Dhanalakshmi, Vb, Soman, K. Pb, V.V, D., R., V., V., S., K.G., S., H.A., A., M., H., D.K., T., A., M., B., J., E., A., &J., B .(2010).*Morphological Analyzer for Telugu Using Support Vector Machine*,Communication and information science.Vol.101P.430-433.
Retrievedfrom:https://www.amrita.edu/publication/Morphological-Analyzer-telugu-using-support-vector-machine

S. Rajendran.(2006). *Parsing In Tamil - Present State of Art*.In journals of Language in India.Vol. 6:8, August.Retrieved from:

https://www.researchgate.net/publication/258665032_Parsing_in_Tamil_Present_State_of_Art.

Sahoo, K. (2004).*A two level Morphological processing of Oriya*. Lecture Compendium: Symposium on Indian Morphology.Phonology Language Engineering.March (19-24)

Sahoo,K. (2004).*A two level processing in Odia*.P.34to37.Retrieved from: http://www.ciil-ebooks.net/html/simple/total.pdf

Sahoo,K.(2016). *Multi-Verb Constructions - Parsing with a Deterministic Finite State Automaton.*International Journal of Computational Linguistics Research.Vol.7.Retrievedfrom:http://www.dline.info/jcl/fulltext/v7n1/jclv7n1_3.pdf.

Sethi,D.P. (2013). Morphological Analyzer for Sambalpuri Odia dialect inflected verbal forms. Vol. 3, Issue 10. *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*. Retrieved from:http://www.ijarcsse.com/docs/papers/Volume_3/10_October2013/V3I10-0267.pdf

Sethi,D.P. (2014).A survey on Odia computational Morphology. Vol. 3, Issue. 3, (pp. 623-625). *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*. Retrieved from:http://ijarcet.org/wp-content/uploads/IJARCET-VOL-3-ISSUE-3-623-625.pdf

Shaji.A,&Sindhu,P.Morphological Analyzer for Malayalam: A Literature Survey Vol.107.No.14.P.24.*International Journal of Computer Applications (0975 – 8887).* Retrieved from:http://research.ijcaonline.org/volume107/number14/pxc3900231.pdf

Swain,P.K.,Mohanty,S., &Santi, P.K.(2005). Second Symposium on Indian Morphology, Phonology &Language Engineering, Kharagpur: Indian Institute of Technology.P.86-89

V. Pa Abeera, Aparna, Sa, Rekha, R. Ua, M Kumar, A., Dhanalakshmi, Va, Soman, K. Pa,&Rajendran,Sb.(2012).*Morphological Analyzer for Malayalam using machine learning*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 6411LNCS.P.252-254. Retrievedfrom:https://www.amrita.edu/publication/Morphological-Analyzer-malayalam-using-machine-learning

Computational Morphology and Sanskrit.P.15
Retrieved fromhttp://sanskrit.jnu.ac.in/rstudents/mphil/muktanand/Chapter%201.pdf

Tapaswi, N., and S. Jain.(2013) "*Knowledge representation of grammatical constructs of Sanskrit Language and modular architecture of ParGram*",International Conference on Advances in Technology and Engineering (ICATE)

Manji Bhadra."*Sanskrit Analysis System (SAS)"*, (2009) Lecture Notes in Computer Science.

Internet sources:

http://sanskrit.jnu.ac.in/rstudents/mphil/subhash.pdf [Accessed on10-Nov-2015]

http://ijarcet.org/wp-content/uploads/IJARCET-VOL-3-ISSUE-3-623-625.pdf[Accessed on18-Sep-2014]
http://research.ijcaonline.org/volume107/number14/pxc3900231.pdf[Accessed on 05-May-2016]
http://www.lrec-conf.org/proceedings/lrec2010/pdf/874_Paper.pdf [Accessed on16-Dec-2012]
http://aclweb.org/aclwiki/index.php?title=Morphology_software[Accessed on24-Apr-2016]

http://sanskrit.jnu.ac.in/rstudents/mphil/muktanand/Chapter%201.pdf[Accessed on 05-Nov-2011]

https://www.amrita.edu/publication/Morphological-Analyzer-malayalam-using-machine-learning[Accessed on19-Jul-2016]

http://www.ciil-ebooks.net/html/simple/total.pdf[Accessed on 3-Jul-2016]

http://sanskrit.jnu.ac.in/rstudents/mphil/diwakar-mani.pdf [Accessed on 10-Jul-2014]

https://nlp.fi.muni.cz/polish-Morphology-Analyzer/[Accessed on 19-Jul-2016]

http://sanskrit.uohyd.ernet.in/scl/Morph/index.html[Accessed on 26-Dec-2013]

http://aclweb.org/anthology/W/W15/W15-32.pdf[Accessed on 22-Apr-2016]

http://anthology.aclweb.org/C/C14/C14-2009.pdf[Accessed on15-May-2015]

http://www.science.gov/topicpages/t/tunisian+arab+population.html[Accessed on 16-Oct-2015]

http://www.slideshare.net/ijnlc/kridanta-analysis-for-sanskrit[Accessed on 27-Feb-2016]

http://www.arpnjournals.org/jeas/research_papers/rp_2015/jeas_1115_2876.pdf[Accessed on 03-Jul-2016]

http://aclweb.org/anthology/P/P14/P14-2042.pdf[Accessed on10-Mar-2016]

http://ltc.amu.edu.pl/book/papers/LRL-13.pdf[Accessed on 28-May-2016]

http://www.proclass.in/category/1181[Accessed on19-May-2015]

http://www.martindalecenter.com/Language_3_India.html[Accessed on11-Oct-2015]

http://www.lrec-conf.org/proceedings/lrec2014/pdf/593_Paper.pdf [Accessed on15-May-2016]

http://www.lotpublications.nl/Documents/180_fulltext.pdf[Accessed on 22-Mar-2016]

http://dblp.dagstuhl.de/pers/hc/d/Dhanalakshmi:V= [Accessed on 30-May-2016]

http://www.definitions.net/definition/dative%20case[Accessed on 27-Dec-2013]

http://www.dline.info/jcl/fulltext/v7n1/jclv7n1_3.pdf[Accessed on24-Apr-2016]

http://www.lrec-conf.org/proceedings/lrec2014/workshops/LREC2014Workshop-WILDRE%20Proceedings.pdf [Accessed on20-Jul-2015]

http://aclweb.org/anthology/P13-2[Accessed on 13-Apr-2016]

http://www.visitodisha.org/Culture[Accessed on3-Dec-2015]

http://en.wikipedia.org/wiki/White-winged_chough[Accessed on26-May-2016]

http://tomcat.apache.org/[Accessed on 29-May-2016]

http://www.oracle.com/technetwork/javajavaee/jsp/index.html [Accessed on 23-May-2016]

Dissertations

Behera, P.(2015).Odia POS Tagging Corpora Suitability of Statistical Models. M.phil.dissertation,Center for Linguistics,JNU.,New Delhi

Banerjee, E. (2011). *Intra-phrasal Agreement within the Hindi Noun Phrase: A Grammar Checker Approach*. M.Phil. Dissertation, Centre for Linguistics, J.N.U., New Delhi

Chaudhuri,N.K.(2006).Developing a Computational Framework for the Verb Morphology of Great Andamanese. M.phil.dissertation,Center for Linguistics,JNU.,New Delhi

Majhi, T. D. (2007). *Descriptive Oriya Morphology in the Paninian Model*.Ph.D. Thesis, Centre for Linguistics, JNU, New Delhi.

Nioding,A.(2011).Issues and Challenges in Marathi Word Segmenter: With a special focus on Sandhi.M.phil.dissertation.Center for Linguistics,JNU,New Delhi

Subash.(2006).Machine Recognition and Morphological Analysis of subanta padas.M.phil.dissertationSpecial center for Sanskrit studies, JNU,New Delhi