Automatic Identification and Analysis of Verb Groups in Hindi

Thesis submitted to

Jawaharlal Nehru University

in partial fulfilment of the requirements for the award of the Degree of

DOCTOR OF PHILOSOPHY

NARAYAN KUMAR CHOUDHARY



Centre for Linguistics

School of Language, Literature & Culture Studies,

Jawaharlal Nehru University, New Delhi, INDIA

2011

DECLARATION BY THE CANDIDATE

Date 19th August, 2011

This thesis titled **"Automatic Identification and Analysis of Verb Groups in Hindi"** submitted by me for the award of the degree of Doctor 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.

(Narayan Kumar Choudhary)

Ph. D. Student

Centre for Linguistics,

SLL&CS,

JNU

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

Date: 19th August, 2011

CERTIFICATE

This thesis titled "***Automatic Identification and Analysis of Verb Groups in Hindi***" submitted by **Mr. Narayan Kumar Choudhary** to the Centre for Linguistics, School of Language, Literature and Culture Studies, Jawaharlal Nehru University, New Delhi, for the award of the degree of **Doctor 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 Doctor of Philosophy.

(PROF. PRAMOD KS PANDEY)
SUPERVISOR

(DR. GIRISH NATH JHA)
CO-SUPERVISOR

(PROF. FRANSON MANJALI)
CHAIRPERSON

DEDICATION

to the service
of humanity
and
the less resourced lanaguages
that need to be enabled with the new technology

ACKNOWLEDGEMENTS

No work is complete by the effort of just one person. This thesis is no exception to it. Efforts of a lot of people either directly or through their contribution at a third party entity have been utilized in writing this thesis. It is not possible to acknowledge all of such people in this small write-up. However, in particular I would like to mention as follows.

First of all, I would like to thank my supervisor Prof. Pramod Pandey who has been a constant support and always had total faith in me. Without his honest advice I would not have completed this in the time frame I have. His faith in me excels that of my own. I had planned to deregister after devoting four years into this thesis and was determined to deregister, fearing that I might not complete it in time. But he believed in me and encouraged me to take the extension of one year. The result is presented. Thanks a lot Sir!

Dr. Girish Nath Jha, my co-supervisor, has been an inspiration for many students of linguistics to work in the area of computational linguistics and look for a career in it. I have always tried to follow his path. Despite being a student from the literature background, it was through his inspiration and support that I dared to come to the field of language computing. I always felt him at the back when in need. My 9 year old association with him and with the computer has paid off huge dividends, both professionally and academically. In the past few years, I saw myself grow up both as a professional and as an academic. Thanks a lot Sir! It has been quite rewarding being with you all these years and I wish I will continue to receive your blessings afterwards as well.

Thanks also go to both Dr. Jha and Microsoft Research India (MSRI) (especially Kalika Bali) for having faith in me while preparing the Hindi corpus in the IL-PoST framework. I gained a lot from the annotations I did for this corpus.

Thanks are due to the various members of the Indian Language Corpora Initiative (ILCI) project as well. My colleages (specially, Priyanka Singh, Avanti Nioding, Ritesh Kumar, Esha Banerjee and Akanksha Bansal, to name a few) and other members at the lab of the Special Centre of Sanskrit Studies, JNU (especially, Diwakar Mishra, Dr. Umesh Kumar Singh and Dr. Vibhuti Jha, to name just a few) have also been an inspiration for me. Association with them made me

believe that there is a lot of work to be done in the area of language technology of Indian languages and this inspired me to devote myself fully to it, sometimes at the cost of deviating a bit from what we call “core linguistics”.

Thanks also go to two companies whom I have consulted on their projects related to Hindi: Keypoint Technologies, Inc. and L10nbridge. My association with them gave me an exposure to the current practices of the corporate sector and the resources they have. This instilled some confidence in me in the statements that I make in the thesis and made me believe that what I am doing is also the need of the day and can be taken up as a valid method by the indsutry as well. Special thanks to Nelson Correia and Roberto De Meo of the Keypoint Technologies, LLC and Ricardo Rodrigues of the L10nbridge.

A lot of thanks goes to Mr. Omar F. Zaidan, a PhD student at Johns Hopkins University who so diligently coded the program in Java presented as the tool in this thesis. At times, it was him who found faults with my logic and suggested a relook and this gave me clues on how to posit my logic in a way that it suits best and gives the best result. Thanks a lot Omar! Though I am not a good programmer with Java, I promise I will try to enhance my skills in it.

I also take this opportunity to thank my parents and other close relatives who silently bore with me while I kept myself busy with academics. They know little about what I am doing, but for sure they know that what I am doing is good and deserves to be done. Thanks to Maa and Papa, I have been very whimsical with my work and you were always so patient. Thanks for supporting me all these while.

Thanks a lot to Rawat-ji, the most jovial man at our department’s office. We always loved his jokes and anecdotes. In case of any official problem, he has always been there to help the students out.

I have spent almost a decade inside the secure and walled campus of Jawaharlal Nehru University. It has been the most prolific time for me to learn in a safe, peaceful environment. During this period, I stayed mainly at two hostels. The hostel mates were a constant support system. I take this opportunity to pay special thanks to Saleem Ahmad, my floormate for last three years. The *chai* sessions at nights with him were always rejuvenating. Thanks a lot to Amit

Ranjan, who used to come all the way from another hostel for late night small talk sessions with *chai* and *maggi*.

CONTENTS

Declaration by the Candidate ii
Certificate iii
Dedication iv
Acknowledgements v
Contents viii
List of Tables xvi
List of Figures xviii
Abbreviations Used xix
Abbreviations Used in Glossing of the example sentences xix
Other Abbreviations Used in the Text xx
Abbreviations Used in Annotation (based on the IL-PoST Framework) xx
1 Introduction 1
1.1 Methods of Research 2
1.2 Research Problem 4
1.3 A Background to the Problem 5
1.4 General Introduction to NLP/Computational Linguistics 6
1.4.1 Sub-Branches of Computational Linguistics 6
1.4.1.1 Text Processing 6
1.4.1.2 Speech Processing 7
1.4.1.3 Discourse Analysis 7
1.4.1.4 Information Retrieval 7
1.4.2 Goals of Computational Linguistics 7
1.4.2.1 Machine Translation 7
1.4.2.1.1 Direct Machine Translation 8
1.4.2.1.2 Rule Based Approach 8
1.4.2.1.3 Corpus-based Approach 10

1.4.2.1.4 Hybrid Approach 11
1.4.2.2 By-products 12
1.4.2.2.1 Dictionaries 12
1.4.2.2.2 Morphological Analysers 12
1.5 Resources for NLP 12
1.5.1 Corpus 12
1.5.2 Annotated Corpus 12
1.5.3 PoS Tagging/Chunking/Analysers 12
1.5.4 Tools 13
1.6 Resources for NLP in Hindi 13
1.6.1 Corpus 13
1.6.2 PoS Taggers/Chunkers/Analysers 14
1.6.3 Tools 16
1.6.4 Lexica 16
1.6.5 WordNet/Neural Networks/Ontologies 16
1.7 Overview of MT/MaT in Hindi 17
1.7.1 AnglaBharati/AnuBharati 17
1.7.2 Shiva and Shakti 17
1.7.3 UNL-based Machine Translation 18
1.7.4 Mantra 18
1.7.5 MaTra 18
1.7.6 Others 18
1.8 The Present Enterprise in the Context 21
1.8.1 Issues in Automatic Identification of a Word Group 21
1.8.2 Annotation Scheme 22
1.8.2.1 Penn Tagset 22
1.8.2.2 IIIT Tagset 22
1.8.2.3 IL-POST Tagset 23
1.8.2.4 Other Tagsets 23
1.8.3 Annotation Standards 23

1.8.3.1 TEI Standards 24
1.8.3.2 EAGLES Guidelines 24
1.8.4 A preliminary comparison of Annotation Schemes 25
1.8.4.1 IIIT Hyderabad Tagset 26
1.8.4.2 IL-POST Framework Tagset for Hindi 29
1.9 Part of Speech Identification 34
1.9.1 Description of available PoS Taggers 34
1.9.1.1 Hidden Markov Model 34
1.9.1.2 Maximum Entropy Markov Model 35
1.9.1.3 Conditional Random Field 35
1.9.1.4 Decision Trees 35
1.9.1.5 Neural Network 36
1.9.1.6 Part of Speech Identification for the present work 36
1.10 Local word grouping (Chunking) 37
1.10.1 Issues in Automatic Identification of a Word Group 37
1.10.2 Approaches for Chunking 37
1.10.3 Approach for Chunking in the present work 37
1.11 Language Encoding 38
2 Hindi Verb Morphology 39
2.1 General Description of Hindi 39
2.1.1 Phonology of Hindi and its orthographic representation 39
2.1.2 Hindi Morphology 44
2.1.2.1 Noun 44
2.1.2.1.1 Nominal Categories 44
2.1.2.1.2 Inflections: Gender 45
2.1.2.1.3 Inflections: Number 45
2.1.2.1.4 Inflections: Case 45
2.1.2.2 Adjectives 47
2.1.2.3 Verbs 47

2.1.2.4 Indeclinables 47
2.1.2.4.1 Adverbs 47
2.1.2.4.2 Pronouns 47
2.1.2.4.3 Demonstratives 55
2.1.2.4.4 Postpositions 55
2.1.2.4.5 Particles 56
2.1.2.4.6 Punctuations 56
2.1.2.4.7 Residual 56
2.2 Hindi Verb Morphology 56
2.2.1 Verb Forms in Hindi 60
2.2.2 Tense, Aspect and Mood (TAM) 61
2.2.2.1 Tense 61
2.2.2.2 Aspect 62
2.2.2.3 Mood 62
2.2.2.4 Agreement 62
2.3 Verb Groups in Hindi 63
2.4 Occurrence of Verb Sequence in Hindi Text 68
2.4.1 Structure of Compound Verbs 69
2.5 Structure of Conjunct Verbs 71
3 Formal Description of the Verb Group in Hindi Text: Identification and Analysis 72
3.1 Template Structures of Verb Groups 72
3.1.1 Definition of a Template 73
3.1.2 Description of Verb Group Templates 73
3.2 TAM Tags 74
3.3 Representation of TAM in Hindi verbs 74
3.4 Verbal Elements in Verb Group 77
3.4.1 One-Word Verb Group 79
3.4.2 TAM Tags for One Word Verb Groups 87
3.5 Periphrastically Marked TAM 88
3.5.1 Representation of Tense 89

3.5.2 Representation of Aspects 91
3.5.2.1 Other Aspects 92
3.5.3 Representation of Moods 96
3.5.3.1 Abilitative\Possibilitative 96
3.5.3.2 Presumptive 97
3.5.3.3 Desiderative and Suggestive 97
3.5.3.4 Obligative 98
3.5.3.5 Permissive 98
3.5.3.6 Other Moods 98
3.5.3.6.1 Permissive Counterfactual 98
3.5.3.6.2 Permissive Suggestive 98
3.5.3.6.3 Permissive Imperative 99
3.5.3.6.4 Optative Probabilitative 99
3.5.3.6.5 Conditional 100
3.5.3.6.6 Contingent 101
3.6 Compound Verb V2 vs. the Aspectual/Modal V2s 102
3.7 Multi-Word Verb Group Templates and their TAM Tags 104
3.7.1 Two-Word Verb Groups 104
3.7.2 Three-Word Verb Group 108
3.7.3 Four-Word Verb Group 111
3.7.4 Five-Word Verb Group 112
3.8 Verb Group in Passive Voice 113
3.8.1 Two-Word Passive VGs 113
3.8.2 Three-Word Passive VGs 114
3.8.3 Four-Word Passive VGs 116
3.9 Verb Group in compound Verb 117
3.9.1 The V2s or the vector verbs 119
3.9.2 Drawing Templates with V2s 119
3.9.2.1 V1 vs V2: Most V2s demand a particular V1 119
3.9.2.2 CV Templates Correspond to Non-CV Templates 123

3.9.2.3 Abstract CV Templates 128
3.9.2.4 Abstract Universal CV Templates 140
3.9.2.5 Templates having CV Potentiality of 'R' 143
3.9.2.6 Realization of Actual CV Templates 151
3.10 Verb Groups (Chunk) Tags 152
4 Mechanical Understanding of the Verb Sequence in Hindi Texts: Implementation and Evaluation 154
4.1 Logic to Identify the Verb Groups 155
4.1.1 Parts of Speech features of a Verbal Word 156
4.1.2 Template Defintions 157
4.1.3 Tag Variables 158
4.1.4 Word Variables 159
4.2 Identifying verb groups: a deterministic pattern matching task 161
4.3 A glimpse of the Tool: "Hindi Verb Group Finder" 164
4.3.1 Back-End 164
4.3.2 Front-End 165
4.4 Evaluation 169
4.4.1 First Test Run 169
4.4.2 Error Analysis in the First Test Run 170
4.4.2.1 Reasons for Orphan VAUXes 170
4.4.2.2 Reasons for Unknown VGs Found 171
4.4.2.3 Incorrect Marking of the Verb Groups 175
4.4.3 The Treatment of Verbs in the IL-PoST Framework for Hindi 176
4.4.3.1.1 Change in the Aspect Slot 177
4.4.3.1.2 Change in the Mood Slot 178
4.4.3.1.3 Change in the Finiteness Slot 178
4.4.4 Corrections Made in the Test Data 179
4.4.4.1 Correcting the Orphan VAUXes 179
4.4.4.2 Removing the Progressive Markings and Placing the Perfective Markings 179
4.4.4.3 Removing the Habitual Markings and Placing the Imperfect Markings 179

4.4.4.4 Corrections in the Finiteness Slots 180
4.4.5 Other Changes 180
4.4.5.1 The Verb *cah* and its Annotations 180
4.4.5.2 Single Template Denoting Two VGs? 181
4.4.5.3 Perfect Form vs. Optative Form 182
4.4.5.4 Conjunctive Participle 183
4.4.5.5 Imperfective Participle 185
4.4.5.6 Verbal Nouns vs. Infinitives 186
4.4.5.7 The Verbal Nouns: Should we have it? 187
4.4.5.8 The *wala* Particle 187
4.4.5.9 The word *lekər* 189
4.4.5.10 The word *cəlte* 190
4.4.5.11 The Use of the Word *ho* 190
4.4.6 Second Run of the Tool with the LDC Corpus 192
4.4.6.1 Unknown Phrases Found in the Second Run 194
4.4.6.1.1 Perfective Forms of Verbs Playing as Nouns 194
4.4.6.1.2 The Verbal Word Contains a Punctuation Mark 196
4.4.6.1.3 Compound Verbs Functioning as Verbal Nouns 197
4.4.6.1.4 The VG contains a word which is confusing and marked wrongly 197
4.4.6.1.5 Spelling/Grammatical Errors in the VGs 198
4.4.6.1.6 Verbal Nouns with a template VR_pft+VINF 199
4.4.6.1.7 Template Missing 200
4.4.6.1.8 VG Conjugated Midway with a Punctuation Mark 201
4.4.6.1.9 A CV with the First Form as ‘pft’ 201
4.4.6.1.10 Two VGs marked as one 202
4.4.6.1.11 A CV with three words 202
4.4.6.1.12 Construction of Conjunct Participle with *ke* Instead of *kər* 203
4.4.6.1.13 Nouns/Adjectives of the Conjunct Verbs Marked as VM 204

4.4.6.2 Incorrect Marking of VGs 204
4.4.6.2.1 Annotation Errors 205
4.4.6.2.2 Ambiguities arising out of Template Definitions 207
4.4.7 Final Run of the Tool with the LDC Corpus and the Example Test Set 210
5 Conclusions 213
5.1 Summary 213
5.2 Motivation for the current Work 215
5.3 The Logic and functioning of the Tool 216
5.4 Evaluation 218
5.4.1 Error Analysis 219
5.4.1.1 Faulty Annotations 219
5.4.1.2 Logical Failures 219
5.5 Final Results 221
5.6 Future Developments 221
Appendix 1: Template Definitions 223
Appendix 2: Word Variables 258
Appendix 3: Tag Variables 284
Appendix 4: Example Sentences and with their gloss 285
Appendix 5: Example Sentences with their Annotations in the IL-PoST framework 408
References 434

LIST OF TABLES

Table 1: PoS tagger and chunker in Hindi 16
Table 2: Major MT projects involving Hindi 20
Table 3: Parts of Speech Tags of Hindi in the IIITH tagset 28
Table 4: Chunk Tags in the IIITH Tagset 29
Table 5: Category Types of Hindi in IL-POST Framework 30
Table 6: Atrributes and Values of Hindi in the IL-POST framework 32
Table 7: Consonants of Hindi as represented through the Devanagari Alphabet 40
Table 8: Vowels of Hindi as represented through the Devanagari Alphabet 40
Table 9: Devanagari Letters, their Unicode Values and corresponding IPA symbols 43
Table 10: Nominal Categories 45
Table 11: Nominal Inflection 47
Table 12: Different forms of the Personal Pronouns 50
Table 13: Emphatic Forms of the Personal Pronouns 52
Table 14: Non-Personal Pronouns in Hindi 55
Table 15: Different Morphological Forms of a verb 60
Table 16: Common V2s in Complex Verbs 70
Table 17: Common V2s occurring in double words 71
Table 18: Examples of Conjunct Verbs 71
Table 19: General Tense and Aspect Templates in Hindi 76
Table 20: General Mood Templates in Hindi 77
Table 21: Summary of the functions of Hindi Verbs 80
Table 22: Verb Forms Occurring as One-Word Verb Group 83
Table 23: TAM Templates for One Word Verb Group 84
Table 24: One Word Verb Groups and their TAM Tags 88
Table 25: Aspects as Identified and Used in the TAM Tags 96
Table 26: Moods Marked in the TAM Tags 102
Table 27: Two-Word VG Templates 107
Table 28: Summary of Two-Word VG Templates 107
Table 29: Three-Word VG Templates 111
Table 30: Summary of Three-Word VG Templates 111
Table 31: Four-Word VG Templates 112
Table 32: Five Word VG Templates 113
Table 33: Two-Word Passive VG Templates 114
Table 34: Three-Word Passive VG Templates 116
Table 35: Four-Word Passive VG Templates 117
Table 36: V1 V2 Parity 123

Table 37: Potential CV templates 140
Table 38: Abstract Universal CV Templates 142
Table 39: Abstract CV Templates 150
Table 40: Tags of the verbs in the input text 157
Table 41: Defining Templates for the VG Finder Tool 158
Table 42: Tag Variables and their Values 159
Table 43: Word Variables 160
Table 44: Pseudo Code for the VG finder tool 164
Table 45: Directory structure of the VG Finder Tool 165
Table 46: Result Summary of the First Run 170
Table 47: Sentences with Orphan VAUXes found in the LDC Corpus 171
Table 48: Sentences with Unknown VGs Found in the LDC Corpus 172
Table 49: Examples of Incorrect VG Markings in the First Test Run 176
Table 50: Removing the Habitual Markings 180
Table 51: Results of the Second Run with the LDC Corpus 192
Table 52: Summary of Unknown VGs Found in the Second Run 194
Table 53: Incorrect VG Markings in the First 1000 Sentences of the LDC Corpus 205
Table 54: Results of the Final Run with the LDC Corpus 210

LIST OF FIGURES

Figure 1: Front-End of the VG Finder Tool 166
Figure 2: Output of the VG Finder Tool retaining the Input PoS tags 167
Figure 3: Output of the VG Finder Tool without the Input PoS tags 168

ABBREVIATIONS USED

ABBREVIATIONS USED IN GLOSSING OF THE EXAMPLE SENTENCES

Abbreviation	Meaning
1	first person
2	second person
3	third person
ABL	ablative
ACC	accusative
ADJ	adjective
ADV	adverb
ASS	associative
AUX	auxiliary
BEN	benefactive
CFCT	counterfactual
COND	conditional
CONJ	conjunction
DAT	dative
DES	desiderative
EMPH	emphatic
ERG	ergative
FSG	feminine singular
FUT	future tense
GEN	genitive
IMP	imperative
IMPF	imperfective
INF	infinitive
INS	instrumental
MSG	masculine singular
MSGH	masculine singular honorific
N	noun
OBL	oblique
OPT	optative
PCPL	participle
PFT	perfective
PL	plural

PRN	pronoun
PRS	present tense
PRT	particle
PST	past tense
REFL	reflexive
SG	singular
SGH	singular honorific
SUGG	suggestive
VR	verb root

OTHER ABBREVIATIONS USED IN THE TEXT

Abbreviation	Meaning
VG	Verb Group
TAM	Tense, Aspect and Mood
GNP	Gender, Number and Person
GNPH	Gender, Number, Person and Honorificity
NLP	Natural Language Processing
CV	Compound Verb
IPA	International Phonetic Alphabet
Gl.	Gloss
EGl.	Gloss in English
Exxx	Example Number 'xxx'
LDC	Linguistic Data Consortium
IIITH	International Institute of Information Technology, Hyderabad
MSRI	Microsoft Research India
V1	First Verb of the Compound Verb
V2	Second Verb of Compound Verb
pas	passive
CID	Corpus Identification Number (i.e. the sentence number in the LDC corpus, this might not be the same ID as given in the LDC corpus because the LDC corpus underwent some modifications and this might has changed original sentence ID).

ABBREVIATIONS USED IN ANNOTATION (BASED ON THE IL-POST FRAMEWORK)

Please refer to section 1.8.4.2 (page 29-34) in the first chapter and read it in conjunction with the chnages made in the IL-PoST framework as shown in Table 40 (page 156-157).

1 INTRODUCTION

Parts of speech have been considered to be one of the basic blocks of language structure. Any grammatical analysis, particularly of syntactic and semantic nature, takes the word as the base to start from. Identification of different parts of speech is the first thing that a child gets to learn in the first grammar classes of any language. Even in deeper levels of linguistic analysis, the common language exemplar consists of 'glosses' which include the grammatical category information of the word along with other grammatical information such as morpho-syntactic features.

In computational linguistics, the task of tagging has been considered as a basic task that helps in several higher level computational analysis of a language. Tagging stands for assigning a tag or label for a given input 'token' or word. Thus part of speech tagging stands for assigning each of the input words with a 'tag' or descriptor which may be similar to what in general linguistics is known as the 'gloss' of the word, however, minus the meaning in any particular language (commonly English or the language of the book or research paper etc. where the example sentence of the language being talked about is discussed).

Parts-of-speech (PoS) tagging is different from 'glossing' in the way that, first of all, the goal of the former is to prepare a corpus with such morpho-syntactic information for each of the words in a text. The goal of the latter i.e. glossing on the other hand is to make things understandable to human researchers and get into the exact meaning of the word being conveyed by the term in the source text. PoS tagging is done primarily for works undertaken through computers i.e. to make data available to computers so that other operations could be done over such data which now has the PoS information for each of the words in the corpus.

A PoS tagged corpus is used for several purposes including higher level linguistic research work, information retrieval, speech processing and, of course, higher level parsing such as marking boundaries of syntax level phrases or just a particular word group in a sentence and its boundaries.

While PoS tagging is a basic work, the work of just the next level up in computational linguistics (CoLing) or natural language processing (NLP) is to have some sort of higher level information. The immediate concern is to have the phrase level information which are the building blocks of syntax. This information may include information such as noun phrase, verb phrase, adjective phrase, adpositional phrase etc.

The syntactic relations between words can be represented in different ways. One can follow one of several different approaches to this effect. Among the most dominant of the approaches as applied in the research works of computational linguistics for parsing are the context free grammar and subsequent improvements on it.

While such kinds of works have been done on different languages, Hindi came off late. It's not that no such work has taken place for Hindi, but as stated in detail later in this chapter, all the works that have taken place follow a particular approach and have a varying degree of results with regard to accuracy and other measures of evaluations.

The present work as undertaken by this researcher takes up the task of identifying the 'verb group' in a given Hindi text corpus. Identifying a word group such as 'verb group' can be equated with the task of identifying a phrase and its boundary in a sentence. If we parse the whole of a sentence then we can have a tree structure in the model of the generative grammar. However, the task that the present approach aspires to accomplish is to find just one of the grammatico-syntactic categories in a given sentence. So, the problem for this work is more of natural language understanding (NLU) than natural language generation or others.

Identification of a verb group includes the identification of different parts of speech that play the role of denoting time, action or state in a sentence. That is, it involves the identification of tense, aspect and mood as well as the agreement markers, if any, in a sentence.

1.1 METHODS OF RESEARCH

There are different approaches to both tagging and local word group identification. Local word group identification is called chunking (Abney, S. 1991). I must state here that tagging is a prerequisite for chunking because for any chunker engine the input has to be tagged accurately. I have discussed the methods of tagging and different approaches available for this task below in this chapter. The different approaches applied for tagging and chunking Hindi text has also been

discussed below. From the discussion that follows, I find out that all the taggers and chunkers available at present use the stochastic and heuristic methods to achieve results reported up to 97%.

The high score of 97% as achieved by the stochastic methods sounds good. But there has been counter claims that the ambiguities missed by such taggers can in fact be improved through the linguistic approach (Samuelsson and Voutilainen, 1997).

I want to build on this linguistic approach which basically uses linguistic rules to arrive at the exact tags and chunk boundaries for a given sentence. This approach is also generally called rule based approach.

In fact the first approaches to the tasks of tagging or annotation was rule based. For example TAGGIT (Greene and Rubin, 1971) was the first large scale tagging engine that used a total of 3300 rules to disambiguate a corpus that had to be assigned a tag from a tagset having a total of 71 tags. Despite the large number of rules, this tagger was able to disambiguate only 77 percent words, the rest being done manually (this effort gave the first annotated corpus of English – a million-word Brown Corpus). Then the statistical methods came in such tasks which have overwhelmed the scene since the start of the 80's. Another rule-based approach came in the 90's with the approach being called the constraint grammar (Karlsson, F., 1990; Karlsson et.al. 1995). The Constraint Grammar (CG) approach as applied to English showed that a further improvement is possible in the current results of the stochastic taggers. It should be noted here that after a certain stage of reaching the results using the stochastic methods, improving upon it even by bits is very difficult. And no statistical tagger has so far been able to reach anything above 99%. On the contrary, the rule based approach of constraint grammar for tagging has been reported to achieve results as high as 99.7% (Samuelsson and Voutilainen, 1997).

Another rule based tagger is transfer based learning (TBL) (Brill, E., 1992; 1995;) which uses a generative model of sentence to generate a syntactic structure and assign tags based on the classification. The TBL approach has also been suggested to be used for parsing wherein Brill "proposes starting with a uniformly right-branching parse and learning rules for rotating local trees in order to improve the fit to a training corpus."

Out of these two major rule-based deterministic methods of tagging and parsing, none has been used for Hindi. Out of these two, of particular mention is the CG approach, which, according to its proponents, is more suitable to languages that are more inflecting in nature. Though CG approach has been used for English as well, it started first with Russian and Slavic languages. According to the authors this deterministic approach is much more suitable to languages which have richer morphology to account for.

My approach for the purpose of chunking starts with an assumption that there is a good tagger available which gives a hundred percent correct PoS tagged data which then forms as the input to the tool I build. This tool will find out the verb groups in such a tagged data. The tagged input data has to be in the form of the IL-PoST schema.

The researcher had helped create such an annotated corpus for the Microsoft Research India Ltd. This corpus (Bali, K. et.al., 2010) is now available for download at Linguistic Data Consortium's (LDC's) website[1]. For a point of reference, I will refer to this data IL-PoST data.

The IL-PoST data has multilayered morpho-syntactic information about a word which can directly be used for the purpose of chunking. There are other tagsets available as well. But the researcher feels that this IL-PoST corpus is directly suitable for this purpose. Comparisons of other tagsets with the IL-PoST have been made in a subsequent chapter.

Running the tool which will have rules defined to capture various verb groups will give us the verb groups. These verb groups will have information about the tense, aspect and mood (TAM) of the sentences (or clauses within it). It is notable here that the TAM cannot be found at the level of PoS because this information can also get spread over other verbal elements in the sentence which form a verb group.

1.2 RESEARCH PROBLEM

The results of the statistical chunkers depend on various factors. The most important among them is the correct PoS input for each of input word (or token). For Hindi, the taggers developed so far are all statistical with various results accuracy. All of these taggers used so far use a tagset

[1] http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2010T24

accessed: 15 July, 2011.

known as IIIT or ILMT tagset. A comparison between these two tagsets follows in another section.

Statistical taggers are though good to start within the tasks of linguistic analysis involving large corpora they should not be always needed and should not be blindly applied even to cases where a certainty can be achieved through other methods. Several practitioners of computational linguistics have indicated time and again that statistics should be used only in cases where no other options are available to achieve better accuracy. A maxim often cited (e.g. Abney, S. 1996) in this regard being: “There are lies, damned lies and statistics”.

The aim of this thesis is simply to investigate the application of a rule-based technique for the task of parsing a particular grammatico-syntactic category in a given sentence of a particular inflecting language known as Hindi, the fifth most widely spoken language of the world[2].

This thesis though will be the first attempt to build a tool for local word grouping through a rule-based method with no statistical inputs required. For this tool, we assume that annotated corpus will be the input and this corpus will have the necessary PoS information of each of the word (particularly for the words that are verbs).

1.3 A BACKGROUND TO THE PROBLEM

Parsing has been stated above to be one of the basic works in CoLing. Parsing has been studied for several languages and there are tools readily available to do the parsing for different languages (e.g. English, Chinese, French, German, etc.).

As stated earlier, both tagging and parsing are done with a purpose and the purpose might change based on needs of the application being developed. Thus there are shallow parsers, dependency parsers, context-free parser and unification-based parsing etc. A shallow parser can start first with building small chunks and then go on to build larger chunks based on it (Abney, S. 1996).

Other hierarchical structure, such as identification of the syntactic relations of subject and object, can be built upon this shallow parsed data. There are different approaches to build higher level

[2] http://www.ethnologue.org/ethno_docs/distribution.asp?by=size

accessed: 15 July, 2011.

parsers. Dependency Grammar (Melcuk, 1987) parsing is done in terms of a set of head-modifier dependency links between words where the links are labelled with the grammatical function (e.g. subject or object) of the modifying word with respect to the head. Within dependency grammar there are different approaches to parsing, e.g. Link Grammar system (Grinberg, Lafferty and Sleator, 1995), Functional Dependency Grammar system (Tapanainen & Jarvinen, 1997).

In the past four to five years, several attempts have been made to develop such tools for Hindi as well. A detailed overview of this has been given elsewhere in this chapter. Tagging and chunking techniques as well as other resources required for such tasks have been discussed in the following sections.

1.4 GENERAL INTRODUCTION TO NLP/COMPUTATIONAL LINGUISTICS

Linguistics has evolved in multifarious dimensions over the short period of its development as a discipline. It has had several branches which later had sub-branches penetrating into various other disciplines. Under the general umbrella of Applied Linguistics comes what the linguists preferably call Computational Linguistics (COLING hereafter). While rooted in the general search of seeking knowledge lying behind the phenomena of human language, COLING aspires to capture and model the knowledge gained as such towards developing language technology for various purposes, including the technology to be used by the common people.

With history of about 60 years since the digital technology came into use, the attempt has been on to represent human linguistic knowledge through the machines. Of course, no machine can equal the intelligence of human, but a simulation of that intelligence is always worth the effort as it makes communication, the central goal of language, easier and faster.

1.4.1 Sub-Branches of Computational Linguistics

As NLP/COLING is a part of the applied sciences, it has diversified into several domains. While its core remains the text processing, it has found wide applications in other domains some of which has been recounted as shown below.

1.4.1.1 Text Processing

In text processing the input is the text in digital form of bytes. It extracts major resource from linguistic descriptions of languages and builds applications through it.

1.4.1.2 Speech Processing

In speech processing the input is speech in the digital audio form. It extracts major resource from the phonetics and phonology of the language(s) concerned.

1.4.1.3 Discourse Analysis

Here the input may be either text or speech. Although at present, for discourse analysis text is the major input.

1.4.1.4 Information Retrieval

Information retrieval has been quite old as of now and has already achieved a level of accuracy. But linguistics can enrich it by giving a human touch to it. Information Retrieval tools that draw on linguistics aim to extract information from a large source of data through queries written in natural language. Linguistic knowledge gained through query analysis can help achieve newer heights in this field.

1.4.2 Goals of Computational Linguistics

Computational Linguistics aims to build or help build computer applications that are based on natural language as used by the human beings. The underlying goal behind each of endeavour of COLING is natural language understanding (NLU). Natural language understanding can be seen as a task in itself as it aspires to build systems that understand the natural language in its own terms. This system with the understanding of the natural language then can be deployed to perform tasks with the intelligence expected of a human being.

1.4.2.1 Machine Translation

One of the most important and initial goals of NLP/COLING has been machine translation i.e. translation from one human language to another using machines. So much so that there have been several attempts in many of major languages of the world and many of them are now publicly available. Although machine translation in the general domain still has lots of issues, considerable success has been achieved in domain specific translations.

The present MT technology research work applies a variety of approaches to achieve the goal of machine translation, depending on the factors like language pairs involved. However, for a classification purpose, the different approaches can be grouped into four approaches (Siddiqui and Tiwary, 2008; Hutchins, 1993) as described below.

1.4.2.1.1 Direct Machine Translation

This is the earliest approach thought of and applied. It bases itself on a word by word matching of the source and the target text, taking almost everything from the respective bilingual dictionary. However, this approach did not prove good enough as language is more than just dictionary and it needed to capture the grammatical nuances and cases of disambiguation, among other things.

1.4.2.1.2 Rule Based Approach

In rule based approach to machine translation, an intermediate layer based on the input language text is created before translating into the target language. This intermediate layer is the desired abstraction of the input text and draws heavily on the grammatical rules of both the source and target language, along with parallel lexicons.

Rule based approach in itself also has several other approaches, the two main being: Transfer-based machine translation and Interlingua machine translation. The transfer based approach functions in three stage model: abstract representation of source language analysis, transfer into abstract representation of the target language and generation or synthesis into the target language text. Examples of projects based on the transfer approach include ARIANE and EUROTRA, both of the 1980's. In the interlingua approach a language neutral representation (called 'interlingua') is evolved which represents both the source and the target language, then generation into the target language is done through this interlingual representation. The interlingua based projects of the 1980's like the DLT (Sadler, 1989) and Rosetta, though rejected the transfer model, also included the rule-base component. The DLT adopted a modified form of Esperanto as its interlingua while Rosetta experimented with the representations and the integration of Montague semantics. The KBMT89 (Goodman and Nirenburg, 1991) was also based on models of interlingua and used a knowledge base to achieve greater success, especially in specific domains such as computers. Some of the other projects based interlingua are: NEC's PIVOT system (Okumura et al. 1991) and ULTRA (Farwell and Wilks, 1991).

Other approaches within Rule-based framework are: Constraint-Based Formalisms (Sadler and Arnold, 1992, Sharp and Streiter, 1992) and Principles-based MT. Embracing unification and constraint based formalisms like Lexical Functional Grammar (KAPLAN & J.BRESNAN, 1982) and Functional Unification Grammar (Kay, 1984) and later into Generalized Phrase Structure

Grammar (GPSG) (Gazdar et al., 1985) and Head-driven Phrase Structure Grammar (HPSG) (Pollard and Sag, 1994) the constraint based approaches give the facility of bi-directionality in bilingual translation because of its inherent reversibility. Deriving on the Rosetta experience, several projects started on this constraint based approach, namely CRITTER and UNITRAN.

Principles-based MT bases itself on the Principles and Parameters framework (Chomsky and Lasnik, 1993) as manifest in the Government-Binding Theory (Chomsky, 1981/1993) of syntax. It functions on the premise that all the languages derive its grammar from one underlying universal language and differences between the languages are the result of diverse settings in the parameters at different levels of linguistic analysis. This approach was used, among others, in UNITRAN, an interlingual MT system that had two of its basic components which applied this approach to process the acceptance and production of the grammatically correct sentences and for deriving an underlying conceptual representation to match appropriate target-language structures (Dorr 1990, Dorr 1992).

In rule based approach to machine translation, an intermediate layer based on the input language text is created before translating into the target language. This intermediate layer is the desired abstraction of the input text and draws heavily on the grammatical rules of both the source and target language, along with parallel lexicons.

Rule based approach in itself also has several other approaches, the two main being: Transfer-based machine translation and Interlingua machine translation. The transfer based approach functions in three stage model: abstract representation of source language analysis, transfer into abstract representation of the target language and generation or synthesis into the target language text. Examples of projects based on the transfer approach include ARIANE and EUROTRA, both of the 1980's. In the interlingua approach a language neutral representation (called 'interlingua') is evolved which represents both the source and the target language, then generation into the target language is done through this interlingual representation. The interlingua based projects of the 1980's like the DLT (Sadler, 1989) and Rosetta, though rejected the transfer model, also included the rule-base component. The DLT adopted a modified form of Esperanto as its interlingua while Rosetta experimented with the representations and the integration of Montague semantics. The KBMT89 (Goodman and Nirenburg, 1991) was also based on models of interlingua and used a knowledge base to achieve greater success, especially

in specific domains such as computers. Some of the other projects based interlingua are: NEC's PIVOT system (Okumura et al. 1991) and ULTRA (Farwell and Wilks, 1991).

Other approaches within Rule-based framework are: Constraint-Based Formalisms (Sadler and Arnold, 1992, Sharp and Streiter, 1992) and Principles-based MT. Embracing unification and constraint based formalisms like Lexical Functional Grammar (KAPLAN & J.BRESNAN, 1982) and Functional Unification Grammar (Kay, 1984) and later into Generalized Phrase Structure Grammar (GPSG) (Gazdar et al., 1985) and Head-driven Phrase Structure Grammar (HPSG) (Pollard and Sag, 1994) the constraint based approaches give the facility of bi-directionality in bilingual translation because of its inherent reversibility. Deriving on the Rosetta experience, several projects started on this constraint based approach, namely CRITTER and UNITRAN.

Principles-based MT bases itself on the Principles and Parameters framework (Chomsky and Lasnik, 1993) as manifest in the Government-Binding Theory (Chomsky, 1981/1993) of syntax. It functions on the premise that all the languages derive its grammar from one underlying universal language and differences between the languages are the result of diverse settings in the parameters at different levels of linguistic analysis. This approach was used, among others, in UNITRAN, an interlingual MT system that had two of its basic components which applied this approach to process the acceptance and production of the grammatically correct sentences and for deriving an underlying conceptual representation to match appropriate target-language structures (Dorr 1990, Dorr 1992).

1.4.2.1.3 Corpus-based Approach

The availability of large corpus of text since the 80's marked a departure in the area of language study throughout. Analysis of corpus for different purposes gave new linguistic insights and facilitated new approaches to emerge. In MT and related areas, two new approaches that emerged with corpus studies are that of statistical MT and example-based MT.

Based on a large corpus of the parliamentary debates in both English and French, the IBM Candide research (Brown et al. 1988, 1990) used stochastic methods to align the sentences and calculate the probabilities that any word in a sentence corresponding to two, one or zero words in the translated sentence in the other language. Its initial results of 48% was later improved upon

by various statistical techniques and along with supplying data on inflectional morphology and syntactic transformations to deal with discontinuous structures

Proposed first by Nagao (1984) the example-based machine translation (EBMT) uses a parallel corpus of the source and the target language. Based on learning by analogy, an EBMT system takes the input of a source sentence and looks for a similar sentence in the source language from the corpus (the example-base) and generates the translation with adaptations, if any required, into the target language. Some examples of systems developed on example-based approaches are ATR project known as Transfer-Driven MT ((Sumita et al. 1990; Furuse & Iida 1992, 1992a) for spoken language translation, research at Hitachi (Kaji et al. 1992) and at Kyoto (Utsuro et al. 1992).

Another approach generating through the corpus is that of ‘connectionist’ (Hutchins, 1993). Corpus or resources are being created with links at different levels of analysis. Examples are like WordNet (Miller, G. 1995), VerbNet (Levin, B. 1993), PropBank (Kingsbury, P., 2002) and FrameNet (Ruppenhofer, J et al, 2006) etc. Most of the connectionist approach attempt to capture the semantic meaning and use it for various NLP tasks, including MT, such as disambiguation at various levels.

1.4.2.1.4 Hybrid Approach

In the modern approach to machine translation, we continuously get to see instances of systems built with a mix of several techniques, be it features from Rule-based or statistical approaches. This is yielding better results for most of such engines.

To add a few words more towards the MT, it would be appropriate to say that, MT is now a reality. The goal is not to render a perfect translation, as envisaged earlier by many. The goal rather has now shifted to the use of such reliable transfer of information. It is now more specific towards users’ needs and has found use in niche areas like specified domains such as bioinformatics, scientific communications, technical writing etc.

1.4.2.2 By-products

While machine translation still remains a goal to achieve, the other products of this endeavour has already achieved new heights and enriched the language resources as never before. Some of the resources that are developed as by-products can be grouped as below.

1.4.2.2.1 Dictionaries

Dictionaries used to be a very time consuming task to prepare. With the help of computational lexicography, not only preparing dictionaries has become easier but also it has facilitated newer approaches towards these endeavours. Now most of the languages have lexicons of considerable size readily available, thanks to corpora available in many languages.

1.4.2.2.2 Morphological Analysers

Morphological Analysis of words based on the linguistic rules derived for a particular language has been a good resource that also caters to several NLP processes.

1.5 RESOURCES FOR NLP

For NLP tasks of the present day, it is necessary to have some resources for the language(s) concerned. Some of the resources that are needed for this kind of research are as follows.

1.5.1 Corpus

Corpus is a must have resource these days for almost any kind of research in NLP. This corpus can be either text or speech, depending on the need of the research work. Corpus plays the pivotal role to kick-start the hypothesis as it provides primary source of data.

1.5.2 Annotated Corpus

For NLP tasks, particularly that of text processing, the corpus has to be annotated using a specified scheme. This annotated corpus is given a representation that can capture the surface level nuances of it. The surface level nuances here mean the basic grammatical nuances i.e. part of speech, which are universal in any language. Such annotated corpus is then used for further processing at several other stages of processing the text.

1.5.3 PoS Tagging/Chunking/Analysers

Annotation though is at first done by the human linguists/language experts, recent developments showed that it can also be done through statistical methods.

For annotation a corpus has to have a consistent set of ‘tags’, i.e. markers, attached to each word in the corpus that can be utilised for further language processing. The first such corpus, the Brown Corpus of English, had a defined tagset that captured various grammatical categories. Other tagsets were also developed later to capture the nuances of other languages as well as deeper grammatical categories and subcategories of English. In a rule based approach, the next task after PoS annotation is that of grouping of phrases within a sentence like noun phrase (NP), verb phrase (VP) etc. This process is often termed as ‘chunking’ in the NLP literature.

Analysers are another resource which basically analyzes the words for its inflectional derivational forms at the morphological levels. These are often called morphological analysers. They can either produce different morphological forms given a stem or root word or can analyse a word for the inflectional and derivational markers inherent in it.

1.5.4 Tools

The basic tools for NLP research are part-of-speech taggers, phrase boundary markers (chunkers) and morphological analysers. For the automation of all these tasks, different approaches have been applied as mentioned above under the MT discussion. There are different tools available, both language generic and language specific. Examples are Stanford PoS Tagger (Toutanova, K. and Manning, Christopher, D., 2000), Conditional Random Field (CRF) based tagger, FreeLing (Atserias, Jordi et.al., 2006) etc.

1.6 RESOURCES FOR NLP IN HINDI

When NLP research in Hindi and other Indian languages started in the 1980’s, there were scarce resources. Over the years, a few good resources have been created for basic NLP undertakings in the Indic languages and efforts are still on to make it richer.

1.6.1 Corpus

By corpus here we mean a corpus of text in the electronic format. Such a kind of corpus can be said to have started very late in Indian languages. Corpora in Hindi are still far and few between. The first major breakthrough in the corpus development was made under the aegis of CIIL-KHS-EMILLE combined. This corpus, commonly called CIIL-EMILLE corpus, has about two million words of different genres of text in Hindi. KHS also did some annotation of this corpus based on the CIIL old tagset. However, not much detail of this corpus is available outside the KHS

premises as they prepared it particularly for pedagogical purposes and for internal use. (That is what was told to the researcher when the researcher approached for this annotated corpus.)

The problem of corpora for research purposes is now not a great problem as the main hurdle of Indian languages being available on computers is solved with the advent of Unicode in almost all the major languages of India. Most of the regional newspapers have an online edition. Besides, there are a lot of other sources of text in Hindi such as blogs in Indian languages and localized versions of many of the India-based websites. However, to have balanced corpus, one still needs to compile it to give proper representation to different domains.

1.6.2 PoS Taggers/Chunkers/Analysers

The need for part of speech tagging in Indian languages was understood lately, particularly when the statistical methods in NLP started taking a prominent role. Approaches earlier to this were mostly through lexicons and paradigms.

When this researcher started on his Ph.D. programme in 2006 there were few automatic PoS tagger available for Hindi. But in the last three years, several new techniques have been deployed to enrich the performance of such a tagger.

Chunking or grouping of words in a sentence with a phrase marker tag is the next step in the computational understanding of a natural language. The task of chunking depends pretty much on the tagger, its accuracy and features captured at the part of speech level. As these taggers were developed, chunking was also done with several of them. The present researcher's task is also of this chunking level. The table below gives a summary of the PoS tagging and chunking done so far in Hindi.

Publicised @	**Authors**	**Title in Brief**	**Brief Description**	**Institute Involved**
ICON-2003	Pradipt Ranjan Ray et. al.	POS tagging and LWG Techniques for NLP Parsing in Hindi	A rule based alogorithm for PoS tagging and chunking	IIT Kharagpur

IJCLP-2005	Akshay Singh	HMM based chunker for Hindi	A statistical tool	IIIT Hyderabad
NLP-AI 2006	Aniket Dalal et. al.	Hindi PoS Taggin and Chunking: A Maximum Entropy Approach	A statistical tool developed	IIT Bombay
IJCAI-2007	S. C. Pammi and K. Prahallad	PoS Tagging and Chunking using Decision Forests	A Statistical Tool Developed	IIIT Hyderabad and CMU, USA
SPSAL-2007	Himanshu Agrawal	PoS Tagging and Chunkiing for Indian Languages	A statistical approach based on CRF	IIIT Hyderabad
ICON-2007	Aniket Dalal et. al.	Feature Rich PoS Tagger for Morphologically Rich Languages: Experiences in Hindi	A statistical Maximum entropy Markov model	IIT Bombay
ICON-2008	M. Shrivastava and P. Bhattacharya	Hindi PoS Tagger Using Naive Stemming	An statistical HMM based affix propelled PoS tagger	IIT Bombay
ICON-2009	Ankur Parikh	Part of Speech Tagging using Neural Network	A novel approach in PoS tagging using Neural	IIIT, Hyderabad

			Network	

Table 1: PoS tagger and chunker in Hindi

Morphological analysers are useful in NLP tasks. But of late, this has gone in the background as statistical measures came into the fore in such works. Nonetheless, the tasks performed by the morphological analysers cannot be ignored. In fact the procedures used in the PoS tagging most of the time use morphological analysis of one kind or another to get closer and better results for morphologically rich languages as the Indic languages are.

The early NLP system of Anusaaraka had a component in its system that did the morphological analysis to get to the PoS tag and local word group of word(s) in an input sentence.

1.6.3 Tools

Not many linguistic analysis tools are available for Indian languages specifically. Although, there are several tools that can be deployed to do the analysis at different levels. For the purpose of this researcher, the tools used were general tools available in the Windows environment, some of which are: MS Word, MS Excel, Notepad and a few other tools to do Unicode Hindi converting and other general text processing works.

1.6.4 Lexica

Lexica of different kinds can be deployed for various NLP tasks. In Hindi, which still faces resource sparseness, lexical resources are far and a few between. Several major projects are however underway that aim to capture various aspects of the language.

With the tools available, now it is easy to extract a word list from a corpus. The task then remains to validate such words for correctness, particularly in terms of spelling (Choudhary, N., 2011).

1.6.5 WordNet/Neural Networks/Ontologies

Lexical resources are always a great help, particularly in solving higher end of the problems faced in NLP. Semantics has long been away from the basic tasks of NLP. But as one moves ahead in building NLP tools such as a bilingual communication system, semantics comes to the fore without which ambiguities posed cannot be handled properly. To cater to such needs.

Semantics can also play a good role at some points to solve the problems of the PoS tagging, chunking and morphological analysis.

To capture the semantics of a language, one has to capture the semantic information attached to the words of the languages. Several frameworks to capture the semantic features at the word level came up in the last one and a half decade. In India, the first of such an endeavour was the development of WordNet (Bhattacharya et. al., 2008) for Hindi at IIT Bombay. Another similar approach for Hindi is that of TreeBank (Bhatt et.al., 2009) wherein a simultaneous development of dependency structure and phrase structure tree banks for Hindi and Urdu, as well as a Prop-Bank is done. PropBanks specifically store the argument structure of a verb.

1.7 OVERVIEW OF MT/MAT IN HINDI

Machine Translation efforts started in India in the 90's with pioneering efforts at IIT Kanpur. The first of such an effort was that of Anusaarak (Bharati et. al., 1995, Bharati et.al., 2001). Anusaarak started as a language accessor for Indian languages. Based on the Paninian grammar, the Anusaarak maps source text to the target text. It works for several pairs of Indian languages, including Hindi as a major language.

1.7.1 AnglaBharati/AnuBharati

AnglaBharati was one of the first machine translation projects started at IIT Kanpur. This is a pseudo-interlingua based MaT system based mainly on pattern directed approach using context free grammar like structures for analyzing source language sentences. Another system that also started at the same setup was called Anubharati. This was an example-based MT system. Later, both these two approaches were put together to bring better results for translation in 2004 with the release of AnglaBharati-II and AnuBharati-II. The design of the system has now been planned to extend to other Indian languages as well under the AnglaBharati Mission initiative which involves several institutions across the country.

1.7.2 Shiva and Shakti

International Institute of Information Technology (IIIT), Hyderabad and Indian Institute of Science (IISc), Bangalore along with Carnegie Mellon University, USA took upon developing

machine translation technology for Indian language and the result was the release of Shakti[3] and Shiva[4] machine translation systems in 2004. These systems have constantly been under development since then and several releases have come up. Shakti is a Hybrid system that uses a rule-base while Shiva is an example based system.

1.7.3 UNL-based Machine Translation

A UNL as interlingua based Machine translation system[5] (Dave et. al., 2002) has been under development at IIT Bombay since past ten years and the results are shown online. This system is based on the interlingua where the interlingua is the Universal Networking Language (UNL), a knowledge representation language in information retrieval applications created at the United Nations University, Tokyo and developed at UNDL Foundation, in Geneva, Switzerland.

1.7.4 Mantra

Mantra[6] machine translation system has been developed by CDAC and is currently deployed at the Rajya Sabha (Upper Parliament) of India to translate gazettes, parliament proceedings and notifications from English to Hindi. The system is based on the tree-adjoining grammar (Joshi and Schabes, 1992) formalism. The work to expand its domain is on.

1.7.5 MaTra

MaTra[7] is a Machine aided Translation system from English to Hindi. It has a structured editor that creates a framed representation of the input sentence with the verb as the pivot which has a template associated with it.

1.7.6 Others

Besides the above mentioned five major projects that have already produced the systems for translation, there are several others that have worked on machine translation systems from or to Hindi. The table below lists some of the major contributions in this area. The web-based machine

[3] http://shakti.iiit.ac.in/~shakti/ accessed 12 April, 2010

[4] http://shiva.iiit.ac.in/ accessed 12 April, 2010

[5] http://www.cfilt.iitb.ac.in/machine-translation/ accessed 12 April, 2010

[6] *www.cdac.in/html/aai/**mantra**.aspx* accessed 12 April, 2010

[7] *www.cdacmumbai.in/**matra*** accessed 12 April, 2010

translation is now common. The most prominent as a service among all these has been the new entrant Google's translation system. Google, a major global web service provider, offers translation from one language to another many of the major languages of the world. It added Hindi two years back and is available as an online web-service[8].

Name of the Project and Institution	Language Pair/Direction	Domains	Approach	Strategy	Year Initiated
Anglabharati-II (IIT-K and C-DAC, N)	Eng-IL (Hindi)	General (Health)	Hybridized, Transfer/Rules, Example-based	Post-edit	2004
Anubharati-II IIT Kanpur, CDAC-Noida	Hindi-English	General	Hybridized Example-based	Post-edit	1995
Mantra (C-DAC, Bombay	Eng-IL (Hindi)	Government Notifications	Transfer/XTAG	Pre-edit	1999
MaTra (C-DAC, M)	Eng-IL (Hindi)	General (News)	Transfer/Frames	Pre-edit	2004
Shakti, IIIT Hyderabad, IISc Bangalore,	Eng-IL (Hindi)	General	Hybrid, Rule-based	Post-edit	2004

[8] http://translate.google.com/ accessed 15 July, 2011

CMU, USA					
Shiva, IIIT Hyderabad, IISc Bangalore, CMU, USA	Eng-IL (Hindi)	General	Hybrid, Example based	Post-edit	2004
UNL MT (IIT-B)	English, Hindi, Marathi	General	Interlingua/UNL	Post-edit	2005
Tamil Anusaaraka (AU-KBC, C)	IL-IL (Tamil-Hindi)	General (Children)	LWG, mapping PG	Post-edit	1995 onwards
MAT (Jadavpur University)	Eng-IL (Hindi)	News Sentences	Transfer/Rules	Post-edit	2005?
Anuvaadak (Super Infosoft)	Eng-IL (Hindi)	General	NA	Post-Edit	
StatMT (IBM)	Eng-IL	General	Statistical, Example-based	Post-edit	???
Google Translate	English-Hindi and other languages and vice versa	General	NA	Post-edit	2008?

Table 2: Major MT projects involving Hindi

1.8 THE PRESENT ENTERPRISE IN THE CONTEXT

This present work aims at identification of 'verb group' in any given Hindi text. By a verb group here is meant any verbal element in a sentence that functions as the verb, either main or auxiliary. By verb group, we do not intend to include predicate which would also include other elements like object or adpositional phrases occurring in the sentence.

In any kind of linguistic analysis of any text in a natural language, the importance of word groups, which form one syntactic unit within a sentence, is crucial. The linguistic community is already pretty much familiar with the importance of the Phrase Structure rules that break a sentence into different phrases like noun phrase, verb phrase, adjective phrase, adpositional phrase etc. These phrases contain within themselves the lexical categories or parts of speech like noun, verb, adjective, preposition or postposition, adverbs, particles, determiners etc.

A verb group is similar to the idea of a phrase or a syntactic category. However, unlike the phrase structure grammar as proposed by Chomsky (1957), wherein a Verb Phrase contains everything else other than the subject, here, I propose to mean only the verbal elements in a sentence. This task of identification of verb group is useful in NLP or even any kind of other linguistic analysis that involves identification of verbal elements and their position in a sentence. One cannot get to know what is the tense, aspect and mood of a sentence of clause until one processes the whole of the verb group because all of these information is spread over several elements that need to ascertained first. A verb in Hindi also takes agreement markings for gender, number and person (GNP) and might also contain information on honorificity. The GNP of a verb group is also needed if one seeks to build a context sensitive spell checker for Hindi.

Several works have already gone into the enquiry of the identification of verb group in Hindi since 2004. The taggers and chunkers already developed, as mentioned above, does show their results in identifying the verb group.

Below I discuss the works that involve in the verb group identification in a given Hindi text.

1.8.1 Issues in Automatic Identification of a Word Group

As said above, a word group contains words which in linguistic terms belong to one or the other lexical categories i.e. they are parts of speech. To group words into syntactic categories means that first these lexical categories are identified. Here arises the need for parts of speech

identification. The solution for this problem is the part-of-speech tagger. Parts of speech taggers are tools that automatically identify the lexical categories of the words in the input sentence/text and marks them with a ‘tag’ or label as specified in the annotation scheme.

So, the major steps that are involved before a verb group is identified can be classified into two broader categories of developing an annotation scheme and then developing a tool to identify parts of speech for each of the words. These two are major tasks and can be said to be the basic tasks in any kind of linguistic computational analysis of a given text.

1.8.2 Annotation Scheme

Annotation schemes are the detailed guidelines following which the words in the corpus to be annotated have to be tagged. The major annotation schemes which have been used to annotate large corpus or have been in use for quite some time are noted below.

1.8.2.1 Penn Tagset

Among all the tagsets used to annotate corpus for NLP tasks as well as theoretical linguistic analysis, the Penn Tagset (Santorini, B., 1990) has been the most used for English. The Penn Tagset, based on the Brown Corpus tagset, has a total of 36 simplified tags. As reported in Marcus, P. et.al. (1993), earlier works that involved having a tagset in English, had a higher number tags to capture the lexical and syntactic features of English. To cite a few, the Brown Corpus (Francis, W., 1964; Francis, W. and Kucera, 1982) had 87 simple tags, the LOB (Lancaster-Oslo/Bergen) uses about 135 tags, the Lancaster UCREL about 165 tags and the London-Lund Corpus of Spoken English about 197 tags. Unlike the other previous works of tagsets which aimed “the ideal of providing distinct codings for all classes of words having distinct grammatical behaviour (Garside, Leech and Sampson, 1987. pp.167), the Penn Treebank modified the Brown Corpus tagset heavily ‘by pairing it down considerably’ ‘to eliminate redundancy by taking into account both lexical and syntactic information’.

1.8.2.2 IIIT Tagset

The first tagset for Hindi and other Indian languages was also based on the Penn Treebank tagset. Developed at the IIIT Hyderabad through two workshops involving linguists and NLP practitioners, the tagset that came out has been documented in the guidelines for annotating Indian languages (Bharti et.al. 2006). Like the Penn Treebank, here also the focus is on

simplification of tags and on the lexical categories. The IIIT tagset has a total of 26 tags which cover the lexical categories. A detailed study of this tagset along with another major tagset has been presented in the ensuing chapter.

1.8.2.3 IL-POST Tagset

In 2008, another initiative towards this basic task of PoS annotation of Hindi was taken by Microsoft Research India, (MSRI) Bangalore when they started on developing a tagset to annotate the corpus of Indian languages. This also involved several workshops of linguists and NLP practitioners to decide upon the tagset design for Indian languages. This tagset, as reported in Baskaran et.al. (2008), follows the EAGLES (discussed below) guidelines and is hierarchical. A detailed description of this annotation scheme is presented in the ensuing section (§1.8.4.2). It has been named as IL-POST (Indian Language – Part of Speech Tagset).

1.8.2.4 Other Tagsets

There are only these two tagsets for Hindi that have been so far purposefully developed for NLP and theoretical linguistics tasks. Both of these tagsets have been developed not exclusively for Hindi, but considering the Indian languages in mind. Some features may be missing or even redundant in both of these tagsets. But then one always can modify to allow for the few deviations in grammatical categories found across the languages. Important among others is a tagset developed for Urdu, which to us linguists is just another variety of the same language that is called Hindi. This tagset developed at the Lancaster University (Andrew Hardie, 2003) follows the EAGLES guidelines and is hierarchical.

1.8.3 Annotation Standards

As seen above, the tagsets developed for Hindi are quite a few. Annotating corpus is a labour intensive task. When one task is completed, it is recommended that the corpus annotated should be used for another task as well. Due to a lack of standards, many a time one annotated corpus becomes useless when used with a different approach. To fill this gap and save the labour, standards for such tasks was felt. In linguistics, we already have quite a few glossing rules. The most important among them is the Leipzig-rules which this researcher has used in his previous works for annotating the linguistic features of different languages. Annotating natural language text for computing purposes can be seen a little different from this linguistic glossing. There have

been some standards since the 1990's defined by international consortiums which ensure that such tasks follow a universal rule so that mapping from annotated corpus to another is easier.

1.8.3.1 TEI Standards

The TEI (Text Encoding Initiative), established in 1987 "to develop, maintain, and promulgate hardware- and software-independent methods for encoding humanities data in electronic form"9, has been a pioneering in this regard. The TEI basically recommends a method to encode the corpora in an XML format that is compliant to TEI guidelines. XML is a technique through which digital texts are archived. So, the TEI has specific guidelines on corpus annotation and linguistic annotation (Burnard, Lou and Bauman, S., eds., 2008) . It recommends how a corpus, its subparts like text portions, paragraph, lines, words and the features of the text like locale, setting, author-investigator information as well as the XML encoding of the PoS tagged information attached to a word should be archived.

1.8.3.2 EAGLES Guidelines

The Expert Advisory Group on Language Engineering Standards (EAGLES)[10] is an initiative of the European Commission which aims to accelerate the provision of standards for:

- Very large-scale language resources (such as text corpora, computational lexicons and speech corpora);
- Means of manipulating such knowledge, via computational linguistic formalisms, mark up languages and various software tools;
- Means of assessing and evaluating resources, tools and products.

Although aimed mainly at the European Union languages, this guideline has been applied on Hindi as well which also is member of the Indo-Aryan branch of the Indo-European language. The Lancaster tagset for Urdu and the IL-POST follow the EAGLES guidelines (Leech and Wilson, 1999). The EAGLES guidelines provide hierarchical scheme for part of speech annotation of words in terms of grammatical categories, their features, attributes and values. To accommodate various languages, the morpho-syntactic features are categorised into obligatory, recommended and optional. The major word classes like noun, pronoun, verb, and adjective are

[9] http://www.tei-c.org/About/history.xml 14 April, 2010

[10] http://www.ilc.cnr.it/EAGLES96/home.html 14 April, 2010

termed as categories. There are thirteen such categories. Among the recommended features are the major subtypes of the categories such as proper/common for noun, degree for adjective etc. Optional features/attributes include features like gender number and person, etc. It is a flexible guideline which allows any grammatical feature to be incorporated within the framework and also at the same time be easy to be mapped to other tagsets.

1.8.4 A preliminary comparison of Annotation Schemes

As noted above, for Hindi and also for Indian languages in general, we have two tagsets that have been used widely to annotate corpus. These two are the IIIT Hyderabad tagset and the IL-POST tagset. Before we do a comparison, it would be apt to have a brief look at the issues that occur in PoS tagset design, some of which though has already come up above. These issues can be summed up in four categories:-

- Elaborate tagset or Basic Tagset
- Flat Tagset vs. Hierarchical Tagset
- Which features to include at the PoS level?

The first issue is whether we need an elaborate tagset or just the basic, as the one with the Penn tagset would suit the purpose. This question depends on some factors like what we need. Tagsets like the IIIT tagset are called coarse tagset as they cover only basic categories and shallow lexical-syntactic features. The aim here is to reduce load on machine learning and also confusion that might arise in the manual annotation of text. The finer details like inflectional and derivational forms can be identified through a morphological analyzer. The elaborate or fine-grained tagsets like the IL-POST aim to capture finer linguistic features attached with words at the lexical level.

Another issue is whether a tagset is a flat tagset or hierarchical. Most of the earlier tagsets, like the Penn tagset and thereby the IIITH tagsets, are flat tagsets. Flat tagsets do not distinguish between the category and its sub-types and mark the words with just one tag which includes all the features put together. Unlike the flat tagsets the hierarchical tagsets are layered so that types and subtypes and their attributes and values can be accessed separately. This layered setting also makes it easier to map corpora annotated in other schemes to this one.

The features to include during the part of speech annotation are another issue. Many people propose that as the parts of speech are basically lexical items it should be restricted to the lexical level only. However, in running text or even otherwise, an annotator is often faced with situations when the lexical items contain both lexical and syntactic information. However, this differs from language to language.

Coming back to the context of Hindi, as stated above, we already have two major tagsets: the IIIT Hyderabad tagset and the IL-POST. The present researcher has also been involved with the development of the latter tagset as an annotator for Hindi through which a corpus of collected from the Web Duniya website was manually tagged over a period of four months. As the researcher is writing this, there is a running debate and discussion on as to which among these two tagsets should be preferred for Indian languages. And the result so far has been in favour of the IL-POST due to the inherent advantages it offers, namely the flexibility, hierarchical, and the ability to map onto other tagsets/annotation schemes.

1.8.4.1 IIIT Hyderabad Tagset

The tagset developed at the IIIT Hyderabad is a flat tagset with tags for both parts of speech and local word groups (chunks). The following tables show the tags of this tagset.

No.	Category	Type	Tag name
1.1	Noun	Common	NN
1.2		Spatio-Temporal	NST
2.		Proper	NNP
3.1	Pronoun		PRP
3.2	Demonstrative		DEM
4	Verb	Main Verb	VM

5		Auxiliary Verb	VAUX
6	Adjective		JJ
7	Adverb		RB
8	Postposition		PSP
9	Particles		RP
10	Conjuncts		CC
11	Question Words		WQ
12.1	Quantifiers		QF
12.2	Cardinal		QC
12.3	Ordinal		QO
12.4	Classifier		CL
13	Intensifier		INTF
14	Interjection		INJ
15	Negation		NEG
16	Quotative		UT
17	Sym		SYM
18	Compounds		*C
19	Reduplicative		RDP
20	Echo		ECH
21	Unknown		UNK

Table 3: Parts of Speech Tags of Hindi in the IIITH tagset

The parts of Speech tags in the IIITH framework (Bharati, A. et.al., 2006) is fixed at 26 (plus a few more tags as the compounds can have a noun, a verb, an adjective or an adverb to increase the number to 30). An example sentence of how a corpus annotated with the IIITH tagset would look like is as follows:

E378	पटना	में	उसने	एक	अच्छा	घर	देखा	।
IPA	pəʈna:	mẽ:	ʊsne:	e:k	əccʰa:	gʰər	de:kʰa:	|
TAG	NNP	PSP	PRP	QC	JJ	NN	VM	SYM
Eng.	He saw a good house in Patna.							

The above example sentence when annotated in a corpus would look like:

पटना_NNP में_PSP उसने_PRP एक_QC अच्छा_JJ घर_NN देखा_VM ।_SYM

The IITH tagset also denotes the tags for the local word groups (chunk) label. The following table shows this list.

No	Chunk Type	Tag Name
1	Noun Chunk	NP
2.1	Finite Verb Chunk	VGF
2.2	Non-finite Verb Chunk	VGNF
2.3	Infinitival Verb Chunk	VGINF
2.4	Verb Chunk (Gerund)	VGNN
3	Adjectival Chunk	JJP
4	Adverb Chunk	RBP
5	Chunk for Negatives	NEGP

No	Chunk Type	Tag Name
6	Conjuncts	CCP
7	Chunk Fragments	FRAGP
8	Miscellaneous	BLK

Table 4: Chunk Tags in the IIITH Tagset

1.8.4.2 IL-POST Framework Tagset for Hindi

For the tasks of analysis in the present research work, this researcher has gone for the IL-POST tagset. Works are already done towards the development of the Hindi specific tagset. Below I give a brief outline of the IL-POST tagset for Hindi[11].

In the Hindi tagset, there are ten major word classes (categories). These are the obligatory categories i.e. every word class belong to one of these ten classes. These are as shown in the following table:

No.	Category	Type	Tag (Top Level)
1	Noun	Common Noun	NC
2		Proper Noun	NP
3		Verbal Noun	NV
4		Spatio-Temporal	NST
5	Verbs	Main Verb	VM
6		Auxiliary Verb	VAUX
7	Pronoun	Pronominal Pronoun	PPR
8		Reflexive Pronoun	PRF
9		Reciprocal Pronoun	PRC
10		Relative Pronoun	PRL

[11] MSRI India, Bangalore already had a guideline prepared for Hindi annotation. Most of the description as given in this section is from this guideline. This guideline comes with LDC corpus available available online at the this link: http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2010T24

11		Wh- Pronoun	PWH
12	Nominal Modifier	Adjective	JJ
13		Quantifiers	JQ
14	Demonstrative	Absolutive Demonstratives	DAB
15		Relative Demonstrative	DRL
16		Wh- Demonstrative	DWH
17	Adverb	Manner Adverb	AMN
18		Location Adverb	ALC
19	Adposition	Postposition	PP
20	Particles	Coordinating Particles	CCD
21		Subordinative Particles	CSB
22		Classifier Particles	CCL
23		Interjection Particles	CIN
24		Other Particles	CX
25	Punctuation	Punctuation	PU
26	Residual	Foreign Word	RDF
27		Symbol	RDS
28		Other	RDX

Table 5: Category Types of Hindi in IL-POST Framework

Compared to the IIIT tagset, the top level tags assigned to a word in IL-POST framework is 28 as against 26 of the former. This number can be increased or decreased easily if there is a need to add new categories or their types. These top level tags are the first level tags of any word annotated in this framework. Other tags such as attributes and their values are attached with a dot (.) mark separator. As obvious, one category type can get more than one attribute and their corresponding values. For example, a noun takes the attributes of gender, number and person (GNP). These attributes, although not grammatically hierarchical, can be arranged in an ordered manner so as to appear hierarchical. This fixed order of appearance of these features on a tag is necessary to facilitate access to them in a systemic way. The attributes and its values as applied to various category types above are as follows:

No.	Attribute	Total No.	Value	Tag

		of Values		
1	Gender	2	Masculine	mas
2			Feminine	fem
3	Number	2	Singular	sg
4			Plural	pl
5	Person	3	First	1
6			Second	2
7			Third	3
8	Tense	3	Present	prs
9			Past	pst
10			Future	fut
11	Case	2	Direct	dir
12			Oblique	obl
13	Case Marker	11	Ergative	erg
14			Accusative	acc
15			Instrumental	ins
16			Dative	dat
17			Genitive	gen
18			Sociative	soc
19			Locative	loc
20			Ablative	abl
21			Benefective	bnf
22			Vocative	voc
23	Aspect	3	Simple	smp
24			Progressive	prg
25			Perfect	prf
26	Mood	3	Declarative	dcl
27			Imperative	imp
28			Habitual	hab
29	Finiteness		Finite	fin

30			Non-Finite	nfn
31			Infinitive	ifn
32	Distributive		Yes	y
33			No	n
34	Definiteness		Yes	y
35			No	n
36	Emphatic		Yes	y
37			No	n
38	Negative		Yes	y
39			No	n
40	Distance		Proximal	prx
41			Distal	dst
42	Incl/Excl		Inclusive	inl
43			Exclusive	exl
44	Honorificity		Yes	y
45			No	n
46	Numeral		Ordinal	ord
47			Cardinal	crd
48			Non-numeral	nnm

Table 6: Atrributes and Values of Hindi in the IL-POST framework

Besides these attributes, in the IL-POST framework, there is also a mechanism of marking a feature as null or not applicable. So, for example, all pronouns have the attribute of number and its value as singular and plural. But Hindi also has words that are inherently singular or plural. In such cases the number attribute will not be applicable. To facilitate the avoidance of such redundant features within an attribute set assigned for a category-type, the mechanism of inserting a ‘0’ is present. The presence of ‘0’ at the slot of an attribute value indicates that the attribute is not applicable for that particular word. In cases of doubt for an attribute value, the annotator can mark it as ‘x’. Thus the two notations of ‘0’ and ‘x’ can be inserted into all the types of attributes.

To exemplify how a tag to a word looks like, the following sentence has been tagged in this framework.

E378	पटना	में	उसने	एक	अच्छा	घर	देखा	।
IPA	pəʈna:	meː̃	usne:	e:k	əccʰa:	gʰər	de:kʰa:	|
TAG	NP.mas.sg.dir.loc	PP.0.0.loc	PPR.0.sg.3.obl.erg.n.n.0.n	JQ.0.sg.dir.n.crd	JJ.mas.sg.dir	NC.mas.sg.dir.0	VM.mas.sg.3.pst.pft.dcl.fin.n	PU
Eng.	He saw a good house in Patna.							

The same sentence as it appears in a corpus after being annotated with the parts of speech in the IL-POST framework looks like as follows:

पटना\NP.mas.sg.dir.loc में\PP.0.0.loc उसने\PPR.0.sg.3.obl.erg.n.n.0.n

एक\JQ.0.sg.dir.n.crd अच्छा\JJ.mas.sg.dir घर\NC.mas.sg.dir.0

देखा\VM.mas.sg.3.pst.pft.dcl.fin.n ।\PU

Here a tag has been separated from the actual word with a backward slash '\'. The tag starts first with the main word class i.e. category type tag and followed by the tags applicable to the category type. Each of the 28 category types mentioned above has a fixed list of attributes associated with it which has to be marked with a value or optionally left unmarked if not applicable. So, for example, a category of proper noun (NP) takes four attributes, ordered in this arrangement:

Gender.Number.Case.Case Marking

All the proper nouns in a corpus annotated in this framework will follow this as a rule. In other words, if a word that has the top level tag of NP will mandatorily be followed by four slots separated by dot '.' which will have the values of four attributes in the order of gender, number,

case and case marking. On a side note, one may notice here that case marking has been attached to the lexical category of noun which, in Hindi, inherently does not have a case marker attached to it. In the MSRI tagset, this has been facilitated to also capture the context. In the case above, the proper noun is followed by a locative case marker, which lends its case marking to the preceding noun.

1.9 PART OF SPEECH IDENTIFICATION

Part of speech identification in the context of NLP means its automatic identification. To do this automatic identification, a tool has to be built that does such identification. Such a tool is called Part of Speech tagger.

1.9.1 Description of available PoS Taggers

A brief overview of available parts of speech for Hindi has already been given in Table 1 above. Here, I present the methodology/approach these taggers have reportedly used to do the annotation.

All the taggers mentioned in Table 1, i.e. all the taggers available and known to this researcher for Hindi are statistical. All require an annotated corpus to be trained first before being able to tag new texts, except Shrivastava et.al. (2008). Besides an annotated corpus, these methods may also require other resources like lexicon, dictionaries or some other augmenting linguistic knowledge base. An initial work (Ray et.al., 2003) proposed a rule based approach towards PoS tagging. Consequent works used a variety of techniques, mostly statistical, to do the task of PoS annotation. One thing common among all these taggers is that they all use either the IIITH tagset or a variety of this flat tagset. The accuracy results for these taggers reported have been varied at the PoS level. Various approaches, including statistical and linguistic resources, have been used in these taggers. The statistical techniques used are Hidden Markov Model (HMM) (Singh, A. et.al., 2005, Shrivastava, M. & Bhattacharya, P., 2008), Maximum Entropy Markov Model (Dalal et.al., 2006, 2007), Decision Forests (Pammi et.al., 2007), Conditional Random Field (CRF) (Agrawal, H., 2007), and Neural Network (Parikh, A., 2009).

1.9.1.1 Hidden Markov Model

Hidden Markov Model (HMM) (Rabiner, 1989) is a Markov Model in which the states are hidden. A Markov Model is a probabilistic process over a finite set, $\{S_1, ..., S_k\}$, usually called its

states. Each state-transition generates a character from the *alphabet* of the process. Using this model one finds out the probability of a given state.

HMM has been used in several disciplines where finding out the probability of a particular phenomenon over a large, often ambiguous, dataset. PoS tagging has been one of such tasks where HMMs have been found useful. For Hindi, the HMM model has been reportedly used by Akshay Singh et.al. (2005) and Manish Shrivastava and Pushpak Bhattacharya (2008). Akshay's work makes a chunker for Hindi using this model. Shrivastava's work reports of building a PoS tagger for Hindi using 'naive stemming'. It reports of using a suffix list only as the linguistic resource to break down the input words into smaller units before assigning the PoS labels. The reported accuracy is 93.12%.

1.9.1.2 Maximum Entropy Markov Model

Maximum Entropy Markov Model (Berger, L. A. et.al., 1996; Ratnaparakhi, A. 1996, 1997) is another statistical model used in machine learning and NLP tasks. Dalal et.al. (2006, 2007) uses this statistical method along with other linguistic resources like annotated corpus to build a 'feature rich' PoS tagger for Hindi. The reported average accuracy of the tagger is 94.38% over evaluation with a corpus of 15.5 thousand words.

1.9.1.3 Conditional Random Field

Conditional Random Field is yet another statistical algorithm applied for Indian languages including Hindi. Two reports as in Agrawal, H. and A. Mani (2006) and Avinesh PVS and Gali, K. (2007) have used CRF as the statistical algorithm for PoS tagging and chunking in Hindi. Both the works use CRF algorithm implementation package called "CRF++: Yet Another CRF Package"[12]. While the former uses the CRF alone for both tagging and chunking, the latter (Avinesh et.al.) enhance it with Transformation Based Learning (Brill, E. 1995). The reported accuracy at the PoS level is between 83-93%.

1.9.1.4 Decision Trees

Decision Trees are other statistical approach to decide upon a particular goal given a set of other possible goals. Decision trees algorithm has been used for PoS tagging task in various languages. In Hindi, the first such approach to help disambiguate cases of multiple possible tags for a word

[12] http://chasen.org/~taku/software/CRF++/ accessed 11 March, 2010

was used by Singh, S. et.al. (2006). In this work, a morphologically driven tagger has been shown to achieve a high degree of PoS annotation results. But this morphology driven taggers leaves several cases ambiguous. To solve this ambiguity, a decision tree model was used to disambiguate such cases. The CN2 decision tree algorithm (Clark, P. and T. Niblett, 1989) has been applied for this work. The accuracy achieved over a corpus is 93.45%.

Another decision tree based tagger for Hindi has been shown in Pammi et.al. (2007). This differs from the above work in the way that it does not use a morphological tagger and has multiple decision trees, called decision forests. It also deploys the syllabic structure of Hindi in breaking the words into units to get to the root of the input words. The accuracy results for this tagger for Hindi have been reported to go as high as 69.92.

1.9.1.5 Neural Network

The Concept of neural network has also been applied recently in PoS tagging in Hindi. This experiment, as reported in Parikh, Ankur (2009), uses linguistic resources of dictionary and Hindi WordNet and reports of achieving a result as high as the HMM based taggers. The advantage that it proposes to have is less time spent in machine learning than HMM or CRF based taggers.

1.9.1.6 Part of Speech Identification for the present work

When this researcher started this work, there were two to four taggers reported for Hindi. It was assumed that one of these taggers will be used for this purpose. But as the time has passed and Hindi NLP has progressed from infancy to its adolescence, several other taggers using varied methods applied in this area have come up. All the available Hindi taggers as discussed above are either statistical throughout or use some components that involves statistical decision making. However, statistical method is not the only method of doing automatic PoS annotation. It has been argued several times that rule based methods would not yield good results in PoS tagging for Indian languages as they are morphologically richer compared to other languages such as English. Some of the taggers available for Hindi, although, does use rule base to some extent (eg. Singh, S. et.al., 2006).

No attempt so far has been reported for a completely rule-based method, using linguistic rules, for PoS annotation in Hindi. This researcher had first plans to develop a rule-based PoS tagger

for Hindi. But later he had to drop it because it required some lexical resources which were beyond the purview of the proposed research and there was not enough time to develop such resources. This researcher is now to use the annotated data instead as input for the LWG tool he is to produce as an output of this thesis.

1.10LOCAL WORD GROUPING (CHUNKING)

Chunks can be defined as a syntactically correlated group of words which typically 'consists of single content word surrounded by a constellation of function words' (Abney, 1991). Identifying these chunks is a major step in natural language parsing as these represent the major units of a sentence.

1.10.1 Issues in Automatic Identification of a Word Group

Word groups are to be identified and grouped together along with a tag that represents the group. The tags and their labels might vary from one language to another language or from one annotation scheme to another annotation scheme. The IIITH tagset as described above has total of 11 chunk types and subtypes as mentioned in Table 4 above.

The tags for the verb groups might have various types of information about the group. The information extracted for a verb might also depend on how this information is to be used. Thus for example, a discourse analyzer might need to know just the TAM of a sentence or clauses within it. A context sensitive speller, however, might also need the agreement markings to be encoded on the VG so that if any errors occurred, it could be suggested.

In my approach, I limit myself to capturing the TAM only. If agreement markings are also needed, this can be easily tweaked on the current solution I have shown in the later chapters.

1.10.2 Approaches for Chunking

Chunking depends on the output rendered by PoS tagger. Like the various PoS tagger methods discussed above, chunking for Hindi also has received a varied approach. The most common among them are the HMM and CRF. No rule based approach has been attempted here as either.

1.10.3 Approach for Chunking in the present work

The present work has limited itself only to mark the chunks of the verbal group. So, based on the results of the tagger, the chunker module, which for the present researcher is a purely rule-based

tool, will mark the chunk boundaries in any given input sentence. The details of how it works, the rules and implementation has been given in the ensuing chapters.

But to summarize the approach I have taken, it would be enough to say here at the outset that the approach involves first searching for the underlying structure of each of the verb groups found in Hindi and then build an abstract of them. This exercise results into a templatic structure running into hundreds of such templates. These templates are then given an appropriate tag that identifies the TAM of template and also contains some other information.

1.11 LANGUAGE ENCODING

The issue of language encoding in an implementation-based task has long remained an issue. Before the advent of Unicode in Indic scripts, several mechanisms were used to represent Hindi using available character encodings or Roman. Transliteration schemes like ITRANS were devised. Hindi also has several fonts like Shusha, Krutidev, etc. For encoding Hindi throughout the thesis and in the implementation in the corpus, lexicon, paradigms etc, this researcher has used Unicode for the obvious reasons that Unicode is a standard method to represent characters universally in different languages which remains valid across various electronic platforms.

Only for the examples that will occur in Hindi, this researcher will be giving an IPA encoding. For glosses, we use the abbreviations that are tags in the IL-POST framework. This is different from the standard practice in linguistics as this researcher himself has earlier used Leipzig Glossing Rules[13].

[13] http://www.eva.mpg.de/lingua/resources/glossing-rules.php accessed 12 March, 2011

2 HINDI VERB MORPHOLOGY

In this chapter I present an overview of the Hindi morphology in general and verb morphology in particular. A detailed account of verb morphology is needed to decide upon what are the features that can be captured through this exercise.

2.1 GENERAL DESCRIPTION OF HINDI

Hindi is a language with about 500 million native speakers. Amidst several dialects, it also has a standard variety which at the structural level encompasses all the varieties and is intelligible to the native speakers of all the varieties. This standard variety is often represented in the national Hindi media houses and propagated by the government of India as the standard. Throughout this thesis, we stick to this variety of Hindi. This is a rich language with a rich literary heritage and good grammatical tradition.

Several grammars have been written for Hindi. Among the descriptive grammars of Hindi, the prominent are Kellogg, S. H. (1893/1990); Bahl, K. (1967); McGregor, R.S. (1972); Guru, K. (1978); Vajpeyi, K. (1979); Bhatia, Tej K. (1987); Singh, R. and Agnihotri, R.K. (1997); Montaut, Annie (2004), Kachru, Y. (2006); Agnihotri, R.K. (2006) and Koul, O.N. (2008). Besides these descriptive grammars dealing with the language on the whole, there are also several works covering different aspects, issues and nuances of the language both in different frameworks and in general. Here I present just a very brief summary of the features of Hindi as a natural language.

Hindi is a language of the Indo-Aryan branch of the Indo-European family. Typologically, Hindi is an SOV language with a modifier-modified structure and post-positional case marking.

2.1.1 Phonology of Hindi and its orthographic representation

According to Kachru, Y. (2006) Hindi has ten vowels and 38 consonants, including five borrowed consonants. Table 7 below shows consonants as it appears in the Devanagari alphabet as used in Hindi (shown here in IPA). The greyed out symbols shown in this table are not considered as phonemes in Hindi but there are still separate character symbols in Devanagari to show them and they are used commonly.

Place of Articulation →		Bilabial		Labio-Dental		Alveo-Dental		Alveolar		Retroflex		Palatal		Velar		Uvular	
Stricture ↓		Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd	Vl	Vd
Stop	Voice	p	b			t	d			ʈ	ɖ	c	ɟ	k	g	(q)	(ɢ)
	Asp.	p^h	b^h			t^h	d^h			$ʈ^h$	$ɖ^h$	c^h	$ɟ^h$	k^h	g^h		
Fricative				f			z	s		ʂ		ʃ				(x)	
Nasal			m				n				ɳ		ɲ		ŋ		
Lateral									l								
Trill									r		ɽ						
Semi-Vowel				w											y		h

Table 7: Consonants of Hindi as represented through the Devanagari Alphabet

The ten vowels of Hindi are shown below in Table 8.

	Front		Central		Back	
HIGH	i:					u:
		ɪ			ʊ	
MID	e:					o:
		ɛ:			ɔ:	
LOW						
			ə	a:		

Table 8: Vowels of Hindi as represented through the Devanagari Alphabet

Devanagari letters of the alphabet can be mapped with their IPA equivalents as shown in Table 9 below. Unicode values for the Devanagari characters have also been given.

	Main Nagari Alphabet Letters				Main Nagari Vowel Diacritics			Other Nagari Diacritics/Characters		
	Nagari	**Unicode Value**	**IPA**	**Phonemic Status?**	**Diacritics**	**Unicode Value**	**Phonemic Status?**	**Diacritic/Symbols**	**Unicode Value**	**Phonemic Status?**
Vowels	अ	U+0905	ə	Y	NA	NA	NA	ॅ	U+0901	Y
	आ	U+0906	a:	Y	ा	U+093E	Y	ं	U+0902	Y
	इ	U+0907	i:	Y	ि	U+093F	Y	ॉ	U+0949	??
	ई	U+0908	ɪ	Y	ी	U+0940	Y	ः	U+0903	N
	उ	U+0909	u:	Y	ु	U+0941	Y	़	U+94D	N
	ऊ	U+090A	ʊ	Y	ू	U+0942	Y	ऋ	U+090B	N
	ए	U+090F	e:	Y	े	U+0947	Y	ॠ	U+0960	N
	ऐ	U+0910	ɛ:	Y	ै	U+0948	Y	ऌ	U+090C	N
	ओ	U+0913	o	Y	ो	U+94B	Y	ॡ	U+0961	N
	औ	U+0914	ɔ:	Y	ौ	U+94C	Y	ढ़	U+095D	??
Con	क	U+0915	k	Y				़	U+093C	N

	क़	U+0958	q	?				ॐ	U+0950	N
	ख	U+0916	k^h	Y				ऽ	U+093D	N
	ख़	U+0959	x	?				।	U+0964	N
	ग	U+0917	g	Y				॥	U+0965	N
	ग़	U+095A	ɢ	?				॰	U+0970	N
	घ	U+0918	g^h	Y				ृ	U+0943	N
	ङ	U+0919	ŋ	Y				ॄ	U+0944	N
	च	U+091A	c	Y				ॢ	U+0962	N
	छ	U+091B	c^h	Y				ॣ	U+0963	N
	ज	U+091C	ɟ	Y				ॅ	U+0945	N
	ज़	U+095B	z	Y				ॆ	U+0946	N
	झ	U+091D	$ɟ^h$	Y				ॊ	U+94A	N
	ञ	U+091E	ɲ	Y						
	ट	U+091F	ʈ	Y						
	ठ	U+0920	$ʈ^h$	Y						
	ड	U+0921	ɖ	Y						
	ड़	U+095C	ɖ	Y						
	ढ	U+0922	$ɖ^h$	Y						
	ण	U+0923	ɳ	Y						

	त	U+0924	t	Y						
	थ	U+0925	t^h	Y						
	द	U+0926	d	Y						
	ध	U+0927	d^h	Y						
	न	U+0928	n	Y						
	प	U+092A	p	Y						
	फ	U+092B	p^h	Y						
	फ़	U+095E	f	Y						
	ब	U+092C	b	Y						
	भ	U+092D	b^h	Y						
	म	U+092E	m	Y						
	य	U+092F	y	Y						
	र	U+0930	r	Y						
	ल	U+0932	l	Y						
	व	U+0935	w	Y						
	श	U+0936	ʃ	Y						
	ष	U+0937	ʂ	Y						
	स	U+0938	s	Y						
	ह	U+0939	h	Y						

Table 9: Devanagari Letters, their Unicode Values and corresponding IPA symbols

As seen in Table 9 above, there are more letters in the Devanagari alphabet than the phonemes. Besides, the syllabic property of the script takes the number of characters higher. Not only this, the formation of ligatures of consonants to represent the consonant clusters which behave as a single character poses even greater challenge in computing.

2.1.2 Hindi Morphology

Morphologically Hindi is an inflecting language. Inflections with the nouns and verbs contribute the greatest number of words in its vocabulary. All the major content class words, namely, noun, verb and adjective inflect. Derivational formation with all these classes of words is also very rich. At the formal level, the Hindi word classes can be divided into two categories: declinable and indeclinable. Declinable class of words are noun, verb and adjective. All the other word classes, namely, adverb, pronoun, conjunctions, interjections and particles etc., can be put under the category of indeclinable. The inflectional morphology applies to the class of declinable words.

2.1.2.1 Noun

A single word noun delimited by space can fall into four distinct semantic categories: common nouns, proper nouns, spatio-temporal nouns and verbal nouns. Nominal inflections take predominantly on common nouns. Other categories rarely take the same inflections. Nouns in Hindi inflect for number and case. Gender is inherent in all the nouns and all the nouns must fall into either masculine or feminine gender. A two way distinction of singular and plural for the common noun is made in the number. Default is the singular while plural is made through regular suffixes. A three way distinction is made for case in the nouns - direct, which is the default word form, oblique and vocative. Case can also be identified through regular suffix marking.

2.1.2.1.1 Nominal Categories

The four categories with examples are shown in the table below.

Nominal Category	Examples	Inflections taken
Common Noun	घर /gʰər/(home), विद्यालय /wɪdja:ləj/ (school), गाय /ga:j/ (cow) etc.	Number, Case (Gender)

Proper Noun	दिल्ली /dɪlli:/ (Delhi), राम /ra:m/ (Ram), हिमालय /hɪma:ləj/ (Himalaya) etc.	None
Spatio-Temporal Nouns	आज /a:ɟ/ (today), कल /kəl/ (tomorrow), दक्षिण /dəkʃɪn/ (South), उत्तर /ʊttər/ (North) etc.	Number (only in some cases)
Verbal Nouns	करना /kərna:/, करनी /kərni/, करने /kərne:/ (infinitive forms of the verb 'do') etc.	Number case and gender (in agreement with the argument)

Table 10: Nominal Categories

2.1.2.1.2 Inflections: Gender

All the nouns in Hindi must have a gender pre-assigned to them. Although gender is said to be inherent or natural in all the Hindi nouns, in some cases the morphology, particularly the word ending, can predict the gender of the nominal category. Identifying the gender of a noun given a random word in Hindi is a very difficult task as many of the time it is arbitrary. The rules that are there to identify it can often lead into ambiguities.

2.1.2.1.3 Inflections: Number

There is a two-way distinction of number in all the common noun. A noun, mostly common noun, can be either singular or plural. The default form is singular while plural can be made using suffixes. The suffixes that make a noun plural depend on three features –word ending, gender and case.

2.1.2.1.4 Inflections: Case

Three cases are marked directly on the nouns. These are direct, oblique and vocative. The default word form may have the direct case marked on it.

A paradigm can be generated to account for the word formation involving the change in gender, number and case. The following table shows how a noun changes its form from singular to plural in different cases. With this table, it is also clear that without the information on gender, it would be difficult to generate these plural and case forms at the lexical level. However, using the word

endings, one can sure get to the root of the word by just removing the suffix and adding some additional measures.

Gender	Case	Default Word Ending	Singular Form Word Ending	Plural Form Word Ending
Masculine	Direct	_a		_e
	Oblique		_e	_o˜
	Vocative			_o
Feminine	Direct			_e˜
	Oblique			_o˜
	Vocative			_o
Masculine	Direct	_i		_ ɸ
	Oblique			_ ɪyo˜
	Vocative			_ɪyo
Feminine	Direct			_ɪya˜
	Oblique			_ɪyo˜
	Vocative			_ɪyo
Masculine	Direct	_ɪ		_ɸ
	Oblique			_ɪyo˜
	Vocative			_ɪyo
Feminine	Direct			_ɪya˜
	Oblique			_ɪyo˜
	Vocative			_ɪyo
Masculine	Direct	_ʊ		_ ɸ
	Oblique			_ uo˜
	Vocative			_uo
Masculine/ Feminine	Direct	_C		_Ce˜
	Oblique			_Co˜
	Vocative			_Co

Table 11: Nominal Inflection

2.1.2.2 Adjectives

Adjectives, at the formal level, can be divided into two groups, namely adjectives proper (or attributive) and quantifiers. Morphologically, adjectives inflect for gender, number and case. They also have comparative and superlative forms; however these are formed mostly periphrastically in Hindi. Only the adjectives that have /_a/ ending change their forms inflectionally for gender, number and case. In such cases, the /_a/ ending, which for adjectives is by default masculine, singular and in direct case form, change to two forms ending in /_i/ and /_e/. These /_i/ and /_e/ ending adjectives respectively denote the feminine and the plural/oblique form. It is notable that the adjectives that have taken the feminine form do not change further for number or case. That is, feminine adjectives having an /_i/ ending remain the same for both the numbers and both the cases.

2.1.2.3 Verbs

The category of verb and its various forms in Hindi has been separately described below in section 2.2 in detail.

2.1.2.4 Indeclinables

2.1.2.4.1 Adverbs

Morphologically, adverbs can be classified into three categories- basic, derived and complex. Basic adverbs are invariable and do not change morphologically. The derived adverbs consist of adverbs derived from noun, pronoun, adjective or verb (Kachru, Y., 2006). Complex adverbs may consist of postpositional phrases and various subordinate clauses that act as adverbs in the sentence.

2.1.2.4.2 Pronouns

Pronouns in Hindi have distinct forms for the categories of person, number, gender, honorificity, proximity and animacy. These are also inflected for case. There are seven classes of pronouns, namely: personal, demonstrative, relative, correlative, indefinite and interrogative. Table 12 below shows all the different word forms of personal pronouns.

			Direct	**Oblique**	**Genitive**	**Accusative/Dative**		**Locative**	**Ablative/Sociative**
						Form 1	Form 2		
1Sg		Mas	मैं /mɛ̃:/	मुझ /mʊjʰ/	मेरा /me:ra:/	मुझको /mʊjʰko:/	मुझे /mʊjʰe:/	मुझमें /mʊjʰme:/	मुझसे /mʊjʰse:/
		Fem			मेरी /me:ri:/				
		Neu			मेरे /me:re:/				
1Pl		Mas	हम /həm/	हम /həm/	हमारा /həma:ra:/	हमको /həmko:/	हमें /həmẽ:/	हममें /həməmẽ:/	हमसे /həmse:/
		Fem			हमारी /həma:ri:/				
		Neu			हमारे /həma:re:/				
2sg		Mas	तू/तूं /tu:/~/tũ:/	तुझ /tʊjʰ/	तेरा /te:ra:/	तुझको /tʊjʰko:/	तुझे /tʊjʰe:/	तुझमें /tʊjʰmẽ:/	तुझसे /tʊjʰse:/
		Fem			तेरी /te:ri:/				

		Neu			तेरे /te:re:/				
		Mas	तुम /tʊm/	तुम /tʊm/	तुम्हारा /tʊmha:ra:/	तुमको /tʊmko:/	तुम्हें /tʊmhẽ:/	तुममें /tʊməmẽ:/	तुमसे /tʊmse:/
		Fem			तुम्हारी /tʊmha:ri:/				
		Neu			तुम्हारे /tʊmha:re:/				
2sg/pl		Mas	आप /a:p/	आप /a:p/	आपका /a:pka:/	आपको /a:pko:/		आपमें /a:pmẽ:/	आपसे /a:pse:/
		Fem			आपकी /a:pki/				
		Neu			आपके /a:pke:/				
3sg	Proximal	Mas	यह /jəh/	इस /ɪs/	इसका /ɪska:/	इसको /ɪsko:/	इसे /ɪse:/	इसमें /ɪsmẽ:/	इससे /ɪsəse:/
		Fem			इसकी /ɪski:/				
		Neu			इसके /ɪske:/				

	Distal	Mas	वह /wəh/	उस /ʊs/	उसका /ʊskaː/	उसको /ʊskoː/	उसे /ʊseː/	उसमें /ʊsmẽː/	उससे /ʊsəseː/
		Fem			उसकी /ʊskiː/				
		Neu			उसके /ʊskeː/				
3pl	Proximal	Mas	ये /je/	इन /ɪn/	इनका /ɪnkaː/	इनको /ɪnkoː/	इन्हें /ɪnhẽː/	इनमें /ɪnmẽː/	इनसे /ɪnseː/
		Fem			इनकी /ɪnkiː/				
		Neu			इनके /ɪnkeː/				
	Distal	Mas	वे /we/	उन /ʊn/	उनका /ʊnkaː/	उनको /ʊnkoː/	उन्हें /ʊnhẽː/	उनमें /ʊnmẽː/	उनसे /ʊnseː/
		Fem			उनकी /ʊnkiː/				
		Neu			उनके /ʊnkeː/				

Table 12: Different forms of the Personal Pronouns

The locative and the ablative/instrumental forms of the personal pronouns shown in the table above might also appear separately, but many a times the case markers appear together with the oblique form of the pronoun and that is why these forms have also been included here to appear in the list of pronominals. Besides these forms of the personal pronouns, there is an emphatic form that is morphologically marked with the pronouns. Table 13 below lists these emphatic forms of the personal pronouns.

			Oblique	**Genitive**	**Acc./Dat.**	**Locative**	**Abl/Soc.**

1Sg		Mas	मुझी /mʊjʰi/		मुझीको /mʊjʰi:ko:/	मुझीमें /mʊjʰi:mẽ:/	मुझीसे /mʊjʰi:se:/
		Fem					
		Neu					
1Pl		Mas	हमीं /həmĩ:/		हमींको /həmĩ:ko:/	हमींमें /həmĩ:me:/	हमींसे /həmĩse:/
		Fem					
		Neu					
2sg/pl		Mas	तुझी /tʊjʰi:/		तुझीको /tʊjʰi:ko:/	तुझीमें /tʊjʰi:mẽ:/	तुझीसे /tʊjʰi:se:/
		Fem					
		Neu					
		Mas	तुम्हीं /tʊmhĩ:/		तुम्हींको /tʊmhĩ:ko:/	तुम्हींमें /tʊmhĩ:mẽ:/	तुम्हींसे /tʊmhĩ:se:/
		Fem					
		Neu					
		Mas					
		Fem					
		Neu					
3sg	Proximal	Mas	इसी /ɪsi:/	इसीका /ɪsi:ka:/	इसीको /ɪsi:ko:/	इसीमें /ɪsi:mẽ:/	इसीसे /ɪsi:se:/
		Fem		इसीकी /ɪsi:ki:/			
		Neu		इसीके /ɪsi:ke:/			
	Distal	Mas	उसी /ʊsi:/	उसीका /ʊsi:ka:/	उसीको /ʊsi:ko:/	उसीमें /ʊsi:mẽ:/	उसीसे /ʊsi:se:/

		Fem		उसीकी /ʊsi:ki:/			
		Neu		उसीके /ʊsi:ke:/			
3pl	Proximal	Mas	इन्हीं /ɪnhĩ:/	इन्हींका /ɪnhĩ:ka:/	इन्हींको /ɪnhĩ:ko:/	इन्हींमें /ɪnhĩ:mẽ:/	इन्हींसे /ɪnhĩ:se:/
		Fem		इन्हींकी /ɪnhĩ:ki:/			
		Neu		इन्हींके /ɪnhĩ:ke:/			
	Distal	Mas	उन्हीं /ʊnhĩ:/	उन्हींका /ʊnhĩ:ka:/	उन्हींको /ʊnhĩ:ko:/	उन्हींमें /ʊnhĩ:mẽ:/	उन्हींसे /ʊnhĩ:se:/
		Fem		उन्हींकी /ʊnhĩ:ki:/			
		Neu		उन्हींके /ʊnhĩ:ke:/			

Table 13: Emphatic Forms of the Personal Pronouns

Other forms of pronouns are listed in Table 14 below.

	Gender	**Direct**	**Oblique**	**Genitive**	**Accusative/Dative**		**Locative**	**Ablative/ Sociative**
					Form1	**Form2**		
Relative Singular	Mas	जो /ɟo/	जिस /ɟɪs/	जिसका /ɟɪska:/	जिसको /ɟɪsko:/	जिसे /ɟɪse:/	जिसमें /ɟɪsmẽ:/	जिससे /ɟɪsəse:/
	Fem			जिसकी /ɟɪski:/				

	Neu			जिसके /jɪske:/				
Relative Plural	Mas	जो /jo/	जिन /jɪn/	जिनका /jɪnka:/	जिनको /jɪnko:/	जिन्हें /jɪnhẽ:/	जिनमें /jɪnmẽ:/	जिनसे /jɪnse:/
	Fem			जिनकी /jɪnki:/				
	Neu			जिनके /jɪnke:/				
Indefinite Singular	Mas	कोई /koi:/	किसी /kɪsi:/	किसीका /kɪsi:ka:/	किसीको /kɪsi:ko:/		किसीमें /kɪsi:mẽ:/	किसीसे /kɪsi:se:/
	Fem			किसीको /kɪsi:ko:/				
	Neu			किसीके /kɪsi:ke:/				
Indefinite Plural	Mas	कुछ /kʊcʰ/	किन्हीं /kɪnhĩ:/	किन्हींका /kɪnhĩ:ka:/	किन्हींको /kɪnhĩ:ko:/			
	Fem			किन्हींकी /kɪnhĩ:ki:/				
	Neu			किन्हींके /kɪnhĩ:ke:/				
Interrogative Animate Singular	Mas	कौन /kɔ:n/	किस /kɪs/	किसका /kɪska:/	किसको /kɪsko:/	किसे /kɪse:/	किसमें /kɪsmẽ:/	किससे /kɪsəse:/
	Fem			किसकी /kɪski:/				

	Neu			किसके /kɪske:/				
Interrogative Animate Plural	Mas	कौन /kɔ:n/	किन /kɪn/	किनका /kɪnka:/	किनको /kɪnko:/	किन्हें /kɪnhẽ:/	किनमें /kɪnmẽ:/	किनसे /kɪnse:/
	Fem			किनकी /kɪnki:/				
	Neu			किनके /kɪnke:/				
Interrogative Inanimate	Mas	क्या /kja:/	किस /kɪs/	किसका /kɪska:/	किसको /kɪsko:/	किसे /kɪse:/	किसमें /kɪsmẽ:/	किससे /kɪsəse:/
	Fem			किसकी /kɪski:/				
	Neu			किसके /kɪske:/				
Other Interrogative Pronouns		कब /kəb/						
		कहाँ /kəhã:/						
		कितना/कितनी/कितने /kɪtna:/~/kɪtni:/~/kɪtne:/						
		कैसा/कैसी/कैसे /kɛ:sa:/~/kɛ:si:/~/kɛ:se:/						
		क्यों/क्यों/क्यूं/क्यूँ /kyõ:/~/kyũ:/						

Other Interrogative Indefinite Pronoun	कभी /kəbʰi:/						
	कहीं/कहीँ /kəhĩ:/						
Other Relative Pronouns	जब /ɟəb/						
	जहाँ /ɟəhã/						
	जितना/जितनी/जितने /ɟɪtna:/~/ɟɪtni:/~/ɟɪtne:/						
	जैसा/जैसी/जैसे /ɟɛ:sa:/~/ɟɛ:si:/~/ɟɛ:se:/						
	ज्यों/ज्यों/ज्यूं/ज्यूँ /ɟyõ:/~/ɟyũ:/						
Other Indefinite Relative Pronoun	जभी /ɟəbʰi:/						

Table 14: Non-Personal Pronouns in Hindi

2.1.2.4.3 Demonstratives

Demonstrative pronouns are the same as the personal pronouns in the third person proximal and distal categories. Yet they have to be given a separate treatment because they behave like modifiers with the nouns and nominal categories and indicate distance and direction.

2.1.2.4.4 Postpositions

Postpositions in Hindi are of two types, basic and complex. Kachru, Y. (2006) lists a total of seven postpositions of Hindi. These are: ने /ne:/, को /ko:/, से /se:/, का /ka:/, के /ke:/, में/mẽ:/ and

पर /pər/. Complex postpositions are formed with two words. The participating words that can form a postposition with one of these basic postpositions, predominantly with 'के' /ke:/, can be a separate word in itself. Although together with these basic postposition they always have the meaning of a postposition. Examples of the second word constituents forming a complex postposition are लिए /lɪje:/, बाद /ba:d/, सामने /sa:mne:/, बगल /bəgəl/ etc.

2.1.2.4.5 Particles

Particles are basically words that do not decline. Examples are ही /hi:/, तो /to:/, भी /bʰ:i/, और /ɔ:r/, तथा /tətʰa:/, etc.

2.1.2.4.6 Punctuations

Punctuation marks in Hindi is almost the same as in English except for the symbol of full stop for which Hindi uses a single vertical bar '।' called '*danda*'. Another symbol might be the double *danda* '॥'the symbol used particularly to mark the end of the second line of a couplet in a verse.

2.1.2.4.7 Residual

The residual category contains words that are not proper words like characters from non-Hindi alphabets, foreign words, symbols and others.

2.2 HINDI VERB MORPHOLOGY

The verbs form the main part of this research work. In this section I am dealing with mainly the verbal morphology, more so of the inflectional kind. Kachru, Y. (2006) states that morphologically, verbs have the following forms expemplified with three verbs रुक /rʊk/ "to stop", खा /kʰa:/ "to eat", and पी /pi:/ "to drink":

Root	as in	रुक /rʊk/, खा /kʰa:/, पी /pi:/ etc.
Infinitive	as in	रुकना /rʊkna:/, खाना /kʰa:na:/, पीना /pina:/ etc.ʰ

Imperfect Participle as in रुकता /rʊkta:/, खाता /kʰa:ta:/, पीता /pita:/etc.

Perfect Participle as in रुका /rʊka:/, खाया /kʰa:ja:/, पिया /pɪja:/ etc.

Causative as in रोका /roka:/, खिलाया /kʰɪla:ja:/, पिलाया /pɪla:ja:/ etc.

Second-Causative as in रुकवाया /rʊkwa:ya:/, खिलवाया /kʰɪlwa:ya:/, पिलवाया /pɪlwa:ya:/ etc.

Besides these forms of a verb, all the root forms of the verbs inflect for tense, aspect and mood (TAM) and for agreement features of gender, number, person and honorificity (GNPH). TAM and GNPH markings are done morpho-phonemically through suffixes. The only derivation that takes place at the lexical level is the formation of causatives and second causatives (also called double-causatives). These causatives, when formed, can take the normal TAM and GNP markings, so much so that they behave more like verb roots. In this thesis, we deal with the causatives as if they were verb roots.

Given a verb root, (including the causative constructions), we find that there can be a total of 25 forms of it (including the verb root). This has been exemplified below in Table 15 with examples of four such verbs (namely कर /kər/ “to do”, खा /kʰa:/ “to eat”, गिना /gɪna:/ “to cause (someone or something) to count”, and तड़प /təɖəp/ “to languish”.

Sl. No.	**Verb Forms**	**Ex. 1**	**Ex. 2**	**Ex. 3**	**Ex. 4**	**Tag Sequence (as in the IL-PoST Hindi corpus of LDC)**
1	Root / Imperative 2SG/2PL	कर /kər/	खा /kʰa:/	गिना /gɪna:/	तड़प /təɖəp/	VM.0.0.0.0.0.0.nfn.0 / VM.0.0.2.0.0.imp.nfn.0

2	Infinitive MSG	करना /kərna:/	खाना /kʰa:na:/	गिनाना /gɪna:na:/	तड़पना /təɖəpna:/	VM.mas.sg.0.0.0.0.ifn.0
3	Infinitive FSG	करनी /kərni/	खानी /kʰa:ni:/	गिनानी /gɪna:ni:/	तड़पनी /təɖəpni:/	VM.fem.sg.0.0.0.0.ifn.0
4	Infinitive PL	करने /kərne:/	खाने /kʰa:ne:/	गिनाने /gɪna:ne:/	तड़पने /təɖəpne:/	VM.0.pl.0.0.0.0.ifn.0
5	Imperfect MSG	करता /kərta:/	खाता /kʰa:ta:/	गिनाता /gɪna:ta:/	तड़पता /təɖəpta:/	VM.mas.sg.0.0.sim.hab.fin.n
6	Imperfect FSG	करती /kərti:/	खाती /kʰa:ti:/	गिनाती /gɪna:ti:/	तड़पती /təɖəpti:/	VM.fem.sg.3.0.sim.hab.fin.n
7	Imperfect FSGH	करतीं /kərtĩ:/	खातीं /kʰa:tĩ:/	गिनातीं /gɪna:tĩ:/	तड़पतीं /təɖəptĩ:/	VM.fem.0.0.0.sim.hab.fin.y
8	Imperfect PL	करते /kərte:/	खाते /kʰa:te:/	गिनाते /gɪna:te:/	तड़पते /təɖəpte:/	VM.0.pl.3.0.sim.hab.fin.y
9	Perfect MSG	किया /kɪja/	खाया /kʰa:ja:/	गिनाया /gɪna:ja:/	तड़पा /təɖpa:/	VM.mas.sg.0.0.pft.dcl.fin.n
10	Perfect FSG	की /ki:/	खायी /kʰa:ji:/	गिनायी /gɪna:ji:/	तड़पी /təɖpi:/	VM.fem.sg.0.0.pft.dcl.fin.n
11	Perfect FSGH	कीं /kĩ:/	खायीं /kʰa:jĩ:/	गिनायीं /gɪna:jĩ:/	तड़पीं /təɖpĩ:/	VM.fem.0.3.0.pft.dcl.fin.y
12	Perfect PL / [14]Optative 3SG	करे, किए /kəre:/, /kɪe:/	खाये /kʰa:e:/	गिनाये /gɪna:je:/	तड़पे /təɖpe:/	VM.0.pl.0.0.pft.dcl.fin.y

[14] The IL-POST tagset for Hindi that we used to annotate the corpus for the MSRI did not have the mood marker of optative. It is included in the imperative tag itself. However, this researcher feels that to disambiguate some cases

13	Future 3MSG	करेगा /kəre:ga:/	खाएगा /kʰa:e:ga:/	गिनाएगा /gɪna:e:ga:/	तड़पेगा /təɖpe:ga:/	VM.mas.sg.3.fut.0.dcl.fin.n
14	Future 2MSGH	करिएगा, कीजिएगा /kərie:ga:/, /kiji̇e:ga:/	खाइएगा /kʰa:ɪe:ga:/	गिनाइएगा /gɪna:ɪe:ga:/	तड़पिएगा /təɖəpɪe:ga:/	VM.0.sg.0.fut.0.dcl.fin.y
15	Future 3FSG	करेगी /kəregi:/	खाएगी /kʰa:e:gi:/	गिनाएगी /gɪna:e:gi:/	तड़पेगी /təɖpe:gi:/	VM.fem.sg.3.fut.0.dcl.fin.n
16	Future 3FSGH	करेंगी /kərẽ:gi:/	खाएंगी /kʰa:ẽgi:/	गिनाएंगी /gɪna:ẽ:gi:/	तड़पेंगी /təɖpẽ:gi:/	VM.fem.0.0.fut.0.dcl.fin.y
17	Future PL	करेंगे /kərẽ:ge:/	खाएंगे /kʰa:ẽ:ge:/	गिनाएंगे /gɪna:ẽ:ge:/	तड़पेंगे /təɖpẽ:ge:/	VM.0.0.0.fut.0.dcl.fin.y
18	Future 2MSGNH	करोगे /kəro:ge:/	खाओगे /kʰa:o:ge:/	गिनाओगे /gɪna:o:ge:/	तड़पोगे /təɖpo:ge:/	VM.mas.sg.2.fut.0.dcl.fin.n
19	Future 2FSGNH	करोगी /kəro:gi:/	खाओगी /kʰa:o:gi:/	गिनाओगी /gɪna:o:gi:/	तड़पोगी /təɖpo:gi:/	VM.fem.sg.2.fut.0.dcl.fin.n
20	Future 1MSG	करूँगा /kərũ:ga:/	खाऊँगा /kʰa:ũga:/	गिनाऊँगा /gɪna:ũ:ga:/	तड़पूँगा /təɖpũ:ga:/	VM.mas.sg.1.fut.0.dcl.fin.0
21	Future 1FSG	करूँगी /kərũ:gi:/	खाऊँगी /kʰa:ũgi:/	गिनाऊँगी /gɪna:ũgi/	तड़पूँगी /təɖpũ:gi:/	VM.fem.sg.1.fut.0.dcl.fin.0

where the optative form is similar to at least two other forms, we will need to give a separate label for optative as well. At the same time there are other features too that need to be tweaked for this purpose. It has been discussed later.

22	Optative 1SG	करूँ /kərũː/	खाऊँ /kʰaːũː/	गिनाऊँ /gɪnaːũ/	तड़पूँ /təɖpũː/	VM.0.sg.1.0.0.imp.fin.0
23	Imperative Plural Honorific / Optative Plural Honorific	करें /kərẽː/	खायें /kʰaːẽː/	गिनायें /gɪnaːẽ/	तड़पें /təɖpẽː/	VM.0.0.0.0.0.imp.fin.y
24	Imperative Honorific	करिए, कीजिए /kərɪːeː/, /kiɟɪeː/	खाईए /kʰaːieː/	गिनाईए /gɪnaːie/	तड़पिए /təɖəpɪeː/	VM.0.0.2.0.0.imp.fin.y
25	Imperative 2SGNH	करो /kəroː/	खाओ /kʰaːoː/	गिनाओ /gɪnaːo/	तड़पो /təɖpoː/	VM.0.sg.2.0.0.imp.fin.n

Table 15: Different Morphological Forms of a verb

Besides these twenty five forms of a verb there are a few verbs which do not fall under the paradigm of the twenty five forms listed in Table 15 above. The most common among them are the copula verbs. These are a small set of eight verbs: **है** /hɛː/ "3.SG.PRS", **हैं** /hɛ̃ː/ "PL.PRS", **हूं/हूँ** /hũː/ "1.SG.PRS", **हो** /hoː/ "2.SG/PL.PRS", **था** /tʰaː/, "3.SG.M.PST', **थे** /tʰeː/ "SG/PL.PST", **थी** /tʰiː/ "3.SG.F.PST" and **थीं** /tʰĩː/ "3.FSG/FPL.PST" which can function both as main verb and as an auxiliary. There is another set of verbs, usually called V2, that also occur both as main verb and as an auxiliary in a sentence. Examples of such verbs are **रह** /rəh/ "to live", **सक** /sək/ "can/be able to", **चुक** /cʊk/ "to get finished/completed", **पड़** /pəɖ/ "to lie/put" etc.

2.2.1 Verb Forms in Hindi

As seen above, Hindi verbs inflect for TAM and GNPH agreement. These two are discussed below.

2.2.2 Tense, Aspect and Mood (TAM)

2.2.2.1 Tense

Only future tense is marked morphologically in Hindi. Past and present are denoted through aspect and auxiliaries respectively. The marker for future tense in Hindi can be summed up with the suffix of /_ega/ and its variations according to GNPH agreement. We have found that in future, a verb root can have 9 different forms, based on the GNP agreement changes (as shown in rows 13 to 21 in Table 15 above).

One may however, argue that past is denoted through the verbal form itself. What I call as the aspect marker may also appear as the past marker when it comes along as the single verbal word in the sentence. For example, in the sentence below

E1	वह	घर	गया	
IPA	wəh	g^hər	gəja:	
Gl	3SG	home	go.PFT.MSG	
EGl	He went home.			

the tense of the sentence is in past and it is denoted by the single verb 'गया' appearing in the sentence. This might lead one to believe that this is the past marker in Hindi, however I argue against it by saying that it in fact is only the completion marker, an aspectual feature. This argument can be verified by the following example:

E2	वह	घर	गया	है	
IPA	wəh	g^hər	gəja:	hɛ:	
Gl	3SG	home	go.PFT.MSG	be.PRS.SG	
EGl	He has gone home.				

The sentence above also uses the same form of the verb as used in the previous example sentence. However, in this sentence the aspect is clearly marked for perfect. In the previous sentence either there is no aspect or it is what is commonly called simple aspect (Kachru, 2006).

2.2.2.2 Aspect

Morphologically two aspects are marked in Hindi verbs: perfective and imperfective. These aspects are marked morpho-phonemically with the use of suffix /-a/ and /-ta/ respectively for perfective and imperfective. These suffixes again change for the GNPH agreements and get a total of 4 different forms each (as shown in rows 9-12 and 5-8 respectively in Table 15 above).

The imperfective marker as denoted here by the suffix of /-ta/ and its GNPH variants has also been called as the habitual marker. Habitual is basically a mood which also is denoted by the same verbal form in Hindi. Although, it is not always necessary that every sentence having an imperfective verb form playing a part in it would be habitual. Moreover, the habitual mood is a syntactic property and should not be used for a morphological construct. Therefore, I would here prefer to use the term of imperfective for such constructs.

2.2.2.3 Mood

Among the moods, only imperative/optative mood is marked morphologically on the verb roots in Hindi. Verb roots by default can be used as imperative-marked verb. The other forms of the imperative include the honorific form and other GNPH change. The total number of forms these two moods bring about is four. It is notable that if we include the verb root and the perfect plural form of the verbs, the total different forms we get for mood goes up to 6.

2.2.2.4 Agreement

Gender, number and person of the subject gets reflected on the verb in the form of agreement. That is, a Hindi verb in a sentence has to agree for these three features of its subject argument.

Gender is marked on all the forms of the verb except for verb root and imperative mood and brings about a morphological change in the verb root. Number is marked on all the verb forms except in the verb root. Person is marked selectively on some forms. Besides these three, some honorificity is also marked on the verb forms, particularly in the second person and the third person cases. Honorificity in plural is the same as in singular so much so that sometimes it is

confusing to see a plural agreement marking attached with a verb having singular subject argument.

2.3 VERB GROUPS IN HINDI

The term ‘verb group’ is broadly used to denote the lexical element(s) that represent the verbal content in a sentence. ‘Complex predicate’ or ‘complex verb’ is the more common term used in the literature.

By Verb Group here is meant all the elements that form the part of the verb complex. Verb group is a part of a sentence that is recognized as a chunk in morphological parsing (Markantonatoul, Stella et. al. 2005; Watanabe, Taro et. al. 2003). The Indo-Aryan languages have a specific kind of verb structure where the verbal content in a sentence is spread over several terms. That is, a verb in Hindi can have several elements together and still be one semantic category representing verb (and thereby the associated tense, aspect and mood information). Let’s for example see the following sentences:

E3	वह	गया
IPA	wəh	gəja:
Gl	3SG	go.PFT.MSG
EGl	He went.	

E4	वह	जाएगा
MB	wəh	ja:-e:ga:
Gl	3SG	go-FUT.3MSG
EGl	He will go.	

E5 वह जाता है

IPA wəh ja:-ta: hɛ:

Gl. 3SG go.IMPF.MSG be.PRS.SG

EGl He goes.

E6 वह घर जा रहा है

IPA wəh gʰər ja: rəh-a: hɛ:

Gl. 3SG home go.VR live-PFT.MSG be.PRS-SG

EGl He is going home.

E7 वह घर जा ही रहा है

IPA wəh gʰər ja: hi: rəh-a: hɛ:

Gl. 3SG home go.VR PRT live-PFT.MSG be.PRS.SG

EGl He is just going home.

E8 वह घर जा भी रहा है

IPA wəh gʰər ja: bʰi: rəh-a: hɛ:

Gl. 3SG home go.VR PRT live-PFT.MSG be.PRS.SG

EGl He is even *going* home.

E9	वह	घर	जा	ही	तो	रहा	है
IPA	wəh	gʰər	ja:	hi:	to	rəh-a:	hɛ:
Gl.	3SG	home	go.VR	PRT	PRT	live-PFT.MSG	be.PRS.SG
EGl	He is just going home only.						

E10	वह	घर	चला	गया	है
IPA	wəh	gʰər	cəl-a:	gəja:	hɛ:
Gl.	3SG	home	walk-PFT.MSG	go.PFT.MSG	be.PRS.SG
EGl	He has gone home.				

E11	उसने		किताब	फाड़	दी	है
IPA	ʊs-ne:		kɪta:b	pʰa:ɖ	di:	hɛ:
Gl.	3SG.OBL-ERG		book	tear.VR	give.PFT	be.PRS.SG
EGl	He has torn the book.					

In the first and second sentences above (E3 and E4) the verb group consists of one element i.e. “gəja:/jaega”. In the third (E5), it has two elements i.e. “jata hɛ:”. In the fourth (E6), it has three i.e. “ja rəha: hɛ:”. In the fourth and fifth (E7 & E8), it consists of four i.e. “ja hi: rəha: hɛ:” and “ja bʰi: rəha: hɛ:” respectively. In the sixth (E9), it consists of five i.e. “ja hi: to: rəha: hɛ:”. In the seventh and eighth (E10 & E11) the verb group consists of three each but it has a different semantic structure from the earlier four sentences i.e. it consists of compound verbs.

In the above six sentences, we see that a verb group may have different constituent elements yet it has one synthetic semantic content i.e. they form the part of the verb group.

See further examples:

E12	वह	गुस्से	से	लाल	हो	गया
IPA	wəh	gʊsse:	se:	lal	ho:	gəja:
Gl.	3SG	anger.N	with.ASS	red	become.VR	go.PFT.SG
EGl	He reddened with anger.					

E13	उसने	तुम्हें	मूर्ख	बना	दिया
IPA	ʊs-ne:	tumhẽ	muːrkʰ	bəna:	dɪja:
Gl.	3SG.OBL.ERG	2SG.ACC	fool	make.VR	give.PFT.SG
EGl	He fooled you.				

In the examples E12 and E13 above, the verb groups are “lal ho gəja:” and “muːrkʰ bəna: dɪja:” respectively. In the first example, the verb has been constructed via serialization of an adjective and a verb while in the second one the verb group is made of noun and verb serialization. In the first example, three elements of the verb group stands for a single concept ‘redden’ while in the second it stands for “to fool”. This kind of verb construction is called conjunct verb.

There is still another kind of verbal construction that occurs as conjunctive participle, also sometimes called *converb*. The examples are:

E14	वह	खाकर	बोला
IPA	wəh	kʰaː-kər	bol-a:
Gl.	3SG	eat-PFT.PCPL	speak-PFT.SG
EGl	He spoke after eating.		

E15	तुम	मन	लगा	कर	काम	करो
IPA	tʊm	mən	ləga:	kər	ka:m	kər-o:
Gl.	2SG	heart	apply.VR	PFT.PCPL	work	do-IMP
EGl	Work heartedly!					

Given the examples above we see that a verb group may constitute of several elements which can be listed as under:

- Main verb,
- Auxiliary verb,
- A V2 (also called explicator, Vector and light verb),
- Conjunct Verb,
- Conjunctive Participle, and
- Particles

Besides these separate, lexically distinct elements, the usual TAM categories also involve which can be marked at any of the first three elements above i.e. the main verb, auxiliary verb and the V2.

The internal structure of the verb group in Hindi thus is complex. This complexity has been commonly classified into two types: compound verbs and conjunct verb. Compound verbs consist of two verbs as exemplified in E10 and E11. Typical structure for a compound verb is V+V2 where the V2 carries all the TAM and agreement markings while the V1 remains in the root form. Conjunct Verbs are typically formed with a nominal or an adjective followed by a verb where again the verb, of course, carries the TAM and agreement markings.

2.4 OCCURRENCE OF VERB SEQUENCE IN HINDI TEXT

Besides the single word verbs that can occur in a sentence, there is a majority of structure that appear in complex forms. Complex verbs have been categorised into compound and conjunct. Accounting for conjunct verbs would involve analysis of two categories besides the verb, namely nouns and adjectives. Apart from this, it also involves semantics. In this research endeavour the present researcher is concerned with complex verbs and simple verbs, leaving conjunct verbs to be treated separately or at a later stage in continuation with this research.

The verbal elements in a verb group can consist of just one verbal word as the main verb as in example sentences E3 and E4 or it may contain two or more verbal words consisting of one main verb and the others as auxiliary. So the first major classification of verbal words that occurs at the lexical level is main verb vs. auxiliary verb. In Hindi, the list of verbs that can occur as auxiliary verb is also greater in number than many other languages like English, thanks to the serial verb constructions abound in Hindi and many other languages. However, the list of verbs that can function as auxiliary is always limited. Given a serial verb construction as for example in the following construction:

E16	वह	महीनो	से	पढ़ता	चला	आ	रहा	है
IPA	wəh	məhiːnoː	seː	pəɖʰ-taː	cəl-aː	a	rəh-aː	hɛː
Gl.	3SG	months.OBL	PP.ASS	read-IMPF.MSG	walk-PFT.MSG	come.VR	live-PFT.MSG	be.PRS.SG

Egl. He has gone on reading for months.

We find that a sentence can have several verbal elements put together and still denote a single verbal process i.e. it still has one main verb while the others function as auxiliary denoting other features like tense, aspect and mood. In the sentence above we find that the main verb / pəɖʰ/ ‘to read’ is followed by four other words with verbal intent and act as auxiliaries, carrying different TAM and agreement markings. Also notable is the fact that these words separately contain different aspects and tense i.e. the TAM and agreement is spread over all these elements.

2.4.1 Structure of Compound Verbs

The general structure of compound verb as stated above is V1+V2. In this structure, it is the V2 that takes the TAM and the agreement markings while V1 remains in the root form. V2 has also been called with several other terms such as ‘vector verb’ and ‘light verb’. V2 is a small set of words that has been listed variously in different descriptions of Hindi verb structure. There is no unanimity on how many light verbs in Hindi exist. Hook (1974) reports that there are studies that list these verbs starting from having a total number of 8 to 61.

The following verbs playing as V2 are commonly regarded as auxiliary verbs that represent aspectual or modal values: /rəh/, /sək/, /cʊk/, /pa:/, /ləg, /de:/, /ja:/ and /kar/. The total number of words that I have verified to get included in the V2 list as auxiliaries has been listed in Table 16 below:

Sl. No.	V2	IPA	Lexical Meaning	Sl. No.	V2	IPA	Lexical Meaning
1	आ	a:	to come	32	निकाल	nɪka:l	to bring out/to take out
2	उठ	ʊʈh	to get up/ to stand	33	पकड़	pəkəɖ	to catch
3	उड़	ʊɖ	to fly	34	पटक	pəʈək	to dash/to throw down
4	खप	kʰəp	to get consumed/to get finished	35	पड़	pəɖ	to lay
5	खा	kʰa:	to eat	36	पहुँच	pəhʊ̃c	to reach
6	गिर	gɪr	to fall	37	पा	pa:	to get/find
7	गिरा	gɪra:	to make (sth/smn) fall	38	फँस	pʰə̃s	to get caught/trapped
8	गुजर	gʊjər	to pass	39	फिर	pʰɪr	to turn/wander
9	घसीट	gʰəsi:ʈ	to pull	40	फेंक	pʰẽ:k	to throw
10	घुस	gʰʊs	to enter	41	बन	bən	to get made/ to get done
11	घेर	gʰe:r	to surround	42	बस	bəs	to settle down/to inhabit
12	चढ़	cəɖʰ	to climb	43	बीत	bi:t	to go by/to elapse
13	चल	cəl	to walk	44	बैठ	bɛ:ʈʰ	to sit
14	चुक	cʊk	to finish	45	भाग	bʰa:g	to run away/to flee

15	छोड़	cʰoːɖ	to leave	46	भिड़	bʰɪɖ	to collide/attack
16	जा	jaː	to go	47	भेज	bʰeːɟ	to send
17	झपट	jʰəpəʈ	to pounce	48	मंगा	mə̃gaː	to order
18	टपक	ʈəpək	to drip	49	मर	mər	to die
19	टूट	ʈuːʈ	to break	50	मार	maːr	to strike
20	ठहर	ʈʰəhər	to stop	51	मिट	mɪʈ	to get obliterated
21	डट	ɖəʈ	to stand firm	52	मिल	mɪl	to meet
22	डाल	ɖaːl	to put	53	रख	rəkʰ	to keep
23	डूब	ɖuːb	to sink	54	रह	rəh	to remain/stay
24	ढा	ɖʰaː	to run	55	लग	ləg	to begin/to get felt/to touch/to stick
25	दिखा	dɪkʰaː	to show	56	ला	laː	to bring
26	दे	deː	to give	57	लिपट	lɪpəʈ	to cling
27	देख	deːkʰ	to see	58	ले	leː	to take
28	दौड़	dɔːɖ	to run	59	लेट	leːʈ	to lie down
29	धमक	dʰəmək	to thump	60	सक	sək	can/to be able to
30	धर	dʰər	to catch	61	सुना	sʊnaː	to tell/to relate
31	निकल	nɪkəl	to come out				

Table 16: Common V2s in Complex Verbs

Besides these single words, there are verb sequences also that may occur as part of the compound verb, forming a total of three word compound verb constructions. Table 17 shows some of such sequences. It is notable here that in such sequences, we have adjectives like /kʰəɖaː/ or adverb like /baːhər/ as first component of such sequence while the following word is the normal V2. This is a property typical to the conjunct verbs (described below).

Sl. No.	**V2 Sequence**	**IPA**	**English Meaning**
1	खड़ा करना	kʰəɖaː kərnaː	to make (something/someone) stand up
2	खड़ा होना	kʰəɖaː hoːnaː	to stand up

3	बाहर करना	ba:hər kərna:	to get (something/someone) out
4	ले चलना	le: cəlna:	to take along / to carry
5	ले जाना	le: ja:na:	to carry

Table 17: Common V2s occurring in double words

2.5 STRUCTURE OF CONJUNCT VERBS

Conjunct verbs are formed of a nominal or an adjective followed by a verb. These two together semantically denote an action or a process or a state. Table 18 shows some examples of such constructions. Their appearances in sentences are as in E12 and E13. Given our current strategy, the PoS tags to these words should be given correctly. However, at present, we consider their identification at the verb group level as a challenge as it involves three grammatical categories as well as their semantic analysis.

Sl. No.	Conjunct Verb	Construction Type	English Meaning
1	लाल होना /la:l ho:na:/	Adjective + Verb	to redden
2	आज्ञा देना /a:gya: de:na:/	Noun + Verb	to allow
3	साफ करना /sa:pʰ kərna:/	Adjective + Verb	to clean
4	पता चलना /pəta: cəlna:/	Noun + Verb	to come to knowledge

Table 18: Examples of Conjunct Verbs

3 FORMAL DESCRIPTION OF THE VERB GROUP IN HINDI TEXT: IDENTIFICATION AND ANALYSIS

In this chapter we try to find out all the verb groups that can be found in Hindi. We start with the understanding of the grammatical structure of the verb groups and try to make a template for each of the verb groups that can possibly be formed in Hindi. We bring it down to an abstract level where these templates are represented through a rule which can be deterministically found out with the morpho-syntactic tags a PoS tagger is supposed to assign each of the verbal elements in a text.

We know that in Hindi the intent of verb may be spread over a number of verbal words in a sentence. It is through the combination of all these verbal elements that the tense, aspect and mood of a sentence can be ascertained. The information of voice is also to be found out this way.

3.1 TEMPLATE STRUCTURES OF VERB GROUPS

As we move ahead with finding out different structures of verb groups, we draw a template which can deterministically be mapped to a verb group tag. This tag will subsequently have the information of TAM represented by such a template.

By template we simply mean the representation of a verb group in an abstract form. For example in the example sentence E6 as shown above and also repeated here:

E6	वह	घर	जा	रहा	है
IPA	wəh	gʰər	ɟa:	rəh-a:	hɛ:
Gl.	3SG	home	go.VR	live-PFT.MSG	be.PRS-SG
EGl	He is going home.				

The verb group /ɟa: rəha: hɛ:/ is to be represented by the following template:

VR+rah_pft+prs_aux

3.1.1 Definition of a Template

A template is a serialization of variables which represent the morphological form of both the main verb and the auxiliary verb(s), if any, appearing in a verb group.

A verb group template in Hindi may be punctuated by a particle. That is, it may occur before any of the verbal word occurring inside the verb group. Therefore, when we define a template, we exclude the particles (/hi:/, /to:/, /bʰi:/, etc.).

3.1.2 Description of Verb Group Templates

A template is formed of variables. These variables are in fixed numbers. The number of such a variable depends on specific language. As we move ahead, we will find out many such variables in Hindi.

Given that auxiliary verbs in Hindi (and for many other languages) are limited in number and can be deterministically found out, we can set the auxiliary verbs as having a fixed set of words. For example, in Hindi present tense is marked with the help of present auxiliaries. The present auxiliaries in Hindi are /hɛ:/, /hɛ̃:/, /ho:/ and /hũ:/. Thus we can posit a variable like prs_aux which will have all these four words as its values. Similarly, pst_aux could be a variable with its values being /tʰa:/, /tʰi:/, /tʰe:/ and /tʰĩ:/.

So, we have two types of variables. One variable is to be found out by the morphological structure of the verbal word. This information can be extracted from the PoS tags assigned to each of the verbal words. For example, a verbal word /ja:/ is a verb root (VR). This can be found out by the tags assigned to it. In the IL-POST schema, such a word will have the tag of ‘nfn’ (non-finite). The variables which have to be found out by the tags assigned to it, we call them ‘tag-variables’.

The other type of variable has a set of fixed values. These are variables that we call word variables. For example the prs_aux, pst_aux, ja_pft, cuk_fut etc. are word variables and have a fixed set of verbal words as their variables.

These tag-variables and word-variables are the variables a verb group template (VG Template) is constituted of. Following are some examples of the VG templates which have been used to find out the verb group in an annotated text:

VR_impf+ho_pft

VR_pft+ho_pft

VR_impf+prs_aux

VINF+prs_aux

VR_pft+rah_impf+pst_aux

3.2 TAM TAGS

Once we have identified a template, a TAM tag has to be assigned to it. As stated earlier, the goal of this research work is to identify the verb groups and find out the tense, aspect and mood of the verb groups. We do this by assigning a TAM tag for each of the verb groups identified.

The general structure of the TAM tag that this researcher has come up with is as follows:

VG.Tense.Aspect.Mood

In the tag structure shown above, the period '.' plays as the delimiter between the categories. The first two letters of 'VG' is the identifier showing that it is a verb group. The next would be an identifier for tense, then aspect and then mood.

For example, the template of 'prs_aux' will have the TAM tag of 'VG.prs.sim.dcl' wherein 'prs' stands for present, 'sim' stands for simple aspect and 'dcl' stands for the declarative mood.

3.3 REPRESENTATION OF TAM IN HINDI VERBS

We have seen that a Hindi verb root can have a maximum of 25 morphological forms (barring the copula verbs). These forms do not represent all the tense, aspect and mood that a sentence can indicate through its verbs. That is, there are quite a lot of other TAM categories that are represented periphrastically, through the help of auxiliary verbs. The TAM elements that are expressed through any single word of these forms are limited to the following:

Tense: Future or non-future

Aspect: perfect or imperfect

Mood: imperative or optative

We see that future tense is explicitly marked morphologically for Hindi. The past and the present tense are marked periphrastically or understood by default connotation. Thus, for example, in a sentence having just one word as the verbal element, the perfective marked verb can also be understood as the past tense. Because Hindi marks only future tense morphologically at the lexical level, Hindi tense system is understood to be Future-Non-Future.

Among the aspects, only perfect and imperfect are explicitly marked on the verb. Other aspects, namely progressive, inceptive, durative etc., are marked periphrastically with the help of other verbs (second verbs or V2s).

Among the moods also, only two moods of imperative and optative are marked morphologically. Other moods, such as counterfactual, presumptive and contingent are marked periphrastically through V2s.

Therefore, we need to find out how the other tenses, aspects and moods are marked in Hindi. We need to find out all of the TAMs that might be indicated through the verb groups in Hindi.

The best way would be to start with a basic list of all the TAM markings denoted in Hindi. It would be apt here to give a summary of the tense, aspect and mood system in Hindi. The following two tables are an attempt to summarise the TAM system in Hindi.

Sl. No.	**ASPECT**	**TENSE**		
		Present	**Past**	**Future**
1	**simple**	prs_aux	(pst_aux	VR_fut
2	**Imperfect**	VR_impf+ prs _aux	VR_impf+ pst_ aux	
3	**Perfect**	VR_pft + prs _aux	VR_pft + pst_ aux	VR_pft + ho_fut
4	**Progressive**	VR + rah_pft + prs _aux	VR + rah_pft + pst_ aux	VR + rah_pft+ho_fut

5	**INCEPTIVE**	1.) VINF+lag_impf +prs _aux 2.) VINF+lag_pft+prs _aux	1.) VINF+lag_impf+ pst_ aux 2.) VINF+lag_pft+pst_ aux 3.) VINF+lag_pft	VINF+lag_fut
6	**CONTINUATIVE**	1.) VR_impf+rah_impf +prs _aux 2.) VR_impf+rah_pft +prs _aux	1.) VR_impf+rah_impf_pft +pst _aux 2.) VR_impf+rah_impf +pst _aux 3.) VR_impf+rah_pft	VR_impf+rah_fut
7	**DURATIVE**	1.) VR_impf+ja_impf +prs _aux 2.) VR_impf+ja_pft +prs _aux 3.) VR_impf+ja_prog+prs _aux 4.) VR_impf+cal_pft+a+rah_ pft+prs_aux	1.) VR_impf+ja_impf+ pst _aux 2.) VR_impf+ja_pft+ pst _aux 3.) VR_impf+ja_prog+ pst _aux 4.) VR_impf+ja_pft 5.) VR_impf+cal_pft+a+rah _pft+pst_aux	VR_impf+ja_fut
8	**FREQUENTATIVE**	VR_pft+kar_impf+prs _aux	1.) VR_pft+kar_impf+pst _aux / 2.) VR_pft+kar_impf	VR_pft+kar_impf+ho_fut

Table 19: General Tense and Aspect Templates in Hindi

Mood (Non-Declarative)	**TENSE**		
	Present	**Past**	**Future**
imperative	VR		
optative	VR_opt		

presumptive			1.) VR_impf + ho_fut 2.) VR_impf+cal_pft+a+rah_pft+ho_fut
counterfactual		VR_impf+ho_impf VR_impf	
contingent			VR_impf+ho_fut

Table 20: General Mood Templates in Hindi

Table 19 gives a summary of the different conjugations of all the three tense and eight aspects (as noted in Kachru (2006)) in the declarative/indicative mood. In this table we provide the templates for all the possible conjugations of the verb group in an abstract form.[15] Table 20 shows the conjugations with different moods other than declarative. The declarative gets covered with Table 19. From both the tables above, we get a total of 44 TAM templates for verb groups in Hindi.

3.4 VERBAL ELEMENTS IN VERB GROUP

In our tagset, i.e. the IL-POST framework, we posit that only one verb can have a VM tag in a simple sentence (or a clause within sentence). All the other verb forms such as infinitive and participial constructions when not playing as the main verb in the sentence should be tagged according to its function in the category. For example in the following sentences

E17	मुझे	रात	में	खाना	अच्छा	लगता	है
IPA	mʊjʰe:	ra:t	mẽ:	kʰa:na:	ɘccʰa:	ləg-ta:	hɛ:
Gl.	1SG.DAT	night	PP.LOC	eat.INF	good.ADJ	feel-IMPF.MSG	be.PRS.SG
EGl	I like eating at night.						

[15] One might note here that other than /kər/, /ləg/, /rəh/, /ho:/ and /ja:/ as we see in Table 19 above, there are several other V2s that have the value of aspect and add a different dimension to the meaning of the aspect. We will deal with them separately under the compound verb.

E18	मैं	खाकर	सो	गया
IPA	mɛ̃:	kʰa:-kər	so	gəja:
Gl.	I	eat-PFT.PCPL	sleep.VR	go.PFT.SG
EGl.	I slept after eating.			

E19	वह	खाते-खाते	सो	गया
IPA	wəh	kʰa:te:-kʰa:te:	so	gəja:
Gl.	3SG	eat.IMPF.PCPL	sleep.VR	go.PFT.SG
EGl.	He slept while eating.			

The functions of the underlined verbal category words are respectively infinitival, participial and adverbial. We consider the case of /kʰa:na:/ in the first sentence above as verbal noun and the cases of /kʰa:kər/ and /kʰa:te:-kʰa:te:/ as cases of adverbial use of the verb. So, the tags assigned to these occurrences of the verbal elements in our case would not be verbal (i.e. they would not be marked as VM or VAUX).

However, in the BIS annotation guidelines, these occurrences of verb groups are marked as verb groups which must be marked as VM or VAUX as necessary. In such a case, these have to be given a VG tag as well. Even in such a situation, the VGs can be captured and respective tags could be given to these verb groups.

So what we have for the verbal group is the verb or sequence of the verbal words that play as the main verb in the sentence. We have noted that a verb group in Hindi can be punctuated only by a few indeclinable words called as particles. This also proves that we can exhaustively find out such verb sequences and mark them for their TAM features. These templates along with their TAM tags would then be reused to mark any similar occurrences matching the template structure.

The question now is how to find out all such VG templates possible in Hindi. There have been earlier attempts to define such templates/sequences in descriptive books (e.g Guru, 1978). However, there has not been any attempt to capture all of such templates/sequences exhaustively, mainly because there has not been any need for such an attempt. Even for computational purposes, this was never attempted because there is no approach that would need it. Most of the new approaches use one or the other statistical method to identify chunks/word groups. There have also been a few attempts to capture some other elements occurring inside the larger category of verb groups, within the paradigm of statistical methods. Chakrabarti, D. et.al. (2008) and Begum, R. et.al. (2011) are some such attempts that try to find out the lexical compound verbs and conjunct verbs respectively. Both of these works are inspired from needs that may not be the inspiration of the work presented here by this researcher.

For the reasons of computing, we will start with classifying the verb groups based on the number of verbal words that play as the constituent in it. An initial observation on Hindi verb structure reveals that we can have verb groups having just one verbal word to verb groups having up to five words in a sequence. This sequence can be broken by particles like /hi:/, /to:/, /bʰi:/, /nəhĩ:/ etc. For the time being we will set aside the cases of these particles and focus mainly on the verbal sequences.

3.4.1 One-Word Verb Group

In this group we include all the verb groups that have only one verbal element playing as the main verb in the input sentence. In Table 15 in the previous chapter we have seen that a verb root can have at most 25 morphological forms representing different TAM and GNPH markings. A summary of those 25 forms has been provided below in an abstract form:

Abstract Form	**No. of GNPH Forms**	**Co-occurring form**	**Occurrence as a single word VG**
Verb Root	1	Imperative	Yes
Infinitive	3		No
Imperfect	4		Yes
Perfect	4	Optative	Yes
Future	9		Yes
Optative (1sg)	1		Yes
Imperative/Optative 2person plural honorific	1		Yes

Imperative	2		Yes

Table 21: Summary of the functions of Hindi Verbs

In the table above I have brought down the 25 forms of a verb root into the 8 abstract categories. This abstraction is based on the tense, aspect and mood only and does not include the gender, number, person and honorificity (GNPH) markers on these verbs. The number of GNPH forms for each of the abstract form indicates the various inflectional forms that a verb root can take. The terms used for the abstract forms are commonly used in literature and do not require elaboration here.

Except for the infinitives, all the other seven forms can be used as the main verb in a verb group having a single word. (However, there are instances when even the infinitives can work as main verb, as in the following sentence:

E20	जरा	मुझे	पानी	<u>देना</u>
IPA	ɟərɑ:	mʊɟʰe:	pɑ:ni	<u>de:nɑ:</u>
Gl.	little-ADV	1SG.ACC	water	give.INF.IMP
EGl	Give me some water.			

Although, here the infinitive should come with the imperative marker and this form does not inflect for GNPH. This use of infinitives may be treated as non-finite.)

There are two forms that are lexically ambiguous in their functions. The first one of these two is the verb in its root form which can also function as the imperative. Thus in the following two sentences, the verbal word /rʊk/ occurs as verb root and imperative respectively:

E21	पानी	बहना	<u>रुक</u>	जाएगा
IPA	pɑ:ni	bəhnɑ:	<u>rʊk</u>	ɟɑ:-e:gɑ:
Gl.	water	flow.INF	stop.VR	go-FUT.MSG
EGl	The water will stop flowing.			

E22	रुक	मैं	भी	साथ	आऊंगा	
IPA	rʊk	mɛ̃:	bʰi:	sa:tʰ	a:-ũ:ga:	
Gl.	stop	1SG	also.PRT	with.ASS	come-FUT.1MSG	
EGl	Stop, I will also come along.					

This ambiguity however can be resolved quite easily because we know that there cannot be any verb group having just a verb root. That is, if a verb group has only one verb and this verb is in its root form, then it must be in the imperative form.

The other lexically and functionally ambiguous verb form is the perfect form which can also play as an optative. For example,

E23	मेरी	इच्छा	है	कि	वह	खाना	खाए
IPA	me:ri:	ɪccʰa:	hɛ:	kɪ	wəh	kʰa:na:	kʰa:-e:
Gl.	1SG.GEN	wish.N	be.PRS.SG	that.CNJ	3SG	food	eat.OPT
EGl	I wish that he eats food.						

E24	उन्होंने	फल	खाए	
IPA	unhõ:ne:	pʰəl	kʰa:e:	
Gl.	he/PL.ERG	fruit.N	eat-PFT.PL	
EGl	They ate fruits.			

In these two sentences above, the surface level form of the verb / kʰa:e:/ though looks the same, they have different functions. In the first sentence, it is an optative while in the second one, it is simple past. Its disambiguation process may be difficult if the PoS tag assigned to them is not

correct. Therefore, the correctness of the PoS is crucial in such cases on the part of the annotator/tagger. (This is of importance because tagging may be confusing to an annotator if s/he is not clear. In later chapter I show the errors found in the manually annotated data. Many of these errors can be attributed to as the mistakes on the part of the annotator and lack of clear guidelines.)

Barring the copular forms of the verbs and the counterfactual mood constructions, we have a total of 18 + 1 forms that can occur as the single verb group. Given a verb root as रह /rəh/ "to live/remain", these would be as follows:

Sl. No.	Morphological Form Type	Actual Forms of the Root	Subject Prediction	Aspect	Mood	Tense
1	Root / Imperative 2SG/2PL	/rəh/	2.0.0.0	???	imperative	???
2	Perfect MSG	/rəha:/	1/2/3.sg.mas.0	perfective	Dcl	past
3	Perfect FSG	/rəhi:/	1/2/3.sg.fem.0	perfective	Dcl	past
4	Perfect FSGH	/rəhĩ:/	1/2/3.0.0.hon	perfective	Dcl	past
5	Perfect PL / Optative 3SG	/rəhe:/	1/2/3.0.0.hon	perfective	Dcl	past
6	Future 3MSG	/rəhe:ga:/	3.sg.mas.0	simple	Dcl	future
7	Future 3MSGH	/rəhi:ega:/	2.0.0.hon	simple	Dcl	future
8	Future 3FSG	/rəhe:gi:/	3.sg.fem.0	simple	Dcl	Future
9	Future 3FSGH	/rəhẽ:gi:/	1/2/3.pl.fem.0	simple	Dcl	Future

10	Future PL	/rəhẽ:ge:/	3.0.fem.hon	simple	Dcl	Future
11	Future 2MSGNH	/rəho:ge:/	2.sg.mas.0	simple	Dcl	Future
12	Future 2FSGNH	/rəho:gi:/	2.sg.fem.0	simple	Dcl	Future
13	Future 1MSG	/rəhũ:ga:/	1.sg.mas.0	simple	Dcl	Future
14	Future 1FSG	/rəhũ:gi:/	1.sg.fem.0	simple	Dcl	Future
15	Optative 1SG	/rəhũ:/	1.sg.0.0	???	optative	???
16	Imperative Plural Honorific	/rəhẽ:/	1/2/3.0.0.hon	???	imp/opt	???
17	Imperative Honorific	/rəhɪe:/	2.0.0.hon	???	imperative	???
18	Imperative 2SGNH	/rəho:/	2.sg.0.0	???	imperative	???

Table 22: Verb Forms Occurring as One-Word Verb Group

Each of these 19 forms of a verb plus all the eight copular forms of the verb /ho:/ along with its existential past contingent (counterfactual) mood constructions with /ho:ta:/, /ho:ti:/, /ho:te:/, /hotĩ/ (e.g. as in E25 below:

E25	वह	आज	प्रधानमंत्री	<u>होतीं</u>
IPA	wəh	a:ɟ	prəd^{h}a:nməntri:	<u>ho:-tĩ:</u>
Gl.	she	today	prime minister	be.CFCT
EGl	She would have been prime minister today.			

) can constitute a verb group all by itself. If we talk in templates, then we will have just the following seven templates for the one word verb group, without the agreement markers.

	ASPECT	TENSE		
ASPECT		**Present**	**Past**	**Future**
	Simple	prs_aux	pst_aux	VR_fut
	imperfect			
	Perfect		VR_pft	
	Progressive			
	Inceptive			
	Continuative			
	Durative			
	Frequentative			
MOOD	**Imperative**	VR_imp		
	Optative	VR_opt		
	Presumptive			
	Counterfactual		VR_impf	
	Contingent			

Table 23: TAM Templates for One Word Verb Group

All of the seven templates shown above in Table 23 can stand on their own. They constitute of a single verbal word and denote a state, action or process. In the sentences they can appear as the main verb and may not require any auxiliary to denote their verbal content.

We can add one more template into this same category. However, this is not very common and grammatical literature often does not provide this. Reconsider the example sentence E20 given above and also repeated here:

E20	जरा	मुझे	पानी	देना
IPA	jəra:	mʊjʰe:	pa:ni	de:na:
Gl.	little-ADV	1SG.ACC	water	give.INF.IMP
EGl	Give me some water.			

The meaning of the sentence is clearly in imperative, even though the form of the main verb denoting this intent is in infinitive. This kind of imperative formation may also be considered as common. To capture this, we can add a template as follows

VINF_imp

This template says that if there is an infinitive verb occurring as VM and if it has the tag of imperative, consider it as an imperative sentence. Please note that it is important that the tag for such occurrence of infinitives must have the imperative identifying tag.

Another one word template that can be added is in the desiderative mood marker. The verb चाहिए /ca:hɪe:/ though seems to be in the honorific imperative mood form, its function is totally different. It is used mainly to show the desiderative mood. Thus in the following sentence, it is the main verb and the only verbal word in the sentence and it shows the desire of the subject:

E26	उसे	किताब	चाहिए
IPA	ʊse:	kɪta:b	ca:hɪje:
Gl.	3SG.OBL.ACC	book.N	want-DES
EGl	He wants book.		

The tense as expressed in this sentence is fuzzy. It can be taken to be in the present tense, which actually it is. However, morphologically there is no mark of present tense in it.

Hindi sentence might have verbal words which might not be the main verb or part of the main complex verb. For example consider the following sentences:

E27 वह <u>खाकर</u> किताब पढ़ा

IPA wəh <u>kʰaːkər</u> kɪtaːb pəɖʰaː

Gl. 3SG eat.VR.PCPL book read-PFT

EGl. He read book after eating.

E28 उसको <u>खाते</u> देख वह बोला

IPA ʊs-koː <u>kʰaː-teː</u> deːkʰ wəh bol-aː

Gl. 3SG.OBL-ACC eat-IMPF see.VR 3SG speak-PFT.MSG

EGl. Seeing him eat, he spoke.

E19 वह <u>खाते-खाते</u> बोला

IPA wəh <u>kʰaːteː-kʰaːteː</u> bolaː

Gl. 3SG eat-IMPF.PCPL speak.PFT.MSG

EGl. He spoke while eating.

E29 सुरक्षा बहाल <u>करना</u> सरकार की जिम्मेदारी है

IPA surəkʃaː bəhaːl <u>kərnaː</u> sərkaːr kiː jɪmmeːdaːriː hɛː

Gl. security maintain do.INF government GEN responsibility be.PRS.3SG

EGl. Maintaining security is the government’s responsibility.

In the first and the second sentences above, the function of the verbal elements /kʰa:kər/ and /de:kʰ/ do not form part of the main verb complex (Though upon analysis, one can argue that there are two clauses in it and thus there are two main verb complexes in these two sentences, and therefore they should also be marked as main verbs). These two verbal elements are mainly showing a manner and can be argued that they are adverbial in nature. For PoS annotation, many a times these elements are marked as verbal element. And in such cases, we need to capture them as well. Because they are a single word, we can assign a VG tag to them.

For the third example sentence given above also the same rule applies at the semantic level. That is, the semantic content of the verbal word /kʰa:te:-kʰa:te:/ is that of manner. This construction is also very common in this sense and this construction occurs mostly in reduplicated form. Such occurrences can be called participial constructions. We denote it as VR_pcpl and give it a template status with the same name.

In the fourth example sentence given above, the verbal word /kərna:/ is in infinitive form. It also is a common construction and many a times is considered as a nominal verb because its function in the sentence is like a noun. It can be debated whether this word be tagged as a verb or a noun. If it is tagged as a verb, it has to be given a VG tag as well because it occurs separately from the main verb complex. The template for such occurrences could be simply VINF.

Thus we have a total of 12 templates for one word verb group.

3.4.2 TAM Tags for One Word Verb Groups

We have identified a total of 12 templates that are one word verb groups. These templates along with their TAM tags and example sentences are shown in the table below:

T. No.	VG Template	TAM Tag	Example Sentence
1	prs_aux	VG.prs.sim.dcl	E30
2	pst_aux	VG.pst.sim.dcl	E31
3	VR_fut	VG.fut.sim.dcl	E32
4	VR_pft	VG.pst.pft.dcl	E33

5	VR_imp	VG.0.0.imp	E34
6	VR_opt	VG.0.0.opt	E35
7	VR_impf	VG.pst.impf.cgnt	E36
8	VR_kar	VG.cp.pft	E27
9	VR_pcpl	VG.ip.impf	E19
10	VINF_imp	VG.0.0.imp	E37
11	VINF	VG.inf	E38
12	cahiye	VG.0.0.des	E26

Table 24: One Word Verb Groups and their TAM Tags

In the table above, the first column (T. No.) is the serial number of the template identified. This number is merely a notation and purely for reference purposes. In the second column the template identified has been defined and stated in an abstract form. The third column gives the TAM tag that will be attached to the respective template after the template has been identified. The fourth column gives a Hindi example sentence.

The assignment of TAM tags for these templates is self-explanatory and can be verified by the simple word forms as these are one word verb groups.

For the verb groups that do not function as the main verb in the sentence, namely, templates 8, 9 and 11, the normal VG format has not been applied. These tags do not have the three dot delimited slots filled. Rather they have a different structure unique to their own.

3.5 PERIPHRASTICALLY MARKED TAM

As we have seen in the one-word verb group discussion above, all the TAM categories are not expressed through the one-word verb group. There is a lot more that are left. Hindi uses the verb serialization feature to indicate various other tense, aspect and mood categories. Below, I present a discussion on periphrastically marked TAM categories.

3.5.1 Representation of Tense

The tense system in Hindi is morphologically Future-Non-Future. The reason for this is that Hindi marks only future tense morphologically. Barring the copular constructions and the use of the perfect form of a verb root as simple aspect, both the past and present tense are marked periphrastically and not morphologically.

In Hindi TAM is often represented through compound verbs. The most common example would be the representation of past and present tense in non-copular constructions. The copula verbs are the only forms of verbs which have clear present and past forms. All other verbs have to be indicated their past and present tense intent through the help of these past and present copula verbs. This also means that these two categories of verbs have to be auxiliary verbs whenever they appear in a multi-word verb group. Let's represent the auxiliary occurrences of these verbs by two short names, which we also utilize in our TAM templates, namely 'prs_aux' and 'pst_aux'. The category of 'prs_aux' contains present forms of the copular verb, namely /hɛ:/, /hɛ̃:/, /ho:/ and /hũ:/. The category of pst_aux contains the past forms of the copular verb i.e. /tʰa:/, /tʰe:/, /tʰi:/ and /tʰĩ:/.

If we keep the main verb function of the copular verb away, no sentence in Hindi can be formed in the present tense without the help of the copular auxiliary. Thus we get the first two word verb group in Hindi wherein the first verbal word in the verb group is a non-copular main verb and the other one is auxiliary 'prs_aux'.

One makes an aspectual construction (perfect and imperfect) of a verb and adds the prs_aux to get a two word verb group in Hindi that is in present tense. It has already been shown that a verb can have at most 25 morphological forms. Only two aspects, namely perfect and imperfect are expressed through the forms from among these 25 forms.

E2	वह	घर	गया	है
IPA	wəh	gʰər	gəja:	hɛ:
Gl	3SG	home	go.PFT.MSG	be.PRS.3SG
EGl	He has gone home.			

and

E1	वह	घर	<u>गया</u>	
IPA	wəh	gʰər	<u>gɔja:</u>	
Gl	3SG	home	go.PFT.MSG	
EGl	He went home.			

In the E2 above the perfective construction is used along with the present tense marker prs_aux. In the E1, the perfective construction is used in a standalone way and denotes the past tense. However, like in E2, a copular auxiliary in the past tense can also be used in the second sentence so as to have the following form:

E39	वह	घर	गया	था
IPA	wəh	gʰər	gɔja:	tʰa:
Gl	3SG	home	go.PFT.MSG	be.PST.MSG
EGl	He had gone home.			

This E39 can be argued to mean the same as the E1 one above. There seems to be no change in the actionsart of the verb group. However, it is usually said that the second sentence is simple in aspect while the third sentence is in perfect aspect.

Marking the tense in the TAM Tags

When we identify a verb group, we give them all the TAM tags. The tense tag obviously will be three: that of present (prs), past (pst) and future (fut). However, there will be tags which will have its tense slot marked as ‘0’ indicating that the tense in this verb group is null. There are many templates which will have a ‘0’ tense marking. For example, the sentences which are in imperative or optative mood may be argued not to have a tense marked intrinsically. Thus the

following sentence being in optative can be argued to be in present, but because it is optative, the tense marking may be skipped.

E35	वह	किताब	पढ़े
IPA	wəh	kıta:b	pəɖʱe:
Gl.	3SG	book	read.OPT
EGl.	(I wish) He should read book.		

3.5.2 Representation of Aspects

We have noted that Hindi has only two aspects denoted morphologically i.e. imperfective and perfective. We have seen their use in template numbers 7 and 4 respectively above. However, there are other aspects noted periphrastically in Hindi. The most important among the aspects marked periphrastically is the progressive. The progressive aspect is marked with the use of the second verb /rəh/ in its perfective forms (/rəha:/, /rəhi:/, /rəhe:/, /rəhĩ:/). Thus the following sentences are in progressive aspect:

E40	वह	किताब	पढ	रही	है	
IPA	wəh	kıta:b	pəɖʱ	rəh-i:	hɛ:	
Gl.	3SG	book.N	read.VR	live-PFT.3FSG	be.PRS.SG	
EGl	She is reading a book.					

E41	वह	किताब	पढ	रहा	था	
IPA	wəh	kıta:b	pəɖʱ	rəh-a:	tʰa:	
Gl.	3SG	book.N	read.VR	live-PFT.MSG	be.PST.MSG	
EGl	He was reading a book.					

E42	तुम	किताब	पढ	रहे	होगे
IPA	tʊm	kɪta:b	pəɖʰ	rəh-e:	ho:-ge:
Gl.	2SG	book.N	read.VR	live-PFT.PL	be-FUT.2SG
EGl	You will be reading a book.				

The three example sentences above are respectively in present, past and future tense. The aspect for all of them is progressive.

3.5.2.1 Other Aspects

Kachru (2006) notes four other aspects, namely inceptive, continuative, durative and frequentative. These are all denoted using a compound verb structure. ‘These are not as general in distribution as imperfect, perfect and progressive.’ That is, they have some restrictions and cannot occur as freely as the progressive aspect.

The following sentences are respectively examples of inceptive, continuative, durative and frequentative (modelled after Kachru’s (ibid) formation rules):

E43	वह	किताब	पढने	लगा
IPA	wəh	kɪta:b	pəɖʰne:	ləg-a:
Gl.	3SG	book.N	read.INF	begin-PFT.MSG
EGl	He started reading a book.			

E44	वह	किताब	पढता	रहा
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-a:
Gl.	3SG	book.N	read-IMPF.MSG	live-PFT.3MSG

EGl He kept on reading a book.

E45 वह किताब पढ़ता गया

IPA wəh kɪta:b pəɖʱ-ta: gəja:

Gl. 3SG book.N read-IMPF.MSG go.PFT-MSG

EGl He went on reading a book.

E46 वह किताब पढ़ा करता

IPA wəh kɪta:b pəɖʱ-a: kər-ta:

Gl. 3SG book.N read-PFT.MSG do-IMPF.MSG

EGl He used to read a book.

Kachru (ibid) also notes other aspects under the verbal paradigm category: presumptive, contingent and past contingent, examples respectively of each:

E47 उसने किताब पढ़ी होगी

IPA ʊs-ne: kɪta:b pəɖʱ-i: ho:-gi:

Gl. 3SG.OBL-ERG book.N read-PFT.FSG be-FUT.3FSG

EGl He would have read the book.

E48 वह किताब पढ़ा हो

IPA wəh kɪta:b pəɖʱ-a: ho:

Gl. 3SG book.N read-PFT.MSG be.OPT

EGl He should have read the book.

E49 उसने किताब पढ़ी होती

IPA ʊs-ne: kɪta:b pəɖʰ-i: ho:-ti:

Gl. 3SG.OBL-ERG book.N read-PFT.FSG be-IMPF.FSG

EGl He could have read the book.

One will note as one goes ahead that these same categories also appear under the moods. At times it appears that the distinction between the aspectual category and the modal category is getting blurred.

Also of note is the fact that two aspects can come bounded together as in the following sentence:

E50 वह किताब पढ़ने लगता था

IPA wəh kɪta:b pəɖʰne: ləg-ta: tʰa:

Gl. 3SG book.N read.INF begin-IMPF.MSG be.PST.MSG

EGl He used to start reading the book.

In this sentence the inceptive marking V2 /ləg/ appears in the imperfective form along with the past tense marking auxiliary verb. If we consider inceptive as one aspect being denoted by the V2 /ləg/ then the past imperfect aspect marking is also attached to it, because it is in the imperfective form. Thus both the inceptive and imperfect is marked in this verb group. Similar is the case with many other verb groups where we have more than two aspects being marked either morphologically or through the help of two different V2s.

E51 वह किताब पढ़ता जाता है

IPA	wəh	kıta:b	pəḍʰ-ta:	ɟa:-ta:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	go-IMPF.MSG	be.PST.SG
EGl	He keeps on reading the book.				

→ denotes imperfective and durative aspects

E52	वह	बैठा	रहा	
IPA	wəh	bɛ:ʈʰ-a:	rəh-a:	
Gl.	3SG	sit-PFT.MSG	live-PFT.MSG	
EGl	He kept on sitting.			

→ denotes perfect and durative aspects

E53	वह	किताब	पढता	रहा	था
IPA	wəh	kıta:b	pəḍʰ-ta:	rəh-a:	tʰa:
Gl.	3SG	book.N	read-IMPF.MSG	live-PFT.MSG	be.PST.MSG
EGl	He had kept on reading the book.				

→ denotes imperfect and perfect aspects simultaneously

E54	वह	बैठा	रहता	है
IPA	wəh	bɛ:ʈʰ-a:	rəh-ta:	hɛ:
Gl.	3SG	sit-PFT	live-IMPF.MSG	be.PRS.SG
EGl	He keeps on sitting.			

→ denotes perfect and imperfect aspects simultaneously

The total aspects identified in various verb groups are as follows:

No.	Aspect Name	Abbreviated Form as used in TAM Tag
1	Perfect	pft
2	Null	0
3	Continuative	cnt
4	Durative	dur
5	Imperfect	impf
6	Imperfect Continuative	impf_cnt
7	Imperfect Durative	impf_dur
8	Imperfect Frequentative	impf_frq
9	Imperfect Inceptive	impf_ince
10	Imperfect Perfective	impf_pft
11	Imperfect-Perfect-Continuative	impf_pft_cnt
12	Perfect Continuative	pft_cnt
13	Perfect Durative	pft_dur
14	Perfect Frequentative	pft_frq
15	Perfect-Imperfect-Continuative	pft_impf_cnt
16	Perfect Inceptive	pft_ince
17	Progressive	prog
18	Progressive Durative	prog_dur
19	Simple	sim
20	Simple Inceptive	sim_ince

Table 25: Aspects as Identified and Used in the TAM Tags

3.5.3 Representation of Moods

Only two modal categories are marked morphologically in Hindi. Declarative is the default mood marking. Contingent mood marking may be considered as one of the mood marked through the imperfective form of the verb roots.

Other moods are marked periphrastically. The moods marked periphrastically in Hindi are noted below.

3.5.3.1 Abilitative\Possibilitative

The V2 /sək/ is used to express ability or possibility. This V2 can come in all its form and still denote ability or possibility. It does not occur as a main verb in any construction. Another V2 /pa:/ is also used to show the same senses. However, /pa:/ has its main verb functions as well where it means ‘to get or find’.

3.5.3.2 Presumptive

The presumptive is marked through the V2 /ho:/ in its future forms. Presumptive is always in the future tense.

3.5.3.3 Desiderative and Suggestive

Desiderative is always marked through the invariable form of /ca:hɪje:/. It actually has two meanings. When this word occurs as the main verb in a sentence, then it gives the sense of 'want', as in:

E55	मुझे	किताब	चाहिए	
IPA	mʊjʰe:	kɪta:b	ca:hɪje:	
Gl.	1SG.DAT	book.N	want.DES	
EGl	I want a book.			

E56	उसे	किताब	चाहिए	थी
IPA	ʊse:	kɪta:b	ca:hɪje:	tʰi:
Gl.	3SG.OBL.DAT	book.N	want.DES	be.PST.FSG
EGl	He wanted a book.			

Another sense that it gives is when it is used with an infinitival verb preceding it, as in:

E57	उसे	जाना	चाहिए
IPA	ʊse:	ja:na:	ca:hɪje:
Gl.	3SG.OBL.DAT	go.INF	want.DES
EGl	He should go.		

In this sentence it does not function as the main verb. The main verb here is /ja:na:/ ‘to go’ while /ca:hɪje/ “should” is working simply as an auxiliary verb. Here the lexical sense of the verb is lost and the meaning it gives is that of suggestion. Whenever the verb /ca:hɪje/ occurs in this sense, it requires that the preceding main verb be in its infinitival form.

3.5.3.4 Obligative

Obligative construction is formed with an infinitive form of the verb followed by the V2 of /ho:/ or /pəɖ/. These two V2s can subsequently take all aspect markers and can appear in all the tenses. The difference between /ho:/ and /pəɖ/ is that while the former denotes internal motivation the latter denotes external motivation.

3.5.3.5 Permissive

The permissive is formed with V2 /de:/. This V2 can further inflect for all the tenses and aspects. It requires that its preceding main verb be in infinitive form.

3.5.3.6 Other Moods

It is possible that a verb group denotes two moods put together. The following categories are an example of that.

3.5.3.6.1 Permissive Counterfactual

Consider the following example:

E58	काश	तुमने	उसे	जाने	दिया	होता
IPA	ka:ʃ	tʊmne:	ʊse:	ja:ne:	dɪja:	ho:-ta:
Gl.	wish.INJ	2SG.ERG	3SG.OBL.ACC	go.INF	de.PFT.MSG	be-IMPF.MSG
EGl	(I wish) you had let him go.					

In the sentence above, the first V2 /dɪja:/ denotes the permissive intent while the third V2 /hota/ shows the counter factuality.

3.5.3.6.2 Permissive Suggestive

Consider the following example:

E59	तुम्हें	यह	काम	उसे	करने	देना	चाहिए
IPA	tʊmhẽ:	jəh	ka:m	ʊse:	kərne:	de:na:	ca:hɪje
Gl.	2SG.ACC	this	job	3SG.OBL.ACC	do.INF	give.INF	should
EGl	You should let him do this job.						

In this sentence, while the first V2 /de:na:/ denotes the permission the second V2 /ca:hɪje/ is suggestive in nature. Both of these two V2s come together to denote a single mode of permissive suggestive.

3.5.3.6.3 Permissive Imperative

Similarly in the following examples:

E60	उसे	खाने	दो	
IPA	ʊse:	$k^ha:ne:$	do:	
Gl.	3SG.OBL.ACC	eat.INF	give.IMP	
EGl	Let him eat.			

E61	यदि	उसे	भूख	लगी	तो	खाने	देना
IPA	jədɪ	ʊse:	$b^hu:k^h$	ləg-i:	to	$k^ha:ne:$	de:na:
Gl.	if	3SG.OBL.DAT	hunger.N	begin-PFT.FSG	then.CONJ	eat.INF	give.INF.IMP
EGl	If he feels hungry then (you) let him eat.						

The V2 /de:/ occurs in imperative form/sense and thus denotes both permissive and imperative sense.

3.5.3.6.4 Optative Probabilitative

Another double mood marking verb group can be witnessed in the following sentence:

E62	ताकि	मैं	उसे	दिल्ली	ला	सकूं
IPA	ta:kɪ	mɛ̃:	ʊse:	dɪlli:	la:	səkũ:
Gl.	so that.CONJ	1SG	3SG.OBL.ACC	Delhi	bring.VR	can.OPT
EGl	So that I could bring him/her to Delhi.					

Here the V2 /sək/ comes in its optative form. Because /sək/ as V2 is an abilitative/possibilitative marker in itself, it denotes two moods put together here in this sentence.

3.5.3.6.5 Conditional

Conditional is mostly denoted by a non-verbal entity, usually through the conditional marking conjunctions. However, there is one template that expresses this conditionality just through its verb group structure:

E63	फिर	तो	वह	पढ	चुका
IPA	pʰir	to:	wəh	pəɖʰ	cʊk-a:
Gl.	then.CONJ	PRT	3SG	read.VR	finish-PFT.MSG
EGl	It’s unlikely that he would have read then.				

This construction must have the perfective form of the V2 /cʊk/ and the VG must end with it. Please note that a verb group cannot end with the perfective form of V2 /cʊk/ until it denotes a conditionality, which might also get expressed through other conjunctions (e.g. /pʰir to/ as in the example above). Thus the following example sentence is incorrect in itself:

*E64	वह	पढ	चुका
IPA	wəh	pəɖʰ	cʊk-a:
Gl.	3SG	read.VR	finish-PFT.MSG
EGl	He had read.		

3.5.3.6.6 Contingent

Contingent mood expresses ‘a possibility, whether habitual, present, past, or future, and whether assumed to be completed or not” (ibid). The V2 /ho:/ is used to express thus mood e.g.:

E65	कहीं	उसने	किताब	पढी	हो	तो
IPA	kəhĩ:	ʊs-ne:	kɪta:b	pəɖʰ-i:	ho:	to:
Gl.	if-PRT	3SG.OBL-ERG	book.N	read-PFT.FSG	be.CFCT	then
EGl	If he had the book read then?					

E66	वह	अब	तक	किताब	पढ	चुका	हो
IPA	wəh	əb	tək	kɪta:b	pəɖʰ	cʊk-a:	ho:
Gl.	3SG	now	till	book.N	read.VR	finish-PFT.MSG	be.OPT
EGl	He should have read the book by now.						

Past can also be denoted by a contingent mood, in which case it is called past contingent (also called counterfactual). Noteworthy in past construction is that there is no use of perfective or overt past marking in such construction. Past contingent is usually marked with the imperfective form of V2 /ho:/ (/hota/, /hoti/, /hotĩ/, /hote/), thus the following sentences denote a counterfactual mood:

E67	काश	वह	पढा	होता
IPA	ka:ʃ	wəh	pəɖʰ-a:	ho:-ta:
Gl.	wish.INJ	3SG	read-PFT.MSG	be-IMPF.MSG
EGl	I wish he had read (it).			

E68	अब	तक	तो	तुम	किताब	पढ	चुके	होते
IPA	əb	tək	to:	tʊm	kɪta:b	pəɖʰ	cʊk-e:	ho:-te:
Gl.	now	till	PRT	2SG	book.N	read.VR	finish-PFT.PL	be-IMPF.PL
EGl	By now you would have read the book.							

The following table gives a summary of all the moods that can be captured through the templates.

Mood Names	Mood Tags
Imperative	imp
Optative	opt
Abilitative	abil
Declarative	dcl
Permissive Imperative	perm_imp
Permissive Optative	perm_opt
Counterfactual	cfct
Desiderative/Suggestive	sugg
Permissive Desiderative	perm_sugg
Optative Probabilitative	opt_prob

Table 26: Moods Marked in the TAM Tags

3.6 COMPOUND VERB V2 VS. THE ASPECTUAL/MODAL V2S

One very important thing to note here is the structure of the verb group in it. The importance of this construction comes to the fore when we see the verb groups having a compound verb structure. Compound verb is often defined as a sequence of verb + verb where the second verb takes all the TAM marking while the first verb remains in the root form.

So, let's take the following example sentence which has a compound verb structure:

E69	उसने		किताब	पढ	डाली	है
IPA	ʊs-ne:		kɪta:b	pəɖʰ	ɖa:l-i:	hɛ:
Gl.	3SG.OBL-ERG		book.N	read.VR	put-PFT.FSG	be.PRS.SG

EGl I wish he had read (it).

Here we see that the verb /ɖa:li/ occurs in the same position as /rəhi:/ in the progressive aspect formation as in the following sentence:

E40	वह	किताब	पढ़	रही	है
IPA	wəh	kıta:b	pəɖʱ	rəh-i:	hɛ:
Gl.	3SG	book.N	read.VR	live-PFT.3FSG	be.PRS.SG
EGl	She is reading a book.				

The difference is that there is no aspectual marker name attached to /ɖa:li/ constructions and the like while /rəhi:/ has a particular aspect name attached to it.

Here the VG template is VR+VR_pft+prs_aux

If we take this template then both the progressive aspect marker V2 rah_pft and the above V2 dal_pft of the compound verb would get covered with the same template. However, it would also mean that we would need to give a single VG tag to both of these verb groups which would be wrong. To deal with such a situation, we need to identify each of V2 constructions (given that the number V2s are limited, it is rather more reasonable).

Another observation which raises a question is that why not consider all the V2s as aspect or mood markers. Given that all the V2s occurring in the verb group loses their lexical meaning and effects an aspectual or modal altercation in the meaning of the verb group, this is rather a valid question. While much of the literature report that the compound verb V2s do bring about a change in the aspectual value of the verb group, no name has been given to many of such V2s. However, there is a wide range of literature available that discusses or classifies the various functions of these V2s. (e.g. Abbi, 1991, 1992, Abbi and Gopalkrishnan, 1992; Hook, 1974, to name a few). For example the V2s like /ɖa:l/, /mər/, /ma:r/, /ʈəpək/, /pʰẽ:k/ etc. show a semantic change in the manner of the action (expressed through V1) being done.

We may think of a name for the change these V2s bring in to the aspect or mood of the VG. However, it is not possible all the time to use a name, especially when aspect is already there

while the V2s add only a shade in the meaning. Thus, for the cases where an aspect is invariably encoded by a V2 (e.g. progressive gets encoded invariably by the V2 /rəh/ in its perfective form), it is proper to use a name. But for the cases where only a shade has been added, we may simply denote this by use of a terminology that indicates that it is a compound verb (I use 'cv' as the prefix in all such TAM tags).

3.7 MULTI-WORD VERB GROUP TEMPLATES AND THEIR TAM TAGS

The multi-word groups are variously described in literature. There are several terms used to indicate this phenomenon. The terms like complex predicate, compound verbs, serial verbs or even complement construction are in use. This feature is common in many human languages around the globe.

Multi-word verb groups are the verb groups which have more than one word constituting a verb group. These elements together indicate their TAM meaning. Any change in these constituents, either in order of occurrence or removal of one or more of such words would alter the meaning thus expressed.

In this sub-section I present the various periphrastically constructed VG templates and give them a TAM tag following the guidelines presented in the discussion above. For the sake of convenience and clear classification, the templates that involve compound verb constructions and passive constructions are discussed separately in the subsequent sections.

The templates presented here have been drawn based on the paradigmatic deductions involving different possible conjugations of the TAM.

3.7.1 Two-Word Verb Groups

The following table lists all the two word verb groups occurring in Non-compound verb and non-passive constructions. There are a total of 63 templates that I have found.

T. No.	VG Template	TAM Tag	Example Sentence
13	VR_impf+prs_aux	VG.prs.impf.dcl	E70
14	VR_pft+prs_aux	VG.prs.pft.dcl	E71
15	VINF+prs_aux	VG.prs.sim.oblg	E72

16	VINF+wala+prs_aux	VG.prs.sim.apprx	E73
17	cahiye+pst_aux	VG.pst.0.des	E56
18	VR_impf+pst_aux	VG.pst.impf.dcl	E74
19	VR_pft+pst_aux	VG.pst.pft.dcl	E75
20	VINF+lag_pft	VG.pst.pft_ince.dcl	E43
21	VR_impf+rah_pft	VG.pst.impf_pft_cnt.dcl	E44
22	VR_impf+ja_pft	VG.pst.pft_dur.dcl	E45
23	VR_pft+kar_impf	VG.pst.pft_frq.dcl	E46
24	VR_impf+ho_impf	VG.pst.impf.cfct	E76
25	VR+sak_pft	VG.pst.pft.abil	E77
26	VR+pa_pft	VG.pst.pft.abil	E78
27	VINF+de_pft	VG.pst.pft.perm	E79
28	VINF+par_pft	VG.pst.pft.oblg	E80
29	VINF+ho_impf	VG.pst.impf.oblg	E81
30	VINF+pst_aux	VG.pst.sim.oblg	E82
31	VR_impf+ban_pft	VG.pst.pft.abil	E83
32	VR_pft+rah_pft	VG.pst.pft_cnt.dcl	E52
33	VINF+wala+pst_aux	VG.pst.sim.apprx	E84
34	VINF+de_impf	VG.pst.impf.perm	E85
35	VINF+pa_pft	VG.pst.pft.abil	E86
36	VR_impf+cal_pft	VG.pst.pft.dcl	E87
37	VR_impf+rah_impf	VG.pst.dur.cgnt	E88
38	VR_pft+ho_impf	VG.pst.impf.cgnt	E89
39	VR_pft+rakh_pft	VG.pst.pft.dcl	E90

40	VR_pft+ho_fut	VG.fut.pft.dcl	E91
41	VINF+lag_fut	VG.fut.sim_ince.dcl	E92
42	VR_impf+rah_fut	VG.fut.cnt.dcl	E93
43	VR_impf+ja_fut	VG.fut.dur.dcl	E94
44	VR_pft+kar_fut	VG.fut.impf_frq.dcl	E95
45	VR_impf+ho_fut	VG.fut.impf.prsm	E96
46	VR_impf+ho_prsm_aux	VG.fut.impf.cgnt	E97
47	VR+sak_fut	VG.fut.sim.abil	E98
48	VR+pa_fut	VG.fut.sim.abil	E99
49	VINF+de_fut	VG.fut.sim.perm	E100
50	VINF+par_fut	VG.fut.sim.oblg	E101
51	VINF+ho_fut	VG.fut.sim.oblg	E102
52	VR_pft+rah_fut	VG.fut.pft_cnt.dcl	E103
53	VINF+wala+ho_fut	VG.fut.sim.apprx	E104
54	VR_impf+VR_opt	VG.0.0.opt	E105
55	VR_pft+rah_opt	VG.0.pft_cnt.opt	E106
56	VR_pft+rah_imp	VG.0.pft_cnt.imp	E107
57	VR+sak_imp	VG.0.0.abil	E108
58	VR_pft+kar_imp	VG.0.0.imp	E109
59	VINF+de_imp	VG.0.0.perm_imp	E110
60	VINF+de_opt	VG.0.0.perm_opt	E111
61	VINF+cahiye	VG.0.0.sugg	E112
62	VINF+de_inf	VG.0.0.perm_imp	E113
63	VR_impf+cal_imp	VG.0.0.imp	E114

64	VR_impf+cal_opt	VG.0.0.opt	E115
65	VR_impf+ho_opt	VG.0.0.cfct	E116
66	VR_impf+rah_opt	VG.0.0.opt	E117
67	VR_pft+ho_opt	VG.0.0.cfct	E118
68	VR_pft+VINF_imp	VG.0.0.imp	E119
69	VINF+wala	VG.gp	E120
70	VR+kar	VG.cp.pft	E121
71	VR_impf+ho_pft	VG.ip.impf	E122
72	VR_pft+ho_pft	VG.ip.pft	E123
73	VR_pft+VINF	VG.nv.pft	E124
74	VR+rah_pft	VG.ip.prog	E125
75	VR_impf+ja_impf	VG.ip.impf.dur	E126

Table 27: Two-Word VG Templates

The table above can be summarized in the following the table.

Tense Markers	Number of templates	Template Sl. Nos.
Present Tense Templates	4	13-16
Past Tense Templates	23	17-39
Future Templates	14	40-53
Zero Tense Templates	15	54-68
Non Main Verb Templates	8	69-75

Table 28: Summary of Two-Word VG Templates

The summary table above shows that among the two word verb groups, there is a total of 4 templates that are in present tense, 23 templates show past tense and 14 templates show future tense. There are 15 templates that do not have any tense assigned to their TAM tags. These templates are in a special mood where the tense can be construed either as null or at best present tense. Otherwise, the tense would depend on the larger context of the text such VGs appear in. There are other 7 templates which might occur as non-main verbs within a sentence.

3.7.2 Three-Word Verb Group

The following table gives an overview of all the three-word templates captured.

T. No.	VG Template	TAM Tag	Example Sentence
76	VR_impf+ban_pft+prs_aux	VG.prs.impf.abil	E127
77	VR+cuk_pft+prs_aux	VG.prs.pft.dcl	E128
78	VR+rah_pft+prs_aux	VG.prs.prog.dcl	E129
79	VINF+lag_impf+prs_aux	VG.prs.impf_ince.dcl	E130
80	VINF+lag_pft+prs_aux	VG.prs.pft_ince.dcl	E131
81	VR_impf+rah_impf+prs_aux	VG.prs.impf_cnt.dcl	E132
82	VR_impf+rah_pft+prs_aux	VG.prs.impf_pft_cnt.dcl	E133
83	VR_pft+rah_impf+prs_aux	VG.prs.pft_impf_cnt.dcl	E134
84	VR_impf+ja_impf+prs_aux	VG.prs.impf_dur.dcl	E51
85	VR_impf+ja_pft+prs_aux	VG.prs.pft_dur.dcl	E135
86	VR_pft+kar_impf+prs_aux	VG.prs.impf_frq.dcl	E136
87	VR+sak_impf+prs_aux	VG.prs.impf.abil	E137
88	VR+sak_pft+prs_aux	VG.prs.pft.abil	E138
89	VR+pa_impf+prs_aux	VG.prs.impf.abil	E139
90	VR+pa_pft+prs_aux	VG.prs.pft.abil	E140
91	VINF+de_impf+prs_aux	VG.prs.impf.perm	E141
92	VINF+de_pft+prs_aux	VG.prs.pft.perm	E142
93	VINF+par_impf+prs_aux	VG.prs.impf.oblg	E143
94	VINF+par_pft+prs_aux	VG.prs.pft.oblg	E144
95	VINF+ho_impf+prs_aux	VG.prs.impf.oblg	E145
96	VR_impf+a_pft+prs_aux	VG.prs.impf_dur.dcl	E146
97	VR_impf+ho_inf+cahiye	VG.prs.sim.sugg	E147

98	VR_pft+ho_impf+prs_aux	VG.prs.impf.dcl	E148
99	VR_pft+ho_pft+prs_aux	VG.prs.pft.dcl	E149
100	VR_pft+par_pft+prs_aux	VG.prs.pft.dcl	E150
101	VR_pft+rakh_impf+prs_aux	VG.prs.pft_dur.dcl	E151
102	VR_impf+ban_pft+pst_aux	VG.pst.impf.abil	E152
103	VR+cuk_pft+pst_aux	VG.pst.pft.dcl	E153
104	VR+rah_pft+pst_aux	VG.pst.prog.dcl	E41
105	VINF+lag_impf+pst_aux	VG.pst.impf_ince.dcl	E50
106	VINF+lag_pft+pst_aux	VG.pst.pft_ince.dcl	E154
107	VR_impf+rah_impf+pst_aux	VG.pst.impf_cnt.dcl	E155
108	VR_impf+rah_pft+pst_aux	VG.pst.impf_pft_cnt.dcl	E53
109	VR_pft+rah_impf+pst_aux	VG.pst.pft_impf_cnt.dcl	E156
110	VR_impf+ja_impf+pst_aux	VG.pst.impf_dur.dcl	E157
111	VR_impf+ja_pft+pst_aux	VG.pst.pft_dur.dcl	E158
112	VR_pft+kar_impf+pst_aux	VG.pst.impf_frq.dcl	E159
113	VR+sak_impf+pst_aux	VG.pst.impf.abil	E160
114	VR+sak_pft+pst_aux	VG.pst.pft.abil	E161
115	VR+pa_impf+pst_aux	VG.pst.impf.abil	E162
116	VR+pa_pft+pst_aux	VG.pst.pft.abil	E163
117	VINF+de_impf+pst_aux	VG.pst.impf.perm	E164
118	VINF+de_pft+pst_aux	VG.pst.pft.perm	E165
119	VINF+par_impf+pst_aux	VG.pst.impf.oblg	E166
120	VINF+par_pft+pst_aux	VG.pst.pft.oblg	E167
121	VINF+ho_impf+pst_aux	VG.pst.impf.oblg	E168

122	VINF+de_pft+ho_impf	VG.pst.pft.perm_cfct	E169
123	VINF+cahiye+pst_aux	VG.pst.0.sugg	E170
124	khada+rah+ja_pft	VG.pst.dur.dcl	E171
125	VR_impf+cal_pft+ja_pft	VG.pst.pft_cnt.dcl	E172
126	VR_impf+rah+ja_pft	VG.pst.pft_dur.dcl	E173
127	VR_pft+ho_pft+pst_aux	VG.pst.pft.dcl	E174
128	VR_pft+par_pft+pst_aux	VG.pst.pft.dcl	E175
129	VR_impf+ban_fut	VG.fut.impf.abil	E176
130	VR+cuk_pft+ho_fut	VG.fut.pft.dcl	E177
131	VR+rah_pft+ho_fut	VG.fut.prog.prsm	E178
132	VR_pft+kar_impf+ho_fut	VG.fut.impf_dur.dcl	E179
133	VR+sak_pft+ho_fut	VG.fut.pft.abil	E180
134	VR+pa_pft+ho_fut	VG.fut.pft.abil	E181
135	VINF+de_pft+ho_fut	VG.fut.pft.perm_cfct	E182
136	VINF+par_pft+ho_fut	VG.fut.pft.oblg	E183
137	VINF+ho_impf+ho_fut	VG.fut.pft.oblg	E184
138	VINF+rah_pft+ho_fut	VG.fut.prog.oblg_prsm	E185
139	VINF+de_impf+ho_fut	VG.fut.impf.perm_cfct	E186
140	VINF+lag_pft+ho_fut	VG.fut.pft_ince.prsm	E187
141	VINF+lag_impf+ho_fut	VG.fut.impf_ince.prsm	E188
142	VR_impf+rah_pft+ho_fut	VG.fut.impf_pft.prsm	E189
143	VR+rah_pft+ho_cfct	VG.0.prog.cfct	E190
144	VINF+de_inf+cahiye	VG.0.0.perm_sugg	E191
145	VR_impf+ja+rah_pft	VG.ip.prog.dur	E192

146	VR+ja+rah_pft	VG.ip.prog	E193

Table 29: Three-Word VG Templates

The following table provides a summary of the three word VG templates. There are a total of 71 VG templates in this group. Here, we find that past tense dominates in the three word VGs with a total number of 27 templates, followed by present (26 templates) and future (14). Five templates have not been assigned any tense tags while we still have two templates that are verb groups not playing as main verbs.

Tense Markers	Number of templates	Template Sl. Nos.
Present Tense Templates	26	76-101
Past Tense Templates	27	102-128
Future Templates	14	129-142
Zero Tense Templates	2	143-144
Non Main Verb Templates	2	145-146

Table 30: Summary of Three-Word VG Templates

3.7.3 Four-Word Verb Group

The four word VGs are found mostly in the progressive aspect and its variants. I find a total of 26 templates in various TAM categories that require a total of four words to denote its TAM. These are listed in the table below. There are a total of 9 templates each in present and past tense. The future tense has a total of only 7 templates because the abilitative/probabilitative mood construction is not possible here.

T. No.	VG Template	TAM Tag	Example Sentence
147	VR_impf+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl	E194
148	VR+sak+rah_pft+prs_aux	VG.prs.prog.abil	E195
149	VR+pa+rah_pft+prs_aux	VG.prs.prog.abil	E196
150	VINF+de+rah_pft+prs_aux	VG.prs.prog.perm	E197
151	VINF+par+rah_pft+prs_aux	VG.prs.prog.oblg	E198
152	VINF+par+sak_impf+prs_aux	VG.prs.impf.abil	E199
153	VR_impf+a+rah_pft+prs_aux	VG.prs.dur.dcl	E200
154	VR_pft+a+rah_pft+prs_aux	VG.prs.dur.dcl	E201

155	VR_pft+rah+sak_impf+prs_aux	VG.prs.pft_dur.abil	E202
156	VR_impf+ja+rah_pft+pst_aux	VG.pst.prog_dur.dcl	E203
157	VR+sak+rah_pft+pst_aux	VG.pst.prog.abil	E204
158	VR+pa+rah_pft+pst_aux	VG.pst.prog.abil	E205
159	VINF+de+rah_pft+pst_aux	VG.pst.prog.perm	E206
160	VINF+par+rah_pft+pst_aux	VG.pst.prog.oblg	E207
161	VINF+par+sak_impf+pst_aux	VG.pst.impf.abil	E208
162	VR_impf+a+rah_pft+pst_aux	VG.pst.dur.dcl	E209
163	VR_pft+a+rah_pft+pst_aux	VG.pst.dur.dcl	E210
164	VR_pft+rah+sak_impf+pst_aux	VG.pst.pft_dur.abil	E211
165	VR_impf+ja+rah_pft+ho_fut	VG.fut.impf_dur.prsm	E212
166	VR+sak+rah_pft+ho_fut	VG.fut.prog.abil_prsm	E213
167	VR+pa+rah_pft+ho_fut	VG.fut.prog.abil_prsm	E214
168	VINF+de+rah_pft+ho_fut	VG.fut.prog.perm_prsm	E215
169	VINF+par+rah_pft+ho_fut	VG.fut.prog.oblg_prsm	E216
170	VR_impf+a+rah_pft+ho_fut	VG.fut.dur.dcl	E217
171	VR_pft+a+rah_pft+ho_fut	VG.fut.dur.dcl	E218

Table 31: Four-Word VG Templates

3.7.4 Five-Word Verb Group

Five word verb groups are far and few between. However, one example has already been given in the previous chapter. This occurs only in the progressive durative aspect demarcation. There are two V2 sequences that can be added to an already progressive template. These two V2 sequences are /cəla: ja:/ and /cəla: a:/. In such a five word VG construction these two V2 sequences require that the main verb be in the imperfective form. The other tense markings can change as usual. Thus we get six more templates in the five word VG as noted in the table below.

T. No.	VG Template	TAM Tag	Example Sentence
172	VR_impf+cal_pft+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl	E219
173	VR_impf+cal_pft+ja+rah_pft+pst_aux	VG.pst.prog_dur.dcl	E220
174	VR_impf+cal_pft+a+rah_pft+prs_aux	VG.prs.prog_dur.dcl	E221
175	VR_impf+cal_pft+a+rah_pft+pst_aux	VG.pst.prog_dur.dcl	E222
176	VR_impf+cal_pft+ja+rah_pft+ho_fut	VG.fut.impf_dur.cfct	E223
177	VR_impf+cal_pft+a+rah_pft+ho_fut	VG.fut.impf_dur.cfct	E224

Table 32: Five Word VG Templates

3.8 VERB GROUP IN PASSIVE VOICE

Passive voice construction can be viewed as a transformation of what occurs in the active voice. However, not all the templates we have drawn so far will have a passive construction. This is so because not all the aspects and moods will have a corresponding passive construction. For example the sentences in imperative mood cannot transform into passive voice.

Below we note all the passive constructions possible in the non-compound verb construction. I have found a total of 69 templates that can be categorized under the non-CV passive constructions. The impersonal constructions are also counted in this category. We categorize these templates again into the number of words these VGs have.

The passive construction in Hindi requires the passive marking verb /ja:/. This makes the passive construction a multi-word affair. Therefore there cannot be any VG in passive voice with a single verbal word.

3.8.1 Two-Word Passive VGs

Two-word passive VGs are the passive transformations of those one-word active VGs that can have a passive construction. There is no present tense in this group of templates. Past tense has a total of 2 templates and future has just one. One template covers the optative mood. The other two templates are non main-verb templates. Thus there are a total of 6 templates in this group.

T. No.	VG Template	TAM Tag	Example Sentence

178	VR_pft+ja_pft	VG.pas_pst.pft.dcl	E225
179	VR_pft+ja_impf	VG.pas_pst.impf.cgnt	E226
180	VR_pft+ja_fut	VG.pas_fut.sim.dcl	E227
181	VR_pft+ja_opt	VG.pas_0.0.opt	E228
182	VR_pft+ja_inf	VG.pas_nv.pft	E229
183	VR_pft+ja_inf+wala	VG.pas_gp	E230

Table 33: Two-Word Passive VG Templates

3.8.2 Three-Word Passive VGs

Three-word passive VGs include all the tenses. There are a total of 35 VG templates that come under it. I have also included the constructions which I call as approximative (templates 188, 201 and 210). These are VGs where the /wa:la:/ particle gives a sense of approximation i.e. the action is yet to begin. We consider them as three word VG because we nullify the particles when counting the number of words in the VG. There are also six VG templates that have no tense marked for them (templates 212-217). These templates are in moods that make the tense category fuzzy and the mood marking itself should be enough to avoid this fuzziness. We also have one template (T. No. 218) which does not occur as main verb in a sentence.

T. No.	VG Template	TAM Tag	Example Sentence
184	VR_pft+ja_impf+prs_aux	VG.pas_prs.impf.dcl	E231
185	VR_pft+ja_pft+prs_aux	VG.pas_prs.pft.dcl	E232
186	VR_pft+ja+rah_pft+prs_aux	VG.pas_prs.prog.dcl	E233
187	VR_pft+ja_inf+prs_aux	VG.pas_prs.sim.oblg	E234
188	VR_pft+ja_inf+wala+prs_aux	VG.pas_prs.sim.apprx	E235
189	VR_pft+ja_impf+pst_aux	VG.pas_pst.impf.dcl	E236
190	VR_pft+ja_pft+pst_aux	VG.pas_pst.pft.dcl	E237
191	VR_pft+ja+rah_pft+pst_aux	VG.pas_pst.prog.dcl	E238
192	VR_pft+ja_inf+lag_pft	VG.pas_pst.pft_ince.dcl	E239

193	VR_pft+ja_inf+lag_fut	VG.pas_pst.pft_ince.dcl	E240
194	VR_pft+ja_impf+rah_pft	VG.pas_pst.pft_cnt.dcl	E241
195	VR_pft+ja_impf+rah_fut	VG.pas_pst.pft_cnt.dcl	E242
196	VR_pft+ja_pft+ho_impf	VG.pas_pst.impf.cfct	E243
197	VR_pft+ja+sak_pft	VG.pas_pst.pft.abil	E244
198	VR_pft+ja_inf+ho_impf	VG.pas_pst.impf.oblg	E245
199	VR_pft+ja_inf+pst_aux	VG.pas_pst.sim.oblg	E246
200	VR_pft+ho_impf+pst_aux	VG.pas_pst.pft_impf_cnt.dcl	E247
201	VR_pft+ja_inf+wala+pst_aux	VG.pas_pst.sim.apprx	E248
202	VR_pft+ja+sak_impf	VG.pas_pst.impf.cgnt	E249
203	VR_pft+ja_pft+ho_fut	VG.pas_fut.pft.dcl	E250
204	VR_pft+ja+rah_pft+ho_fut	VG.pas_fut.prog.prsm	E251
205	VR_pft+ja_impf+ho_fut	VG.pas_fut.impf.prsm	E252
206	VR_pft+ja+sak_fut	VG.pas_fut.sim.abil	E253
207	VR_pft+ja_impf+ho_opt	VG.pas_fut.impf.cgnt	E254
208	VR_pft+ja_inf+ho_fut	VG.pas_fut.sim.oblg	E255
209	VR_pft+ho_impf+ho_fut	VG.pas_fut.pft_cnt.dcl	E256
210	VR_pft+ja_inf+wala+ho_fut	VG.pas_fut.sim.apprx	E257
211	VINF+de_pft+ja_fut	VG.pas_fut.sim.perm	E258
212	VR_pft+ja_impf+rah_opt	VG.pas_0.0.opt	E259
213	VR_pft+ja_pft+ho_opt	VG.pas_0.pft.cfct	E260
214	VR_pft+ja_inf+cahiye	VG.pas_0.0.sugg	E261
215	VR_pft+ja_inf+ho_opt	VG.pas_0.0.opt	E262
216	VR_pft+ja+sak_imp	VG.pas_0.0.opt_prob	E263

217	VR_pft+ja+sak_inf	VG.pas_0.0.inf	E264
218	VR_pft+ja+rah_pft	VG.pas_ip.dur	E265

Table 34: Three-Word Passive VG Templates

3.8.3 Four-Word Passive VGs

A total of 26 VG templates require four verbal words as their constituent. Present and past have 9 templates each. Future has a total of 6 templates. There are also two templates that have zero tense because they are in counterfactual mood.

Some of the templates may be disputed here. These are the templates whose example starts with a question mark. It may be hard to find a real world example matching these templates. I have included these templates also here so that just in case if we find such a template in the real text, it could be captured.

T. No.	VG Template	TAM Tag	Example Sentence
219	VR_pft+ja+cuk_pft+prs_aux	VG.pas_prs.pft.dcl	E266
220	VR_pft+ja_inf+lag_impf+prs_aux	VG.pas_prs.impf_ince.dcl	E267
221	VR_pft+ja_inf+lag_pft+prs_aux	VG.pas_prs.pft_ince.dcl	E268
222	VR_pft+ja_impf+rah_pft+prs_aux	VG.pas_prs.impf_cnt.dcl	E269
223	VR_impf+ja_impf+rah_pft+prs_aux	VG.pas_prs.pft_dur.dcl	E270
224	VR_pft+ja+sak_impf+prs_aux	VG.pas_prs.impf.abil	E271
225	VR_pft+ja+sak_pft+prs_aux	VG.pas_prs.pft.abil	E272
226	VR_pft+ja_pft+ho_impf+prs_aux	VG.pas_prs.pft_impf_cnt.dcl	E273
227	VR_pft+ja_impf+rah_impf+prs_aux	VG.pas_prs.pft_cnt.dcl	E274
228	VR_pft+ja_impf+rah_impf+pst_aux	VG.pas_pst.pft_cnt.dcl	E275
229	VR_pft+ja+cuk_pft+pst_aux	VG.pas_pst.pft.dcl	E276
230	VR_pft+ja_inf+lag_impf+pst_aux	VG.pas_pst.pft_ince.dcl	E277
231	VR_pft+ja_inf+lag_pft+pst_aux	VG.pas_pst.impf_ince.dcl	E278

232	VR_pft+ja_impf+rah_pft+pst_aux	VG.pas_pst.impf_cnt.dcl	E279
233	VR_pft+ja+sak_impf+pst_aux	VG.pas_pst.impf.abil	E280
234	VR_pft+ja+sak_pft+pst_aux	VG.pas_pst.pft.abil	E281
235	VR_pft+ja_inf+cahiye+pst_aux	VG.pas_pst.0.sugg	E282
236	VR_pft+ja_pft+ho_impf+pst_aux	VG.pas_pst.pft_impf_cnt.dcl	E283
237	VR_pft+ja+cuk_pft+ho_fut	VG.pas_fut.pft.dcl	E284
238	VR_pft+ja+sak_pft+ho_fut	VG.pas_fut.pft.abil	E285
239	VR_pft+ja_inf+lag_pft+ho_fut	VG.pas_fut.pft_ince.prsm	E286
240	VR_pft+ja_inf+lag_impf+ho_fut	VG.pas_fut.impf_ince.prsm	E287
241	VR_pft+ja_impf+rah_pft+ho_fut	VG.pas_fut.impf_pft.prsm	E288
242	VR_pft+ja_pft+ho_impf+ho_fut	VG.pas_fut.pft_cnt.prsm	E289
243	VR_pft+ja+rah_pft+ho_opt	VG.pas_0.prog.cfct	E290
244	VR_pft+ja+sak_pft+ho_opt	VG.pas_0.0.cfct	E291

Table 35: Four-Word Passive VG Templates

3.9 VERB GROUP IN COMPOUND VERB

In our brief introduction to the compound verb in Hindi we noted that it has the structure V1+V2 where V1 is always VR while V2 takes the TAM and agreement markings. If this is the definition of a compound verb, there are several entries in the templates we have drawn so far that should count as compound verb. For example, templates 25, 26 involving the V2s of सक /səkˑ/ and पा/pa:/ as in

E292	वह	दौड़	सका	
IPA	wəh	dɔ:ɖ	səkˑ-a:	
Gl.	3SG	run.VR	can-PFT.MSG	
EGl	He could run.			

E293	वह	चल	नहीं	पाया	
IPA	wəh	cəl	nəhĩ:	paːjaː	
Gl.	3SG	walk.VR	NEG	get-PFT.MSG	
EGl	He could not walk.				

may also be counted as compound verbs if we consider just the structure. Besides these two V2s, there some other V2s that follow a verb stem in the templates we have drawn so far. These V2s are /cʊk/ and /rəh/. Some other templates that might be considered as having a compound verb (CV) structure are 113-116, 130-131 where the verb group starts with a verb stem and takes one of these V2s as in the following sentences.

E129	वह	किताब	पढ	रहा	है	
IPA	wəh	kɪtaːb	pəɖʰ	rəh-aː	hɛː	
Gl.	3SG	book.N	read.VR	live-PFT.3MSG	be. PRS	
EGl	He is reading the book.					

E128	वह	किताब	पढ	चुका	है
IPA	wəh	kɪtaːb	pəɖʰ	cʊk-aː	hɛː
Gl.	3SG	book.N	read.VR	finish-PFT.3MSG	be. PRS
EGl	He has read the book.				

A major question for this researcher is whether it is needed to identify if a verb group contains a CV or not. Given the goal of this research is to identify the verb groups and find out the TAM information in them, it appears that that even if we do not mark a VG for a CV or non-CV, the

goal is fulfilled. Our goal is achieved if we identify all the verb groups along with their TAM information whether it contains a compound verb or not.

However, I have stated above in section 3.6 that we will use the prefix of ‘cv’ to indicate that a VG contains a CV. I add to it that we will use this ‘cv’ prefix only in cases that where the V2s are not the same as the ones that have appeared in the aspect and mood discussion above. That is barring the templates we have drawn so far and which might be interpreted as having compound verbs.

3.9.1 The V2s or the vector verbs

In compound verbs, the second verb has been called with several names such as ‘explicator’, vector verbs, second verb and light verb. I have been constantly referring to it as V2 and will continue to do so. This V2 is always a closed set. There is no unanimity upon exactly how many V2s are there in Hindi. Hook (1974) tries to provide a list of the various V2 suggested by several scholars. Drawing on all of the V2s as described in Hook, the researcher has presented a list of the common V2s in the previous chapter (see Table 16 and 17). There are a total of 61 such verbs. Not all of these verbs are listed as V2s by every scholar. However, this list is merely a combination of all the V2s listed by every writer. That is why this list is also perhaps the biggest possible for the V2s in Hindi.

3.9.2 Drawing Templates with V2s

In our approach we have been enumerating what we call the auxiliaries. We have also said that auxiliaries are limited in number. All the V2s are also auxiliaries in this sense. This means that we have to come up with a template for all the conjugations that the V2s allow. Suppose we consider a compound verb construction as the verb root, it would mean that we would get all the forms of the V2s that can appear as a single verb.

3.9.2.1 V1 vs V2: Most V2s demand a particular V1

It is also noticeable that the V2s do not come with all the stems as V1s. Except for a few, the V2s have some specific requirements for V1s. For example the pairs of V1+V2 like

*सो गिरा /so: gɪra:/ “fell asleep”

*रो धमका /ro: dʰəmka:/ “cried out”

*पढ़ दौड़ा /pəɖʱ dɔ:ɖa:/ “read ...” etc.

as compound verb are wrong. Because we are simply enumerating the templates, it is not possible to enumerate the V1s for each of the V2s as it will involve a semantic analysis. So, if an unwanted pair of compound verb occurs in the text, it will yet be recognized as a template. However, we believe that the input text and its annotation are perfect (more on annotation in the next chapter). The following table gives all the V2s considered in this work along with a few examples that they can occur with. We provide this template to basically show that a compound verb can be formed with the combination of these two verbs in various templates that we draw based on the previous templates we have drawn so far.

V2s	**Example of V1s attached with**	**V2s**	**Example of V1s attached with**
आ /a:/ “come”	उतर /ʊtər/ “get down”, हो /ho:/ “be/become”,	निकाल /nɪka:l/ “take out”	सोच /soc/ “think”, कर /kər/ “do”,
उठ /ʊʈʰ/ “get up”	कांप /kãp/ “shiver”, जल /ɟəl/ “burn”,	पकड़ /pəkəɖ/ “catch”	आ /a:/ “come”, जा /ja:/ “go”, धर /dʰər/ “hold”,
उड़ /ʊɖ/ “fly”	चल /cəl/ “climb”, ले /le:/ “take”,	पटक /pəʈək/ “dash”	जा /ja:/ “go”, दे /de:/ “give”, ला /la:/ “bring”,
खप /kʰəp/ “finish”	आ /a:/ “come”, जा /ja:/ “go”,	पड़ /pəɖ/ “lay”	रो /ro:/ “cry”, बोल /bol/ “speak”,,
खा /kʰa:/ “eat”	मार /ma:r/ “kill/beat”, काट /ka:ʈ/ “cut”,	पहुँच /pəhʊ̃c/ “reach”	आ /a:/ “come”, जा, /ja:/ “go”, ले /le:/ “take”,
गिर /gɪr/ “fall”	आ /a:/ “come”, जा /ja:/ “go”,	पा /pa:/ “get”	पहुंच /pəhũ:c/ “reach”, पढ़ /pəɖʱ/ “read”, बिता /bɪta:/ “pass”, कर /kər/ “do”,
गिरा /gɪra:/ “make	मार /ma:r/	फँस /pʰə̃s/ “get	आ /a:/ “come”,

fall”	“kill/beat”, दे /de:/ “give”,	trapped”	जा /ja:/ “go”,
गुजर /gʊjər/ “come”	हो /ho:/ “be/become”, कर /kər/ “do”,	फिर /pʰɪr/ “turn/wander”	चल /cəl/ “climb”,
घसीट /gʰəsiʈ/ “pull”	ला /la:/ “bring”,	फेंक /pʰẽ:k/ “throw”	उखाड़ /ʊkʰa:ɖ/ “uproot”, दे /de:/ “give”,
घुस /gʰʊs/ “enter”	आ /a:/ “come”, जा, /ja:/ “go”,	बन /bən/ “form”	आ /a:/ “come”, चढ़ /cəɖʰ/ “climb”,
घेर /gʰe:r/ “surround”	आ /a:/ “come”, जा, /ja:/ “go”,	बस /bəs/ “settle down”	आ /a:/ “come”, जा /ja:/ “go”, चल /cəl/ “climb”,
चढ़ /cəɖʰ/ “climb”	आ /a:/ “come”, जा, /ja:/ “go”,	बीत /bi:t/ “pass”	ले /le:/ “take”,
चल /cəl/ “walk”	हो /ho: / “be/become”, मार /ma:r/ “kill/beat”,	बैठ /bɛ:ʈʰ/ “sit”	बन /bən/ form, आ /a:/ “come”, मार /ma:r/ “kill/beat”, ले /le:/ “take”,
चुक /cʊk/ “finish”	बाँध /bã:dʰ/ “tie”, हो /ho:/ “be/become”, खा /kʰa:/ “eat”, उठा /ʊʈʰa:/ “raise”	भाग /bʰa:g/ “run away”	निकल /nɪkəl/ “come out”, धर /dʰər/ “catch”,
छोड़ /cʰoɖ/ “leave”	बना /bəna:/ “make”,	भिड़ /bʰɪɖ/ “collide”	आ /a:/ “come”, जा, /ja:/ “go”, चल /cəl/ “climb”, दौड़ /dɔ:ɖ/ “run”
जा /ja:/ “go”	हो /ho:/ “be/become”, छूट /cʰu:ʈ/ “release”, खा /kʰa:/ “eat”,	भेज /bʰe:ɟ/ “send”	कहला /kəhla:/ “get told”, बुला /bʊla/ “call”, मंगा /mə̃ga/ “call for”, लिख /lɪkʰ/ “write”
झपट /jʰəpəʈ/	आ /a:/ “come”,	मंगा /mə̃ga:/ “call	भेज /bʰeɟ/ “send”,

“pounce”	ले /le:/ “take,	for”	
टपक /ʈəpək/ “drip”	आ /a:/ “come”, पहुंच /pəhũ:c/ “reach”,	मर /mər/ “die”	आ /a:/ “come”, लड़ /ləɖ/ “fight”, चल /cəl/ “climb”, कर /kər/ “do”,
टूट /ʈu:ʈ/ “break”	आ /a:/ “come”, जा /ja:/ “go”,	मार /mar:/ “kill/beat”	छान /cʰa:n/ “sift”, दे /de:/ “give”, दौड़ा /dɔ:ɖa/ “chase”,
ठहर /ʈʰəhər/ “stop”	आ /a:/ “come”, जा /ja:/ “go”,	मिट /mɪʈ/ “get obliterated”	मर /mər/ “die”, जल /ɟəl/ “burn”,
डट /ɖəʈ/ “stand firm”	आ /a:/ “come”, जा /ja:/ “go”,	मिल /mɪl/ “get found”	आ /a:/ “come”, जा, /ja:/ “go”, लिख /lɪkʰ/ “write”,
डाल /ɖa:l/ “put”	पढ़ /pəɖʰ/ “read”, मार /mar/ “kill”,	रख /rəkʰ/ “keep”	हो /ho:/ “be/become”, आ /a:/ “come”, पाल /pa:l/ “raise”,
डूब /ɖu:b/ “sink”	ले /le:/ “take”,	रह /rəh/ “live”	सो /so:/ “sleep”,, मिल /mɪl/ “/meet/find”,
ढा /ɖʰa:/ “run”	चढ़ /cəɖʰ/ “climb”,	लग /ləg/ “begin/feel”	आ /a:/ “come”, जा, /ja:/ “go”, चूम /cu:m/ “kiss”, लग /ləg/ “to begin/get felt/to touch/to stick "
दिखा /dɪkʰa:/ “show”	जा /ja:/ “go”,	ला /la:/ “bring”	भर /bʰər/ “fill”, चढ़ा /cəɖʰa:/ “help climb”,
दे /de:/ “give”	चल /cəl/ “climb”,	लिपट /lɪpəʈ/ “cling”	आ /a:/ “come”, जा, /ja:/ “go”,
देख /de:kʰ/ “see”	कर /kər/ “do”,	ले /le:/ “take”	सो /so:/ “sleep”, मिल /mɪl/ “meet/find”, हो /ho / “be/become”, कर /kər/ “do”,
दौड़ /dɔ:ɖ/ “run”	ले /le:/ “take,	लेट /le:ʈ/ “lie	आ /a:/ “come”,

	आ /a:/ “come”, जा /ja:/ “go”, चढ़ /cəɖʰ/ “climb”,	down”	जा, /ja:/ “go”,
धमक /dʰəmək/ “thump”	आ /a:/ “come”, जा /ja:/ “go”,	सक /sək/ “can/be able to”	उड़ /ʊɖ/ “fly”, बदल /bədəl/ “change”,
धर /dʰər/ “catch”	कर /kər/ “do”, पकड़ /pəkəɖ/ “catch”,	सुना /sʊna:/ “tell/narrate”	कह /kəh/ “say”,
निकल /nɪkəl/ “come out”	आ /a:/ “come”, बढ़ /bəɖʰ/ “move ahead”,		

Table 36: V1 V2 Parity

3.9.2.2 CV Templates Correspond to Non-CV Templates

We have already drawn templates for verbs that are not a compound construction. Given that a compound verb may also behave more like a main verb where the V2 takes all the TAM markings, can we simply apply the same templates as drawn above so far to fit into these compound verbs as well? The only change that will take place in the templates will be the addition of the V1 in its stem form at the beginning. For example for the one word VG templates we will have a two word VG templates which will be a compound verb construction.

Thus for the template numbers 1-12, we can check whether we will have template for each of them. For example templates 1 and 2 are copular sentences and we will not have a corresponding template in CV construction. Template 9 which is an imperfective participial construction will also not have a corresponding CV template. For example, we can say /bɛʈʰəte-bɛ:ʈʰəte/ but we cannot say /mar bɛʈʰəte-bɛ:ʈʰəte/. Template 12 is also exempt from the compound verb constructions because it functions itself as a single main verb. If we want to use this modal with a compound verb construction, it will no longer be a desiderative. Rather it will turn into what we call suggestive here. This word also requires that if it comes in a VG along with another verbal word, the verb preceding it must be a –nV ending verb. That is, it should be in its infinitival construction. Thus if we use it in a compound verb construction, the structure would be as follows: V1+V2_inf+cahıye. This makes a three word verb group.

Template 8 with a conjunct participle such as /kʰa:kər/, /pəɖʰkər/ is a fuzzy case. The constructions like the following might be put in for example:

?*E294	वह	किताब	पढ़	डालकर		लिखने	बैठा
IPA	wəh	kıta:b	pəɖʰ	ɖa:l-kər		lıkʰne:	bɛ:ʈʰ-a:
Gl.	3SG	book.N	read.VR	put-PFT.PCPL		write.INF	sit-PFT.MSG
EGl	Having read the book he sat to write.						

?*E295	उपाय	सोच	निकालकर		वह	बोला
IPA	ʊpa:j	so:c	nıka:l-kər		wəh	bol-a:
Gl.	solution.N	thought.N	take out-PFT.PCPL		3SG	speak-PFT.MSG
EGl	Having found the solution he spoke.					

?*E296	कमरा	छान	मारकर		वह	थक	गया
IPA	kəmra:	cʰa:n	ma:r-kər		wəh	tʰək	gəja:
Gl.	room.N	sift.N	kill-PFT.PCPL		3SG	tire.VR	go.PFT.MSG
EGl	Having sifted through the room he tired away.						

The sentences above are rather odd or ungrammatical. We would prefer that these are not included. However, the following cases are a strong pointer and may not be called as fuzzy.

E297	जो	मार	खाकर	खड़ा	रहे	वही	आगे	चलता	है
IPA	ɟo:	ma:r	kʰa:-kər	kʰəɖ-a:	rəh-e:	wəh-i:	a:ge:	cəl-ta:	hɛ:

Gl.	who	beat.VR	eat-PFT.PCPL	stand-PFT.MSG	live-PFT.MSG	he-EMPH	ahead	walk-PFT.MSG	be.PRS.SG

EGl Those who keep standing even after being beaten, move ahead.

E298	चकाचक	कपड़े	पर	जा	<u>बैठकर</u>
IPA	cəka:cək	kəpɖe:	pər	ɟa:	<u>bɛ:ʈʰ-kər</u>
Gl.	shining	clothes.N	LOC	go.VR	sit-PFT.PCPL
EGl	… having gone, sitting on the shining clothes …				

Thus we cannot say that the corresponding CV template of template number 8 is impossible to have. Actually we can have this template but only in some select cases of the V2s. Given that the above two examples of conjunctive participle construction in compound verb is valid, we may retain more of such templates by analogy.

Template number 10 is the imperative use of the infinitive form of the verb. This template can be used for all the CVs as well. The following examples should verify this:

E299	तू	<u>उतर</u>	<u>आ</u>
IPA	tu:	<u>ʊtər</u>	<u>a:</u>
Gl.	2SG	climb down.VR	come-VR
EGl	You climb down!		

E300	दुश्मन	को	चारों	ओर	से	<u>जा</u>	<u>घेरो</u>
IPA	dʊʃmən	ko:	carõ:	o:r	se:	<u>ɟa:</u>	<u>gʰe:r-o:</u>

Gl. enemy ACC four-PL.OBL side ASS ɟa:.VR surround-IMP

EGl … Surround the enemies from all four sides …

E301	जिसको	चाहे	<u>ला</u>	<u>घसीटो</u>
IPA	ɟɪs-ko:	ca:he:	<u>la:</u>	<u>gʰəsi:ʈ-o:</u>
Gl.	who-ACC	want-PFT.PL	bring.VR	pull-IMP
EGl	… Pull any one you want …			

E302	मुझे	भी	<u>दिलवा</u>	<u>दो</u>
IPA	mʊjʰe:	bʰi:	<u>dɪlwa:</u>	<u>do:</u>
Gl.	1SG.ACC	also.PRT	cause to give.VR	give-IMP
EGl	Get me also given.			

E303	कभी	भी	<u>आ</u>	<u>धमको</u>
IPA	kəbʰi:	bʰi:	<u>a:</u>	<u>dʰəmk-o:</u>
Gl.	any time	also	come.VR	thump-IMP
EGl	Get here any time.			

Template 11 is about the nominal or gerundive use of the infinitive form (-nV ending) of the verb. We find that this construction can also be equally applied to all of the V2s. Below are a few examples:

E304 बसों के सड़कों पर उतर आने से

IPA bəsõ: ke: səɖkõ: pər ʊtər a:ne: se:

Gl. buses GEN roads LOC come down.VR come.INF INS

EGl … due to buses coming out on the roads …

E305 ले उड़ने की इजाजत न दें

IPA le: uɖne: ki: ɪja:jət nə dẽ:

Gl. take.INF fly.INF GEN permission.N NEG give-IMP

EGl … do not permit to fly away with …

E306 हजार दीपक के एक साथ जल उठने से

IPA həja:r di:pək ke: e:k sa:tʰ jəl ʊɭʰne: se:

Gl. thousand.N lamp.N GEN one with burn.VR rise.INF ASS

EGl … with a thousand lamps lighting up …

E307 पुरस्कार कैटरीना की झोली में आ गिरने को बेकरार है

IPA pʊrəska:r kɛ:ʈri:na: ki: ʄʰo:li: mẽ: a: gɪrne: ko: be:kra:r hɛ:

Gl. prize.N Katrina GEN bag LOC come.VR fall.INF ACC eager hɛ:.PRS.SG

EGl The prize is eager to fall into Katrina’s bag.

E308	सूरज	को	मार	गिराने	की	कहानी
IPA	su:rəj	ko:	ma:r	gıra:ne:	ki:	kəha:ni:
Gl.	Suraj.N	ACC	mar.VR	fall.INF	GEN	story.N
EGl	The story of killing down Suraj.					

Given the above examples, we conclude that template 11 can be applied to almost all of the compound verb constructions possible and a template can be drawn for each of the V2s we have listed.

For the rest of the five templates in the one word verb group i.e. templates 3, 4, 5, 6 and 7, a corresponding CV template can be formed validly. Thus corresponding with the first word verb group templates, we have the following templates in compound construction:

Template Name	**TAM Tag**
V1+V2_fut	VG.cv_fut.sim.dcl
V1+V2_pft	VG.cv_pst.sim.dcl
V1+V2_imp	VG.cv_0.0.imp
V1+V2_opt	VG.cv_0.0.opt
V1+V2_impf	VG.cv_pst.impf.cgnt
V1+V2_kar	VG.cp.pft
V1+V2_inf	VG.cv_0.0.imp
V1+V2_inf	VG.cv_inf

3.9.2.3 Abstract CV Templates

The templates we have drawn thus above are in abstract form. Note the use of V2 directly in the template. If we use a variable of V2, then its values will be all the V2s we are considering to occur in the compound verb constructions. Also note that there is a restriction among the V1s a V2 can take. If we implement these abstract forms, we will lose those restrictions and virtually

any V1 can occur with any V2. These V2s are also categorized as auxiliaries in a verb group. They are given a tag of auxiliary in annotation. If a word functions as auxiliary, we list it and put it within one of the several categories we have defined. Therefore, to get the proper template structure, we need to replace these V2 occurrences in the templates above with actual V2 variables which are word variables. This way we can restrain the templates to overshoot the possible verb groups it can apply to. However, this does increase the number of templates we will have in the CV VG templates. For example, for the template V1+V2_pft we will have a template for each of the V2s listed. That is a total of 61 templates. Similar would be the case for the following templates: V1+V2_impf; V1+V2_fut; V1+V2_imp; and V1+V2_imp. For these five abstract templates we will get a total of 61*5=305 CV templates.

This was an account of only the CV VGs corresponding to the 12 one word VG templates we had drawn above (T.Nos. 1-12). We have a total of 177 templates in the non-CV and non passive VGs. The 67 passive templates might also get CV constructions of their own. That means we have a total of 244 VG templates which can be considered for having a possible CV construction. I start first with a summary of the CV constructions possible in the non-Passive templates. We will consider the total of 177 such templates and try to fit in any of the CVs inside it and see if the template remains the same (that is the change taking place is just of an addition of a V1 and the V2 takes the place of the first word of the template, inheriting its formal morphological property). The following table is an outcome of a subjective study. This researcher checked the possibility of a CV construction based on the V1-V2 parity table given above and also checked the viability of such a construction by checking for the exact combination of the most common conjugation with such a CV on a prominent search engine. I might not be correct with the markings in the CV potentials of each of the templates. The ‘CV Potential’ column in the following table uses three notations the meanings of which are as follows:

N	=	No corresponding CV construction possible for this template.
Y	=	All of the V2s can have a corresponding CV template
R	=	Not all (i.e. restricted) V2s can have a corresponding CV template

T.No.	VG Template	TAM Tag	CV Potential	Example Sentence
1	prs_aux	VG.prs.sim.dcl	N	E30
2	pst_aux	VG.pst.sim.dcl	N	E31
3	VR_fut	VG.fut.sim.dcl	Y	E32
4	VR_pft	VG.pst.sim.dcl	Y	E33
5	VR_imp	VG.0.0.imp	Y	E34
6	VR_opt	VG.0.0.opt	Y	E35
7	VR_impf	VG.pst.impf.cgnt	Y	E36
8	VR_kar	VG.cp.pft	R	E27
9	VR_pcpl	VG.ip.impf	N	E19
10	VINF_imp	VG.0.0.imp	Y	E37
11	VINF	VG.inf	Y	E38
12	cahiye	VG.0.0.des	N	E26
13	VR_impf+prs_aux	VG.prs.impf.dcl	Y	E70
14	VR_pft+prs_aux	VG.prs.pft.dcl	Y	E71
15	VINF+prs_aux	VG.prs.sim.oblg	R	E72
16	VINF+wala+prs_aux	VG.prs.sim.apprx	R	E73
17	cahiye+pst_aux	VG.pst.0.des	N	E56
18	VR_impf+pst_aux	VG.pst.impf.dcl	Y	E74
19	VR_pft+pst_aux	VG.pst.pft.dcl	Y	E75
20	VINF+lag_pft	VG.pst.pft_ince.dcl	R	E43
21	VR_impf+rah_pft	VG.pst.impf_pft_cnt.dcl	Y	E44
22	VR_impf+ja_pft	VG.pst.pft_dur.dcl	R	E45
23	VR_pft+kar_impf	VG.pst.pft_frq.dcl	R	E46

24	VR_impf+ho_impf	VG.pst.impf.cfct	R	E76
25	VR+sak_pft	VG.pst.pft.abil	R	E77
26	VR+pa_pft	VG.pst.pft.abil	R	E78
27	VINF+de_pft	VG.pst.pft.perm	R	E79
28	VINF+par_pft	VG.pst.pft.oblg	R	E80
29	VINF+ho_impf	VG.pst.impf.oblg	R	E81
30	VINF+pst_aux	VG.pst.sim.oblg	R	E82
31	VR_impf+ban_pft	VG.pst.pft.abil	R	E83
32	VR_pft+rah_pft	VG.pst.pft_cnt.dcl	R	E52
33	VINF+wala+pst_aux	VG.pst.sim.apprx	R	E84
34	VINF+de_impf	VG.pst.impf.perm	R	E85
35	VINF+pa_pft	VG.pst.pft.abil	Y	E86
36	VR_impf+cal_pft	VG.pst.pft.dcl	R	E87
37	VR_impf+rah_impf	VG.pst.dur.cgnt	N	E88
38	VR_pft+ho_impf	VG.pst.impf.cgnt	Y	E89
39	VR_pft+rakh_pft	VG.pst.pft.dcl	N	E90
40	VR_pft+ho_fut	VG.fut.pft.dcl	Y	E91
41	VINF+lag_fut	VG.fut.sim_ince.dcl	R	E92
42	VR_impf+rah_fut	VG.fut.cnt.dcl	R	E93
43	VR_impf+ja_fut	VG.fut.dur.dcl	R	E94
44	VR_pft+kar_fut	VG.fut.impf_frq.dcl	Y	E95
45	VR_impf+ho_fut	VG.fut.impf.prsm	Y	E96
46	VR_impf+ho_prsm_aux	VG.fut.impf.cgnt	Y	E97
47	VR+sak_fut	VG.fut.sim.abil	R	E98

48	VR+pa_fut	VG.fut.sim.abil	R	E99
49	VINF+de_fut	VG.fut.sim.perm	R	E100
50	VINF+par_fut	VG.fut.sim.oblg	Y	E101
51	VINF+ho_fut	VG.fut.sim.oblg	R	E102
52	VR_pft+rah_fut	VG.fut.pft_cnt.dcl	R	E103
53	VINF+wala+ho_fut	VG.fut.sim.apprx	Y	E104
54	VR_impf+VR_opt	VG.0.0.opt	Y	E105
55	VR_pft+rah_opt	VG.0.pft_cnt.opt	R	E106
56	VR_pft+rah_imp	VG.0.pft_cnt.imp	R	E107
57	VR+sak_imp	VG.0.0.abil	R	E108
58	VR_pft+kar_imp	VG.0.0.imp	Y	E109
59	VINF+de_imp	VG.0.0.perm_imp	Y	E110
60	VINF+de_opt	VG.0.0.perm_opt	Y	E111
61	VINF+cahiye	VG.0.0.sugg	Y	E112
62	VINF+de_inf	VG.0.0.perm_imp	Y	E113
63	VR_impf+cal_imp	VG.0.0.imp	Y	E114
64	VR_impf+cal_opt	VG.0.0.opt	Y	E115
65	VR_impf+ho_opt	VG.0.0.cfct	Y	E116
66	VR_impf+rah_opt	VG.0.0.opt	Y	E117
67	VR_pft+ho_opt	VG.0.0.cfct	Y	E118
68	VR_pft+VINF_imp	VG.0.0.imp	N	E119
69	VINF+wala	VG.gp	Y	E120
70	VR+kar	VG.cp.pft	N	E121
71	VR_impf+ho_pft	VG.ip.impf	N	E122

72	VR_pft+ho_pft	VG.ip.pft	R	E123
73	VR_pft+VINF	VG.nv.pft	R	E124
74	VR+rah_pft	VG.ip.prog	R	E125
75	VR_impf+ja_impf	VG.ip.impf.dur	R	E126
76	VR_impf+ban_pft+prs_aux	VG.prs.impf.abil	N	E127
77	VR+cuk_pft+prs_aux	VG.prs.pft.dcl	N	E128
78	VR+rah_pft+prs_aux	VG.prs.prog.dcl	R	E129
79	VINF+lag_impf+prs_aux	VG.prs.impf_ince.dcl	R	E130
80	VINF+lag_pft+prs_aux	VG.prs.pft_ince.dcl	R	E131
81	VR_impf+rah_impf+prs_aux	VG.prs.impf_cnt.dcl	R	E132
82	VR_impf+rah_pft+prs_aux	VG.prs.impf_pft_cnt.dcl	R	E133
83	VR_pft+rah_impf+prs_aux	VG.prs.pft_impf_cnt.dcl	N	E134
84	VR_impf+ja_impf+prs_aux	VG.prs.impf_dur.dcl	R	E51
85	VR_impf+ja_pft+prs_aux	VG.prs.pft_dur.dcl	R	E135
86	VR_pft+kar_impf+prs_aux	VG.prs.impf_frq.dcl	R	E136
87	VR+sak_impf+prs_aux	VG.prs.impf.abil	R	E137
88	VR+sak_pft+prs_aux	VG.prs.pft.abil	R	E138
89	VR+pa_impf+prs_aux	VG.prs.impf.abil	R	E139
90	VR+pa_pft+prs_aux	VG.prs.pft.abil	R	E140
91	VINF+de_impf+prs_aux	VG.prs.impf.perm	R	E141
92	VINF+de_pft+prs_aux	VG.prs.pft.perm	R	E142
93	VINF+par_impf+prs_aux	VG.prs.impf.oblg	R	E143
94	VINF+par_pft+prs_aux	VG.prs.pft.oblg	R	E144
95	VINF+ho_impf+prs_aux	VG.prs.impf.oblg	R	E145

96	VR_impf+a_pft+prs_aux	VG.prs.impf_dur.dcl	R	E146
97	VR_impf+ho_inf+cahiye	VG.prs.sim.sugg	R	E147
98	VR_pft+ho_impf+prs_aux	VG.prs.impf.dcl	R	E148
99	VR_pft+ho_pft+prs_aux	VG.prs.pft.dcl	R	E149
100	VR_pft+par_pft+prs_aux	VG.prs.pft.dcl	N	E150
101	VR_pft+rakh_impf+prs_aux	VG.prs.pft_dur.dcl	N	E151
102	VR_impf+ban_pft+pst_aux	VG.pst.impf.abil	N	E152
103	VR+cuk_pft+pst_aux	VG.pst.pft.dcl	N	E153
104	VR+rah_pft+pst_aux	VG.pst.prog.dcl	N	E41
105	VINF+lag_impf+pst_aux	VG.pst.impf_ince.dcl	R	E50
106	VINF+lag_pft+pst_aux	VG.pst.pft_ince.dcl	R	E154
107	VR_impf+rah_impf+pst_aux	VG.pst.impf_cnt.dcl	R	E155
108	VR_impf+rah_pft+pst_aux	VG.pst.impf_pft_cnt.dcl	R	E53
109	VR_pft+rah_impf+pst_aux	VG.pst.pft_impf_cnt.dcl	R	E156
110	VR_impf+ja_impf+pst_aux	VG.pst.impf_dur.dcl	R	E157
111	VR_impf+ja_pft+pst_aux	VG.pst.pft_dur.dcl	R	E158
112	VR_pft+kar_impf+pst_aux	VG.pst.impf_frq.dcl	R	E159
113	VR+sak_impf+pst_aux	VG.pst.impf.abil	N	E160
114	VR+sak_pft+pst_aux	VG.pst.pft.abil	N	E161
115	VR+pa_impf+pst_aux	VG.pst.impf.abil	N	E162
116	VR+pa_pft+pst_aux	VG.pst.pft.abil	N	E163
117	VINF+de_impf+pst_aux	VG.pst.impf.perm	R	E164
118	VINF+de_pft+pst_aux	VG.pst.pft.perm	R	E165
119	VINF+par_impf+pst_aux	VG.pst.impf.oblg	R	E166

120	VINF+par_pft+pst_aux	VG.pst.pft.oblg	R	E167
121	VINF+ho_impf+pst_aux	VG.pst.impf.oblg	R	E168
122	VINF+de_pft+ho_impf	VG.pst.pft.perm_cfct	R	E169
123	VINF+cahiye+pst_aux	VG.pst.0.sugg	Y	E170
124	khada+rah+ja_pft	VG.pst.dur.dcl	N	E171
125	VR_impf+cal_pft+ja_pft	VG.pst.pft_cnt.dcl	R	E172
126	VR_impf+rah+ja_pft	VG.pst.pft_dur.dcl	N	E173
127	VR_pft+ho_pft+pst_aux	VG.pst.pft.dcl	R	E174
128	VR_pft+par_pft+pst_aux	VG.pst.pft.dcl	R	E175
129	VR_impf+ban_fut	VG.fut.impf.abil	N	E176
130	VR+cuk_pft+ho_fut	VG.fut.pft.dcl	R	E177
131	VR+rah_pft+ho_fut	VG.fut.prog.prsm	R	E178
132	VR_pft+kar_impf+ho_fut	VG.fut.impf_dur.dcl	R	E179
133	VR+sak_pft+ho_fut	VG.fut.pft.abil	R	E180
134	VR+pa_pft+ho_fut	VG.fut.pft.abil	R	E181
135	VINF+de_pft+ho_fut	VG.fut.pft.perm_cfct	Y	E182
136	VINF+par_pft+ho_fut	VG.fut.pft.oblg	Y	E183
137	VINF+ho_impf+ho_fut	VG.fut.pft.oblg	R	E184
138	VINF+rah_pft+ho_fut	VG.fut.prog.oblg_prsm	R	E185
139	VINF+de_impf+ho_fut	VG.fut.impf.perm_cfct	R	E186
140	VINF+lag_pft+ho_fut	VG.fut.pft_ince.prsm	R	E187
141	VINF+lag_impf+ho_fut	VG.fut.impf_ince.prsm	R	E188
142	VR_impf+rah_pft+ho_fut	VG.fut.impf_pft.prsm	R	E189
143	VR+rah_pft+ho_cfct	VG.0.prog.cfct	R	E190

144	VINF+de_inf+cahiye	VG.0.0.perm_sugg	Y	E191
145	VR_impf+ja+rah_pft	VG.ip.prog.dur	R	E192
146	VR+ja+rah_pft	VG.ip.prog	N	E193
147	VR_impf+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl	R	E194
148	VR+sak+rah_pft+prs_aux	VG.prs.prog.abil	R	E195
149	VR+pa+rah_pft+prs_aux	VG.prs.prog.abil	R	E196
150	VINF+de+rah_pft+prs_aux	VG.prs.prog.perm	R	E197
151	VINF+par+rah_pft+prs_aux	VG.prs.prog.oblg	R	E198
152	VINF+par+sak_impf+prs_aux	VG.prs.impf.abil	R	E199
153	VR_impf+a+rah_pft+prs_aux	VG.prs.dur.dcl	R	E200
154	VR_pft+a+rah_pft+prs_aux	VG.prs.dur.dcl	R	E201
155	VR_pft+rah+sak_impf+prs_aux	VG.prs.pft_dur.abil	N	E202
156	VR_impf+ja+rah_pft+pst_aux	VG.pst.prog_dur.dcl	R	E203
157	VR+sak+rah_pft+pst_aux	VG.pst.prog.abil	N	E204
158	VR+pa+rah_pft+pst_aux	VG.pst.prog.abil	N	E205
159	VINF+de+rah_pft+pst_aux	VG.pst.prog.perm	R	E206
160	VINF+par+rah_pft+pst_aux	VG.pst.prog.oblg	R	E207
161	VINF+par+sak_impf+pst_aux	VG.pst.impf.abil	R	E208
162	VR_impf+a+rah_pft+pst_aux	VG.pst.dur.dcl	R	E209
163	VR_pft+a+rah_pft+pst_aux	VG.pst.dur.dcl	R	E210
164	VR_pft+rah+sak_impf+pst_aux	VG.pst.pft_dur.abil	N	E211
165	VR_impf+ja+rah_pft+ho_fut	VG.fut.impf_dur.prsm	R	E212
166	VR+sak+rah_pft+ho_fut	VG.fut.prog.abil_prsm	N	E213
167	VR+pa+rah_pft+ho_fut	VG.fut.prog.abil_prsm	N	E214

168	VINF+de+rah_pft+ho_fut	VG.fut.prog.perm_prsm	R	E215
169	VINF+par+rah_pft+ho_fut	VG.fut.prog.oblg_prsm	R	E216
170	VR_impf+a+rah_pft+ho_fut	VG.fut.dur.dcl	R	E217
171	VR_pft+a+rah_pft+ho_fut	VG.fut.dur.dcl	R	E218
172	VR_impf+cal_pft+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl	N	E219
173	VR_impf+cal_pft+a+rah_pft+prs_aux	VG.prs.prog_dur.dcl	N	E220
174	VR_impf+cal_pft+ja+rah_pft+pst_aux	VG.pst.prog_dur.dcl	N	E221
175	VR_impf+cal_pft+a+rah_pft+pst_aux	VG.pst.prog_dur.dcl	N	E222
176	VR_impf+cal_pft+a+rah_pft+ho_fut	VG.fut.impf.cfct	N	E223
177	VR_impf+cal_pft+ja+rah_pft+ho_fut	VG.fut.impf.cfct	N	E224
178	VR_pft+ja_pft	VG.pas_pst.pft.dcl	R	E225
179	VR_pft+ja_impf	VG.pas_pst.impf.cgnt	R	E226
180	VR_pft+ja_fut	VG.pas_fut.sim.dcl	Y	E227
181	VR_pft+ja_opt	VG.pas_0.0.opt	R	E228
182	VR_pft+ja_inf	VG.pas_nv.pft	R	E229
183	VR_pft+ja_inf+wala	VG.pas_gp	N	E230
184	VR_pft+ja_impf+prs_aux	VG.pas_prs.impf.dcl	R	E231
185	VR_pft+ja_pft+prs_aux	VG.pas_prs.pft.dcl	R	E232
186	VR_pft+ja+rah_pft+prs_aux	VG.pas_prs.prog.dcl	R	E233
187	VR_pft+ja_inf+prs_aux	VG.pas_prs.sim.oblg	R	E234
188	VR_pft+ja_inf+wala+prs_aux	VG.pas_prs.sim.apprx	R	E235
189	VR_pft+ja_impf+pst_aux	VG.pas_pst.impf.dcl	R	E236
190	VR_pft+ja_pft+pst_aux	VG.pas_pst.pft.dcl	R	E237
191	VR_pft+ja+rah_pft+pst_aux	VG.pas_pst.prog.dcl	R	E238

192	VR_pft+ja_inf+lag_pft	VG.pas_pst.pft_ince.dcl	R	E239
193	VR_pft+ja_inf+lag_fut	VG.pas_pst.pft_ince.dcl	R	E240
194	VR_pft+ja_impf+rah_pft	VG.pas_pst.pft_cnt.dcl	R	E241
195	VR_pft+ja_impf+rah_fut	VG.pas_pst.pft_cnt.dcl	R	E242
196	VR_pft+ja_pft+ho_impf	VG.pas_pst.impf.cfct	R	E243
197	VR_pft+ja+sak_pft	VG.pas_pst.pft.abil	R	E244
198	VR_pft+ja_inf+ho_impf	VG.pas_pst.impf.oblg	R	E245
199	VR_pft+ja_inf+pst_aux	VG.pas_pst.sim.oblg	R	E246
200	VR_pft+ho_impf+pst_aux	VG.pas_pst.pft_impf_cnt.dcl	R	E247
201	VR_pft+ja_inf+wala+pst_aux	VG.pas_pst.sim.apprx	R	E248
202	VR_pft+ja+sak_impf	VG.pas_pst.impf.cgnt	R	E249
203	VR_pft+ja_pft+ho_fut	VG.pas_fut.pft.dcl	R	E250
204	VR_pft+ja+rah_pft+ho_fut	VG.pas_fut.prog.prsm	R	E251
205	VR_pft+ja_impf+ho_fut	VG.pas_fut.impf.prsm	R	E252
206	VR_pft+ja+sak_fut	VG.pas_fut.sim.abil	R	E253
207	VR_pft+ja_impf+ho_opt	VG.pas_fut.impf.cgnt	R	E254
208	VR_pft+ja_inf+ho_fut	VG.pas_fut.sim.oblg	R	E255
209	VR_pft+ho_impf+ho_fut	VG.pas_fut.pft_cnt.dcl	R	E256
210	VR_pft+ja_inf+wala+ho_fut	VG.pas_fut.sim.apprx	R	E257
211	VINF+de_pft+ja_fut	VG.pas_fut.sim.perm	R	E258
212	VR_pft+ja_impf+rah_opt	VG.pas_0.0.opt	R	E259
213	VR_pft+ja_pft+ho_opt	VG.pas_0.pft.cfct	R	E260
214	VR_pft+ja_inf+cahiye	VG.pas_0.0.sugg	R	E261
215	VR_pft+ja_inf+ho_opt	VG.pas_0.0.opt	R	E262

216	VR_pft+ja+sak_imp	VG.pas_0.0.opt_prob	R	E263
217	VR_pft+ja+sak_inf	VG.pas_0.0.inf	R	E264
218	VR_pft+ja+rah_pft	VG.pas_ip.dur	R	E265
219	VR_pft+ja+cuk_pft+prs_aux	VG.pas_prs.pft.dcl	R	E266
220	VR_pft+ja_inf+lag_impf+prs_aux	VG.pas_prs.impf_ince.dcl	R	E267
221	VR_pft+ja_inf+lag_pft+prs_aux	VG.pas_prs.pft_ince.dcl	R	E268
222	VR_pft+ja_impf+rah_pft+prs_aux	VG.pas_prs.impf_cnt.dcl	R	E269
223	VR_impf+ja_impf+rah_pft+prs_aux	VG.pas_prs.pft_dur.dcl	N	E270
224	VR_pft+ja+sak_impf+prs_aux	VG.pas_prs.impf.abil	R	E271
225	VR_pft+ja+sak_pft+prs_aux	VG.pas_prs.pft.abil	R	E272
226	VR_pft+ja_pft+ho_impf+prs_aux	VG.pas_prs.pft_impf_cnt.dcl	R	E273
227	VR_pft+ja_impf+rah_impf+prs_aux	VG.pas_prs.pft_cnt.dcl	R	E274
228	VR_pft+ja_impf+rah_impf+pst_aux	VG.pas_pst.pft_cnt.dcl	R	E275
229	VR_pft+ja+cuk_pft+pst_aux	VG.pas_pst.pft.dcl	R	E276
230	VR_pft+ja_inf+lag_impf+pst_aux	VG.pas_pst.pft_ince.dcl	R	E277
231	VR_pft+ja_inf+lag_pft+pst_aux	VG.pas_pst.impf_ince.dcl	R	E278
232	VR_pft+ja_impf+rah_pft+pst_aux	VG.pas_pst.impf_cnt.dcl	R	E279
233	VR_pft+ja+sak_impf+pst_aux	VG.pas_pst.impf.abil	R	E280
234	VR_pft+ja+sak_pft+pst_aux	VG.pas_pst.pft.abil	R	E281
235	VR_pft+ja_inf+cahiye+pst_aux	VG.pas_pst.0.sugg	R	E282
236	VR_pft+ja_pft+ho_impf+pst_aux	VG.pas_pst.pft_impf_cnt.dcl	R	E283
237	VR_pft+ja+cuk_pft+ho_fut	VG.pas_fut.pft.dcl	R	E284
238	VR_pft+ja+sak_pft+ho_fut	VG.pas_fut.pft.abil	R	E285
239	VR_pft+ja_inf+lag_pft+ho_fut	VG.pas_fut.pft_ince.prsm	R	E286

240	VR_pft+ja_inf+lag_impf+ho_fut	VG.pas_fut.impf_ince.prsm	R	E287
241	VR_pft+ja_impf+rah_pft+ho_fut	VG.pas_fut.impf_pft.prsm	R	E288
242	VR_pft+ja_pft+ho_impf+ho_fut	VG.pas_fut.pft_cnt.prsm	R	E289
243	VR_pft+ja+rah_pft+ho_opt	VG.pas_0.prog.cfct	R	E290
244	VR_pft+ja+sak_pft+ho_opt	VG.pas_0.0.cfct	R	E291

Table 37: Potential CV templates

3.9.2.4 Abstract Universal CV Templates

As seen in the table above, there is a total of 37 templates that can have a corresponding CV templates in all the V2s listed. These templates are 3, 4, 5, 6, 7, 10, 11, 13, 14, 18, 19, 21, 35, 38, 40, 44, 45, 46, 50, 53, 54, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 69, 123, 135, 136, 144 and 180. Examples of 3, 4, 5, 6, 7 and 10 we have already seen above in this section. For the other templates, they are verb groups having three or more words in the corresponding VG. Let's take template 13 and 14 for example. The non-CV templates for these two templates are as follows: "VR_impf+prs_aux" and "VR_pft+prs_aux".

Their corresponding CV template will be as follows:

VR+V2_impf+prs_aux

VR+V2_pft+prs_aux

Some of the actual templates realizing out of these abstract templates would be as follows:

Actual Template	**Example Sentence**
VR+baith_impf+prs_aux	वह उठ बैठता है (E374)
VR+de_impf+prs_aux	वह चल देता है (E375)

VR+bhag_pft+prs_aux वह उसे ले भागा है (E376)

VR+dal_pft+prs_aux उसने इसका चटनी बना डाला है (E377)

Other constructions for these can be made for each of the V2s. Below I provide one example of each of the abstract CV templates having CV construction in all the V2s (i.e. having the value 'Y' in the 'CV Potential' column in the table above). I will refer to these 37 templates as 'abstract universal CV templates'.

ATNo.	Abstract CV Template	TAM Tag	CV Potential	Example
A3	V1+V2_fut	VG.cv_fut.sim.dcl	Y	E309
A4	V1+V2_pft	VG.cv_pst.sim.dcl	Y	E310
A5	V1+V2_imp	VG.cv_0.0.imp	Y	E311
A6	V1+V2_opt	VG.cv_0.0.opt	Y	E312
A7	V1+V2_impf	VG.cv_pst.impf.cgnt	Y	E313
A10	V1+V2_inf_imp	VG.cv_0.0.imp	Y	E314
A11	V1+V2_inf	VG.cv_inf	Y	E315
A13	V1+V2_impf+prs_aux	VG.cv_prs.impf.dcl	Y	E316
A14	V1+V2_pft+prs_aux	VG.cv_prs.pft.dcl	Y	E69
A18	V1+V2_impf+pst_aux	VG.cv_pst.impf.dcl	Y	E317
A19	V1+V2_pft+pst_aux	VG.cv_pst.pft.dcl	Y	E318
A21	V1+V2_impf+rah_pft	VG.cv_pst.impf_pft_cnt.dcl	Y	E373
A35	V1+V2_inf+pa_pft	VG.cv_pst.pft.abil	Y	E319
A38	V1+V2_pft+ho_impf	VG.cv_pst.impf.cgnt	Y	E320
A40	V1+V2_pft+ho_fut	VG.cv_fut.pft.dcl	Y	E321
A44	V1+V2_pft+kar_fut	VG.cv_fut.impf_frq.dcl	Y	E322

A45	V1+V2_impf+ho_fut	VG.cv_fut.impf.prsm	Y	E323
A46	V1+V2_impf+ho_prsm_aux	VG.cv_fut.impf.cgnt	Y	E324
A50	V1+V2_inf+par_fut	VG.cv_fut.sim.oblg	Y	E325
A53	V1+V2_inf+wala+ho_fut	VG.cv_fut.sim.apprx	Y	E326
A54	V1+V2_impf+VR_opt	VG.cv_0.0.opt	Y	E327
A58	V1+V2_pft+VR_imp	VG.cv_0.0.imp	Y	E328
A59	V1+V2_inf+de_imp	VG.cv_0.0.perm_imp	Y	E329
A60	V1+V2_inf+de_opt	VG.cv_0.0.perm_opt	Y	E330
A61	V1+V2_inf+cahiye	VG.cv_0.0.sugg	Y	E331
A62	V1+V2_inf+de_inf	VG.cv_0.0.perm_imp	Y	E332
A63	V1+V2_impf+cal_imp	VG.cv_0.0.imp	Y	E333
A64	V1+V2_impf+cal_opt	VG.cv_0.0.opt	Y	E334
A65	V1+V2_impf+ho_opt	VG.cv_0.0.cfct	Y	E335
A66	V1+V2_impf+rah_opt	VG.cv_0.0.opt	Y	E336
A67	V1+V2_pft+ho_opt	VG.cv_0.0.cfct	Y	E337
A69	V1+V2_inf+wala	VG.cv_gp	Y	E338
A123	V1+V2_inf+cahiye+pst_aux	VG.cv_pst.0.sugg	Y	E339
A135	V1+V2_inf+de_pft+ho_fut	VG.cv_fut.pft.perm_cfct	Y	E340
A136	V1+V2_inf+par_pft+ho_fut	VG.cv_fut.pft.oblg	Y	E341
A144	V1+V2_inf+de_inf+cahiye	VG.cv_0.0.perm_sugg	Y	E342
A180	V1+V2_pft+ja_fut	VG.cv_pas_fut.sim.dcl	Y	E343

Table 38: Abstract Universal CV Templates

Given these abstract CV templates, we can possibly have a total of 37*61= 2257 templates for the VGs having a CV in it. And there is more: the templates allowing a CV construction only with a few select V2s. For the templates having the CV potentiality of ‘R’ in Table 37, we have to identify the V2s which allow it and then draw templates for them.

3.9.2.5 Templates having CV Potentiality of 'R'

There are a total of 167 templates with the CV potentiality of 'R'. These are templates where one or more of the V2s can constitute a verb group. However, not all of the V2s can constitute them as it will create a semantic anomaly.

To ease out the discussion, we first provide an example of some of the templates in their CV avatar:

ATNo.	Abstract CV Template	TAM Tag	Example Sentence
A79	V1+V2+rah_pft+prs_aux	VG.cv_prs.prog.dcl	E344
A94	V1+V2_inf+par_impf+prs_aux	VG.cv_prs.impf.oblg	E345
A81	V1+V2_inf+lag_pft+prs_aux	VG.cv_prs.pft_ince.dcl	E350a

The three examples for the three templates above might appear a bit odd to many native speakers. It is also very uncommon to find examples of such constructions in written Hindi. However, in colloquial Hindi, structures like these are used to some extent. These templates can apply with most of the V2s. However, there are a few V2s which cannot fit with these templates. For example, the V2 चुक /cʊk/ cannot have any of these three templates:

*E346	वह	लिख	चुकता	रहा	है
IPA	wəh	lɪkʰ	cʊk-ta:	rəh-a:	hɛ:
Gl.	3SG	write.VR	finish-IMPF.MSG	live-PFT.MSG	be.PRS.SG
EGl	He has kept on completing the writing.				

*E347	उसे	लिख	चुकना	पड़ता	है
IPA	ʊse:	lɪkʰ	cʊkna:	pəɽ-ta:	hɛ:

Gl. 3SG.OBL.ACC write.VR finish.INF lay-IMPF.MSG be.PRS.SG

EGl He has to complete the writing.

*E348	वह	लिख	चुकने	लगा	है
IPA	wəh	lɪkʰ	cʊkne:	ləg-a:	hɛ:
Gl.	3SG	write.VR	finish.INF	begin-PFT.MSG	be.PRS.SG

EGl He has started completing the writing.

There is another confusion with such constructions. For example if we compare the following sentence:

E350a	वह	काट	खाने	लगा	है
IPA	wəh	ka:ʈ	kʰa:ne:	ləg-a:	hɛ:
Gl.	3SG	cut	eat.INF	begin-PFT.PL	be.PRS.SG

EGl He has started cutting and eating.

or

E351a	जंगली	घास	उखाड़	फेंकने	पड़ते	हैं
IPA	jə̃gli:	gʰa:s	ʊkʰa:ɖ	pʰẽ:kne:	pəɖ-te:	hẽ:
Gl.	wild	grass	uproot.VR	throw.INF	lay-IMPF.PL	be.PRS.PL

EGl The wild grass has to be thrown away.

with the sentences below respectively:

E350	वह	काट	कर	खाने	लगा	है

IPA	wəh	ka:ʈ	kər	kʰa:ne:	ləga:		hɛ:
Gl.	3SG	cut	do.VR	eat.INF	begin-PFT.PL		be.PRS.SG
EGl	He has started cutting and eating.						

or

E351	जंगली	घास	उखाड़	कर	फेंकने	पड़ते	हैं
IPA	jə̃gli:	gʰa:s	ʊkʰa:ɖ	kər	pʰẽ:kne:	pəɖ-te:	hɛ̃:
Gl.	wild	grass	uproot.VR	do.VR	throw.INF	lay-IMPF.PL	be.PRS.PL
EGl	The wild grass has to be thrown away.						

there is a probability that the meaning expressed is the same. While in the previous two examples it appears as a single verb group, in the latter corresponding examples it is clear that there are two verb groups in them. Both the uses are valid. But when annotating such constructions for their PoS, it has to be considered that they are annotated properly. In the previous examples, the main verb is /ka:ʈ/ and /ʊkʰa:ɖ/ respectively while in the latter examples, the main verb is /kʰa:ne:/ and /pʰẽ:kne:/.

Which abstract template will be universal and which will be restricted is always a fuzzy question to answer. This is so because of the nature of the compound verb in Hindi. One structure might appear as perfectly fine at one time while the same structure might seem odd at others. The native speaker judgement is also often confused in such cases. Structurally it is possible to have all the V2s have a realization of the abstract templates. However, there are some non-CV templates that will not have any corresponding CV template. The table below is an attempt to capture as many CV templates as possible. It is based on all the non-CV templates we have drawn. We found 40 templates out of the 244 non-CV templates that will not have a corresponding CV template. These include the following Template Nos.:

1, 2, 9, 12, 17, 37, 39, 68, 70, 71, 76, 77, 83, 100, 101, 102, 103, 104, 113, 114, 115, 116, 124, 126, 129, 146, 155, 157, 158, 164, 166, 167, 172, 173, 174, 175, 176, 177, 183 and 223.

The other templates along with their abstract TAM template and corresponding VG Tags are given in the following table.

AT No.	Abstract CV Template	TAM Tag	CV Potential
A3	V1+V2_fut	VG.cv_fut.sim.dcl	Y
A4	V1+V2_pft	VG.cv_pst.sim.dcl	Y
A5	V1+V2_imp	VG.cv_0.0.imp	Y
A6	V1+V2_opt	VG.cv_0.0.opt	Y
A7	V1+V2_impf	VG.cv_pst.impf.cgnt	Y
A8	V1+V2_kar	VG.cv_cp.pft	R
A10	V1+V2_inf_imp	VG.cv_0.0.imp	Y
A11	V1+V2_inf	VG.cv_nv	Y
A13	V1+V2_impf+prs_aux	VG.cv_prs.impf.dcl	Y
A14	V1+V2_pft+prs_aux	VG.cv_prs.pft.dcl	Y
A15	V1+V2_inf+prs_aux	VG.cv_prs.sim.oblg	R
A16	V1+V2_inf+wala+prs_aux	VG.cv_prs.sim.apprx	R
A18	V1+V2_impf+pst_aux	VG.cv_pst.impf.dcl	Y
A19	V1+V2_pft+pst_aux	VG.cv_pst.pft.dcl	Y
A20	V1+V2_inf+lag_pft	VG.cv_pst.sim_ince.dcl	R
A21	V1+V2_impf+rah_pft	VG.cv_pst.impf_sim_cnt.dcl	Y
A22	V1+V2_impf+ja_pft	VG.cv_pst.sim_dur.dcl	R
A23	V1+V2_pft+kar_impf	VG.cv_pst.pft_frq.dcl	R
A24	V1+V2_impf+ho_impf	VG.cv_pst.impf.cfct	R
A25	V1+V2+sak_pft	VG.cv_pst.pft.abil	R
A26	V1+V2+pa_pft	VG.cv_pst.pft.abil	R
A27	V1+V2_inf+de_pft	VG.cv_pst.pft.perm	R
A28	V1+V2_inf+par_pft	VG.cv_pst.sim.oblg	R
A29	V1+V2_inf+ho_impf	VG.cv_pst.impf.oblg	R
A30	V1+V2_inf+pst_aux	VG.cv_pst.sim.oblg	R
A31	V1+V2_impf+ban_pft	VG.cv_pst.pft.abil	R
A32	V1+V2_pft+rah_pft	VG.cv_pst.pft_cnt.dcl	R
A33	V1+V2_inf+wala+pst_aux	VG.cv_pst.sim.apprx	R
A34	V1+V2_inf+de_impf	VG.cv_pst.impf.perm	R
A35	V1+V2_inf+pa_pft	VG.cv_pst.pft.abil	Y
A36	V1+V2_impf+cal_pft	VG.cv_pst.pft.dcl	R
A38	V1+V2_pft+ho_impf	VG.cv_pst.impf.cgnt	Y
A40	V1+V2_pft+ho_fut	VG.cv_fut.pft.dcl	Y
A41	V1+V2_inf+lag_fut	VG.cv_fut.sim_ince.dcl	R
A42	V1+V2_impf+rah_fut	VG.cv_fut.cnt.dcl	R
A43	V1+V2_impf+ja_fut	VG.cv_fut.dur.dcl	R
A44	V1+V2_pft+kar_fut	VG.cv_fut.impf_frq.dcl	Y
A45	V1+V2_impf+ho_fut	VG.cv_fut.impf.prsm	Y

A46	V1+V2_impf+ho_prsm_aux	VG.cv_fut.impf.cgnt	Y
A47	V1+V2+sak_fut	VG.cv_fut.sim.abil	R
A48	V1+V2+pa_fut	VG.cv_fut.sim.abil	R
A49	V1+V2_inf+de_fut	VG.cv_fut.sim.perm	R
A50	V1+V2_inf+par_fut	VG.cv_fut.sim.oblg	Y
A51	V1+V2_inf+ho_fut	VG.cv_fut.sim.oblg	R
A52	V1+V2_pft+rah_fut	VG.cv_fut.pft_cnt.dcl	R
A53	V1+V2_inf+wala+ho_fut	VG.cv_fut.sim.apprx	Y
A54	V1+V2_impf+VR_opt	VG.cv_0.0.opt	Y
A55	V1+V2_pft+rah_opt	VG.cv_0.pft_cnt.opt	R
A56	V1+V2_pft+rah_imp	VG.cv_0.pft_cnt.imp	R
A57	V1+V2+sak_imp	VG.cv_0.0.abil	R
A58	V1+V2_pft+VR_imp	VG.cv_0.0.imp	Y
A59	V1+V2_inf+de_imp	VG.cv_0.0.perm_imp	Y
A60	V1+V2_inf+de_opt	VG.cv_0.0.perm_opt	Y
A61	V1+V2_inf+cahiye	VG.cv_0.0.sugg	Y
A62	V1+V2_inf+de_inf	VG.cv_0.0.perm_imp	Y
A63	V1+V2_impf+cal_imp	VG.cv_0.0.imp	Y
A64	V1+V2_impf+cal_opt	VG.cv_0.0.opt	Y
A65	V1+V2_impf+ho_opt	VG.cv_0.0.cfct	Y
A66	V1+V2_impf+rah_opt	VG.cv_0.0.opt	Y
A67	V1+V2_pft+ho_opt	VG.cv_0.0.cfct	Y
A69	V1+V2_inf+wala	VG.cv_gp	Y
A72	V1+V2_pft+ho_pft	VG.cv_ip.pft	R
A73	V1+V2_pft+VINF	VG.cv_nv.pft	R
A74	V1+V2+rah_pft	VG.cv_ip.prog	R
A75	V1+V2_impf+ja_impf	VG.cv_ip.impf.dur	R
A76	V1+V2+rah_pft+prs_aux	VG.cv_prs.prog.dcl	R
A79	V1+V2_inf+lag_impf+prs_aux	VG.cv_prs.impf_ince.dcl	R
A80	V1+V2_inf+lag_pft+prs_aux	VG.cv_prs.pft_ince.dcl	R
A81	V1+V2_impf+rah_impf+prs_aux	VG.cv_prs.impf_cnt.dcl	R
A82	V1+V2_impf+rah_pft+prs_aux	VG.cv_prs.impf_pft_cnt.dcl	R
A84	V1+V2_impf+ja_impf+prs_aux	VG.cv_prs.impf_dur.dcl	R
A85	V1+V2_impf+ja_pft+prs_aux	VG.cv_prs.impf_pft_dur.dcl	R
A86	V1+V2_pft+kar_impf+prs_aux	VG.cv_prs.impf_frq.dcl	R
A87	V1+V2+sak_impf+prs_aux	VG.cv_prs.impf.abil	R
A88	V1+V2+sak_pft+prs_aux	VG.cv_prs.pft.abil	R
A89	V1+V2+pa_impf+prs_aux	VG.cv_prs.impf.abil	R
A90	V1+V2+pa_pft+prs_aux	VG.cv_prs.pft.abil	R
A91	V1+V2_inf+de_impf+prs_aux	VG.cv_prs.impf.perm	R
A92	V1+V2_inf+de_pft+prs_aux	VG.cv_prs.pft.perm	R
A93	V1+V2_inf+par_impf+prs_aux	VG.cv_prs.impf.oblg	R
A94	V1+V2_inf+par_pft+prs_aux	VG.cv_prs.pft.oblg	R

A95	V1+V2_inf+ho_impf+prs_aux	VG.cv_prs.impf.oblg	R
A96	V1+V2_impf+a_pft+prs_aux	VG.cv_prs.impf_dur.dcl	R
A97	V1+V2_impf+ho_inf+cahiye	VG.cv_prs.sim.sugg	R
A98	V1+V2_pft+ho_impf+prs_aux	VG.cv_prs.impf.dcl	R
A99	V1+V2_pft+ho_pft+prs_aux	VG.cv_prs.pft.dcl	R
A105	V1+V2_inf+lag_impf+pst_aux	VG.cv_pst.impf_ince.dcl	R
A106	V1+V2_inf+lag_pft+pst_aux	VG.cv_pst.pft_ince.dcl	R
A107	V1+V2_impf+rah_impf+pst_aux	VG.cv_pst.impf_cnt.dcl	R
A108	V1+V2_impf+rah_pft+pst_aux	VG.cv_pst.impf_pft_cnt.dcl	R
A109	V1+V2_pft+rah_impf+pst_aux	VG.cv_pst.pft_impf_cnt.dcl	R
A110	V1+V2_impf+ja_impf+pst_aux	VG.cv_pst.impf_dur.dcl	R
A111	V1+V2_impf+ja_pft+pst_aux	VG.cv_prs.impf_pft_dur.dcl	R
A112	V1+V2_pft+kar_impf+pst_aux	VG.cv_pst.impf_frq.dcl	R
A117	V1+V2_inf+de_impf+pst_aux	VG.cv_pst.impf.perm	R
A118	V1+V2_inf+de_pft+pst_aux	VG.cv_pst.sim.perm	R
A119	V1+V2_inf+par_impf+pst_aux	VG.cv_pst.impf.oblg	R
A120	V1+V2_inf+par_pft+pst_aux	VG.cv_pst.pft.oblg	R
A121	V1+V2_inf+ho_impf+pst_aux	VG.cv_pst.impf.oblg	R
A122	V1+V2_inf+de_pft+ho_impf	VG.cv_pst.pft.perm_cfct	R
A123	V1+V2_inf+cahiye+pst_aux	VG.cv_pst.0.sugg	Y
A125	V1+V2_impf+cal_pft+ja_pft	VG.cv_pst.pft_cnt.dcl	R
A127	V1+V2_pft+ho_pft+pst_aux	VG.cv_pst.pft.dcl	R
A128	V1+V2_pft+par_pft+pst_aux	VG.cv_pst.pft.dcl	R
A130	V1+V2+cuk_pft+ho_fut	VG.cv_fut.pft.dcl	R
A131	V1+V2+rah_pft+ho_fut	VG.cv_fut.prog.prsm	R
A132	V1+V2_pft+kar_impf+ho_fut	VG.cv_fut.pft_impf_dur.dcl	R
A133	V1+V2+sak_pft+ho_fut	VG.cv_fut.pft.abil	R
A134	V1+V2+pa_pft+ho_fut	VG.cv_fut.pft.abil	R
A135	V1+V2_inf+de_pft+ho_fut	VG.cv_fut.pft.perm_cfct	Y
A136	V1+V2_inf+par_pft+ho_fut	VG.cv_fut.pft.oblg	Y
A137	V1+V2_inf+ho_impf+ho_fut	VG.cv_fut.pft.oblg	R
A138	V1+V2_inf+rah_pft+ho_fut	VG.cv_fut.prog.oblg_prsm	R
A139	V1+V2_inf+de_impf+ho_fut	VG.cv_fut.impf.perm_cfct	R
A140	V1+V2_inf+lag_pft+ho_fut	VG.cv_fut.pft_ince.prsm	R
A141	V1+V2_inf+lag_impf+ho_fut	VG.cv_fut.impf_ince.prsm	R
A142	V1+V2_impf+rah_pft+ho_fut	VG.cv_fut.impf_pft.prsm	R
A143	V1+V2+rah_pft+ho_cfct	VG.cv_0.prog.cfct	R
A144	V1+V2_inf+de_inf+cahiye	VG.cv_0.0.perm_sugg	Y
A145	V1+V2_impf+ja+rah_pft	VG.cv_ip.prog.dur	R
A147	V1+V2_impf+ja+rah_pft+prs_aux	VG.cv_prs.prog_dur.dcl	R
A148	V1+V2+sak+rah_pft+prs_aux	VG.cv_prs.prog.abil	R
A149	V1+V2+pa+rah_pft+prs_aux	VG.cv_prs.prog.abil	R
A150	V1+V2_inf+de+rah_pft+prs_aux	VG.cv_prs.prog.perm	R

A151	V1+V2_inf+par+rah_pft+prs_aux	VG.cv_prs.prog.oblg	R
A152	V1+V2_inf+par+sak_impf+prs_aux	VG.cv_prs.impf.abil	R
A153	V1+V2_impf+a+rah_pft+prs_aux	VG.cv_prs.dur.dcl	R
A154	V1+V2_pft+a+rah_pft+prs_aux	VG.cv_prs.dur.dcl	R
A156	V1+V2_impf+ja+rah_pft+pst_aux	VG.cv_pst.prog_dur.dcl	R
A159	V1+V2_inf+de+rah_pft+pst_aux	VG.cv_pst.prog.perm	R
A160	V1+V2_inf+par+rah_pft+pst_aux	VG.cv_pst.prog.oblg	R
A161	V1+V2_inf+par+sak_impf+pst_aux	VG.cv_pst.impf.abil	R
A162	V1+V2_impf+a+rah_pft+pst_aux	VG.cv_pst.dur.dcl	R
A163	V1+V2_pft+a+rah_pft+pst_aux	VG.cv_pst.dur.dcl	R
A165	V1+V2_impf+ja+rah_pft+ho_fut	VG.cv_fut.impf_dur.prsm	R
A168	V1+V2_inf+de+rah_pft+ho_fut	VG.cv_fut.prog.perm_prsm	R
A169	V1+V2_inf+par+rah_pft+ho_fut	VG.cv_fut.prog.oblg_prsm	R
A170	V1+V2_impf+a+rah_pft+ho_fut	VG.cv_fut.dur.dcl	R
A171	V1+V2_pft+a+rah_pft+ho_fut	VG.cv_fut.dur.dcl	R
A178	V1+V2_pft+ja_pft	VG.cv_pas_pst.pft.dcl	R
A179	V1+V2_pft+ja_impf	VG.cv_pas_pst.impf.cgnt	R
A180	V1+V2_pft+ja_fut	VG.cv_pas_fut.sim.dcl	Y
A181	V1+V2_pft+ja_opt	VG.cv_pas_0.0.opt	R
A182	V1+V2_pft+ja_inf	VG.cv_pas_nv.pft	R
A184	V1+V2_pft+ja_impf+prs_aux	VG.cv_pas_prs.impf.dcl	R
A185	V1+V2_pft+ja_pft+prs_aux	VG.cv_pas_prs.pft.dcl	R
A186	V1+V2_pft+ja+rah_pft+prs_aux	VG.cv_pas_prs.prog.dcl	R
A187	V1+V2_pft+ja_inf+prs_aux	VG.cv_pas_prs.sim.oblg	R
A188	V1+V2_pft+ja_inf+wala+prs_aux	VG.cv_pas_prs.sim.apprx	R
A189	V1+V2_pft+ja_impf+pst_aux	VG.cv_pas_pst.impf.dcl	R
A190	V1+V2_pft+ja_pft+pst_aux	VG.cv_pas_pst.pft.dcl	R
A191	V1+V2_pft+ja+rah_pft+pst_aux	VG.cv_pas_pst.prog.dcl	R
A192	V1+V2_pft+ja_inf+lag_pft	VG.cv_pas_pst.sim_ince.dcl	R
A193	V1+V2_pft+ja_inf+lag_fut	VG.cv_pas_pst.sim_ince.dcl	R
A194	V1+V2_pft+ja_impf+rah_pft	VG.cv_pas_pst.impf_cnt.dcl	R
A195	V1+V2_pft+ja_impf+rah_fut	VG.cv_pas_pst.pft_cnt.dcl	R
A196	V1+V2_pft+ja_pft+ho_impf	VG.cv_pas_pst.impf.cfct	R
A197	V1+V2_pft+ja+sak_pft	VG.cv_pas_pst.sim.abil	R
A198	V1+V2_pft+ja_inf+ho_impf	VG.cv_pas_pst.impf.oblg	R
A199	V1+V2_pft+ja_inf+pst_aux	VG.cv_pas_pst.sim.oblg	R
A200	V1+V2_pft+ho_impf+pst_aux	VG.cv_pas_pst.pft_impf_cnt.dcl	R
A201	V1+V2_pft+ja_inf+wala+pst_aux	VG.cv_pas_pst.sim.apprx	R
A202	V1+V2_pft+ja+sak_impf	VG.cv_pas_pst.impf.cgnt	R
A203	V1+V2_pft+ja_pft+ho_fut	VG.cv_pas_fut.pft.dcl	R
A204	V1+V2_pft+ja+rah_pft+ho_fut	VG.cv_pas_fut.prog.prsm	R
A205	V1+V2_pft+ja_impf+ho_fut	VG.cv_pas_fut.impf.prsm	R
A206	V1+V2_pft+ja+sak_fut	VG.cv_pas_fut.sim.abil	R

A207	V1+V2_pft+ja_impf+ho_opt	VG.cv_pas_fut.impf.cgnt	R
A208	V1+V2_pft+ja_inf+ho_fut	VG.cv_pas_fut.sim.oblg	R
A209	V1+V2_pft+ho_impf+ho_fut	VG.cv_pas_fut.pft_cnt.dcl	R
A210	V1+V2_pft+ja_inf+wala+ho_fut	VG.cv_pas_fut.sim.apprx	R
A211	V1+V2_inf+de_pft+ja_fut	VG.cv_pas_fut.sim.perm	R
A212	V1+V2_pft+ja_impf+rah_opt	VG.cv_pas_0.0.opt	R
A213	V1+V2_pft+ja_pft+ho_opt	VG.cv_pas_0.pft.cfct	R
A214	V1+V2_pft+ja_inf+cahiye	VG.cv_pas_0.0.sugg	R
A215	V1+V2_pft+ja_inf+ho_opt	VG.cv_pas_0.0.opt	R
A216	V1+V2_pft+ja+sak_imp	VG.cv_pas_0.0.opt_prob	R
A217	V1+V2_pft+ja+sak_inf	VG.cv_pas_0.0.inf	R
A218	V1+V2_pft+ja+rah_pft	VG.cv_pas_ip.dur	R
A219	V1+V2_pft+ja+cuk_pft+prs_aux	VG.cv_pas_prs.pft.dcl	R
A220	V1+V2_pft+ja_inf+lag_impf+prs_aux	VG.cv_pas_prs.impf_ince.dcl	R
A221	V1+V2_pft+ja_inf+lag_pft+prs_aux	VG.cv_pas_prs.pft_ince.dcl	R
A222	V1+V2_pft+ja_impf+rah_pft+prs_aux	VG.cv_pas_prs.impf_cnt.dcl	R
A224	V1+V2_pft+ja+sak_impf+prs_aux	VG.cv_pas_prs.impf.abil	R
A225	V1+V2_pft+ja+sak_pft+prs_aux	VG.cv_pas_prs.pft.abil	R
A226	V1+V2_pft+ja_pft+ho_impf+prs_aux	VG.cv_pas_prs.pft_impf_cnt.dcl	R
A227	V1+V2_pft+ja_impf+rah_impf+prs_aux	VG.cv_pas_prs.pft_impf_cnt.dcl	R
A228	V1+V2_pft+ja_impf+rah_impf+pst_aux	VG.cv_pas_pst.pft_cnt.dcl	R
A229	V1+V2_pft+ja+cuk_pft+pst_aux	VG.cv_pas_pst.pft.dcl	R
A230	V1+V2_pft+ja_inf+lag_impf+pst_aux	VG.cv_pas_pst.pft_ince.dcl	R
A231	V1+V2_pft+ja_inf+lag_pft+pst_aux	VG.cv_pas_pst.impf_ince.dcl	R
A232	V1+V2_pft+ja_impf+rah_pft+pst_aux	VG.cv_pas_pst.impf_cnt.dcl	R
A233	V1+V2_pft+ja+sak_impf+pst_aux	VG.cv_pas_pst.impf.abil	R
A234	V1+V2_pft+ja+sak_pft+pst_aux	VG.cv_pas_pst.pft.abil	R
A235	V1+V2_pft+ja_inf+cahiye+pst_aux	VG.cv_pas_pst.0.sugg	R
A236	V1+V2_pft+ja_pft+ho_impf+pst_aux	VG.cv_pas_pst.pft_impf_cnt.dcl	R
A237	V1+V2_pft+ja+cuk_pft+ho_fut	VG.cv_pas_fut.pft.dcl	R
A238	V1+V2_pft+ja+sak_pft+ho_fut	VG.cv_pas_fut.pft.abil	R
A239	V1+V2_pft+ja_inf+lag_pft+ho_fut	VG.cv_pas_fut.pft_ince.prsm	R
A240	V1+V2_pft+ja_inf+lag_impf+ho_fut	VG.cv_pas_fut.impf_ince.prsm	R
A241	V1+V2_pft+ja_impf+rah_pft+ho_fut	VG.cv_pas_fut.impf_pft.prsm	R
A242	V1+V2_pft+ja_pft+ho_impf+ho_fut	VG.cv_pas_fut.pft_cnt.prsm	R
A243	V1+V2_pft+ja+rah_pft+ho_opt	VG.cv_pas_0.prog.cfct	R
A244	V1+V2_pft+ja+sak_pft+ho_opt	VG.cv_pas_0.0.cfct	R

Table 39: Abstract CV Templates

3.9.2.6 Realization of Actual CV Templates

With the exercise above, we have got a total of 204 abstract CV templates. 37 of these are templates which we say that they can get a total of 61 realizations for the 61 V2s we are considering. This, however, might not be true and there may not be enough textual evidence to prove it. However, that does not deter that it is not possible. For example the universal abstract template A21 has the following structure:

V1+V2_impf+rah_pft

While this template is fine for the following sentence:

E373	वह	किताब	पढ	डालता	रहा
IPA	wəh	kıta:b	pəɖʰ	ɖa:l-ta:	rəh-a:
Gl.	3SG	book	read.VR	put-IMPF.MSG	live.PFT.MSG
EGl	He kept on completing reading the book.				

One cannot say the following for surety even though the structure is just fine as per the rules. It is the semantics of the following sentence that is anomalous and might be summarily rejected by a native speaker.

?*E352	वह	किताब	पढ	सकता	रहा
IPA	wəh	kıta:b	pəɖʰ	sək-ta:	rəh-a:
Gl.	3SG	book	read.VR	can-IMPF.MSG	live-PFT.MSG
EGl	He went on being able to read the book.				

But we have included even these templates because it does not adversely affect the VG identification task. Even if there is one such verb group present in the text, they will be captured through these templates.

These abstract templates can be realized in two ways, depending upon how the computing is done. The default will be to let both the V1 and V2 identified based on tag variables. We can change the abstract V1 and V2 to simply VR so that template 21 would look like this:

VR+VR_impf+rah_pft

However, this approach has its own problems. There may be cases a non-CV template will be mistaken for a CV structure. For example we have a non-CV template with the following structure:

VR+rah_pft+prs_aux

which denotes the progressive aspect. We also have another CV template with the following structure:

VR+V2_pft+prs_aux

These two templates can get mixed or one template can even get ignored all the time for the other so much so that all the valid instance of one of these two templates would get marked wrongly for the other.

We can fix this by making a change in the logic of program so that if there are two matching templates vying for a one verb group, the one with greater number of word variables will be given a priority over the one having more of tag variables.

Another way out is to make the V2s of the templates have word variables. This is the way I have adopted at present. This solution does increase the number of CV templates we have by many folds though.

3.10 VERB GROUPS (CHUNK) TAGS

As stated earlier in this chapter, after we identify the verb group, we need to give it a tag as the boundary marker. A verb group functions the same way semantically as the common meaning of the verb in a sentence. The structure of the TAM Tags has been explained above in section 3.2. The template tables presented above also have their TAM tags given along. I will present only a summary of the TAM tags here.

For each of the VG templates we have drawn, there is a TAM tag. Though each of the VG templates is unique, it is not necessary that the VG tags assigned to them would also be unique. That is we have TAM tags that are shared by two or more of the templates. Among the 244 non-CV templates there are a total of 169 unique TAM tags. 125 templates share the 50 common TAM tags. That is a total of 50 TAM tags are such that they are shared by two or more TAM templates. The rest of the 119 non-CV templates have unique tags. Similarly, among the CV templates, there are a total 153 unique TAM tags for the whole of 204 CV templates. There are 89 templates sharing a total 38 TAM Tags. 115 templates have a unique tag of their own.

4 MECHANICAL UNDERSTANDING OF THE VERB SEQUENCE IN HINDI TEXTS: IMPLEMENTATION AND EVALUATION

In the previous chapter I have given an analysis of the verb group and come up with a templatic structure of all the verb groups possible in Hindi. This exercise has pointed us to some facts which might help in coming up with a good logic to build the identification program. These are as follows:

i. There cannot be a verb group without a main verb.

ii. The main verb is always the first verbal word from the left in the verb group.

iii. There cannot be a verb group consisting only of one or more auxiliaries.

iv. Verbs playing as auxiliaries are limited in number and can be exhaustively counted.

v. A verb group in Hindi must have a minimum of one verbal word and it can have a maximum of five verbal words in it.

Based on this formalization, an abstract template of the verb groups in Hindi can be drawn as follows:

VM(Aux)(Aux)(Aux)(Aux)

This template above says that the first element of a verb group should be a main verb. This main verb is mandatory in a verb group. All the other elements must be auxiliary verbs. A verb group cannot have more than four auxiliary.

In the VG templates we have not included the words that can interrupt a verb group. That is we have not discussed the elements that come in between a verb group. We have though stated

earlier that a particle can come in between a verb group. We assert it here again that apart from particles, there is no other element that can come inside a verb group.

4.1 LOGIC TO IDENTIFY THE VERB GROUPS

The goal of this research is to develop a tool that would identify the verb groups, mark the boundaries for each of the verb groups coming in a sentence and give each of them a TAM tag. I have stated earlier that the input for would be a PoS annotated text.

When we give the input of an annotated sentence/text, such as the following (E238):

किताब\NC.fem.sg.dir.0 पढ़ी\VM.fem.sg.3.pst.pft.dcl.fin.n जा\VAUX.0.0.0.0.0.0.nfn.0

रही\VAUX.fem.sg.0.0.pft.dcl.fin.n थी\VAUX.fem.sg.0.pst.sim.dcl.fin.n

The output should be either:

किताब\NC.fem.sg.dir.0 [पढ़ी\VM.fem.sg.3.pst.pft.dcl.fin.n जा\VAUX.0.0.0.0.0.0.nfn.0

रही\VAUX.fem.sg.0.0.pft.dcl.fin.n

थी\VAUX.fem.sg.0.pst.sim.dcl.fin.n]VG.pas_pst.prog.dcl

or

किताब [पढ़ी जा रही थी]VG.pas_pst.prog.dcl

The first output retains the PoS features from the input text and the second output removes them. The second format of the output generation is easier to read, understand and evaluate.

Once the program finds an input sentence it will look for the words having a VM or VAUX tag. The search starts from right to left and follows the following steps during the course of its execution.

If the first word encountered in the search is a VM, it moves on to look for the other morphological features given in the eight slots a verbal word has in the input. Reading the input means reading the tags. This warrants an understanding of the structure of the input text itself.

4.1.1 Parts of Speech features of a Verbal Word

In the IL-PoS tagset, a verb has to have information on the following features:

i. Main verb or auxiliary verb

ii. Gender Agreement Marking

iii. Number Agreement Marking

iv. Person Agreement Marking

v. Tense Marking

vi. Aspect Marking

vii. Mood Marking

viii. Finiteness Marking, and

ix. Honorificity marking

The order of features encoded for each of the verbs remains the same as stated above. That is, the first feature marked tells whether the verb is a main verb or an auxiliary verb. The second, third and fourth features encodes information about the GNP agreements marked on the verb. It is the fifth, sixth, seventh and the eigth features that would of interest to us. These have information about the tense, aspect, mood and the finiteness of the verb. There is a closed set of values these eight slots can have. These values are as presented in the table below:

Feature	Values
Verbal Category	VM, VAUX
Gender	mas, fem, neu, 0, x
Number	sg, pl, 0,x
Person	1, 2, 3, 0, x
Tense	prs, pst, fut, 0, x
Aspect	sim,impf,pft,0,x

Mood	dcl,imp,opt,hab,0,x
Finiteness	fin,nfn,ifn,0,x
Honorificity	n,y,0,x

Table 40: Tags of the verbs in the input text

Thus if there is a verb, its features will be one of these. There is no duplicate value used for any features. That is a constant value is for only one feature.

Thus, when a verbal word is encountered reading the input sentence from the right, the first thing to be checked is whether the verb is a main verb or an auxiliary. If the first verb encountered is a main verb, it must be a one word verb group. When the main verb is identified, its features are checked for the TAM information contained in the tags.

One word verb groups or for that matter any of the main verbs are identified and given a tag based on tags of the input text. This identification is based on the tag variables defined. It has been indicated in the previous chapter that we use three files that contain what we call template definitions, tag-variables and word variables. The VG templates we have drawn consist of tag variables and word variables. Thus, it would be apt to give a greater description of how we define them.

4.1.2 Template Defintions

Template definitions are simply enumeration of the VG templates along with their TAM tag. It is a file containing all the VG templates followed by the TAM tag. The delimiter is a TAB mark. It looks like as follows:

prs_aux	VG.prs.sim.dcl
pst_aux	VG.pst.sim.dcl
VR_fut	VG.fut.sim.dcl
VR_pft	VG.pst.sim.dcl
VR_imp	VG.0.0.imp
VR_opt	VG.0.0.opt
VR_impf	VG.pst.impf.cgnt
VR_kar	VG.cp.pft
VR_pcpl	VG.ip.impf
VINF_imp	VG.0.0.imp
VINF	VG.inf
cahiye	VG.0.0.des

VR_impf+prs_aux	VG.prs.impf.dcl
VR_pft+prs_aux	VG.prs.pft.dcl
VINF+prs_aux	VG.prs.sim.oblg
VINF+wala+prs_aux	VG.prs.sim.apprx
cahiye+pst_aux	VG.pst.0.des
VR_impf+ban_pft+prs_aux	VG.prs.impf.abil
VR+cuk_pft+prs_aux	VG.prs.pft.dcl
VR+rah_pft+prs_aux	VG.prs.prog.dcl
VINF+lag_impf+prs_aux	VG.prs.impf_ince.dcl
VINF+lag_pft+prs_aux	VG.prs.pft_ince.dcl
VR_impf+rah_impf+prs_aux	VG.prs.impf_cnt.dcl
VR_impf+rah_pft+prs_aux	VG.prs.impf_pft_cnt.dcl
VR_pft+rah_impf+prs_aux	VG.prs.pft_impf_cnt.dcl
VR_impf+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl
VR+sak+rah_pft+prs_aux	VG.prs.prog.abil
VR+pa+rah_pft+prs_aux	VG.prs.prog.abil
VINF+de+rah_pft+prs_aux	VG.prs.prog.perm
VINF+par+rah_pft+prs_aux	VG.prs.prog.oblg
VINF+par+sak_impf+prs_aux	VG.prs.impf.abil
VR_impf+a+rah_pft+prs_aux	VG.prs.dur.dcl
VR_pft+a+rah_pft+prs_aux	VG.prs.dur.dcl
VR_impf+cal_pft+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl
VR_impf+cal_pft+a+rah_pft+prs_aux	VG.prs.prog_dur.dcl
…	…
…	…

Table 41: Defining Templates for the VG Finder Tool

4.1.3 Tag Variables

Tag variables are all the variables used in the TAM templates. The values of these variables are to be found within the tags of the verbs. For example the tag variable of VR is to be found by checking if the tag contains the value of ‘nfn’. Similarly, the tag variable of VINF is to be found out by checking if the tag of the verb contains the value of ‘inf’. The table below contains all the tag variables and the values that it must have to get satisfied.

Meaning of the Tag Variable	Tag Variable	Required Tag Values
Main Verb	VM	VM
Auxiliarly Verb	VAUX	VAUX
V1 (of the CV, always in the root form)	V1	nfn
Infinitive Form (-nV ending verb forms)	VINF	ifn
Infinitive form of the verb having an imperative function	VINF_imp	ifn,imp

Verb Root	VR	nfn
Future form of the verb	VR_fut	fut
Imperfect form of the verb	VR_impf	impf
Optative form of the verb	VR_opt	opt
Perfect form of the Verb	VR_pft	pft
Imperative form of the verb	VR_imp	imp
Conjuctive participles	VR_kar	nfn,pft
Imperfective Participle construction of the verb (as in खाते-खाते)	VR_pcpl	impf,nfn

Table 42: Tag Variables and their Values

As is seen, there are a total of 13 tag variables we have. There are a few tag variables that must have two values to get satisfied.

4.1.4 Word Variables

The word variables are the variables which have a particular word in the input text as its value. For example the word variable 'pst_aux' is satisfied if the word is a verb (i.e. VM/VAUX) and equals any of these values: **था/थी/थीं/थे** /tʰa:, tʰi:, tʰĩ:, tʰe:/. There are a lot of word variables we use for the tool. The number for the word variable is high because we include all the words in Hindi that can functionally play the role of auxiliary verb. This includes all the forms of the V2s as well as the copular auxiliaries. Previsously, we have stated that we include a total of 61 V2s. To get an estimate of the word variables, we can have all the forms of the V2s functionally working as the auxiliary. Thus a V2 like **रह** /rəh/ "live/stay/remain" will have seven variables as expemplified below:

rah	रह
rah_fut	रहेगा/रहेगी/रहेंगी/रहेंगे/रहिएगा/रहियेगा/रहोगे/रहोगी/रहूंगा/रहूंगी
rah_imp	रह/रहो/रहें/रहिए/रहिये
rah_impf	रहता/रहती/रहतीं/रहते
rah_inf	रहना/रहनी/रहने
rah_opt	रहूं/रहूँ/रहे/रहें
rah_pft	रहा/रही/रहीं/रहे

We store the word variables also in a text file where the word variables are aligned with their value(s) delimited by a TAB. The following as a snapshot of the word-variables file:

Word Variable	Values
wala	वाला/वाली/वाले
khada	खड़ा/खड़ा
khada_pft	खड़ा/खड़ी/खड़ीं/खड़े/खड़ा/खड़ी/खड़ीं/खड
prs_aux	है/हैं/हो/हूं/हूँ
pst_aux	था/थी/थीं/थे
ho_cfct	हो/हों
cahiye	चाहिए/चाहिये
jane	जाने
particles	नहीं/ना/ही/तो/भी/सा/सी/से/क्या/न/वाला/वाली/वाले
a	आ
a_fut	आएगा/आयेगा/आएगी/आयेगी/आएंगी/आयेंगी/आईएगा/ आईयेगा/आओगे/आओगी/आउंगा/आऊंगा/आउंगी/आऊंगी
a_imp	आ/आओ/आईए
a_impf	आता/आती/आतीं/आते
a_inf	आना/आनी/आने
a_opt	आऊं/आऊँ/आएं/आयें
a_pft	आया/आई/आईं/आए/आये
bach	बच
bach_fut	बचेगा/बचेगी/बचेंगी/बचेंगे/बचिएगा/बचियेगा/बचोगे/बचोगी/बचूंगा/बचूंगी
bach_imp	बच/बचो/बचें/बचिए/बचिये
bach_impf	बचता/बचती/बचतीं/बचते
bach_inf	बचना/बचनी/बचने
bach_opt	बचूं/बचूँ/बचे/बचें
bach_pft	बचा/बची/बचीं/बचे
baith	बैठ
baith_fut	बैठेगा/बैठेगी/बैठेंगी/बैठेंगे/बैठिएगा/बैठियेगा/बैठोगे/बैठोगी/बैठूंगा/बैठूंगी
baith_imp	बैठ/बैठो/बैठें/बैठिए/बैठिये
…	…
…	…

Table 43: Word Variables

4.2 IDENTIFYING VERB GROUPS: A DETERMINISTIC PATTERN MATCHING TASK

Now that we have the templates defined and each of the elements in the templates defined to be found out in a deterministic way (that is we do not have any statistical heuristics playing a role in it), it is a deterministic pattern matching task only that we need to do through the program.

Continueing from the section above, I will begin with the process of how the tool works. I will show how the tool works in different scenarios and captures all of the verb groups based on the templates defined.

The tool starts searching from the right and finds a verbal word. If the verbal word is a main verb, it checks if any of the one word verb group templates matches with it. We have a total of 12 templates for the one word verb group. Three of these templates (prs_aux, pst_aux and cahiye) have word variables. The rest of them are to be found out by tags inside the verbs. It can match with any of these 12 templates. Once it is matched, it marks the boundary (i.e. inserts square brackets at the start and end of verb group boundary).

Multi-word verb groups are also identified in a similar fashion. When the search finds a word that is an auxiliary (i.e. has the VAUX in its category tag), it continues the search and looks for another word preceding it. It is expected that the preceding word should also be a verb. If anything else occurs in between it must be a particle (particle is a word variable). The search ignores the particles coming in between a VAUX and a VM. It continues the search until a main verb is found. This combination of main verb and the auxiliary is then matched with the templates defined and a corresponding tag is given to it at the end of the VG boundary.

The following table attempts to give a simplistic pseudo code for the program.

For each of the input sentence, search from the right.
Find a tag containing VM or VAUX
//Find One word Verb Groups
If the first tag from the right contains VM
And if the word equals any of the values of variable 'prs_aux'
mark the boundary //i.e. insert '[' to the left of the word and ']' to the right of word.
assign the tag of "VG.prs.sim.dcl" //i.e. insert the tag "VG.prs.sim.dcl" after the right

				boundary.				
			And if the word equals any of the values of variable 'pst_aux'					
				mark the boundary				
				assign the tag of "VG.pst.sim.dcl"				
			And if the PoS tag contains "fut"					
				mark the boundary				
				assign the tag of "VG.fut.sim.dcl"				
			And if the PoS tag contains "pft"					
				mark the boundary				
				assign the tag of "VG.pst.sim.dcl"				
			…					
				…				
				…				
		//Identify multiple word Verb Groups						
		If the first tag from the right contains VAUX						
			Check the next word towards the left and continue till a word with its PoS tag containing 'VM' is found.					
			Ignore particles //words that occur as values to the variable of particles					
			Count the distance of the VM from the VAUX i.e. count how many words are there between the VAUX and the VM. ignore the cases of 'particles' coming in between.					
				Now, if the distance between the VAUX and the VM equals 1 i.e. the VM is just after the first found VAUX (ignoring the particles)				
				//This will identify Verb Groups having two words				
				//The next word on the right of the VAUX will be VM				
					Match the following cases:			
						VR_impf+prs_aux //i.e. if the first word equals prs_aux and the PoS tag of the second word contains 'impf'		
							assign "VG.prs.impf.dcl" // assign this tag to the group in their boundary	
						VR_impf+pst_aux //i.e. if the first word equals pst_aux and the PoS tag of the second word contains 'impf'		
							assign "VG.pst.impf.dcl" // assign this tag to the group in their boundary	
						VR_pft+prs_aux		
							assign "VG.prs.pft.dcl" // assign this tag to the group in their boundary	
						//Do the same for the following templates (to be filled in later)		
						.....		
						.....		
				If the distance between the VAUX and the VM equals 2 i.e. the VM is after two VAUXes (ignoring the particles)				
				//This will identify Verb Groups having three words				
				//The VM now will occur after two VAUXes from the right of the sentence boundary				

					Match the following cases:		
						VR+rah_pft+prs_aux //i.e. if the first word equals prs_aux, second word equals rah_pft and the PoS tag of the third word contains 'nfn'	
							assign "VG.prs.prog.dcl" // assign this tag to the group in their boundary
						VR+rah_pft+pst_aux //i.e. if the first word equals pst_aux, second word equals rah_pft and the PoS tag of the third word contains 'nfn'	
							assign "VG.pst.prog.dcl" // assign this tag to the group in their boundary
						VR+rah_pft+ho_fut //i.e. if the first word equals ho_fut, second word equals rah_pft and the PoS tag of the third word contains 'nfn'	
							VG.fut.prog.prsm
					//	Do the same for more of the template-tag pairs (to be filled in later)	
						.....	
						.....	
				If the distance between the VAUX and the VM equals 3			
				//This will identify Verb Groups having four words			
				//The VM now will occur after three VAUXes from the right of the sentence boundary			
					Match the following cases:		
						VR_IMPF+ja+rah_pft+prs_aux //i.e if the first word equals prs_aux, second rah_paft, third is ja and the tag of the fourth contains impf	
							assign VG.prs.prog_dur.dcl
						VR_IMPF+ja+rah_pft+pst_aux //i.e if the first word equals pst_aux, second rah_paft, third is ja and the tag of the fourth contains impf	
							assign VG.pst.prog_dur.dcl
						VR_impf+ja+rah_pft+ho_fut //i.e if the first word equals ho_fut, second rah_paft, third is ja and the tag of the fourth contains impf	
							assign VG.fut.impf_dur.prsm
					//	Do the same for more of the template-tag pairs (to be filled in later)	
						.....	
						.....	
				If the distance between the VAUX and the VM equals 4			
				//This will identify Verb Groups having five words			
				//The VM now will occur after four VAUXes from the right of the sentence boundary			
					Match the following cases:		
					VR_impf+cal_pft+a+rah_pft+prs_aux		

						assign VG.prs.prog_dur.dcl
					VR_impf+cal_pft+a+rah_pft+pst_aux	
						assign VG.pst.prog_dur.dcl
					VR_impf+cal_pft+a+rah_pft+ho_fut	
						assign VG.pst.impf.cfct
					//	Do the same for more of the template-tag pairs (to be filled in later)
						.....
						.....
		//Identify if the sentence has more than one clause				
		If the the sentence has more VMs, repeat the process above.				
		If there is a VAUX but no VM, print the error message “Found an orphan VG”				
		If there is no VM or VAUX in the sentence, print the error message “No VG found in this sentence”				

Table 44: Pseudo Code for the VG finder tool

4.3 A GLIMPSE OF THE TOOL: “HINDI VERB GROUP FINDER”

The tool developed for this work is coded in Java. For the sake of naming, I have named it as “Hindi Verb Group Finder”. For the testing purposes, I had first developed a command line program and then developed it into an applet. The applet can be run through a web browser on any Java enabled computer. This applet version can be hosted on a server as well. For the testing purposes, we have hosted this web based tool at server of the Special Centre for Sanskrit Studies, Jawaharlal Nehru University, New Delhi. The tool can be tested at the following URL:

http://sanskrit.jnu.ac.in/vgt/index.html

and alternatively also on the following one:

www.langlex.com/vgt/index.html

4.3.1 Back-End

At the back end, the tool has a simple structure. All the files needed for the tool is located in the same directory where the applet files and the corresponding web page is available. The table below shows all the files that this directory has along with their functions.

Name of the File	Function
vgfinder.html	// runs the tool when opened in a browser
FindVerbGroupsApplet.class	//Compiled Java file that contains the applet structure
Helper.class	//The compiled Java file containing all the code for the logic

	implementation
tag-variables.txt	//is the tag variable file as explained in section 4.1.3
template-definitions.txt	// is the file containing all the VG templates
word-variables.txt	//is the word variable file containing all the word variables as explained in section 4.1.4

Table 45: Directory structure of the VG Finder Tool

4.3.2 Front-End

At front end, it is a web-page with two text fields. The first text area is allocated for the input text. The input is the annotated Hindi text, preferably one sentence per line. Below this first text field there are two radio buttons which can be selected for the desired format of the input. The "Preserve PoS tags in output" button is selected by default. Once the input sentence is given and the button "Mark VG's" is clicked, the output is generated in the second text area.

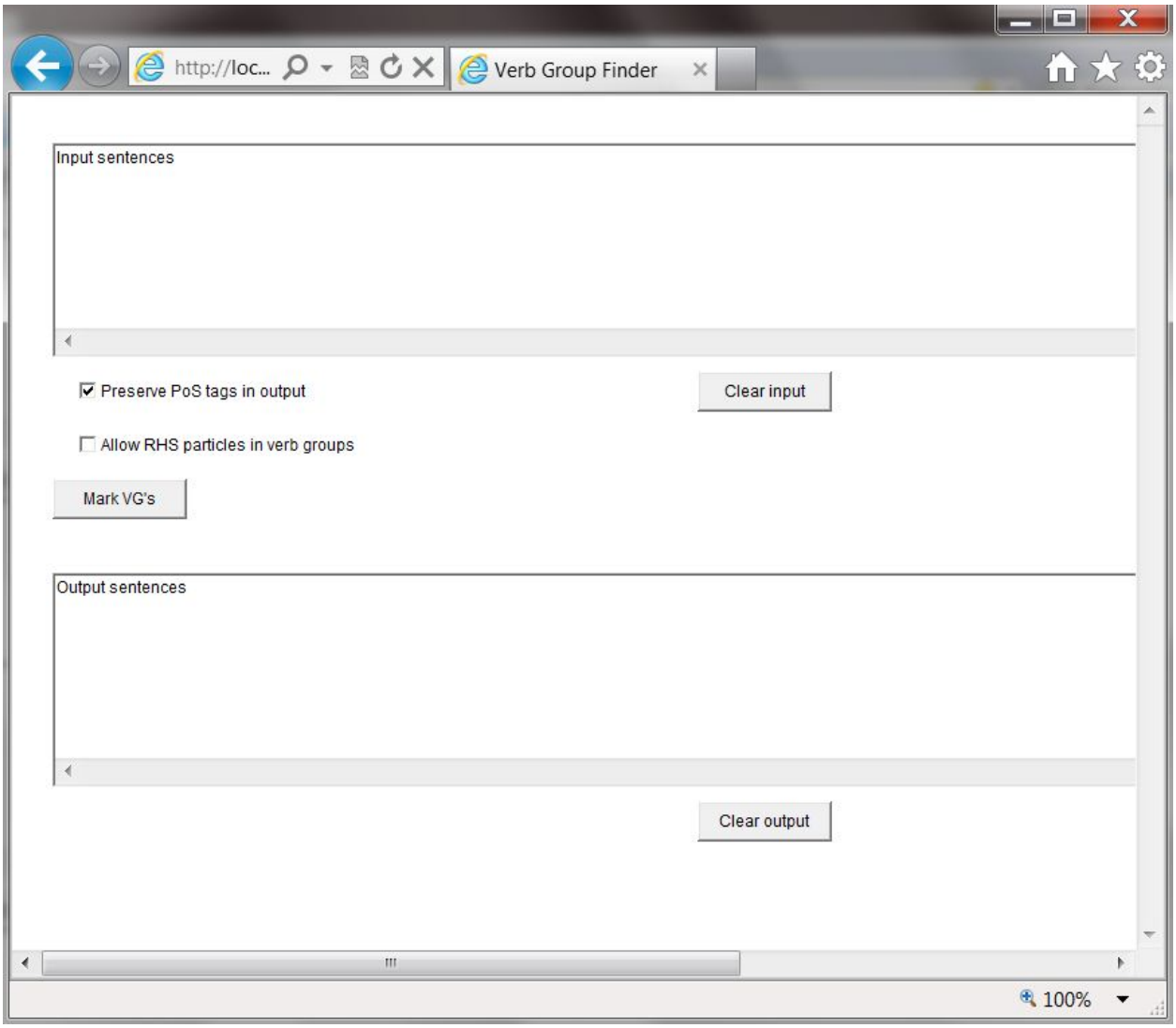


Figure 1: Front-End of the VG Finder Tool

Because the "Preserve PoS tags in output" is selected, the output generated retains the PoS tags coming in the input text. It looks as shown in Figure 2 below:

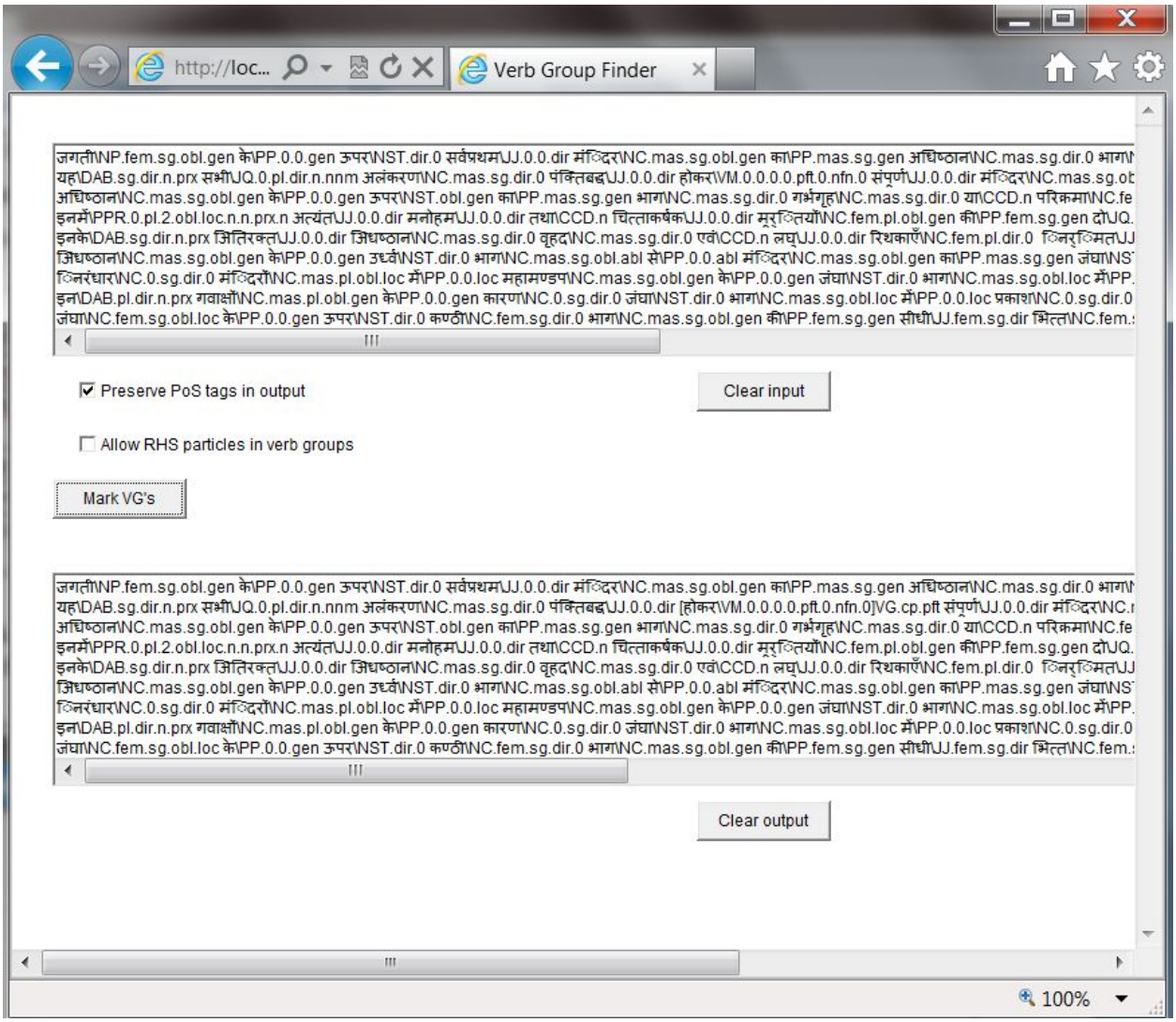


Figure 2: Output of the VG Finder Tool retaining the Input PoS tags

If the “Preserve PoS tags in output” button is unselected, the output will not retain the PoS tags of the input text and will look as shown in Figure 3 below:

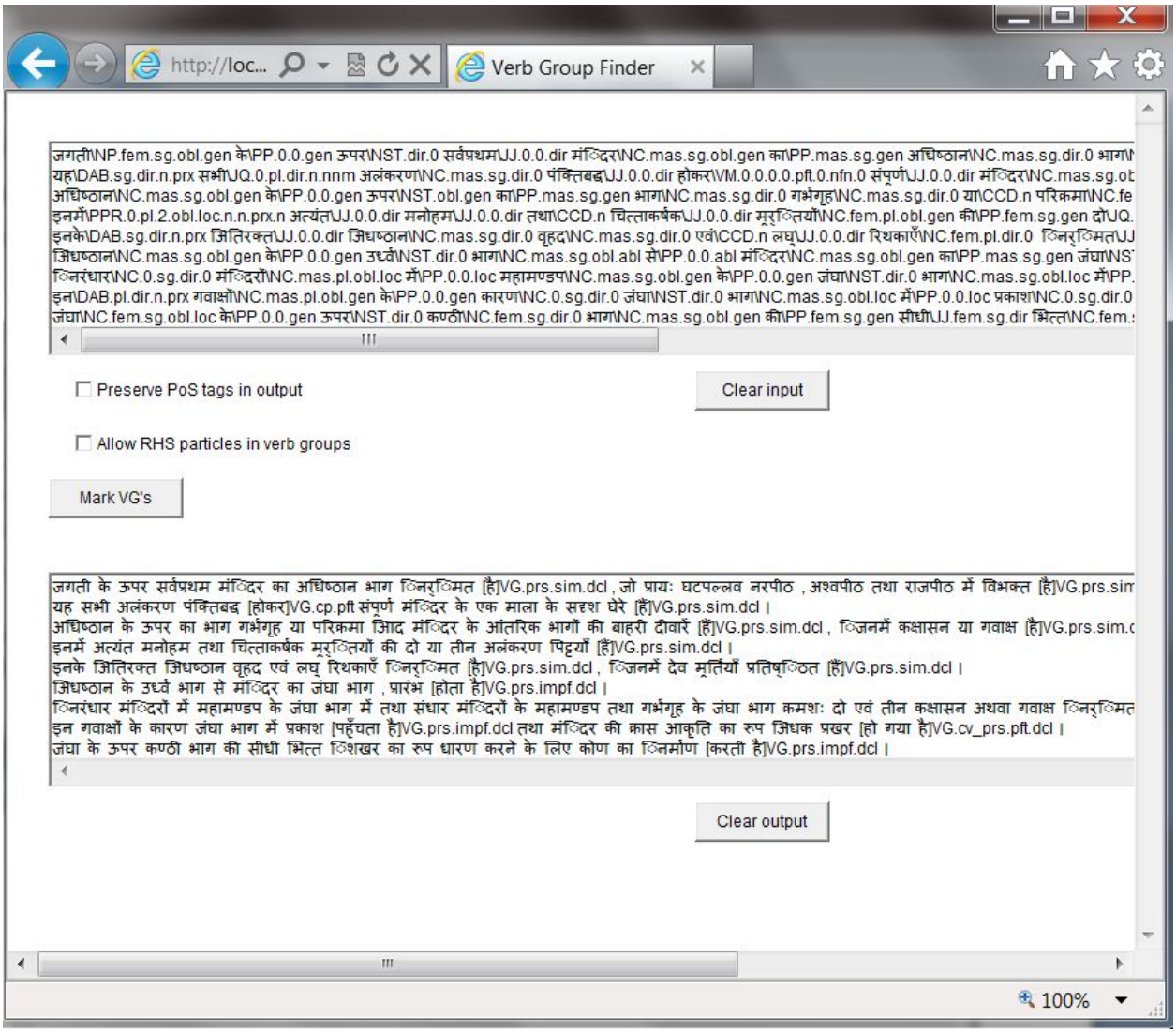


Figure 3: Output of the VG Finder Tool without the Input PoS tags

The other radio button "Allow RHS particles in verb groups" is meant to retain any particles coming at the right hand side of the verb groups. One can keep this always unchecked. This feature has been provided to capture the particles that modifies the VG but comes after it as in

E353	वह	<u>खाते</u>	<u>ही</u>	बोल	पड़ा	"वाह"
IPA	wəh	<u>kʰa:-te:</u>	<u>hi:</u>	bo:l	pəɖ-a:	wa:h
Gl.	3SG	eat-IMPF.PL	only-PRT	speak.VR	lay-PFT.MSG	wow

EGl Just as he ate, he said ‘wow’!

4.4 EVALUATION

As stated earlier, we use as input the data annotated in the IL-PoST framework. There is a corpus already available annotated in this framework (Bali, K. et.al., 2010). We use the whole of this corpus as the input for the tool and see how the results fare with reference to the verb group identification. As per the documentation provided with this corpus, this corpus has a total of 4859 sentences and a total of 98450 words with an average of 20 words per sentence. This corpus is in both text and the xml format. I use the data that is in plain text format. There are a total of 18 text files containing the annotated data in the proper format. There are three other files with the name of e1.txt, e2.txt and e3.txt. I use these files also as input even though their format is somewhat different from the other files.

4.4.1 First Test Run

I combined the whole the 18+3 text files into one single text file. The total number of sentences I found out came out to be different from what has been stated in the documentation provided with the corpus. The total number of sentences I found out to be 4853. This difference might also be because of the combining process I perfomed over these files. There were also instances when some sentences were broken in between. Sometimes this break also occurred in the middle of a word or in the middle of the tags of the word. Such cases were found out and necessary corrections were made so that the sentences in this combined corpus file are aligned properly.

We ran the tool over the whole of this combined corpus. The following table gives an overview of the results for the first run of the tool.

Description of the Results Parameters	Raw Number	Percentage
Sentences with at least one unknown Verb Group found	1020 (Includes 103 instances of orphan VAUXes)	21%
Sentences with at least one orphan auxiliary verb found	1064	22
Total number of sentences with at least one orphan auxiliary verb found	1064+103=1169	24
Total number of sentences having either at least one	1169+1064=2233	46

unknown VG or at least one orphan VG or both		
Total Number of Sentences having no issues	4853-2233=2620	54
Total Number of Sentences that has no VGs	1213	25
Total Instances of Verb Groups Found and tagged	7399	NA
Incorrect Marking of the VGs	Cannot be ascertained (Needs manual checking)	

Table 46: Result Summary of the First Run

4.4.2 Error Analysis in the First Test Run

As seen in the table above, 46% of the sentences have an error at least two types: they contain a verb group which could not be identified and hence marked as 'unknown' or there are verb groups consisting of no VM and hence the VAUXes are left orphan.

A total of 7399 instances of verb groups have been marked. Right now, we do not have any mechanism to automatically test whether these VG markings are correct or not. It has to be checked manually for each of the instances of the VGs marked.

Before we check the VG markings in the first run, I would focus on the reasons why we had so much of unknown VGs and why are there so much of orphan VGs.

4.4.2.1 Reasons for Orphan VAUXes

Theoretically, an auxiliary verb cannot exist on its own in any sentence. This is one of core tenets of annotation guidelines. If an orphan auxiliary has been found in the input data, it means that there is a problem with the annotation. For example the following 5 sentences taken from the LDC corpus has VGs with no VMs present:

C.ID.	Example Output of the Original Annotated Sentence from the LDC Corpus
3	[[FOUND AN ORPHAN VAUX; PLEASE CHECK POS ANNOTATION]] इस\DAB.sg.dir.n.prx प्रकार\NC.0.sg.dir.0 धूप\NC.fem.sg.obl.gen छांव\NC.fem.sg.obl.gen की\PP.fm.sg.gen भी\CX.n सुंदर\JJ.0.0.dir व्यवस्था\NC.fem.sg.dir.0 हो\VAUX.0.0.0.0.pft.0.0.0 गई\VAUX.fem.sg.3.prs.pft.dcl.fin.n [है\VM.0.sg.3.prs.sim.dcl.fin.n]VG.prs.sim.dcl ।\PU
7	[[FOUND AN ORPHAN VAUX; PLEASE CHECK POS ANNOTATION]] अिधष्ठान\NC.mas.sg.obl.gen के\PP.0.0.gen उर्ध्व\NST.dir.0 भाग\NC.mas.sg.obl.abl से\PP.0.0.abl मंिदर\NC.mas.sg.obl.gen का\PP.mas.sg.gen जंघा\NST.dir.0 भाग\NC.mas.sg.dir.0 ,\PU प्रारंभ\JJ.0.0.dir होता\VAUX.mas.sg.3.prs.pft.hab.fin.n [है\VM.0.sg.3.prs.sim.dcl.fin.n]VG.prs.sim.dcl ।\PU
49	[[FOUND AN ORPHAN VAUX; PLEASE CHECK POS ANNOTATION]] प्रमुख\JJ.0.0.dir मीिडया\NC.mas.sg.dir.0 िशिवर\NC.mas.sg.dir.0 अिधकारी\NC.mas.sg.obl.gen एवं\CCD.n उप\JJ.0.0.dir सूचना\NC.fem.sg.dir.0 िनदेशक\NC.mas.sg.dir.0 िशवप्रसाद\NP.mas.sg.obl.gen भारती\NP.mas.sg.obl.gen के\PP.0.0.gen अनुसार\NC.0.sg.dir.0 मेले\NC.mas.sg.obl.loc में\PP.0.0.loc समाचार\NC.mas.sg.dir.0 संकलन\NC.mas.sg.obl.x के\PP.0.0.gen िलए\PP.0.0.x [आए\VM.mas.pl.3.pst.pft.dcl.fin.n]VG.pst.sim.dcl स्थानीय\JJ.0.0.dir ,\PU

	राष्ट्रीय\JJ.0.0.dir ,\PU अंतरराष्ट्रीय\JJ.0.0.dir इलेक्ट्रॉनिक\JJ.0.0.dir एवं\CCD.n प्रिंट\JJ.0.0.dir मीडिया\NC.mas.sg.obl.gen के\PP.0.0.gen प्रितिनिधयों\NC.mas.pl.obl.acc को\PP.0.0.acc मेलाधिकारी\NC.mas.sg.obl.gen ,\PU वरिष्ठ\JJ.0.0.dir पुलिस\NC.fem.sg.dir.0 अधीक्षक\NC.mas.sg.obl.gen ,\PU प्रभारी\NC.mas.sg.obl.gen अिधकारी\NC.mas.sg.obl.gen मीिडया\NC.fem.sg.dir.0 िशिवर\NC.mas.sg.dir.0 के\PP.0.0.gen हस्ताक्षर\NC.mas.sg.dir.0 युक्त\JJ.0.0.dir फोटो\NC.mas.sg.dir.0 लगे\VAUX.mas.pl.3.pst.pft.dcl.fin.n प्रेस\NC.fem.sg.dir.0 पास\NC.mas.sg.dir.0 जारी\NC.0.sg.dir.0 [िकए\VM.mas.pl.3.prs.pft.dcl.fin.n]VG.pst.sim.dcl गए\VAUXmas.pl.3.prs.pft.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n और\CCD.n उनका\PPR.mas.pl.3.obl.gen.n.n.dst.n सम्पूर्ण\JJ.0.0.dir िववरण\NC.mas.sg.dir.0 मीिडया\NC.mas.sg.dir.0 कैम्प\NC.mas.sg.obl.loc में\PP.0.0.loc दर्ज\JJ.0.0.dir [हैं\VM.0.sg.3.prs.sim.dcl.fin.n]VG.prs.sim.dcl ।\PU
126	[[FOUND AN ORPHAN VAUX; PLEASE CHECK POS ANNOTATION]] इंग्लैंड\NP.mas.sg.obl.gen के\PP.0.0.gen यार्कशायर\NP.mas.sg.obl.gen की\PP.fem.sg.gen रहने\NV.dir.0 वाली\CX.n आयरीन\NP.fem.sg.dir.0 यहाँ\ALC.dir.0.n अपने\PRF.mas.sg.obl.gen.n पित\NC.mas.sg.obl.acc को\PP.0.0.acc श्रद्धांजिल\NC.fem.sg.dir.0 अर्िपत\JJ.0.0.dir करने\VAUX.mas.sg.3.prs.pft.dcl.fin.n [आई\VM.fem.sg.3.prs.pft.dcl.fin.n हैं\VAUX.0.sg.3.prs.sim.dcl.fin.y]VG.prs.pft.dcl ।\PU
154	[[FOUND AN ORPHAN VAUX; PLEASE CHECK POS ANNOTATION]] लोस्वेट\NP.fem.sg.dir.0 लोगो\NC.0.pl.obl.acc को\PP.0.0.acc संगम\NP.mas.sg.obl.loc में\PP.0.0.loc स्नान\NC.mas.sg.dir.0 करते\VAUX.mas.sg.3.pst.pft.dcl.fin.n [देख\VM.0.0.0.0.pft.0.nfn.0]VG.cp.pft इस\DAB.sg.dir.n.prx कदर\CX.n प्रभाित\JJ.0.00.dir [हुई\VM.fem.sg.3.pst.pft.dcl.fin.n]VG.pst.sim.dcl की\CSB.n वह\PPR.0.sg.3.dir.0.n.n.dst.n संगम\NP.mas.sg.obl.loc में\PP.0.0.loc डुबकी\NC.fem.sg.dir.0 लगाने\NV.obl.abl से\PP.0.0.abl खुद\NC.0.sg.obl.acc को\PP.0.0.acc रोक\JJ.0.0.dir नहीं\CX.y [पाई\VM.fem.sg.3.pst.pft.dcl.fin.n]VG.pst.sim.dcl ।\PU

Table 47: Sentences with Orphan VAUXes found in the LDC Corpus

As is evident from the example annotated sentences above, these sentences simply do not have any VMs marked in it while there are VAUXes present there. This is theoretically not possible and these are annotation errors present in the LDC corpus. Thus, for all of such occurrences in the 1169 sentences, we have to correct the annotations of such VGs before we put them to test in the tool.

4.4.2.2 Reasons for Unknown VGs Found

There can be various reasons for a VG being not recognized. The most incriminating for this test would be the failure of any of the templates to recognize it. Fortunately, that is the case with only a few of the sentences. The most number of unknown occurrences are due to other reasons.

C.ID	Example Output of the Original Annotated Sentence from the LDC Corpus
29	[[AT LEAST ONE UNKNOWN VG FOUND]] यत्र-तत्र\ALC.dir.0.n कई\JQ.0.pl.dir.n.nnm ईंटे\NC.fem.pl.dir.0 ,\PU मंिदर\NC.mas.sg.obl.gen के\PP.0.0.gen अवशेष\NC.mas.sg.dir.0 तथा\CCD.n भग्नों\NC.mas.pl.obl.gen के\PP.0.0.gen अवशेषों\NC.mas.pl.obl.gen के\PP.0.0.gen आधार\NC.mas.sg.obl.loc पर\PP.0.0.loc [िदखाई\VM.fem.sg.3.prs.pft.dcl.fin.n देते\VAUX.mas.pl.3.prs.pft.hab.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n]UNKNOWN ।\PU
56	[[AT LEAST ONE UNKNOWN VG FOUND]] शुक्ला\NP.mas.sg.obl.acc एवं\CCD.n उसकी\PPR.fem.sg.3.obl.gen.n.n.dst.n पोती\NC.fem.sg.dir.0 िप्रया\NP.fem.sg.obl.acc को\PP.0.0.acc अस्पताल\NC.mas.sg.obl.loc में\PP.0.0.loc भर्ती\JJ.0.0.dir

	[करा\VM.mas.sg.3.prs.pft.dcl.fin.n िदया\VAUX.mas.sg.3.prs.pft.dcl.fin.n गया\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n]UNKNOWN ।\PU
75	[[AT LEAST ONE UNKNOWN VG FOUND]] उन्हें\PPR.0.pl.3.obl.gen.n.n.dst.n आज\ALC.dir.0.n भी\CX.n प्रदूिषत\JJ.0.0.dir जल\NC.mas.sg.obl.loc में\PP.0.0.loc [नहाना\VM.0.0.0.0.pft.0.nfn.0 पड़\VAUX.0.0.0.0.0.0.nfn.0 रहा\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n]UNKNOWN ।\PU
115	[[AT LEAST ONE UNKNOWN VG FOUND]] ठग\NC.mas.sg.dir.0 रेलवे\NC.mas.sg.dir.0 स्टेशन\NC.mas.sg.obl.abl एवं\CCD.n बस\NC.fem.sg.dir.0 अड्डों\NC.mas.pl.obl.abl से\PP.0.0.abl अपना\PRF.mas.sg.obl.gen.n िशकार\NC.mas.sg.dir.0 तलाश\JJ.fem.sg.dir [कर\VM.0.0.0.0.0.0.nfn.0]UNKNOWN वारदात\NC.fem.sg.obl.acc को\PP.0.0.acc अंजाम\NC.mas.sg.dir.0 [दे\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.prs.prg.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n]VG.prs.prog.dcl ।\PU
122	[[AT LEAST ONE UNKNOWN VG FOUND]] उनके\PPR.0.pl.3.obl.gen.n.n.dst.n िशकार\NC.mas.sg.dir.0 कई\JQ.0.pl.dir.n.nnm लोगों\NC.0.pl.obl.acc को\PP.0.0.acc तौिलए\NC.mas.sg.dir.0 या\CCD.n गमछा\NC.mas.sg.dir.0 [लपेटकर\VM.0.0.0.0.pft.0.nfn.0]VG.cp.pft संगम\NP.mas.sg.obl.abl से\PP.0.0.abl [लौटना\VM.mas.sg.3.prs.pft.dcl.fin.n पड़ता\VAUX.mas.sg.3.prs.pft.hab.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n]UNKNOWN ।\PU
134	[[AT LEAST ONE UNKNOWN VG FOUND]] आयरीन\NP.fem.sg.obl.erg ने\PP.0.0.erg [बताया\VM.fin.prs.sim.dcl.1.n.n.n]UNKNOWN िक\CSB.n उनके\PPR.0.sg.3.obl.gen.n.n.dst.y पित\NC.mas.sg.obl.acc को\PP.0.0.acc भारत\NP.mas.sg.obl.abl से\PP.0.0.abl काफी\JQ.0.pl.dir.n.nnm लगाव\NC.mas.sg.dir.0 [था\VM.mas.sg.3.pst.sim.dcl.fin.n]VG.pst.sim.dcl और\CCD.n वे\PPR.0.sg.3.dir.0.n.n.dst.n दोनों\JQ.0.pl.dir.n.ord २००१\RDX के\PP.0.0.gen महाकुंभ\NP.mas.sg.obl.loc में\PP.0.0.loc भी\CX.n यहाँ\ALC.dir.0.n [आए\VM.mas.pl.3.prs.pft.dcl.fin.n थे\VAUX.mas.pl.3.pst.sim.dcl.fin.n]VG.pst.pft.dcl ।\PU
142	[[AT LEAST ONE UNKNOWN VG FOUND]] ओलना\NP.fem.sg.dir.0 कुंभ\NP.mas.sg.obl.gen के\PP.0.0.gen यादगार\JJ.0.0.dir अवसर\NC.mas.sg.obl.acc को\PP.0.0.acc कैमरे\NC.mas.sg.obl.loc में\PP.0.0.loc कैद\NC.mas.sg.dir.0 [कर\VM.0.0.0.0.0.0.nfn.0 रही\VAUX.fem.sg.3.prs.prg.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.y]VG.prs.prog.dcl ताि क\CSB.n इसे\DAB.sg.dir.y.prx अपने\PRF.0.sg.obl.gen.n दोस्तों\NC.0.pl.obl.acc को\PP.0.0.acc [िदखा\VM.mas.sg.3.fut.pft.dcl.fin.n सके\VAUX.0.sg.3.fut.pft.dcl.fin.y]UNKNOWN ।\PU

Table 48: Sentences with Unknown VGs Found in the LDC Corpus

As seen in the first example output sentence above (corpus ID 29), the VG missed is िदखाई देते हैं /dɪkʰaːiː dete hɛ̃ː/. At the first sight, we find that the encoding of the first word is problematic here. The diacritic 'ि' appears separated from the consonant 'द' /d/. However, this should not have been a problem because this is marked as the VM. Being a VM other than prs_aux and pst_aux, it has to be identified by the PoS tags it contains. Thus, we need to look at its PoS tags. We find that the PoS annotation for the first word is fine. The second word देते has a PoS annotation of 'pft' as the aspectual value, which actually is wrong because it is morphologically marked for imperfect aspect. The third word है /hɛː/ is correctly marked. We find that this VG is a compound verb construction. A template such as "V1_pft+V2_impf+prs_aux" or for that matter "VR_pft+de_impf+prs_aux" (because for the CV constructions we specify the V2 as a word

variable, as stated in the previous chapter) should identify this VG. Even though the second verb is wrongly marked with ‘pft’, the tool should be able to identify the VG (because VAUXes are understood not by their PoS tags but by the word variable) if we define a template like “VR_pft+de_impf+prs_aux”. However, such a template would be incorrect in most of the cases. If we define this template then phrases like the following would also get verified:

*खाया देते हैं /kʰaːja: deːte: hɛ̃ː/

*गए देते हैं /gəe: deːte: hɛ̃ː/

*पढ़ाई देते हैं /pəɖ̣ʰaːi: deːte: hɛ̃ː/

If a template verifies phrases like the above, it would be wrong because Hindi does not allow such constructions. We also observe that in Hindi, a phrase like the one coming in C.ID. 29 can best be categorized as exceptional. Only two phrases with this structure can be found in Hindi:

दिखाई /dɪkʰaːiː/+de_XX+xx

सुनाई /sʊnaːiː/+de_XX+xx

That is, the two words of दिखाई /dɪkʰaːiː/ and सुनाई /sʊnaːiː/ demand a particular VG structure for themselves. We need to modify the template definitions or add a new template to capture such constructions in their various conjugations. May be in such cases, we can also make the VM be a word variable to capture just these two words.

In the second example sentence (corpus ID 56) the VG identified as unknown is “करा िदया गया है” /kərɑː dɪjɑː gəjɑː hɛː/. Instantly we can see that the second word in this VG has an encoding problem where the diacritic does not conjugate well with the consonant. This can cause a problem in its identification. Because it is a VAUX and most of the VAUXes are identified based on the word variables. In its corrected form, it would look like दिया /dɪjaː/ which would be a value of the variable ‘de_pft’. If the input is incorrect at this place, it will never find such a value and would render the VG as unknown. If we correct this spelling error and reprocess this

sentence, we still find the VG as unknown. The reason now is that the annotation for the first word is incorrect. The first word in this VG करा /kəra:/ has been marked for ‘pft’ value. It should not have had this value because it actually is a verb stem in the present context. If we correct this and mark the first word as non-finite ‘nfn’, and process the sentence again it is identified as a VG with the tag ‘VG.cv_pas_prs.pft.dcl’ which is the correct identification.

For the third example sentence above (corpus ID 75) the VG is “नहाना पड़ रहा है” /nəha:na: pəɖ rəha: hɛ:/. The problem in this VG is with the PoS annotation of the first word ‘नहाना’ /nəha:na:/ which has been marked as perfect ‘pft’ and non-finite ‘nfn’. This word actually is neither perfect nor non-finite. It is a /-na/ ending infinitive verb and the proper tag for this should have been infinitive ‘ifn’. If we make this change and process the sentence again, the VG is correctly marked with the tag “VG.cv_prs.prog.oblg”.

For the fourth example sentence with corpus ID 115, the VG marked as unknown is ‘कर’ /kər/. It is a part of the conjunct verb where the previous word is a noun (marked wrongly here as adjective (JJ)). Conjunct verbs are not covered in this thesis for the reasons stated earlier. However, the verbal part in conjunct verbs should still be identified. So, the word कर /kər/ here should be identified as conjunctive participle with the tag of “VG.cp.pft”. However, this would happen only if the word is tagged both as non-finite and perfective. Because the only aspect intended in such construction is perfect, it is perfectly fine to mark them as perfective and because it does not play as the main verb in the sentence and is in stem form, it has to be non-finite. If we mark this word as ‘pft’ at the aspectual slot along with the ‘nfn’ tag in the modal slot, it will be identified as “VG.cp.pft”.

The fifth sentence with the corpus ID of 122 has the VG “लौटना पड़ता है” /lɔ:ʈna: pəɖta: hɛ:/. Here the problem is the first word लौटना /lɔ:ʈna:/ which should have been marked for an infinitive verb. If this change is made, the VG is correctly identified as “VG.prs.impf.oblg”. Even though the tag of the second word has some problems here, it will be ignored because they are recognized through the word variables and tags for such words are ignored.

The sentence ID 134 has the VG "बताया" /bəta:ja:/ as unknown. The problem with this VG is clearly with the annotation. It should have been marked for 'pft' which it misses. Additionally, it is marked for present 'prs' and is also in format other than normal (see the last three slots marked for 'n', the order of the tags is not the same). Adding the 'pft' tag at any of the eight slots after 'VM' in the tags can resolve this and the tool will correctly identify this VG as "VG.pst.sim.dcl"

The last sentence (corpus ID 142) in the example set above has the VG "िदखा सके" /dɪkʰa: səke:/ as unidentified. Here again we find that the annotation is pathetic. Both the words have been marked for future 'fut' and perfect 'pft'. None of these two words are in future forms. The first word is also not perfect. It actually is a stem form and should have been marked as non-finite 'nfn'. If we remove the future tags from both the words and make the first word as 'nfn', it will be identified as "VG.cv_0.0.opt".

4.4.2.3 Incorrect Marking of the Verb Groups

There are verb groups which have been marked incorrectly. Only a manual check for each of the verb groups identified can be used to test this. That is why we have not given any number to the incorrect marking row in Table 46 above.

C.ID	Example Output of the Original Annotated Sentence from the LDC Corpus
7	[[FOUND AN ORPHAN VAUX; PLEASE CHECK POS ANNOTATION]] अिधष्ठान\NC.mas.sg.obl.gen के\PP.0.0.gen उर्ध्व\NST.dir.0 भाग\NC.mas.sg.obl.abl से\PP.0.0.abl मंिदर\NC.mas.sg.obl.gen का\PP.mas.sg.gen जंघा\NST.dir.0 भाग\NC.mas.sg.dir.0 ,\PU प्रारंभ\JJ.0.0.dir होता\VAUX.mas.sg.3.prs.pft.hab.fin.n [है\VM.0.sg.3.prs.sim.dcl.fin.n]VG.prs.sim.dcl ।\PU
10	जंघा\NC.fem.sg.obl.loc के\PP.0.0.gen ऊपर\NST.dir.0 कण्ठी\NC.fem.sg.dir.0 भाग\NC.mas.sg.obl.gen की\PP.fem.sg.gen सीधी\JJ.fem.sg.dir भित्त\NC.fem.sg.dir.0 िशखर\NC.mas.sg.obl.gen का\PP.mas.sg.gen रूप\NC.mas.sg.dir.0 धारण\NC.0.sg.dir.0 करने\NV.obl.x के\PP.0.0.gen िलए\PP.0.0.x कोण\NC.mas.sg.obl.gen का\PP.mas.sg.gen िनर्माण\NC.mas.sg.dir.0 [करती\VM.fem.sg.3.prs.pft.hab.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n]VG.prs.pft.dcl ।\PU
457	इसके\DAB.sg.dir.n.prx अंतर्गत\JJ.0.0.dir तक\PP.0.0.loc जहाँ\ALC.dir.0.n प्लेटिनम\NP.0.sg.dir.0 प्लाजा\NP.0.sg.dir.0 १०\RDX मंिजला\JJ.0.0.dir [बनेगा\VM.mas.sg.3.fut.pft.dcl.fin.n]VG.pst.sim.dcl ,\PU वहीं\ALC.dir.0.n प्लेटिनम\NP.0.sg.dir.0 पार्क\NP.0.sg.dir.0 ७\RDX मंजिला\JJ.0.0.dir [होगा\VM.mas.sg.3.fut.pft.dcl.fin.n]VG.pst.sim.dcl ।\PU
466	इनके\DAB.pl.dir.n.prx अितिरक्त\JJ.0.0.dir इन\DAB.pl.dir.n.prx ब्लॉकों\NC.mas.pl.obl.gen के\PP.0.0.gen बेसमेंट\NC.mas.sg.obl.loc में\PP.0.0.loc पार्िकंग\NC.fem.sg.obl.gen की\PP.fem.sg.gen समुिचत\JJ.0.0.dir व्यवस्था\NC.fem.sg.dir.0 [रहेगी\VM.fem.sg.3.fut.pft.dcl.fin.n]VG.pst.sim.dcl ।\PU
482	जब\ALC.dir.0.n हम\PPR.0.pl.1.dir.0.n.n.prx.n सातवीं\JQ.fem.pl.dir.n.ord से\PP.0.0.abl तेरहवी\JQ.fem.pl.dir.n.ord शताब्दी\NC.fem.sg.obl.gen के\PP.0.0.gen स्थापत्य\NC.mas.sg.obl.gen का\PP.mas.sg.gen पर्यवेक्षण\NC.mas.sg.dir.0

	[करते\VM.mas.pl.3.prs.pft.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n]VG.prs.pft.dcl तो\CX.n हम\PPR.0.sg.1.dir.0.n.n.prx.n [पाते\VM.mas.pl.3.prs.pft.hab.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n]VG.prs.pft.dcl िक\CSB.n वह\PPR.0.sg.3.dir.0.n.n.dst.n एक\JQ.0.sg.dir.n.crd नये\JJ.mas.pl.dir राजनीितक\JJ.0.0.dir पिरस्ाथित\NC.fem.sg.obl.gen के\PP.0.0.gen अनुरुप\NC.0.sg.dir.0 [ढल\VM.0.0.0.0.0.0.nfn.0 जाता\VAUX.mas.sg.3.prs.pft.hab.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n]VG.cv_prs.impf.dcl ।\PU

Table 49: Examples of Incorrect VG Markings in the First Test Run

The example output sentences given in the table above show that there are instances where the VG though has been identified and also marked but the marking is wrong. For example in CID 7, we find that even though there is an orphan VAUX, the other verb has been marked. The annotation is wrong. The word that has been left is the VM while the word marked as VM is VAUX. The marking should have considered both words together as a verb group.

C.ID. 10 has a sentence where the last VG "करती है" /kərti: hɛ:/ has been marked wrongly as VG.prs.pft.dcl. Here the VG is not a perfective construction, it is in imperfect aspect. This wrong marking occurred because the first word करती /kərti:/ has been marked as 'pft' in the input tags. It should have been marked with the imperfect tag 'impf' for the correct identification. Please note here that in the IL-PoS tagset, we do not have a tag for the imperfective marking. If an imperfective aspect marked verb occurs it is given a mood tag of 'hab'. I consider this wrong. So, all the VGs which have to be identified with the help of the tag variable 'impf' will not be marked correctly simply because the tag does not exist in the IL-POST framework. Later I give more details on this and other changes that are needed to be made in the IL-PoS framework for this purpose.

Sentences 457 and 466 have VGs marked wrongly for past tense while the VG is clearly in the future tense. This is all because instead of future tense the verb forms have been marked for past.

Sentence 482 contains two VGs which are in imperfect aspect but have been identified as in perfect aspect. Here also the annotation is the culprit. Instead of 'impf', these have been marked as 'pft'.

4.4.3 The Treatment of Verbs in the IL-PoST Framework for Hindi

The LDC corpus that we use as the test data for the tool developed has various types of errors and needs to be corrected. There are also some modifications in the tagset and the annotations

guidelines I will argue here. Starting first with the MSRI guidelines, I will discuss the bulk changes I made in the corpus and justify the reasons for doing this.

The IL-PoST annotation framework is hierarchical in nature. For the verbs, it starts first with major category type and then goes on to tag the GNPH and TAM category in a particular order. This has been discussed in section 4.1.1 in this chapter above. Given the errors we have found and for the needs of the VG tool, I offer here some modifications in the IL-PoST annotation guidelines (this document comes along with the LDC corpus) for the verbal category tags.

4.4.3.1.1 Change in the Aspect Slot

According to the IL-PoST guidelines, three aspects are marked in the aspect slot. The guidelines document confirms that it is only the imperfective and the perfect that are marked morphologically on the verb. The imperfective has been equalled to simple here. The guidelines also recommend marking an aspect, namely the progressive. This aspect is marked through the help of an auxiliary verb रह-V /rəh-V/. For our purposes we do not need this. Please note here the progressive aspect uses the perfective form of the verb 'रह' /rəh/. If this form is marked progressive in the aspect slot, its perfective form will be lost because both perfective and the progressive are aspect markers and appear in the same slot. We can mark only one of them.

There are also other aspects that are marked periphrastically. For example the inceptive aspect is marked with the perfective form of the verb लग /ləg/ as in

E354	वह	रोने	लगा
IPA	wəh	ro:ne:	ləg-a:
Gl.	3SG	cry.INF	begin-PFT.MSG
EGl	He started crying.		

If we follow the guidelines, we should mark this inceptive marked auxiliary also as an aspect. Besides what is called as the progressive marker (i.e. रह-V /rəh-V/) also plays its role in durative, continuative and frequentative aspect markings. For this reason I would suggest that the

progressive auxiliary verb be removed from the IL-PoST guidelines. All the periphrastically marked aspects should be left to be identified at the level of local word grouping.

There is no marking for imperfect in the IL-PoST guideline. The imperfect has been equalled with simple aspect. So both the copular verbs and the imperfective forms are marked as simple. For a greater clarity, I propose that these two varieties be kept distinct. Simple should be left only for the cases where the copular verbs come. For the cases where the morphological form is clearly marked for imperfective (i.e. /-tV/ ending), they should be marked as imperfective. I have added a tag of 'impf' on my own in the aspectual values as can be seen in Table 40 above.

Change Suggested in the IL-PoST tagset for the aspectual slot of Hindi: the aspect attribute of progressive 'prg' be removed. Instead another value for imperfect aspect should be added. I suggest the tag of 'impf' here.

4.4.3.1.2 Change in the Mood Slot

There are three moods recommended to be marked in the IL-PoST guidelines: indicative, imperative and habitual. It is said that these are morphologically present in Hindi. I would like to disagree with this here as well. At first the indicative mood in Hindi is not marked morphologically. Rather it is understood by default if no other mood is marked either morphologically or periphrastically. A VG with an imperfect main verb can be indicative as well. So, the category of habitual appears fallacious here. While imperative is present and is also recommended in the guidelines another mood marked is missing. It is the optative. The optative mood is clearly marked in the first person singular form of the verbs. The optative forms in the other persons might by homonymous with other forms such as imperative and perfective.

Changes Suggested in the Mood slot: the habitual marker tag 'hab' should be removed. It should include the optative mood. I suggest the tag of optative to be 'opt'.

4.4.3.1.3 Change in the Finiteness Slot

The finiteness category as mentioned in the IL-PoST guidelines document is at best confusing. While it says "A finite verb is morphologically inflected for Gender, Number, Person, Tense, Aspect, and Mood." At the same time it gives examples that make these inflected categories as infinite and non-finite. For example all the '-tV' ending verbs have to be marked as infinitive according this guideline and all the verb roots and the '-nV' ending verbs have to be marked as

the non-finites. I disagree with this. Even the LDC corpus does not follow this. There is barely any '-tV' ending verb in the LDC corpus that has been marked as 'ifn'. Almost all of such verbs are marked as 'fin' here.

Changes suggested in the Finiteness Slot: Digressing from the current guidelines, I suggest that it is verb roots themselves that should be marked as non-finite 'nfn' and all the '-nV' ending verbs be marked as infinitive with the tag of 'ifn'.

4.4.4 Corrections Made in the Test Data

In order to evaluate the tool we had to test the tool with the real world examples. The only corpus we have is the LDC corpus. As noted above, this corpus has a lot of error. Without pointing to why these errors occurred in this corpus, I will simply move on to show what changes I made in this corpus before I ran the second test.

4.4.4.1 Correcting the Orphan VAUXes

All the instances of orphan VAUXes have been corrected. In most of the cases the whole of the VGs were marked as VAUX with sentences having no main verb at all. Correcting this itself improved the results by quite a good percentile.

4.4.4.2 Removing the Progressive Markings and Placing the Perfective Markings

As noted above, there was no need of 'prg' tag for the VAUXes. We changed this tag with the 'pft' marking. A simple process of find and replace method did most of the work. We changed all the occurrences of sequences of "pst.prg.dcl" and "prs.prg.dcl" with the sequence of "pst.pft.dcl" and "prs.pft.dcl". For now we left the other cases of prg (such as those occurring in the future marked verbs).

4.4.4.3 Removing the Habitual Markings and Placing the Imperfect Markings

Several replacements were made for the tag sequences which had 'hab' marking. Some of the replacement rules applied have been shown in the table below. A total of about 109 such sequences were replaced.

Sequence Replaced	Sequence Replaced With	No. of Replacements
mas.sg.3.prs.sim.hab.fin.n	mas.sg.3.prs.impf.dcl.fin.n	277
mas.sg.3.prs.pft.hab.fin.n	mas.sg.3.prs.impf.dcl.fin.n	136

mas.pl.3.prs.sim.hab.fin.n	mas.pl.3.prs.impf.dcl.fin.n	123
fem.sg.3.prs.sim.hab.fin.n	fem.sg.3.prs.impf.dcl.fin.n	97
fem.sg.3.prs.pft.hab.fin.n	fem.sg.3.prs.impf.dcl.fin.n	66
mas.pl.3.prs.pft.hab.fin.n	mas.pl.3.prs.impf.dcl.fin.n	44
mas.sg.3.pst.pft.hab.fin.n	mas.sg.3.pst.impf.dcl.fin.n	28
mas.sg.3.pst.sim.hab.fin.n	mas.sg.3.pst.impf.dcl.fin.n	26
mas.pl.3.pst.pft.hab.fin.n	mas.pl.3.pst.impf.dcl.fin.n	13
mas.pl.3.pst.sim.hab.fin.n	mas.pl.3.pst.impf.dcl.fin.n	13
mas.sg.0.0.0.hab.fin.n	mas.sg.0.0.impf.dcl.fin.n	11
mas.sg.0.0.pft.hab.fin.n	mas.sg.0.0.impf.dcl.fin.n	11
mas.sg.0.prs.sim.hab.fin.n	mas.sg.0.prs.impf.dcl.fin.n	5
0.0.0.0.0.hab.fin.n	0.0.0.0.impf.hab.fin.n	4
fem.sg.3.pst.sim.hab.fin.n	fem.sg.3.pst.impf.dcl.fin.n	4
mas.sg.0.pst.pft.hab.fin.n	mas.sg.0.pst.impf.dcl.fin.n	3
...	...	...

Table 50: Removing the Habitual Markings

4.4.4.4 Corrections in the Finiteness Slots

There were several instances where finiteness slot was marked wrongly. Finiteness is a property that has to be marked with each of the verbal words. While the aspect, tense slots may be marked 0, in none of the instance the finiteness slots can be marked as '0' which stands for 'not applicable'. This had to be checked manually for each of the occurrences. And still I am not sure if all of such occurrences have been marked correctly or not.

4.4.5 Other Changes

There were several other changes made. Particularly, there were several instances where the 'pft' was marked wrongly. Some changes have to be made based on the criterion as suggested in the tool. For example the cases of conjunctive participle where a verb root was added with -कर /-kər/ as in खाकर /kʰaːkər/ the tags had to have both 'pft' and 'nfn'. This was not present in many of the cases of the original data. Many changes were made while checking for the problems with the 'unknown' results as given in the output of the tool.

4.4.5.1 The Verb *cah* and its Annotations

The modal auxiliary of चाह /caːh/ might be somewhat confusing. The form चाहिए /caːhɪjeː/ is purely modal, other forms of the stem चाह /caːh/ are not. They function as a normal category verb and should not be marked as VAUX. However, in the LDC corpus, we have many of such

occurrences marked as VAUX instead of VM. This happens mainly because this kind of verb most of the time takes a VINF as object and comes in a series of verb sequence that appears as if the VINF is the VM and /ca:h/ and other verbs coming after it are auxiliary verbs.

For example in the sentence,

E355	मैं	यह	काम	करना	चाहता	हूं
IPA	mɛ̃:	jəh	ka:m	kərna:	ca:h-ta:	hũ
Gl.	I	this	job	do.INF	want-IMPF.MSG	be.PRS.1SG
EGl	I want to do this job.					

the VINF करना /kərna:/ is one VM while चाहता /ca:hta:/ is another VM. However, in the LDC corpus, the second case is marked as VAUX. This has been changed in the corpus for the purpose of a more correct data.

4.4.5.2 Single Template Denoting Two VGs?

Consider this:

E356	उसने	उसे	टोपी	पहना	कर	ठग	लिया
IPA	ʊs-ne:	ʊse:	ʈopi:	pəhna:	kər	ʈʰəg	lɪja:
Gl.	he-OBL-ERG	him	cap	wear.INF	do.VR	cheat.VR	take-PFT.MSG
EGl	Having capped him, he cheated him.						

E357	तू	रोज	टोपी	पहना	कर
IPA	tu:	roːɟ	ʈopi:	pəhna:	kər
Gl.	2SG	everyday	cap	wear.INF	do.VR

EGl You wear a cap everyday!

In the first example sentence above there are two VGs coming in a sequence. Sometimes this has been misinterpreted in the corpus as a single VG and only the first word has been marked as VM all three others being marked as VAUXes. In the first VG in this sentence पहना कर /pəhna: kər/ is a participial construction while the second one ठग लिया /ʈʰəg lıja:/ is CV past perfect construction. It is the second VG that carries the TAM markings while the first VG is simply a modifier of the second one. In the second example sentence however there is only one VG. Both the sentences have the same sequence of verb coming as a VG, the sequence being पहना कर /pəhna: kər/. In both the sentences the second verb कर /kər/ is an auxiliary. However, the tags for each of the occurrences would be different. In the first one the tag would be non-finite 'nfn' and perfect 'pft' because it is a conjunctive participle. In the second occurrence however, the auxiliary कर /kər/ is not non-finite. It is finite and carries the imperative mood with it. This has been missed at some places in the LDC corpus. It has been corrected at places where it came in the way of identification.

4.4.5.3 Perfect Form vs. Optative Form

There are cases where ambiguities come to the fore in doing a morphological annotation. These are the cases when the context plays an important role in annotation.

The perfect 'pft' form and the optative 'opt' form in the third person plural are the same. So in a sentence like the following:

E35	वह	किताब	<u>पढे</u>
IPA	wəh	kıta:b	<u>pəɖʰe:</u>
Gl.	3SG	book	read.OPT
EGl.	(I wish) He should read book.		

the verb form पढे /pəɖʰe:/ is optative/imperative. But in the sentence below,

E358	वे	लोग	किताब	पढ़े
IPA	we	lo:g	kɪta:b	pəɖʱ-e:
Gl.	3PL	people	book	read.PFT.PL
EGl	They read a book!			

the word पढ़े /pəɖʱe:/ is perfective. There are instances in the corpus where the optative forms have been marked wrongly as perfective. This results into either 'Unknown' VG or a wrong TAM tag assignment. Some of such instances have been corrected in the corpus but I cannot say for sure if all such instances have been corrected.

4.4.5.4 Conjunctive Participle

There is no way to find out the conjunct participle constructions occurring in the sentences. Thus in sentences like the following:

E14	वह	खाकर		बोला
IPA	wəh	kʰa:-kər		bol-a:
Gl.	3SG	eat-PFT.PCPL		speak-PFT.SG
EGl	He spoke after eating.			

There is no way to identify the participial constructions at the PoS level. The periphrastic participial construction as in the following sentence

E122	वह	खाते	हुए	बोला
IPA	wəh	kʰa:te:	hʊe	bola:
Gl.	3SG	eat-IMPF	be.PFT	speak-PFT.3MSG
EGl	He spoke while eating.			

can be found out by making a template. But the single word verbal participles cannot be found out. So, it is mandatory to mark the verb participles at the PoS annotation level. I have stated it earlier that the conjunct participle constructions like खाकर /kʰaːkər/ denote a sense of perfectness so a perfective tag ‘pft’ could be assigned to them. And because they do not inflect for GNPH and other TAM, they can be kept in the category of non-finite. Such constructions thus will have just two tags of ‘pft’ and ‘nfn’ and thus can be identified based on the tag variables.

Sometimes conjunct verbs can also occur as having just the VR with no insertion of the word '-कर' /-kər/ as suffixed to it. For example,

E359	उसने	उसे	अपने	पास	बैठ	रखवाली	करने	को	कहा
IPA	ʊs-ne:	ʊse:	əpne:	paːs	bɛːʈʰ	rəkʰwaːliː	kərne:	ko:	kəh-a:
Gl.	3SG.OBL-ERG	3SG.OBL.ACC	own. REFL	near	sit.VR.PCPL	watch.N	do.INF	ACC	say-PFT.MSG

EGl He asked him to watch while sitting near him.

Such cases should be carefully marked as having both ‘nfn’ and ‘pft’ tag and not as ‘nfn’ without pft.

We also have a case of conjunct verb when the construction does not involve कर /kər/

हाथ\NC.mas.sg.dir.loc में\PP.0.0.loc झंडे\NC.mas.pl.0.0 **लिए\VM.0.0.0.0.0.0.nfn.n** आ\VM.0.0.0.0.0.0.nfn.n गए\VAUX.0.pl.3.pst.sim.dcl.fin.n

Here the verb लिए /lɪeː/ has the meaning of लेकर /lekər/ but does not manifest it morphologically here. In such cases, it should be annotated as if it were लेकर /lekər/. If it is annotated as finite perfect form, then it would create a problem.

4.4.5.5 Imperfective Participle

For the imperfective participial verbal constructions like “खाते-खाते” /kʰa:te:-kʰa:te:/ also we can have a similar approach. Such verbal forms are morphologically marked for imperfective as is seen with the ‘-tV’ ending VGs in the following sentences:

E360	वह	खाते-खाते	बोला	
IPA	wəh	kʰa:te:-kʰa:te:	bol-a:	
Gl.	3SG	eat-IMPF.PCPL	speak-PFT.MSG	
EGl	He spoke while eating.			

Thus an imperfective tag ‘impf’ is quite suitable for them. And because they also do not inflect for any other GNPH and TAM, the tag of non-finite ‘nfn’ would fit with them. Thus they can also be identified based on the tags they have. These imperfective participles would also have only two tags of ‘impf’ and ‘nfn’.

Similar treatment can be done to the following occurrences of the imperfective form of the verbs. Though this is different from the construction we found in the previous example sentence.

E361	अपने	सिर	पर	लगाते	नजर	आए	
IPA	əpne:	sɪr	pər	ləga:-te:	nəɟər	a:-e:	
Gl.	own	head	on-LOC	apply-IMPF.PL	view	come-PFT.MSG	
EGl	was seen applying on his/her/their head.						

E362	अपने	बिगड़ते	स्वास्थ्य	की	वजह	से
IPA	əpne:	bɪgəɖ-te:	swa:stʰjə	ki:	vəɟəh	se:
Gl.	own	worsen-IMPF.PL	health	GEN	cause	ASS

EGl … due to his/her/their worsening health …

Here these verbs appear as modifying a noun phrase or a verb group.

4.4.5.6 Verbal Nouns vs. Infinitives

In the IL-PoST framework, we have a category of verbal nouns. This category has been reserved for the cases where a '-nV' ending infinitive form of the verb is marked as noun (NV). For example in the following sentences

E363	राम	का	<u>कहना</u>	<u>है</u>	कि	वह	वहां	नहीं	था
IPA	ra:m	ka:	<u>kəhna:</u>	<u>hɛ:</u>	kɪ	wəh	wəhã:	nəhĩ:	$t^ha:$
Gl.	Ram	GEN	say.INF	be.PRS.SG	that	3SG	there	NEG	be.PST.MSG
EGl	… Ram has to say that he was not there.								

E364	राम	को	<u>कहना</u>	<u>है</u>	कि	वह	वहां	नहीं	था
IPA	ra:m	ko:	<u>kəhna:</u>	<u>hɛ:</u>	kɪ	wəh	wəhã:	nəhĩ:	$t^ha:$
Gl.	Ram	ACC	say.INF	be.PRS.SG	that	3SG	there	NEG	be.PST.MSG
EGl	… Ram has to say that he was not there.								

the word कहना /kəhna:/ in the first sentence is a noun (though one can arguably posit that is is also functioning as a verb) while in the second sentence it is an infinitive verb. However, at places in the corpus these occurrences have been found to be marked as infinitive e.g.

गैर-दलीय\JJ.0.0.dir आधार\NC.mas.sg.obl.gen का\PP.0.0.gen मतलब\NC.mas.sg.0.0 है\VM.0.sg.3.prs.sim.dcl.fin.n स्वतंत्र\JJ.0.0.dir प्रतिनिधियों\NC.mas.pl.obl.ins द्वारा\PP.0.0.ins चुनाव\NC.mas.sg.dir.0 जीतकर\VM.0.0.0.0.pft.0.nfn.0 स्थानीय\JJ.0.0.dir निकायों\NC.mas.pl.obl.loc ,\PU विधानसभाओं\NC.fem.pl.obl.loc या\CCD.0 राष्ट्रीय\JJ.0.0.dir एसेम्बली\NC.fem.sg.obl.loc में\PP.0.0.loc **<u>आना\VM.0.0.0.0.0.0.ifn.0</u>** ।\PU

One should note here that an infinitive cannot stand on its own as a VG. If it comes it has to be either a verbal noun or must have been used in an imperative sense (in which case it will have the 'imp' tag).

4.4.5.7 The Verbal Nouns: Should we have it?

What is called as the verbal noun is also possible in the compound verb constructions. For example in the following two sentences: एक दरार इसी प्रकार मानवीय

999	बांध में एक दरार [हो]VG.prs.sim.dcl जाने की देर [थी]VG.pst.sim.dcl , फिर तो मन की उमंगो ने खुद-ब-खुद काम करना शुरु [किया]VG.pst.sim.dcl ।
1871	इसी प्रकार मानवीय संवेदनाओं के पाखंड और लेखकीय उत्तरदायित्व का दिखावा करने वाले समाज में किसी सजग और चेतनशील व्यक्ति द्वारा लेखन को मजबूरीवश [चलाए]VG.pst.sim.dcl रखना क्या विरक्ति से भी अच्छा [हो सकता है]VG.prs.impf.abil ?

The two verb sequences of हो जाने /ho ja:ne:/ and चलाए रखना /cəla:e: rəkʰna:/ have a nominal function. However, both of these two verb sequences contain two verbal words where the first verb is the main verb while the second one is the auxiliary. Here the problem is that of annotation. The guidelines itself is not very clear for such constructions. One can mark either both the verbs as nominal verbs (NVs) or just the last ones. If the last ones are the only ones marked as NVs while the first ones are marked as a verbal category then the problem will be that of incorrect TAM tag assignment. This is the case with the two sentences (999 and 1871) given as examples above. The cases of compound verbs occurring as NVs would remain elusive as VGs until both the words are marked as verbal elements.

4.4.5.8 The *wala* Particle

The वाला /wa:la:/ particle can be attached to the /–ne:/ ending infinitive verbs. It can both be attached to this word as if it was a suffix and this particle can also stand on its own. Both the forms are considered valid constructions. This construction can then be used in two senses. One is when it is used as an agentivizer and can be considered as a noun or an adjective. In its agentivizer construction the वाला /wa:la:/ particle works mostly as a suffix, attaching with the /-nV/ending verb form. The construction in this sense has been covered in the annotation

guidelines that come with the LDC corpus. However, another sense of it is missing from this guideline document. This construction is also used to indicate a distinct mood— that of approximative. It denotes that the task is just about to begin. For example:

E365	गाड़ी	खुलने	ही	वाली	है
IPA	gaːɖiː	kʰʊlneː	hiː	waːliː	hɛː
Gl.	vehicle	open.INF	only.PRT	PRT	be.PRS.SG
EGl	The train is just about to start.				

In this sentence the VG contains both words खुलने /kʰʊlneː/ and है /hɛː/. The other two words ही /hiː/ and वाली /waːliː/ are just particles coming in between. The main verb here is खुलने /kʰʊlneː/ and not है /hɛː/ which is a VAUX.

As long as the VINF in the वाला /waːlaː/ constructions of the VG intent is marked as NV in the IL-PoST annotation scheme, it will not be marked as a VG in the tool built here. However, if a VG containing a *wala* particle comes in and all the verbal words before and after it are marked as verbs (VM or VAUX), then we need to capture this template. There are a total of 8 such non-CV templates that have been captured. These are template numbers: 16, 33, 54, 70, 184, 189, 202 and 211. The following are two examples of such sentences:

E104	वह	पढ़ने	वाला	होगा
IPA	wəh	pəɖʰneː	waːlaː	hoː-gaː
Gl.	3SG	read.INF	PRT	be-FUT.3MSG
EGl	He will be about to read.			

E366	पढ़ने	वाले	लड़के	को	बुलाओ

IPA	pəɖʰne:	wa:le:	ləɖke:	ko:	bʊlao
Gl.	read.INF	PRT	boy.OBL	ACC	call-IMP
EGl	Call the reading boy.				

In the first sentence above it is in approximative mood. This mood here is effected by the use of the *wala* particle. Note that this *wala* here is different from other particles. The other particles such as /hi:/, /to:/, bʰi:/, /nahĩ:/ etc. actually do not bring a change in the TAM expressed by the VG. This change is brought by this *wala* particle. And that is why they need to be treated differently. The current result of this and similar sentences does not give proper result. For example, the first sentence where the verb sequence पढ़ने वाला होगा /pəɖʰne: wa:l:a ho:ga:/ is treated as a single VG, gives the following result:

वह [पढ़ने वाला होगा]VG.fut.sim.oblg ।

This is incorrect. Because the correct TAM tag for this VG is assigned to be “VG.fut.sim.apprx”. This error has occurred because at present the *wala* inside a VG is treated just as a particle. That means it is ignored if it comes inside a VG and the VG is treated as if it were “पढ़ने होगा”. The first word here matches with the VINF definition and the second word is ho_fut. Thus it matches the template VINF+ho_fut and the resultant TAM tag “VG.fut.sim.oblg” is assigned to it. This of course is incorrect. To resolve this we will have to give the *wala* an equivalent treatment as the verbs. The tool at present cannot handle this and will keep on giving the wrong results until we modify the tool to account for the *wala* as part of the VG.

4.4.5.9 The word *lekər*

The word लेकर /lekər/ looks like it is a conjunct participle made of two verbal words ले /le:/ and कर /kər/. However, it also functions many a times as a postposition. While annotating, it should be noted and its postpositional occurrences should be marked as postpositions only. However, in the LDC corpus, in most of the cases this has been marked as verb which is wrong and such occurrences creates problems in identification.

4.4.5.10 The word *cəlte*

The word चलते /cəlte:/ is though a clear verbal verb, it also functions as a conjunction as in the following sentence:

> संस्था\NC.fem.sg.obl.gen की\PP.fem.sg.gen उदासीनता\NC.fem.sg.obl.gen के\PP.0.0.gen चलते\VM.mas.sg.3.pst.pft.dcl.fin.n

The word चलते /cəlte:/ in the sentence extract from the LDC corpus has been wrongly marked as a VM. Such occurrences should also be annotated properly or it will result into a wrong VG marking.

4.4.5.11 The Use of the Word *ho*

The word हो /ho:/ is used in two senses. One is the simple second person auxiliary tense marking. Another use is that of the optative and counterfactual. In the latter cases, the optative and counterfactual markings should be made clear to differentiate between the temporal and modal use of the word. This should be marked clearly on the verbal tags under the modal slot.

Identifying a prs_aux हो /ho:/ and optative or counterfactual हो /ho:/

While counterfactual mood is many a times expressed periphrastically through the use of conditional words, in some constructions it is expressed through the use of the word हो/हों /ho:/hõ:/ inside VG construction itself. Thus constructions like the following:

E367	पढ़ी	जा	रही	हो
IPA	pəɖʱ-i:	ɟa:	rəhi:	ho:
Gl.	read-PFT.FSG	go.VR	live.PFT.FSG	be.CFCT
EGl	… should be getting read.			

E368	खा	रहा	हो	
IPA	ja:	rəh-i:	ho:	
Gl.	go.VR	live.PFT-FSG	be.CFCT	
EGl	… should be getting eaten.			

E369	नियुक्ति	की	जानी	हो	तो
IPA	nɪjʊktɪ	ki:	ja:ni:	ho:	to:
Gl.	appointment	do.PFT.FSG	go.INF	be.CFCT	then
EGl	… if the appointment has to be done then …				

are counterfactual in the sense of their interpretation. At present, it is being marked as if it were a value of the word variable prs_aux (in agreement with the 2nd person singular subjects). Here the problem is that of homonymy. The same auxiliary verb is coming in two senses. Even though we have assigned two word variables with the values of हो/हों /ho:/hõ:/ to a separate word variable of ho_cfct, the template selection is random. For the sentences like the following:

अगर\CSB.n वह\PPR.0.sg.3.dir.0.n.n.dst.n किताब\NC.fem.sg.dir.0
पढ़ी\VM.fem.sg.1.prs.pft.dcl.fin.n जा\VAUX.0.0.0.0.0.0.nfn.0
रही\VAUX.fem.sg.3.0.pft.dcl.fin.n हो\VAUX.0.0.0.prs.0.opt.fin.0

where the VAUX हो is actually in optative form, the template VR_pft+ja+rah_pft+prs_aux applies before the template “VR_pft+ja+rah_pft+ho_opt” could apply. To resolve this, a change in the logic has to be applied so that the tags of the VAUXes are also taken into account. Please note that the हो /ho:/ which is a value of the word variable prs_aux will not have the tag of opt/cfct in the modal slot. And thus we can resolve this by making the word variable have a specific word with a specific tag as well.

4.4.6 Second Run of the Tool with the LDC Corpus

After the changes we made in the LDC corpus we ran the tool again and the following tables shows the results achieved now.

Description of the Results Parameters	Raw Number	Percentage
Sentences with at least one unknown Verb Group found	112	2.31
Sentences with at least one orphan auxiliary verb found	0	0
Total number of sentences with at least one orphan auxiliary verb found	0	0
Total number of sentences having either at least one unknown VG or at least one orphan VG or both	112	2.31
Total Number of Sentences having no issues	4839-140=4699	97.69
Total Number of Sentences that has no VGs	508	13.23
Total Instances of Verb Groups Found	8226	NA
Incorrect Marking of the VGs	Cannot be ascertained (Needs manual checking)	

Table 51: Results of the Second Run with the LDC Corpus

If we compare the results of this second run as seen in Table 51 above with the first run as shown in Table 46 previously, we find that we have achieved a great lot of improvement. The error rate is a little above 2%, an improvement of 19% from the first run. That is we achieved the identification accuracy of around 98%. The number of sentences that do not have a VG is now 508, down by more than 700. These 508 sentences actually do not have a VG- which means they are either titles, or phrase fragments coming without a VG. The number of VGs identified has also increased by 827. The total number of sentences has also come down from previously 4853 to now 4839, a decrease of 14 sentences. During the course of checking, we found wrong line breaks which actually spread a single sentence onto two lines. However, I must note that I have also seen instances of lines where a single line contains more than one sentences. An instance would be the following sentence (CID 2514) which appears on one line but has more than one sentence here:

आय\NP.fems.sg.obl.gen कर\NP.fems.sg.obl.gen अधिनियम\NP.fems.sg.obl.gen का\PP.mas.sg.gen अध्याय\NC.mas.sg.dir.0 10\RDX दो\JQ.0.sg.dir.n.crd संबद्धों\NC.mas.pl.obl.gen के\PP.0.0.gen बीच\NST.dir.0 या\CCD.n नज़दीकी\NC.fem.sg.dir.0 संबंध\NC.mas.sg.dir.0 रखने\NV.dir.0 वाले\CX.n लोगों\NC.mas.pl.obl.gen के\PP.0.0.gen बीच\NST.dir.0 किसी\DAB.sg.obl.n.0 अनिवासी\NC.fem.sg.obl.soc के\PP.0.0.gen साथ\PP.0.0.soc

संव्यवहारों\NC.mas.pl.obl.ins से\PP.0.0.ins उद्भूत\JJ.0.0.dir होने\NV.dir.0 वाले\CX.n कर\NC.fem.sg.obl.gen के\PP.0.0.gen परिवर्जन\NC.mas.sg.obl.soc से\PP.0.0.soc संबंधित\JJ.0.0.dir विशिष्ट\JJ.0.0.dir प्रावधान\NC.mas.sg.dir.0 उपलब्ध\JJ.0.0.dir कराता\VM.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n .\PU यह\DAB.sg.dir.n.prx संव्यवहार\NC.mas.sg.dir.0 ऐसे\DAB.pl.obl.n.prx लोगों\NC.mas.pl.obl.gen के\PP.0.0.gen बीच\NST.dir.0 आय\NC.mas.sg.obl.gen या\CCD.n आस्तियों\NC.fem.pl.obl.gen के\PP.0.pl.gen अंतरण\NC.mas.sg.obl.gen के\PP.0.0.gen परिणाम\NC.mas.sg.obl.loc में\PP.0.0.loc हो\VM.0.0.0.0.0.0.nfn.0 सकता\VAUX.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n .\PU ‘‘आर्म्स\NP.mas.sg.obl.gen लेंग्थ\NP.mas.sg.obl.gen पर\NP.mas.sg.obl.gen संव्यवहार\NP.mas.sg.obl.gen ’’\PU के\PP.0.pl.gen प्रमुख\JJ.0.0.dir सिद्धांतों\NC.mas.pl.obl.gen के\PP.0.pl.gen माध्यम\NC.mas.sg.obl.ins से\PP.0.0.ins ऐसे\DAB.pl.obl.n.prx नज़दीकी\NC.fem.sg.dir.0 रूप\NC.mas.sg.obl.soc से\PP.0.0.soc संबंध\NC.mas.sg.dir.0 रखने\NV.dir.0 वाले\CX.n लोगों\NC.mas.pl.obl.ins या\CCD.n संबद्ध\JJ.0.0.dir उद्यमों\NC.mas.pl.obl.ins द्वारा\PP.0.0.ins कर\NC.fem.sg.obl.gen के\PP.0.pl.gen परिवर्जन\NC.mas.sg.obl.gen का\PP.mas.sg.gen अध्ययन\NC.mas.sg.dir.0 करने\NV.obl.x के\PP.0.0.gen लिए\PP.0.0.x एक\JQ.0.sg.dir.n.crd समिति\NC.fem.sg.obl.gen की\PP.fem.sg.gen स्थापना\NC.fem.sg.dir.0 की\VM.fem.sg.3.pst.pft.dcl.fin.n गई\VAUX.fem.sg.3.pst.pft.dcl.fin.n थी\VAUX.fem.sg.3.pst.sim.dcl.fin.n .\PU इस\DAB.sg.obl.n.prx समिति\NC.fem.sg.obl.gen की\PP.fem.sg.gen अनुशंसाओं\NC.mas.pl.obl.gen के\PP.0.0.gen आधार\NC.mas.pl.obl.loc पर\PP.0.0.loc ,\PU आय\NC.mas.sg.dir.0 कर\NC.fem.sg.obl.gen के\PP.0.0.gen अंतर्गत\NST.dir.0 आर्म्स\NP.mas.sg.dir.0 लेंग्थ\NP.mas.sg.dir.0 वाले\CX.n संव्यवहारों\NC.mas.pl.obl.soc से\PP.0.0.soc संबंधित\JJ.0.0.dir सिद्धांतों\NC.ma.pl.obl.gen की\PP.fem.sg.gen मान्यता\NC.fem.sg.obl.gen के\PP.0.pl.gen कारण\NC.mas.sg.dir.0 वित्त\NP.mas.sg.obl.gen विधेयक\NP.mas.sg.obl.gen 2001\RDF के\PP.0.0.gen माध्यम\NC.mas.sg.obl.ins से\PP.0.0.ins संशोधन\NC.mas.sg.dir.0 प्रस्तावित\JJ.0.0.dir किए\VM.mas.pl.3.prs.pft.dcl.fin.n गए\VAUX.mas.pl.3.prs.pft.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n .\PU आय\NP.fem.sg.obl.gen कर\NP.fem.sg.obl.gen अधिनियम\NP.fem.sg.obl.gen की\PPfem.sg.gen धारा\NC.fem.sg.obl.acc 92\RDF और\CCD.n 93\RDF को\PP.0.0.acc पुन:\AMN.dir.0.n प्रारूपित\JJ.0.0.dir किया\VM.mas.sg.3.prs.pft.dcl.fin.n गया\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n .\PU

I have not checked all of such instances and also did not try to correct them and bring to one sentence per line as this does not seem to hamper the performance of the VG Finder tool.

4.4.6.1 Unknown Phrases Found in the Second Run

There are still a total of 112 sentences which have at least one VG marked as unknown. We will talk about error type of these occurrences in this section and see if we can resolve them. The following table summarizes the unknown VGs found in the second run:

Type of Problem	No. of Instances Found
Perfective Forms of Verbs Playing as Nouns	29
The verbal word contains a punctuation mark	23
Compound Verbs Functioning as Verbal Nouns	21
The VG contains a word which is confusing and marked wrongly	17
Spelling/Grammatical Errors in the VG	8
Verbal Nouns with a Template VR_pft+VINF	4
Template Missing	2
VG Conjugated Midway with a Punctuation Mark	1
A CV with the first form as pft	1
Two VGs marked as one	1
A CV with three words	1
Construction of Conjunct Participle with के /ke/ Instead of कर /kər/	1
A CV Conjuct participle construction	1
NP/JJ of the Conjunct Verb marked as VM	1

Table 52: Summary of Unknown VGs Found in the Second Run

4.4.6.1.1 Perfective Forms of Verbs Playing as Nouns

There are a few verbs in Hindi which though appear in their perfective form but play as noun and create a conjunct verb. Consider the following sentences coming from the output file:

C.ID.	Example Output Sentence in the Second Run
1863	[[AT LEAST ONE UNKNOWN VG FOUND]] कुछ-कुछ घबराहट [दिखाई दे रही है]UNKNOWN ।
1916	[[AT LEAST ONE UNKNOWN VG FOUND]] वे फिर [कहते हुए]VG.ip.impf [दिखाई देते हैं]UNKNOWN - बेटी [मारोगे तो]VG.fut.sim.dcl अनर्थ [हो जाएगा]VG.cv_fut.sim.dcl ।
2551	[[AT LEAST ONE UNKNOWN VG FOUND]] मैंने अपने मकान को रात के समय लालसा-भरी आंखों से [देखा]VG.pst.sim.dcl - दरवाज़े पर दो लालटेनें [जल रही थीं]VG.pst.prog.dcl और बच्चे इधर-उधर [खेल रहे थे]VG.pst.prog.dcl , हर सफ़ाई और सुथरापन [दिखायी देता था]UNKNOWN ।
4096	[[AT LEAST ONE UNKNOWN VG FOUND]] जिनके हाथों में क्रमशः धनुष , गदा , भाला एवं लड्डू [दिखाई देता है]UNKNOWN ।
789	[[AT LEAST ONE UNKNOWN VG FOUND]] ' अचानक गीती भाभी का धीमा और दृढ़ स्वर [सुनाई

	दिया]UNKNOWN , ' हां हां [कहो न]VG.0.0.imp , तब से चुपचाप , स्क्रिप्ट पर नजर [गड़ाए]VG.pst.sim.dcl [बैठी हो]VG.prs.pft.dcl ।

In the example sentences above we find that the words दिखाई /dɪkʰaːiː/ and सुनाई /sʊnaːiː/ appear in a VG. These words of course are in their perfective forms, which make them finite verbs. However, in these sentences they function as nouns and make a conjunct verb with the next verb दे /deː/ which comes in its auxiliary form. However, these verbs are marked often as verbs.

Also of note is that only दे /deː/ can come as the V2 with these two forms of the verb. It appears that it is more of an idiomatic construction and cannot be part of a general template. For example the following sentences are wrong:

*E370	सुलाई	दे	रहा	है	
IPA	sʊlaiː	deː	rəhaː	hɛː	
Gl.	sleep-PFT.FSG	give.VR	live.INF	be.PRS.SG	
EGl	is letting (someone) sleep.				

*E371	जम्हाई	दे	रहा	है	
IPA	jəmhaːiː	deː	rəh-aː	hɛː	
Gl.	doze-PFT.FSG	give.VR	live.INF	be.PRS.SG	
EGl	… is letting (someone) doze off.				

*E372	पढाई	दे	रहा	है
IPA	pəɖʰaːiː	deː	rəh-aː	hɛː

Gl. read-PFT.FSG give.VR live.PFT.MSG be.PRS.SG

EGl … is letting (someone) read.

This means that we cannot make a general template for such constructions if they are marked as verbs. There are two solutions for such cases. Either we annotate these words as nouns or we create a template specifically covering these verbs (i.e. make a template with all the variables as word variables). This second option is not very promising though as it will involve creating more than half a dozen templates for each of such words. Besides, we know for sure that these words are functioning as nouns. So, if we just mark these verbs as nouns and the consequent V2s as VMs, we will have these VGs identified and tagged properly.

4.4.6.1.2 The Verbal Word Contains a Punctuation Mark

There are quite a lot of verbal words flanked by a punctuation mark either side. See for example:

C.ID.	Example Output Sentence in the Second Run
1456	[[AT LEAST ONE UNKNOWN VG FOUND]] भारतवर्ष की गौरवमयी , महिमामंडित संत परंपरा में एक श्लाघनीय नाम स्वर्णाक्षरों में उल्लिखित [है-]UNKNOWN भारत ज्योति , अणुव्रत अनुशास्ता आचार्य श्री तुलसी ।
1595	भावप्रकाश निघण्टु में [लिखा है-]UNKNOWN
2524	ऐसी एक या अधिक विधियों का संयोग भी [हो सकता है.]UNKNOWN
3040	आयोग के वचारार्थ वषय इस प्रकार [हैः-]UNKNOWN
3056	आयोग के वचारार्थ वषय इस प्रकार [हैः-]UNKNOWN
3196	[[AT LEAST ONE UNKNOWN VG FOUND]] वे [कहते हैं-]UNKNOWN अभिनय रचनात्मकता से जुड़ी विधा [है]VG.prs.sim.dcl , नकल से नहीं ।
3211	इस जोशीले अभिनेता के लिए शेर मुलहिजा [है-]UNKNOWN
3236	भाव प्रकाश निघण्टु में [लिखा है-]UNKNOWN

We had corrected quite a lot of other occurrences earlier. Yet some of such occurrences remained. Please note that it is the task of the pre-processor to remove occurrences of punctuation marks from the words before they are processed. If such punctuation marks are still present, they will be considered as part of the verbs and would not be processed properly. Our word variables have the actual words and there will be a problem of pattern matching because of the punctuation marks and such words will not be recognized as a value of the word variables.

Once we remove the punctuation marks from within these verbal verbs, the identification will occur as expected.

4.4.6.1.3 Compound Verbs Functioning as Verbal Nouns

As said in section 4.4.5.7 the cases of compound verbs occurring as NVs would be either marked wrongly or would not be marked at all. The following examples are the cases which are not identified and marked as unknown:

1074	[[AT LEAST ONE UNKNOWN VG FOUND]] मेरे लिए यह [डूब]UNKNOWN मरने की जगह [है]VG.prs.sim.dcl कि मैं अपनी बीवी से इस तरह की बातें [करूं]UNKNOWN , उसकी ओर से मेरे दिल मे संदेह [पैदा हो]VG.cv_0.0.imp ।
1554	[[AT LEAST ONE UNKNOWN VG FOUND]] सुरक्षाबलों को चोट [पहुँची]VG.pst.sim.dcl आतंकवादियों की गतिविधियों के कारण , जिन्हें वे [रोक]UNKNOWN पाने में सक्षम तो [हैं]VG.prs.sim.dcl मगर शस्त्र विराम की बेड़ियाँ पाँव में होने से वे ऐसा कोई कदम न उठाने को मजबूर [हैं]VG.prs.sim.dcl , जिससे शस्त्र विराम ' बदनाम ' [होता हो]VG.prs.impf.dcl ।
1681	[[AT LEAST ONE UNKNOWN VG FOUND]] परंतु हाँ , मेरी आम की गुठली में बना चित्र देखने के बाद , कोई व्यक्ति आम खाने के बाद गुठली [फेंक]UNKNOWN देने के पूर्व [रुककर]VG.cp.pft यदि एक क्षण के लिए भी उसे गौर से [निहारे तो]VG.pst.sim.dcl मैं मेरी आंशिक सफलता [मान सकता हूँ]VG.prs.impf.abil ।
1682	[[AT LEAST ONE UNKNOWN VG FOUND]] परंतु मेरी टूटी [हुई]VG.pst.sim.dcl चप्पल की ' मोनालिसा ' देखने के बाद अपनी टूटी चप्पल यदि कोई मेरे ही पास [भेज दे तो]VG.cv_0.0.imp फिर सिवाय माथा [ठोक]UNKNOWN लेने के मैं [कर भी क्या सकता हूँ]VG.prs.impf.abil ?
382	[[AT LEAST ONE UNKNOWN VG FOUND]] काफी समय [गुजर]UNKNOWN जाने के बावजूद देवता प्रकट न होने के कारण मायूस व हताश [होकर]VG.cp.pft वह व्यक्ति अपने साथियों के साथ वापस [लौट आया]VG.cv_pst.pft.dcl और बीच राह में ही उसे अपने वे साथी [मिल गए]VG.cv_pst.pft.dcl जिन्होंने उसे देवता प्रकट होने की बात [कही थी]VG.pst.pft.dcl ।

These cases will remain elusive until both the verbs are marked as verbs.

4.4.6.1.4 The VG contains a word which is confusing and marked wrongly

There are a few words which might be confusing to the annotators while annotating. Consider the following examples:

876	[[AT LEAST ONE UNKNOWN VG FOUND]] ' हां क्यों नहीं इन तीन दिनों में आपको चाय के साथ बिस्किट ऑफर [किए तो]VG.pst.sim.dcl आपने **[मना** कर दिया]UNKNOWN हमेशा ।

1592	[[AT LEAST ONE UNKNOWN VG FOUND]] शहरों के बाग-बगीचों के बाहर और खुली जगह में [**पैदा**]UNKNOWN होने वाले पौधों में एक पौधा [होता है]VG.prs.impf.dcl अपामार्ग का जिसे बोलचाल की भाषा में आँधीझाड़ा भी [कहते हैं]VG.prs.impf.dcl ।
2456	[[AT LEAST ONE UNKNOWN VG FOUND]] उन्होंने इन युवाओं का आह्वान] करते हुए[VG.ip.impf [कहा[VG.pst.sim.dcl क भारत के चकत्सकों ने पूरे वव में अपनी वोषज्ञता और सेवा भावना की] **धाक** जमाई है[UNKNOWN ।
3034	[[AT LEAST ONE UNKNOWN VG FOUND]] उन्होंने [कहा[VG.pst.sim.dcl - इस स्थिति में स्थगन उचित नहीं] होगा[VG.fut.sim.dcl , लेकिन हम याचिका में यह अनुरोध] जोड़ सकते हैं[VG.prs.impf.abil कि अर्जुन तथा अन्य राष्ट्रीय पुरस्कारों के चयन हेतु उचित दिशा-निर्देश] **तय** किए जाएँ[UNKNOWN ।
3321	[[AT LEAST ONE UNKNOWN VG FOUND]] ये यहाँ के बरसाती जंगल में जमीन पर पड़ी पतियों को]**चट** करते हैं[UNKNOWN ।
3950	[[AT LEAST ONE UNKNOWN VG FOUND]] अब अपने बेटे अमित को] **लाँच** कर रहे हैं[UNKNOWN ।
920	[[AT LEAST ONE UNKNOWN VG FOUND]] क्षणभर के लिए मेरे मन में जैसे गुस्से का] **उबाल** आ गया[UNKNOWN , लेकिन मैंने तुरंत उसे [दबा दिया[UNKNOWN ।

In the examples above we find the words like मना /məna:/ “forbid”, पैदा /pɛ:da:/ “get born”, धाक /dʰa:k/ “influence”, तय /təj/ “decide”, चट /cəʈ/ “lick away”, लाँच /lɔ̃:c/ “launch” and उबाल /ʊba:l/ “boil” coming as verbs. In fact these words though look like a verb and occur as part of the verb. But these words are part of the conjunct verbs. All of these words are in fact nouns and should be treated as nouns. These words are more like idiom. We cannot treat them like a verb because they do not change their forms like the verbs in Hindi do. Treating them as nouns or adjective would at least make them a part of the conjunct verb and they can be considered as a verb when the conjunct verbs are processed. This work does not take conjunct verbs into its purview and thus they are not identified for the reasons stated earlier. For this reason, these examples will remain unhandled at present.

4.4.6.1.5 Spelling/Grammatical Errors in the VGs

If a VG contains a grammatical or spelling error, especially for the parts of the VGs which are identified from the list of word variables, the VGs will be either marked as unknown or wrongly marked. Consider the following examples:

1746	[[AT LEAST ONE UNKNOWN VG FOUND]] वैज्ञानिक अपने तरीके से , चित्रकार अपने तरीके से तो

	अन्य विधा के धनी अपने-अपने तरीके से प्रकृति को समझने का प्रयास [कर रहें हैं]UNKNOWN और प्रकृति भी अपना पूरा सहयोग [दे रही है]VG.prs.prog.dcl ।
3131	श्री फातमी ने [कहा]VG.pst.sim.dcl क जब तक वकसत वव , वकासाील दाोेें को सर्वाक्षा में अधक सहायता प्रदान करने के प्रत वचनबद्ध नहीं [होता]VG.pst.impf.cgnt , डाकर समझौते के लक्ष्यों को पूरी तरह प्राप्त नहीं [कया जा]UNKNOWN
3133	[[AT LEAST ONE UNKNOWN VG FOUND]] केन्द्रीय मानव संसाधन वकास मंत्री श्री अर्जुन सिंह ने भारतीय राष्ट्रीय सहयोग आयोग के अध्यक्ष के रूप में कार्यभार ग्रहण [कर लया है]UNKNOWN , जबक श्री बी.एस.बासवान को आयोग का नया महासचव नयुक्त [कया गया है]VG.pas_prs.pft.dcl ।

In the first example above (1746) we see that a verb group contains the sequence कर रहें हैं /kər rəhẽ: hɛ̃:/. While the second verb in this VG i.e. रहें (which is an imperative construction) is quite correct in terms of spelling, its conjugation with the other verbs is ungrammatical. That is, there is no template in Hindi that can cover this VG because it is simply ungrammatical. For the second example sentence 3131 the VG marked as unknown has both a spelling error plus the VG is missing something here and is thus ungrammatical. The spelling error here however would not affect the VG identification because it is a main verb and it would be identified based on the tag variables. However, because the whole of this VG does not make any sense and will not match any of the templates defined, it will still be marked as an unknown VG. The third example sentence 3894 above has received an unknown VG marking simply because of a spelling error. Here the second word of VG marked as unknown 'लया' is missing a diacritic and should have been लिया. Once we make this correction, the VG gets identified.

4.4.6.1.6 Verbal Nouns with a template VR_pft+VINF

There are a few instances of verb groups where a gerundive/nominal use of the compound verb consists of a template like VR_pft+VINF. That is, the V1 comes in the perfective form while the second verb occurs in the infinitive form. Together they function as a gerund. Because gerunds are called verbal nouns in the IL-PoST framework and marked as nouns, the care here should have been taken to mark either both the verbs as nouns or both of them as verbs. This does not happen in many of the cases. Some examples as follows:

1852	[[AT LEAST ONE UNKNOWN VG FOUND]] इस ऊर्जाहीनता की स्थिति को विरक्ति का नाम देना सिर्फ़ अपने अस्तित्व को [बनाए]UNKNOWN रखने की छटपटाहट से अधिक कुछ नहीं [कहा जा सकता है]VG.pas_prs.impf.abil ।
1876	[[AT LEAST ONE UNKNOWN VG FOUND]] कुछ के मन में तत्कालीनता में मंदिर [तोड़कर]VG.cp.pft बाबर द्वारा मस्जिद [बनाए]VG.pst.sim.dcl जाने की कुंठा [है]VG.prs.sim.dcl , तो किसी के मन में मस्जिद [तोड़कर]VG.cp.pft धर्मनिरपेक्षता के स्वरूप को [ढहाए]UNKNOWN जाने की कुंठा घर [कर रही है]VG.prs.prog.dcl ।
2271	[[AT LEAST ONE UNKNOWN VG FOUND]] उपभोक्ताओं को जल की आपूर्ति सामान्य रूप से प्रभावकारी मांग सिद्धान्त पर आधारित [होनी चाहिए]VG.0.0.sugg और यह मांग मोटे तौर पर उस सेवा के स्तर के अनुरूप [होनी चाहिए]VG.0.0.sugg जिसे उपभोक्ता एक समुदाय के रूप में चलाने , [बनाए]UNKNOWN रखने और वित्त-पोषित करने के लिए तैयार [हों]VG.pst.sim.dcl ।

Like other cases of compound verbal nouns, these cases would also be left as unknown for the lack of clear guidelines at the PoS level.

4.4.6.1.7 Template Missing

I also found couple of VGs the corresponding templates for whom were missing. The following two are the ones:

4347	[[AT LEAST ONE UNKNOWN VG FOUND]] 4.3 सामान्य परम्परा यह [है]VG.prs.sim.dcl कि प्रोन्नति पर स्थानांतरण [होता है]VG.prs.impf.dcl . कुछ मामलों में इससे संबंधित व्यक्तियों को कठिनाई [हो सकती है]VG.prs.impf.abil . अत: इस मामले को नियुक्ति समिति के निर्णय पर [छोड़ देना चाहिए]UNKNOWN . इस संदर्भ में 'जेएजी (JAG) ' प्रदान जाने को प्रोन्नति नहीं [माना जाएगा]VG.pas_fut.sim.dcl .
2392	[[AT LEAST ONE UNKNOWN VG FOUND]] कुछ विशेषज्ञों का मानना [था]VG.pst.sim.dcl कि पश्चिमी तट से [हटाए गए]VG.pas_pst.pft.dcl यहूदियों के कारण इतना क्रोध [फैलेगा]VG.fut.sim.dcl कि इसराइली मतदाता शेरोन की पार्टी को [हरा देंगे]UNKNOWN लेकिन कदीमा का सबसे बड़ी पार्टी [बनकर]VG.cp.pft उभरना यह सिद्ध [करता है]VG.prs.impf.dcl कि इसराइल की जनता खुद यह [मानने लगी है]VG.prs.pft_ince.dcl कि फिलिस्तीनियों की ज़मीन पर ३८ साल से [चला आ रहा]UNKNOWN उसका कब्जा अवैध और अनैतिक [है]VG.prs.sim.dcl तथा यह अशांति और युद्ध का सबसे बड़ा कारण [है]VG.prs.sim.dcl ।

For the first sentence above (4347), the verb group is “छोड़ देना चाहिए” /cʰoɖ̤ de:na: ca:hɪ jɛ:/. This VG corresponds to the abstract template A62 which says “V1+V2_inf+cahiye”. It should have been realized as one of these templates but somehow we missed (I have not actually used all the possible realizations of all the abstract templates yet while testing this corpus. Only the most common ones were utilized). So, we need to add a template as a realization of the abstract template A62 and add define a template as “VR+de_inf+cahiye”. This will handle the VG missed above.

The other example sentence (2392) above has a VG which in fact is a durative and perfective participle together. We have not defined any templates for this structure yet. So, a template as VR_pft+a+rah_pft should work for this VG. To capture the unknown VGs of the two sentences above, we can define the two VG templates as below and add them into the template definition file.

VR+de_inf+cahiye	VG.cv_0.0.sugg
VR_pft+a+rah_pft	VG.ip.pft_dur

4.4.6.1.8 VG Conjugated Midway with a Punctuation Mark

There may be instances of VGs where two main verbs of what appears like a single VG can share the same occurrence of the VAUXes. The following example should illustrate this.

696	[[AT LEAST ONE UNKNOWN VG FOUND]] जैसे किसी बच्ची को [पुचकार]UNKNOWN , [दुलार रहे हों]VG.0.prog.cfct - ऐसे शब्द और आवाजें मेरे कानों में [पड़ रही थीं]VG.pst.prog.dcl और मेरा पूरा शरीर [पिघल रहा था]VG.pst.prog.dcl ।

In such cases only one main verb and the closest one to the VAUXes will be identified correctly. The other conjoined VMs would be left alone and may be either marked wrongly or marked as unknown. Please note that all the occurrences of VMs are identified as a VG. If no templates match with this VM and its VAUXes, it is marked as ‘unknown’.

4.4.6.1.9 A CV with the First Form as ‘pft’

Consider the following example:

4002	[[AT LEAST ONE UNKNOWN VG FOUND]] पर्यटन मंत्री श्री अनूप मिश्रा के बंगले पर तैनात सुरक्षाकर्मी तब हतप्रभ [रह गए]VG.cv_pst.pft.dcl जब विवाह में शामिल होने के लिए प्रधानमंत्री श्री अटलबिहारी वाजपेयी बंगले से सड़क तक पैदल ही [चले आए]UNKNOWN ।

In the verb group चले आए /cəle: a:e:/, the first verb is in a perfective form. This gives us a template like VR_pft+a_pft. This is not a common pattern of template. For example, the following VGs on this pattern are grammatically wrong:

*रोए आए /ro:e: a:e:/

*मारे आए /ma:re: a:e:/

*खाए आए /kʰa:e: a:e:/

This means that this pattern is lexicalized in some ways. Semantically it has the same sense as that expressed by the template VR+a_pft. However, we cannot always expect an annotator to mark the word चले /cəle:/ as a verb root because it is not a verb root. So, what we can do here is to just define a template particularly to capture this VG:

cal_pft+a_pft VG.cv_pst.pft.dcl

4.4.6.1.10 Two VGs marked as one

There are some instances when two VGs can come serialized one after the other with no intervening non verbal word coming in between. For example in the following sentence the two VGs are किये /kɪe:/ and बैठा था /bɛːʈʰa: tʰa:/. It has been marked wrongly as a single VG.

1097	[[AT LEAST ONE UNKNOWN VG FOUND]] मैं तो अपनी किस्मत का फैसला पहले से ही [किये बैठा था]UNKNOWN ।

4.4.6.1.11 A CV with three words

In the second chapter under section 2.4.1 while enumerating the V2s occurring in the compound verbs, we had listed out a few V2s which can occur in double words (Table 17). One of those V2s was खड़ा करना /kʰəɖa: kərna:/. The VG identified as ‘unknown’ in the example sentence

below uses the root form of this double word V2. However, we also find that if we want to capture the VG as shown here "ला खड़ा कर दिया" /la: kʰəɖa: kər dɪja:/, the existing templates would not fit in. For a template to capture this VG, the template definition has to be V1+V2+de_pft where the V1 stands for ला /la:/ and V2 stands for खड़ा कर /kʰəɖa: kər/. We find that we do not have any such template and this appears to be unique and idiomatic VG where perhaps all of the word has to be defined by word variables. One may however think of the first word in this example VG as a tag variable but I cannot find evidence than any word other than ला /la:/ can stand at this position. So, the best way possible would be to define a template just to capture this very VG through word variables:

laa+khada+kar+de_pft VG.cv_pst.pft.dcl

1505	[[AT LEAST ONE UNKNOWN VG FOUND]] सच यह [है]VG.prs.sim.dcl कि शस्त्र विराम ने कश्मीर को ऐसे रास्ते पर [**ला खड़ा कर दिया**]UNKNOWN , जहाँ सत्ताधारी नेशनल कॉन्फ्रेंस का भविष्य तो दाँव पर [लग ही गया]VG.cv_pst.pft.dcl , कश्मीर में बढ़े मौत के आँकड़ों से सुरक्षाबल भी हतोत्साहित [हुए]VG.pst.sim.dcl , क्योंकि हाथ [बाँध]VG.pst.sim.dcl उन्हें आतंकवादियों के शिकार के लिए [छोड़ गया]VG.cv_pst.pft.dcl ।

4.4.6.1.12 Construction of Conjunct Participle with *ke* Instead of *kər*

Conjunctive participle is usually constructed by adding the word कर /kər/ as a suffix to the verb root or juxtaposing it after the verb root. This type of construction has already been handled in our framework as discussed earlier in this chapter. However, there is also another way of constructing the conjunctive participle by adding just के /ke/. This के /ke/is homonymous with the postposition के /ke:/. But the function of the latter is the same as that of कर /kər/ in making the conjunctive participle, although it is more informal and colloquial in use. Normally this के /ke:/ is used as a suffix as in खाके /kʰa:ke:/, सोके /so:ke:/ etc. In standard practice of writing this form is usually avoided (except in the case of करके /kərke/). When it comes suffixed with the verb root, it is easier to find out. However, we also see a few instances when this के /ke:/ appears separated from its verb root as shown in the sentence below.

4515	[[AT LEAST ONE UNKNOWN VG FOUND]] वह छूटते ही [बोली]VG.pst.sim.dcl , मम्मी , आप हर काम अपने मन से क्यों [करती हैं]VG.prs.impf.dcl , कम से कम फोन [कर के]UNKNOWN [पूछ तो लिया होता]VG.cv_prs.pft.cfct ।

To capture the conjunctive participles made with juxtaposing कर /kər/ we have defined a template (VR+kar) where the second verb is a word variable. We will have to do the same with this construction as well. Thus, we can define a template like:

VR+ke VG.cp.pft

Here we will also have to add the word variable 'ke:' into our word variables list with the value of के.

4.4.6.1.13 Nouns/Adjectives of the Conjunct Verbs Marked as VM

Sometimes we find that a noun or an adjective which are part of a conjunct verb are wrongly PoS marked to be VM. As shown in the example sentence below, this will hamper in identification of the VG and it will be either recognized as 'unknown' or given a wrong tag.

920	[[AT LEAST ONE UNKNOWN VG FOUND]] क्षणभर के लिए मेरे मन में जैसे गुस्से का [उबाल आ गया]UNKNOWN , लेकिन मैंने तुरंत उसे [दबा दिया]UNKNOWN ।

4.4.6.2 Incorrect Marking of VGs

Although we have achieved the VG identification accuracy of above 97%, it is not certain if all the VGs marked here are correct. This can be evaluated only by a manual checking at present. We checked the first 1000 sentences coming in the output file of this second run. The following table shows the summary of the results we found:

Type of Error	No. of Occurrences
Impf marked as pft	50
Fut contains pft resulting in wrong vg identification	28
CV NV containing ho as V1 marked wrongly	14
VR_pft working as JJ taken as VG	13
rah_opt selected over rah_pft denoting prog aspect	6
Single VG containing Two VMs	6
Conjunct Verb V2 marked as Noun resulting in half identification of VG	1

Marked Wrongly due to Missing Elements	1
NV use of VR_pft	1
One VG broken	1
Pft marked in an imperative construction	1
VINF+prs_aux shows obligative while it is not always so	1

Table 53: Incorrect VG Markings in the First 1000 Sentences of the LDC Corpus

We found a total of 125 sentences with at least one error in verb group identification. These errors can be categorized into two categories: errors generated due to erroneous annotation and errors in proper application of the VG template definitions.

4.4.6.2.1 Annotation Errors

Despite all the changes I made in the LDC corpus, there were still some errors left in the annotations of the verbal elements. These are the errors that make a difference in the VG tag allocation. The most numerous among these errors are the perfect marking of the instances of verbs which are morphologically marked for imperfect. Consider the following sentences:

430	इसकी दर्शनीयता [देखते ही]VG.pst.impf.cgnt **[बनती है]VG.prs.pft.dcl** ।
481	जब हम सातवीं से तेरहवी शताब्दी के स्थापत्य का पर्यवेक्षण **[करते हैं तो]VG.prs.pft.dcl** हम [पाते हैं]VG.prs.impf.dcl िक वह एक नये राजनीितक पिरस्िथित के अनुरुप [ढल जाता है]VG.cv_prs.impf.dcl ।
521	पहाड़ी ढ़लान , चढ़ाव और उतार को ध्यान में **[रखते हुए]VG.ip.pft** सारे नगर को टेढ़ा-मेढ़ा इस तरह [बसाया गया]VG.pas_pst.pft.dcl िक मध्यकालीन उदयपुर की योजना में कहीं रास्ते या चौपड़ की व्यवस्था नहीं [दीख पड़ती]VG.cv_pst.impf.cfct ।

In the first example output sentence (CID 430), we find that the second VG “बनती है” /bənti: hɛ:/ has been marked with a ‘pft’ in the aspect slot while the VG itself contains a word in the imperfect form. This happens because the main verb in this VG has been identified based on the tag variables and the tags of the first word contain the ‘pft’ instead of imperfective ‘impf’ in the aspect slot. Thus this annotation has caused the marking of the wrong TAM tag for this VG. Similar is the case for the other two example sentences given above.

Similar is the case with the VGs which are actually in future tense but have been marked as past. Consider the following examples:

196	जो १६ फरवरी महाशिवरात्रि तक **[चलेगा]VG.pst.sim.dcl** ।
203	राज्य सरकार मंगलवार से प्रतापगढ़ जिले के बदरी के निकट शारदा नहर से प्रितिदन ३०० क्यूसेक पानी गंगा में **[छोड़ेगी]VG.pst.sim.dcl** ताकि इलाहाबाद में उसका जलस्तर [बना रह]VG.0.pft_cnt.imp ।
218	यह क्रम शाम तक **[चलेगा]VG.pst.sim.dcl** ।

This problem has occurred because the tags of these words, which have been identified through the tag variables, contain a tag of perfective 'pft' in the aspect slot. Formally, if a verb is morphologically in the future tense form, it cannot have a perfective tag. The aspect slot for such forms should be marked as '0' as per the IL-PoST framework. Although these forms have been marked for the future tense with the 'fut' tag at the aspectual slot, the 'pft' tag overrides it and these VGs are marked as simple past. To get these corrected, we would need to correct the annotation of such instances.

Another type of error is generated in the cases when a single VG having more than one verbal word is marked as having more than one VM. Consider the following examples:

323	- श्री सलीम खान को कांग्रेस नशा विरोधी प्रकोष्ठ का कार्यकारिणी सदस्य **[बनाया]VG.pst.sim.dcl [गया है]VG.prs.pft.dcl** ।
548	रावल जैसल ने अपनी स्वतंत्र राजधानी स्थापित **[की]VG.pst.sim.dcl [थी]VG.pst.sim.dcl** ।
706	सालों-साल वंचित **[रही]VG.pst.sim.dcl [हूं]VG.prs.sim.dcl** मैं ।

As seen in the examples above, we find that a single VG like बनाया गया है /bəna:ja gəja: hɛ:/, की थी /ki: tʰi:/ and रही हूं /rəhi: hũ:/ have been separated and identified as two VGs. This has happened because VAUXes have been marked wrongly as VM. We have already noted that all the instances of VMs are considered as VG while all the instances of VAUXes must be attached to a VM. The annotations of the examples above do not follow this rule and thus results into two separate VGs just because we have templates that could match each of these separated VMs.

There are also other instances of VGs identified wrongly due to annotation anomalies. For example in the following sentence:

445	िडजाइनरों को हमारी व अगली पीढ़ी के बीच सेतु बनाना [है]VG.prs.sim.dcl ।

The word बनाना /bəna:na:/ should have been the part of the VG. But because it has been annotated as a verbal noun, it is not captured.

Similarly, in the following sentence:

154	लोस्वेट लोगो को संगम में स्नान [करते]VG.pst.sim.dcl [देख]VG.cp.pft इस कदर प्रभाि वत [हुई]VG.pst.sim.dcl की वह संगम में डुबकी लगाने से खुद को रोक नहीं [पाई]VG.pst.sim.dcl ।

the last VG marked is missing its main verb which actually is रोक /ro:k/ while the VG covered is actually a VAUX. This has happened because the word रोक /ro:k/ "stop' has been annotated as a noun and not as a verb.

When a verb form is marked for imperative or optative mood construction, it loses its aspectual value and should be marked as null. However, there are instances when a verbal word marked for imperative has also been marked for an aspectual value as seen in the last VG marked in the example output sentence below:

238	यही कारण [है]VG.prs.sim.dcl ि◌ंक दो के ि◌ंविभन्न ि◌ंहस्सो में तीर्थ [बनाए]VG.pst.sim.dcl और धर्म को यात्रा के साथ [जोड़ा गया]VG.pas_pst.pft.dcl ता ि◌ंक अिधक से अिधक संख्या में यात्री तीर्थ यात्राओ पर [ि◌ंनकलें]VG.pst.sim.dcl ।

As can be seen, this last VG here has been marked as a simple past construction while it should have been simply marked as an imperative.

4.4.6.2.2 Ambiguities arising out of Template Definitions

There are instances of the VGs identified wrongly. This has occurred because a single template defined matches more than one type of VG.

The V1 of the Compound Verbs Playing as Verbal Nouns

We have already noted the elusiveness of the compound verbs which are marked as verbal nouns in the IL-PoST framework. The same structures come here also. The difference is that this time the V1s are given a TAM tag based on the tag variables. However, this TAM assignment is

wrong because the V2s of these VGs are always marked as verbal nouns and do not form the part of VG. See the examples below:

229	बसंत पंचमी का मुहरत रात से शुरू **[हो]VG.prs.sim.dcl जाने** के कारण श्रद्धालुओं ने तड़के ही स्नान करना शुरू [कर दिया]VG.cv_pst.pft.dcl और श्रद्धालुओं के आने का क्रम जारी [है]VG.prs.sim.dcl ।
361	इसके लिए शिक्षकों को उचित प्रशिक्षण एवं मार्गदर्शन **[दिया]VG.pst.sim.dcl जाना** भी परमावश्यक [है]VG.prs.sim.dcl ।
369	मूर्ख दिवस **[मनाया]VG.pst.sim.dcl जाना** एक अति प्राचीन प्रथा [है]VG.prs.sim.dcl , लेकिन इसकी प्रारंभिक शुरुआत के संदर्भ में अलग-अलग धारणाएँ [हैं]VG.prs.sim.dcl ।

These VGs have been marked wrongly and will remain so until the second verb in them are als marked as verbal elements.

Perfective Participle Functioning as a modifier marked as VG

This is another kind of case where an adjective has been marked as if it were a simple VG. Consider the following examples:

53	इलाहाबाद में गंगा की रेती पर [बसें]VG.pst.sim.dcl तंबुओं की नगरी में सोमवार रात एक पंडाल में आग [लग गई]VG.cv_pst.pft.dcl िजसमें एक बच्ची समेत दो लोग [झुलस गए]VG.cv_pst.pft.dcl तथा पंडाल में **[रखा]VG.pst.sim.dcl** सारा सामान [जलकर]VG.cp.pft राख [हो गया]VG.cv_pst.pft.dcl ।
172	*जुलूस को देखने के लिए लोग अखाड़ा मार्ग पर **[लगाई गई]VG.pas_pst.pft.dcl** बिलल्यो पर भी चढ़ने से बाज नहीं [आए]VG.pst.sim.dcl ।
185	सुश्री मायावती ने [कहा]VG.pst.sim.dcl िक बसपा से अलग **[हुए]VG.pst.sim.dcl** ३८ िवधायकों के मामले में उच्चतम न्यायालय के सम्भािवत आदेश के बाद मुलायम िसंह यादव सरकार खुद ब खुद [िगर जाएगी]VG.cv_fut.sim.dcl ।
199	संगम को मुख्य केन्द्र [मानकर]VG.cp.pft पूरे मेला क्षेत्र में **[लगे]VG.pst.sim.dcl** पत्थरों पर िकलोमीटर [दर्शाया गया है]VG.pas_prs.pft.dcl िजससे लोगों को दूरी मालूम [हो सकेगी]VG.fut.sim.abil ।
388	नौकर अपने मालिक के शानदार कपड़े [पहनकर]VG.cp.pft आराम कुर्सी पर [बैठकर]VG.cp.pft हुक्मनामे जारी [करता था]VG.pst.impf.dcl और नौकर **[बना]VG.pst.sim.dcl** मालिक उसके हर हुक्म का सादर पालन [करता था]VG.pst.impf.dcl ।
390	यहाँ तक कि मालिक [बने]VG.pst.sim.dcl नौकर को अपने नौकर **[बने]VG.pst.sim.dcl** मालिक को

	दंड देने का भी पूरा अधिकार व छूट [थी]VG.pst.sim.dcl ।
407	यह दो शताब्दी ईसा पूर्व **[बना]VG.pst.sim.dcl** बौद्ध कला का नमूना [है]VG.prs.sim.dcl ।

Here, the perfective participle constructions are part of an NP where it modifies the nominal heads of the NPs where they occur. However, they still have been marked as a VG because they are marked as a verb in their annotations. These cases will also remain so and they can be identified as a modifier only when the noun phrase comes into the fore.

Optative vs. Progressive

Consider the following examples:

46	उत्तरप्रदेश के इलाहाबाद में गंगा , यमुना एवं अदृश्य सरस्वती के संगम पर **[चल रहे]VR.cv_0.0.opt** अर्द्धकुंभ मेले में फर्जी पत्रकारों की भी भरमार [हैं]VG.prs.sim.dcl और वे पत्रकारिता की आड़ में साधु-संतों एवं जन-सामान्य को दिगभ्रमित [कर रहे हैं]VG.prs.prog.dcl ।
96	यहाँ **[चल रहे]VR.cv_0.0.opt** अर्द्धकुंभ में गंगा के प्रदूषित जल से कई लोग बीमार [हो गए हैं]VG.cv_prs.pft.dcl ।
113	धार्मिक नगरी इलाहाबाद में **[चल रहे]VR.cv_0.0.opt** अर्द्धकुंभ मेले में देश के कोने-कोने से [आए]VG.pst.sim.dcl साधु-संतों और श्रद्धालुओं के अलावा देशभर के ठगों ने भी डेरा [डाल रखा है]VG.cv_prs.pft.dcl ।
125	तीर्थ नगरी प्रयाग (इलाहाबाद) में **[चल रहे]VR.cv_0.0.opt** अर्द्धकुंभ में आने वाले विदेशी पर्यटकों में से कुछ के लिए यह महज एक मेला [हो सकता है]VG.prs.impf.abil लेकिन इनमें से कुछ पर्यटक ऐसे भी [हैं]VG.prs.sim.dcl जो यहाँ [आकर]VG.cp.pft अपनी वर्षो की ख्वाहिश पूरी [कर रहे हैं]VG.prs.prog.dcl या अपने किसी अजीज को श्रद्धांजिल अर्पित [कर रहे हैं]VG.prs.prog.dcl ।
175	*निर्मोही अखाडा के जुलूस में सबसे पीछे **[चल रहे]VR.cv_0.0.opt** विदेशी श्रद्धालुओ के एक समूह ने सबका ध्यान आकर्षित [किया]VG.pst.sim.dcl ।

In these example output sentences we see that the VG चल रहे /cəl rəhe:/ is given the tag of optative while they should have been given the TAM tag of perfective participles because they function as nominal modifiers. This has occurred because we have two templates namely VR+rah_pft and VR+rah_opt and both of these templates can map to this sequence of the VG.

The second verbal element in this VG i.e. रहे /rəhe:/ is identified as word variable and both rah_pft and rah_opt have रहे /rəhe:/ as a value. Though the word रहे /rəhe:/ has the tag of opt in it, it does not help because when a word variable is used and mapped to a template, the tags of such words are simply ignored. It is only by chance that these VGs have been applied the VR+rah_opt template and not VR+rah_pft. Had the second template been applied, we would have got the correct TAM Tag assignment.

The solution to this problem can be found at the logical level. The solution to this is that we can stop ignoring the tags for the templates where word variables have been used. Thus, we can capture the 'pft' tag assigned to second verb in the example VGs above and skip marking it as an optative. If we follow this solution, we will need to change the code of the tool developed so far. However, I retain it for later development.

4.4.7 Final Run of the Tool with the LDC Corpus and the Example Test Set

We have made changes in both the annotations of the LDC corpus and the other inputs of the tools. The tool changes include the changes we made in the template definitions and the word variables. We made a final run of the tool over the whole of this changed LDC corpus and the following are the results:

Description of the Results Parameters	Raw Number	Percentage
Sentences with at least one unknown Verb Group found	45	0.92
Sentences with at least one orphan auxiliary verb found	0	0
Total number of sentences with at least one orphan auxiliary verb found	0	0
Total number of sentences having either at least one unknown VG or at least one orphan VG or both	45	0.92
Total Number of Sentences having no issues	4839-(45) = 4762	99.08
Total Number of Sentences that has no VGs	508	13.23
Total Instances of Verb Groups Found	8304	NA

Table 54: Results of the Final Run with the LDC Corpus

As seen in the table above, we have only 45 sentences which have at least one unknown verb groups. This can be brought down to 0 if we make the necessary changes as stated in the sections above in the IL-Post Framework.

Besides these errors, we also find another kind of error which basically is the failure on the part of the approach we followed and the IL-PoST framework combined. We have been hard coding all the patterns of the verb groups based on its morphological properties and then giving them a TAM tag. There are some ambiguous cases which have not been captured here. These are the cases that we have described in section 4.4.6.2.2 above. There are a total of 32 such sentences found in the first one thousand output sentences of the LDC corpus. If we count these also among the errors, it lowers our results by 3.2% to reach 99.08-3.2=95.88.

In the final stage, it would be proper also to offer the results of the example sentences we created for all the templates. There are a total of 245 non-CV templates we have. We have an example sentence for each of these templates. We annotated these sentences in the IL-PoST Framework and tested them on the tool. The following are the results for this test.

Description of the Results Parameters	Raw Number	Percentage
Sentences with at least one unknown Verb Group found	0	0
Incorrect VG markings	13	0
Total Number of Sentences having no issues	245-13 = 232	94.69

As seen in the table above, we have a total of 13 sentences wherein a VG has been marked wrongly. These VGs can be categorized into two types as discussed below. It shows the restriction provided by the tool at present. These errors can be resolved once we make changes in the logic. A hundred percent result can be achieved after these changes are made.

The *wala* constructions

There are eight templates that have the *wala* particle coming in between the VG template. These are 16, 33, 54, 70, 184, 189, 202 and 211. As shown above in section 4.4.5.8, this *wala* particle is different from the other particles. It brings about a change in the TAM expressed through the VGs. We had also noted that these templates are unidentified until we we modify the tool to account for the *wala* as part of the VG. Making a simple word variable would not work here because it is not considered as a verbal word (i.e. does not have either VM or VAUX tag).

Homonymy of the verb form “ho:”

We stated in section 4.4.5.11 above that the cases where the word ‘ho:’ is coming as an optative/counterfactual marker will be taken wrongly as a value of ‘prs_aux’ until we change the logic of our tool so that the tags of the word variables are also taken into account. In fact all the cases of homonymy need this application. This affects a total of five of our templates where ‘ho’ is coming as an optative or counterfactual marker. These templates are 47, 144, 208, 214 and 216.

All the other templates are mapped correctly.

5 CONCLUSIONS

The goal of this research work has been stated to be the automatic identification of the verb groups in Hindi text. In this final chapter I present how well this has been achieved and where are the lacunae still present and how they can be filled up. I have also taken some things as preliminary and taken for granted to this work. The tool developed takes as input the annotated text in a particular format. This is the first and the most important assumption which bear the most important base to the functioning of the logic I have built and presented in the previous chapters. I will start with providing a summary of each of the four chapters and draw conclusions and show the results and insights gained through this exercise.

5.1 SUMMARY

The first chapter is mainly an introduction to this area of linguistic research. It presents an overview of what is called natural language processing (NLP) or computational linguistics. I have related the work done through this thesis with the larger body of the work done in this field. I have also described various methods/approaches that are in use for the work of chunking of a natural language text. I will argue in this chapter later that the approach chosen here is though a bit time consuming but is more rewarding than the others in terms of results.

The second chapter provides a brief morphological set-up Hindi as a language. Because we are concerned here more with the text (as against the speech), I have also given a brief on the structure of the Devanagari script which is used to write Hindi. This chapter gives a brief on the morphology of all the parts of speech categories of Hindi in general and verb morphology in

particular. Only the inflectional morphology is covered here because it's this that is more regular and most prolific in Hindi.

In the third chapter I start with a formal description of what is called the verb group. I first identify it manually through a rule base approach. Here the main guiding force for me has been the grammatical approach. I first analyze the inflectional verb morphology and show how time and tense is represented through verbal morphology of Hindi. We find that not all the tense, aspect and mood that can be expressed in Hindi (or universally) can be indicated through just the verbal morphology of Hindi. We find that it is the verb group, a kind of verb serialization process, which is employed to show the various TAM combinations expressed in Hindi. Thus we come to the task of identifying such unique verb groups that can show a unique TAM category. As move ahead with this task we find that there can be two or more verb groups that can show the same TAM categories. We give a TAM tag to each of the TAM categories found. While we have a small of TAM categories, the representing verb group templates can be more than one. We identify each of these templates and their TAM categories and list them. We also cover what is called the compound verb constructions in Hindi. The most number of verb group templates come from this category, mainly because we have a lot V2s and for each of the V2s, we have to come up with a separate template, wherever possible.

In the fourth chapter, I provide the logic on how to identify the verb groups and mark them. I start with the logic and show how the results will be shown for the testing purposes. We do not have any automated method of evaluation. We evaluate it purely based on what the tool gives us as the output and what can be checked only through a manual testing. If the tool fails to identify a verb group (i.e. gives as output an "unknown" VG), this can be attributed to either a faulty annotation or failure of a corresponding template being present. There can also be errors of

incorrect identification (i.e. allocation of an incorrect TAM tag to a VG). This can be attributed to either faulty annotation or a logical problem. I have shown that a logic can be drawn that will capture the verb groups (minus the verb groups involving a conjunct verb) with a hundred percent accuracy.

5.2 MOTIVATION FOR THE CURRENT WORK

The present work has been inspired by task which is rather considered as a basic task in NLP: local word grouping (LWG). Verb group has been identified as a local word group in the NLP related tasks. As we have seen in the third chapter, the tense, aspect and mood information are expressed more often through more than one word inside a sentence/clause in Hindi. In fact this is true for most of the languages. While the process of verb serialization might not be very common to many languages, the TAM information is expressed through various other verbal elements in most of the languages. Such words are usually referred to as auxiliary verbs, where their literal meaning is lost and they are used to mark a change in the TAM expressed through the clause they appear in.

Capturing the tense, aspect and mood information in a given text can be good for several purposes. First of all, it is very important for a machine translation tool to understand the exact TAM expressed by the source text. If we take a clause as the basic unit, the TAM of the VG expressed in it would be the target. The TAM tags I have shown as captured through this exercise can be equated with different TAM categories. This might be present in many languages. If we know for certain what it conveys, they can be transferred in a better way in the target language. It would of special importance to a rule based MT engine which can deterministically identify the TAM of the source (Hindi) and then transfer the same TAM in the target language through a language generation mechanism of that language.

This exercise has also brought us a templatic structure of verb sequences that can possibly be allowed in Hindi. We can use this knowledge base in a grammar checker developed for Hindi.

5.3 THE LOGIC AND FUNCTIONING OF THE TOOL

Hindi and many other languages show the tense, aspect and mood features through more than one word. This is reflected also in the tagsets designed for Hindi and many other languages. Among the verbal category, we have two sub-categories of verbs: main verb and auxiliary verb. It is clear that we cannot have an auxiliary all by itself- it must come always attached with a main verb. For Hindi, the main verb always precedes the auxiliary verbs. This gives us the first base towards our logic building. The information of main verb or auxiliary verb is always encoded at the PoS tagging level. To find the verb groups, we just need to find out the main verb and then attach the following auxiliary verbs to it. We have found that in Hindi a verb group can be interrupted by a few words that are usually called particles. We ignore these words when identifying the verb groups. After we have identified a verb group, we need to assign it a tag based on the TAM information it contains.

Assignment of a proper tag involves the identification of different conjugations of TAM that can be expressed in Hindi. This has to be done exhaustively. If we miss a combination it will be a bug that can emerge when testing the tool. At this stage we come down to anaylsing both the morphology of the verbs in Hindi and how TAM is represented in it. Languages tend to have some peculiarities of its own when denoting the TAM. We analyse the different tenses, aspects and moods that can possibly be expressed either morphologically of periphrastically in Hindi. The TAM that are marked morphologically, through a single word verb form, are easy to capture. We have captured this in the third chapter in section 3.4.1. We get a total of 12 templates which are single word form. However, these ver forms do not capture everything that can be expressed to indicate the time and tense. There are several other nuances that are expressed periphrastically: i.e. with the help of an auxiliary verb. With the help of various descriptive grammars of Hindi available along with some insights from my native speaker knowledge (I consider myself as a native speaker of Hindi) as well as taking help from a large corpus (mainly internet, as covered in the search indexes of major search engines, predominantly Google) for instances of particular verb sequence, I try to enumerate the various tense, aspect and mood that can possibly be shown in Hindi. This has been described in section 3.5 of the third chapter. This exercise results into quite a lot of new aspects and and moods that are only clubbed together elsewhere in the literature. For example, we get a total 20 unique aspect tags that are incorporated in the TAM tags we assign to the verb groups. We have aspects that are, for

example, a combination of both perfect and imperfect, in that order and also imperfect and perfect in that order. Similarly, among the modal categories, we get ten different mood tags as shown in table Table 26 in the third chapter earlier.

Thus we come down to the task of actually identifying the different verb groups based on the number verbs it has. This has been described section 3.7. Here we enumerate all the possible verb groups that contain more than one verbal word. Of course, the main guideline in finding out such verb groups is the identification of the auxiliary verbs first and then a combination of the main verb which is identified based purely on the tag variables it has. We do not include the passive constructions and compound verb constructions at this stage. The main auxiliary verbs are always a few and countable. These auxiliary verbs can also be seen separately from the V2s that we handle in section 3.9. It is this premise along with the various combination of TAM described in the descriptive linguistic literature of the language that helped in arriving at the various VG templates shown in section 3.7.

In section 3.8 of chapter three we discuss the various passive structures possible in Hindi. We make the non-CV non-passive VG Templates we identified in section 3.7 as the base to start from. We find that not all of the 177 VG templates identified in section 3.7 will have a passive construction. Only 67 of these templates can be passivized. We enumerate this in section 3.8.

For the verb groups involving a compound verb, we have devoted one whole and rather lengthy section of 3.10 of chapter three. We restrain the compound verb constructions to only V1+V2 form where the V1 always remains in its root form. This is also the standard definition of compound verbs in Hindi. We postulate that all the VG templates involving a compound verb will also have a structure that is non-CV. This means we have already covered all the non-CV VG templates and we just need to check if these templates will have a corresponding CV construction. We take the whole of the non-CV VG templates identified so far, including the passive ones, and put a V1 at the start of it and check if we can possibley find even one construction matching this new CV template. Further, as we know that all the V2s cannot co-occur with all the V1s possible, we make a separate study of the V1-V2 parity as shown in sections 3.9.2.1 – 2. We find a total of 204 such templates that can possibly have a realization in their CV form. We call these abstract templates as “abstract CV Templates”. This is so because these are in an abstract form the way they are presented. These abstract CV templates have two

verbal words in it that are to be identified based on their tag variables. The first one is the main verb (V1) and the second one is the V2 that plays as the auxiliary verb. Our current strategy of defining a VG template does not allow this because we have been enumerating all the auxiliaries and defining them based on the word variables. This we do because if we leave them in their abstract form, it will overgenerate and might cause errors of the type of wrong VG identification and allocation of wrong TAM tag. Among these 204 abstract CV templates, there are a total 37 templates that can appear with all the V2s. This is what we call the universal abstract CV templates, as described in section 3.9.2.4. The rest of abstract CV templates apply for select V2s, in a restricted manner.

It is notable here that we are identifying the VG templates based purely on the morphological form of the verbs. There is no statistical heuristics or any kind of semantics involved in it. All the analysis is done *a priori* and set as a rule in the form of template definitions and variables as described in the fourth chapter section 4.1. The step wise description of how the tool itself has been designed is shown in section 4.2, Table 44. The tool developed is mainly to test the functioning of this logic over the real corpus. The design of the tool has been described in 4.3. The description of the version of the tool shown is a web-based tool where one test the functioning of this algorithm over the real data. However, the same test can be done on a command line program as well.

5.4 EVALUATION

The system that we developed to identify the verb groups has been named as the “Hindi Verb Group Finder” or simply “VGFinder” in short. This VGFinder tool has been tested over the largest corpus available in the IL-PoST framework: the LDC Corpus. We ran the tool over this corpus several times, but summarized the results achieved in three steps as presented in the last section (§4.4) of the previous chapter. The evaluation process also provided several improvements in the tool. Though the tool is a rule based engine, the results that it showed with the data gave us insights that improved the tool further. The first run of the tool (§4.4.1) gave us a lot of errors, with approximate accuracy being only a meagre 54% of sentences marked for all the VGs present in them. The error analysis done on output (§4.4.2) at this stage was an eye opener. We found a lot of errors in the input data itself. We also had to make some modifications

in the IL-PoST tagset and make changes in the LDC corpus itself to get the proper results for many of the VGs.

5.4.1 Error Analysis

The errors that were found in the output can be classified into two categories:

Faulty Annotations, and

Logical Failures

5.4.1.1 Faulty Annotations

The major annotation errors were due to poor annotation quality. For example, we found a lot of orphan auxiliaries. Theoretically, there cannot be any orphan auxiliary verb in a sentence. This shows the error on the part of the annotator. We corrected each of the occurrences of the orphan auxiliary. We also had a lot of VGs marked as “unknown” which means though they were identified as a VG, they did not fit into any of the templates defined and therefore no TAM tag were assigned to them and they were marked as “unknown”. There can be several reasons for this. These reasons can be described as falling into two categories as shown below:

The PoS annotation values are marked wrongly or the correct markers are missing

The auxiliary verb has a spelling error.

If the PoS annotation is not as recommended, it can result into errors of several types. First of all, all the orphan auxiliary verbs (§4.4.2.1) are a result of faulty annotation. Most of the unknown VGs found can be a result of the faulty annotation (§4.4.2.2). Faulty annotations might also result in incorrect TAM tag assignment (§4.4.2.3).

If the input text has a spelling error, especially the words that function as the auxiliary verb in the verb group, this can result into an unknown marking of the VG (§4.4.6.1.5).

5.4.1.2 Logical Failures

The errors generating due to shortcoming in the tool can be said to be the logical failures. The tool is based on a logic that is rule based. If we fail to capture a VG because no rule can capture it or the rules defined are inherently ambiguous and create an internal error, it will be considered a logical failure. For example, we have discussed that a rule can have multiple reading with the

present logic that we have implemented in the tool. At present, we identify the auxiliary verbs in the VG purely on their word values. This creates a problem in the cases where the homonymy is found. We have found a case of the verb form हो /ho:/ where this verb form can function both as optative/counterfactual marker and also as a value of prs_aux (§4.4.5.11. As expressed earlier in section 4.4.7, this can be resolved only if we take the tag variables of these auxiliary verbs also into account. In section 4.4.7 we also note other failures that could not be resolved at present but can for sure be resolved with a change in the logic we apply with the tool. This change applies at both the levels of the logic and annotation practices we follow using the IL-PoST framework. Another case of failure is that of the *wala* particles inside a VG. We have discussed in the same section that these cases also require a change in the logic. We can define a separate variable of *wala* and treat them differently than other particles. We can consider them at par with the other auxiliary verbs. And this has to be done at the logical level because they cannot take a verbal word tag (VM or VAUX).

Every language has some idiomatic expressions and idiosyncrasies that cannot be covered through the general rule covering most of the part of the language. In Hindi also we have some such expressions that cannot be covered through the template definitions we have defined so far. So, we have to make some rules particularly to capture them. One such rule is the template defined in section 4.4.6.1.11:

laa+khada+kar+de_pft VG.cv_pst.pft.dcl

Similar are the cases that we discuss in section 4.4.6.1.1. In this section we show the solution for the cases like दिखाई दिया /dɪkʰaːiː dɪja/. The solution that we offer here is to mark the cases like /dɪkʰaːiː/ as a noun instead of a verb. This will need a conscious effort on the part of the annotators. We can also avoid it by defining a separate template for just two main verbs of /dɪkʰaːiː/ and /sʊnaːiː/ that occur in combination of the second verb /deː/.

Cases like these are considered as exceptional and must be addressed in a way specific to the case so that they do not interrupt the general behavior captured.

5.5 FINAL RESULTS

The final results have been shown in section 4.4.7 and we show here the results of the tool with two corpora. One is the real world corpus of the LDC prepared in the IL-PoST format. Here we show that we achieve a result of about 96% accuracy. Another corpus tested contains one example each of all the non-CV templates. The result here reaches a total of about 95%. The results with this test corpus can reach up to 100% if we make some changes suggested in that section. That is, if we just make the necessary changes to resolve the *wala* constructions and disambiguate the homonymy of /ho:/, then we can achieve a hundred percent accuracy in the example corpus.

5.6 FUTURE DEVELOPMENTS

At present the conjunct verbs are not handled the way they should have been. As we know, the conjunct verbs formed with a noun or an adjective combined with a verb. These nominal or adjectival parts of the verbs are not marked as a verb even though they play a role in the verb group. Therefore, if one wanted to know the meaning of the verb group for some tasks like translation or discourse analysis, it would be not be always correct. All the conjunct verb cases will be incomplete because the nominal or adjectival part of the VG is not covered in the VG marked by the tool. As stated earlier, to cover the conjunct verb and bring them also into the ambit of the verb group, we will have to delve into semantics as well. This can be done at a later stage. At present all the verb groups which actually involve a conjunct verb and are a part of it are marked as a normal verb group.

APPENDICES

APPENDIX 1: TEMPLATE DEFINITIONS

Non-CV Templates	
prs_aux	VG.prs.sim.dcl
pst_aux	VG.pst.sim.dcl
VR_fut	VG.fut.sim.dcl
VR_pft	VG.pst.sim.dcl
VR_imp	VG.0.0.imp
VR_opt	VG.0.0.opt
VR_impf	VG.pst.impf.cgnt
VR_kar	VG.cp.pft
VR_pcpl	VG.ip.impf
VINF_imp	VG.0.0.imp
VINF	VG.nv
cahiye	VG.0.0.des
VR_impf+prs_aux	VG.prs.impf.dcl
VR_pft+prs_aux	VG.prs.pft.dcl
VINF+prs_aux	VG.prs.sim.oblg
VINF+wala+prs_aux	VG.prs.sim.apprx

cahiye+pst_aux	VG.pst.0.des
VR_impf+pst_aux	VG.pst.impf.dcl
VR_pft+pst_aux	VG.pst.pft.dcl
VINF+lag_pft	VG.pst.pft_ince.dcl
VR_impf+rah_pft	VG.pst.impf_pft_cnt.dcl
VR_impf+ja_pft	VG.pst.pft_dur.dcl
VR_pft+kar_impf	VG.pst.pft_frq.dcl
VR_impf+ho_impf	VG.pst.impf.cfct
VR+sak_pft	VG.pst.sim.abil
VR+pa_pft	VG.pst.sim.abil
VINF+de_pft	VG.pst.sim.perm
VINF+par_pft	VG.pst.sim.oblg
VINF+ho_impf	VG.pst.impf.oblg
VINF+pst_aux	VG.pst.sim.oblg
VR_impf+ban_pft	VG.pst.sim.abil
VR_pft+rah_pft	VG.pst.pft_cnt.dcl
VINF+wala+pst_aux	VG.pst.sim.apprx
VINF+de_impf	VG.pst.impf.perm
VINF+pa_pft	VG.pst.sim.abil
VR_impf+cal_pft	VG.pst.pft.dcl

VR_impf+rah_impf	VG.pst.dur.cgnt
VR_pft+ho_impf	VG.pst.impf.cgnt
VR_pft+lag_pft	VG.pst.sim_ince.dcl
VR_pft+rakh_pft	VG.pst.pft.dcl
VR_pft+ho_fut	VG.fut.pft.dcl
VINF+lag_fut	VG.fut.sim_ince.dcl
VR_impf+rah_fut	VG.fut.cnt.dcl
VR_impf+ja_fut	VG.fut.dur.dcl
VR_pft+kar_fut	VG.fut.impf_frq.dcl
VR_impf+ho_fut	VG.fut.impf.prsm
VR_impf+ho_prsm_aux	VG.0.0.cfct
VR+sak_fut	VG.fut.sim.abil
VR+pa_fut	VG.fut.sim.abil
VINF+de_fut	VG.fut.sim.perm
VINF+par_fut	VG.fut.sim.oblg
VINF+ho_fut	VG.fut.sim.oblg
VR_pft+rah_fut	VG.fut.pft_cnt.dcl
VINF+wala+ho_fut	VG.fut.sim.apprx
VR_impf+VR_opt	VG.0.0.opt
VR_pft+rah_opt	VG.0.pft_cnt.opt

VR_pft+rah_imp	VG.0.pft_cnt.imp
VR+sak_imp	VG.0.0.abil
VR_pft+kar_imp	VG.0.0.imp
VINF+de_imp	VG.0.0.perm_imp
VINF+de_opt	VG.0.0.perm_opt
VINF+cahiye	VG.0.0.sugg
VINF+de_inf	VG.0.0.perm_imp
VR_impf+cal_imp	VG.0.0.imp
VR_impf+cal_opt	VG.0.0.opt
VR_impf+ho_opt	VG.0.0.cfct
VR_impf+rah_opt	VG.0.0.opt
VR_pft+ho_opt	VG.0.0.cfct
VR_pft+VINF_imp	VG.0.0.imp
VINF+wala	VG.gp
VR+kar	VG.cp.pft
VR_impf+ho_pft	VG.ip.impf
VR_pft+ho_pft	VG.ip.pft
VR_pft+VINF	VG.nv.pft
VR+rah_pft	VG.ip.prog
VR_impf+ja_impf	VG.ip.impf.dur

VR_impf+ban_impf+prs_aux	VG.prs.impf.abil
VR+cuk_pft+prs_aux	VG.prs.pft.dcl
VR+rah_pft+prs_aux	VG.prs.prog.dcl
VINF+lag_impf+prs_aux	VG.prs.impf_ince.dcl
VINF+lag_pft+prs_aux	VG.prs.pft_ince.dcl
VR_impf+rah_impf+prs_aux	VG.prs.impf_cnt.dcl
VR_impf+rah_pft+prs_aux	VG.prs.impf_pft_cnt.dcl
VR_pft+rah_impf+prs_aux	VG.prs.pft_impf_cnt.dcl
VR_impf+ja_impf+prs_aux	VG.prs.impf_dur.dcl
VR_impf+ja_pft+prs_aux	VG.prs.pft_dur.dcl
VR_pft+kar_impf+prs_aux	VG.prs.impf_frq.dcl
VR+sak_impf+prs_aux	VG.prs.impf.abil
VR+sak_pft+prs_aux	VG.prs.pft.abil
VR+pa_impf+prs_aux	VG.prs.impf.abil
VR+pa_pft+prs_aux	VG.prs.pft.abil
VINF+de_impf+prs_aux	VG.prs.impf.perm
VINF+de_pft+prs_aux	VG.prs.pft.perm
VINF+par_impf+prs_aux	VG.prs.impf.oblg
VINF+par_pft+prs_aux	VG.prs.pft.oblg
VINF+ho_impf+prs_aux	VG.prs.impf.oblg

VR_impf+a_pft+prs_aux	VG.prs.impf_dur.dcl
VR_impf+ho_inf+cahiye	VG.prs.sim.sugg
VR_pft+ho_impf+prs_aux	VG.prs.impf.dcl
VR_pft+ho_pft+prs_aux	VG.prs.pft.dcl
VR_pft+par_pft+prs_aux	VG.prs.pft.dcl
VR_pft+rakh_impf+prs_aux	VG.prs.pft_dur.dcl
VR_impf+ban_impf+pst_aux	VG.pst.impf.abil
VR+cuk_pft+pst_aux	VG.pst.pft.dcl
VR+rah_pft+pst_aux	VG.pst.prog.dcl
VINF+lag_impf+pst_aux	VG.pst.impf_ince.dcl
VINF+lag_pft+pst_aux	VG.pst.pft_ince.dcl
VR_impf+rah_impf+pst_aux	VG.pst.impf_cnt.dcl
VR_impf+rah_pft+pst_aux	VG.pst.impf_pft_cnt.dcl
VR_pft+rah_impf+pst_aux	VG.pst.pft_impf_cnt.dcl
VR_impf+ja_impf+pst_aux	VG.pst.impf_dur.dcl
VR_impf+ja_pft+pst_aux	VG.pst.pft_dur.dcl
VR_pft+kar_impf+pst_aux	VG.pst.impf_frq.dcl
VR+sak_impf+pst_aux	VG.pst.impf.abil
VR+sak_pft+pst_aux	VG.pst.pft.abil
VR+pa_impf+pst_aux	VG.pst.impf.abil

VR+pa_pft+pst_aux	VG.pst.pft.abil
VINF+de_impf+pst_aux	VG.pst.impf.perm
VINF+de_pft+pst_aux	VG.pst.pft.perm
VINF+par_impf+pst_aux	VG.pst.impf.oblg
VINF+par_pft+pst_aux	VG.pst.pft.oblg
VINF+ho_impf+pst_aux	VG.pst.impf.oblg
VINF+de_pft+ho_impf	VG.pst.pft.perm_cfct
VINF+cahiye+pst_aux	VG.pst.0.sugg
khada+rah+ja_pft	VG.pst.dur.dcl
VR_impf+cal_pft+ja_pft	VG.pst.pft_cnt.dcl
VR_impf+rah+ja_pft	VG.pst.pft_dur.dcl
VR_pft+ho_pft+pst_aux	VG.pst.pft.dcl
VR_pft+par_pft+pst_aux	VG.pst.pft.dcl
VR_impf+ban_fut	VG.fut.impf.abil
VR+cuk_pft+ho_fut	VG.fut.pft.dcl
VR+rah_pft+ho_fut	VG.fut.prog.prsm
VR_pft+kar_impf+ho_fut	VG.fut.impf_dur.dcl
VR+sak_pft+ho_fut	VG.fut.pft.abil
VR+pa_pft+ho_fut	VG.fut.pft.abil
VINF+de_pft+ho_fut	VG.fut.pft.perm_cfct

VINF+par_pft+ho_fut	VG.fut.pft.oblg
VINF+ho_impf+ho_fut	VG.fut.pft.oblg
VINF+rah_pft+ho_fut	VG.fut.prog.oblg_prsm
VINF+de_impf+ho_fut	VG.fut.impf.perm_cfct
VINF+lag_pft+ho_fut	VG.fut.pft_ince.prsm
VINF+lag_impf+ho_fut	VG.fut.impf_ince.prsm
VR_impf+rah_pft+ho_fut	VG.fut.impf_pft.prsm
VR+rah_pft+ho_cfct	VG.0.prog.cfct
VINF+de_inf+cahiye	VG.0.0.perm_sugg
VR_impf+ja+rah_pft	VG.ip.prog.dur
VR+ja+rah_pft	VG.ip.prog
VR_impf+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl
VR+sak+rah_pft+prs_aux	VG.prs.prog.abil
VR+pa+rah_pft+prs_aux	VG.prs.prog.abil
VINF+de+rah_pft+prs_aux	VG.prs.prog.perm
VINF+par+rah_pft+prs_aux	VG.prs.prog.oblg
VINF+par+sak_impf+prs_aux	VG.prs.impf.abil
VR_impf+a+rah_pft+prs_aux	VG.prs.dur.dcl
VR_pft+a+rah_pft+prs_aux	VG.prs.dur.dcl
VR_pft+rah+sak_impf+prs_aux	VG.prs.pft_dur.abil

VR_impf+ja+rah_pft+pst_aux	VG.pst.prog_dur.dcl
VR+sak+rah_pft+pst_aux	VG.pst.prog.abil
VR+pa+rah_pft+pst_aux	VG.pst.prog.abil
VINF+de+rah_pft+pst_aux	VG.pst.prog.perm
VINF+par+rah_pft+pst_aux	VG.pst.prog.oblg
VINF+par+sak_impf+pst_aux	VG.pst.impf.abil
VR_impf+a+rah_pft+pst_aux	VG.pst.dur.dcl
VR_pft+a+rah_pft+pst_aux	VG.pst.dur.dcl
VR_pft+rah+sak_impf+pst_aux	VG.pst.pft_dur.abil
VR_impf+ja+rah_pft+ho_fut	VG.fut.impf_dur.prsm
VR+sak+rah_pft+ho_fut	VG.fut.prog.abil_prsm
VR+pa+rah_pft+ho_fut	VG.fut.prog.abil_prsm
VINF+de+rah_pft+ho_fut	VG.fut.prog.perm_prsm
VINF+par+rah_pft+ho_fut	VG.fut.prog.oblg_prsm
VR_impf+a+rah_pft+ho_fut	VG.fut.dur.dcl
VR_pft+a+rah_pft+ho_fut	VG.fut.dur.dcl
VR_impf+cal_pft+ja+rah_pft+prs_aux	VG.prs.prog_dur.dcl
VR_impf+cal_pft+a+rah_pft+prs_aux	VG.prs.prog_dur.dcl
VR_impf+cal_pft+ja+rah_pft+pst_aux	VG.pst.prog_dur.dcl
VR_impf+cal_pft+a+rah_pft+pst_aux	VG.pst.prog_dur.dcl

VR_impf+cal_pft+a+rah_pft+ho_fut	VG.fut.impf.cfct
VR_impf+cal_pft+ja+rah_pft+ho_fut	VG.fut.impf.cfct
VR_pft+ja_pft	VG.pas_pst.sim.dcl
VR_pft+ja_impf	VG.pas_pst.impf.cgnt
VR_pft+ja_fut	VG.pas_fut.sim.dcl
VR_pft+ja_opt	VG.pas_0.0.opt
VR_pft+ja_inf	VG.pas_nv.pft
VR_pft+ja_inf+wala	VG.pas_gp
VR_pft+ja_impf+prs_aux	VG.pas_prs.impf.dcl
VR_pft+ja_pft+prs_aux	VG.pas_prs.pft.dcl
VR_pft+ja+rah_pft+prs_aux	VG.pas_prs.prog.dcl
VR_pft+ja_inf+prs_aux	VG.pas_prs.sim.oblg
VR_pft+ja_inf+wala+prs_aux	VG.pas_prs.sim.apprx
VR_pft+ja_impf+pst_aux	VG.pas_pst.impf.dcl
VR_pft+ja_pft+pst_aux	VG.pas_pst.pft.dcl
VR_pft+ja+rah_pft+pst_aux	VG.pas_pst.prog.dcl
VR_pft+ja_inf+lag_pft	VG.pas_pst.sim_ince.dcl
VR_pft+ja_inf+lag_fut	VG.pas_fut.pft_ince.dcl
VR_pft+ja_impf+rah_pft	VG.pas_pst.sim_cnt.dcl
VR_pft+ja_impf+rah_fut	VG.pas_fut.pft_cnt.dcl

VR_pft+ja_pft+ho_impf	VG.pas_pst.impf.cfct
VR_pft+ja+sak_pft	VG.pas_pst.pft.abil
VR_pft+ja_inf+ho_impf	VG.pas_pst.impf.oblg
VR_pft+ja_inf+pst_aux	VG.pas_pst.sim.oblg
VR_pft+ho_impf+pst_aux	VG.pas_pst.pft_impf_cnt.dcl
VR_pft+ja_inf+wala+pst_aux	VG.pas_pst.sim.apprx
VR_pft+ja+sak_impf	VG.pas_pst.impf.cgnt
VR_pft+ja_pft+ho_fut	VG.pas_fut.pft.dcl
VR_pft+ja+rah_pft+ho_fut	VG.pas_fut.prog.prsm
VR_pft+ja_impf+ho_fut	VG.pas_fut.impf.prsm
VR_pft+ja+sak_fut	VG.pas_fut.sim.abil
VR_pft+ja_pft+ho_prsm_aux	VG.pas_fut.impf.cgnt
VR_pft+ja_inf+ho_fut	VG.pas_fut.sim.oblg
VR_pft+ho_impf+ho_fut	VG.pas_fut.pft_cnt.dcl
VR_pft+ja_inf+wala+ho_fut	VG.pas_fut.sim.apprx
VINF+de_pft+ja_fut	VG.pas_fut.sim.perm
VR_pft+ja_impf+rah_opt	VG.pas_0.0.opt
VR_pft+ja_pft+ho_opt	VG.pas_0.pft.cfct
VR_pft+ja_inf+cahiye	VG.pas_0.0.sugg
VR_pft+ja_inf+ho_prsm_aux	VG.pas_0.0.opt

VR_pft+ja+sak_imp	VG.pas_0.0.opt_prob
VR_pft+ja+sak_inf	VG.pas_0.0.inf
VR_pft+ja+rah_pft	VG.pas_ip.dur
VR_pft+ja+cuk_pft+prs_aux	VG.pas_prs.pft.dcl
VR_pft+ja_inf+lag_impf+prs_aux	VG.pas_prs.impf_ince.dcl
VR_pft+ja_inf+lag_pft+prs_aux	VG.pas_prs.pft_ince.dcl
VR_pft+ja_impf+rah_pft+prs_aux	VG.pas_prs.impf_cnt.dcl
VR_impf+ja_impf+rah_pft+prs_aux	VG.pas_prs.impf_dur.dcl
VR_pft+ja+sak_impf+prs_aux	VG.pas_prs.impf.abil
VR_pft+ja+sak_pft+prs_aux	VG.pas_prs.pft.abil
VR_pft+ja_pft+ho_impf+prs_aux	VG.pas_prs.pft_impf_cnt.dcl
VR_pft+ja_impf+rah_impf+prs_aux	VG.pas_prs.pft_cnt.dcl
VR_pft+ja_impf+rah_impf+pst_aux	VG.pas_pst.impf_cnt.dcl
VR_pft+ja+cuk_pft+pst_aux	VG.pas_pst.pft.dcl
VR_pft+ja_inf+lag_impf+pst_aux	VG.pas_pst.impf_ince.dcl
VR_pft+ja_inf+lag_pft+pst_aux	VG.pas_pst.pft_ince.dcl
VR_pft+ja_impf+rah_pft+pst_aux	VG.pas_pst.impf_pft_cnt.dcl
VR_pft+ja+sak_impf+pst_aux	VG.pas_pst.impf.abil
VR_pft+ja+sak_pft+pst_aux	VG.pas_pst.pft.abil
VR_pft+ja_inf+cahiye+pst_aux	VG.pas_pst.0.sugg

VR_pft+ja_pft+ho_impf+pst_aux	VG.pas_pst.pft_impf_cnt.dcl
VR_pft+ja+cuk_pft+ho_fut	VG.pas_fut.pft.dcl
VR_pft+ja+sak_pft+ho_fut	VG.pas_fut.pft.abil
VR_pft+ja_inf+lag_pft+ho_fut	VG.pas_fut.pft_ince.prsm
VR_pft+ja_inf+lag_impf+ho_fut	VG.pas_fut.impf_ince.prsm
VR_pft+ja_impf+rah_pft+ho_fut	VG.pas_fut.impf_pft.prsm
VR_pft+ja_pft+ho_impf+ho_fut	VG.pas_fut.pft_cnt.prsm
VR_pft+ja+rah_pft+ho_opt	VG.pas_0.prog.cfct
VR_pft+ja+sak_pft+opt	VG.pas_0.0.cfct
CV Templates	
VR+VINF	VG.cv_nv
VR+dal_fut	VG.cv_fut.sim.dcl
VR+dal_pft	VG.cv_pst.sim.dcl
VR+dal	VG.cv_0.0.imp
VR+dal_imp	VG.cv_0.0.imp
VR+dal_opt	VG.cv_0.0.opt
VR+dal_impf	VG.cv_pst.impf.cgnt
VR+cuk_pft	VG.cv_pst.impf.cnd
VR+ja_pft	VG.cv_pst.sim.dcl
VR+de_pft	VG.cv_pst.sim.dcl

VR+par_pft	VG.cv_pst.sim.dcl
VR+sak_impf	VG.cv_0.0.abil
VR+ja_impf+prs_aux	VG.cv_prs.impf.dcl
VR+de_imp	VG.cv_0.0.imp
VR+le_imp	VG.cv_0.0.imp
VR+ja_fut	VG.cv_fut.sim.dcl
VR+le_pft	VG.cv_pst.sim.dcl
VR+le_fut	VG.cv_fut.0.dcl
VR+pa_impf	VG.cv_pst.impf.cfct
VR+VINF_imp	VG.cv_0.0.imp
VR+ja_opt	VG.cv_0.0.opt
VR+de_fut	VG.cv_fut.sim.dcl
VR+pahunch_pft	VG.cv_pst.sim.dcl
VR+a_pft	VG.cv_pst.sim.dcl
VR+par_impf	VG.cv_pst.impf.cfct
VR+uth_pft	VG.cv_pst.sim.dcl
VR+sak_opt	VG.cv_0.0.opt
VR+ja_impf	VG.cv_pst.impf.cfct
VR+le_opt	VG.cv_0.0.opt
VR+laa_pft	VG.cv_pst.sim.dcl

VR+cal_pft	VG.cv_pst.sim.dcl
VR+rakh_imp	VG.cv_0.0.imp
VR+nikal_pft	VG.cv_pst.sim.dcl
VR+dikh_pft	VG.cv_pst.sim.dcl
VR+chadh_pft	VG.cv_pst.sim.dcl
VR+gir_pft	VG.cv_pst.sim.dcl
VR+phans_pft	VG.cv_pst.sim.dcl
VR+kar_pft	VG.cv_pst.sim.dcl
VR+kha_pft	VG.cv_pst.sim.dcl
VR+giraa_pft	VG.cv_pst.sim.dcl
VR+gujar_pft	VG.cv_pst.sim.dcl
VR+ghasit_pft	VG.cv_pst.sim.dcl
VR+cah_pft	VG.cv_pst.sim.dcl
VR+chod_pft	VG.cv_pst.sim.dcl
VR+tapak_pft	VG.cv_pst.sim.dcl
VR+dikha_pft	VG.cv_pst.sim.dcl
VR+dekh_pft	VG.cv_pst.sim.dcl
VR+daud_pft	VG.cv_pst.sim.dcl
VR+dhamak_pft	VG.cv_pst.sim.dcl
VR+dhar_pft	VG.cv_pst.sim.dcl

VR+nikaal_pft	VG.cv_pst.sim.dcl
VR+patak_pft	VG.cv_pst.sim.dcl
VR+phenk_pft	VG.cv_pst.sim.dcl
VR+bach_pft	VG.cv_pst.sim.dcl
VR+ban_pft	VG.cv_pst.sim.dcl
VR+bas_pft	VG.cv_pst.sim.dcl
VR+bit_pft	VG.cv_pst.sim.dcl
VR+baith_pft	VG.cv_pst.sim.dcl
VR+baitha_pft	VG.cv_pst.sim.dcl
VR+bhag_pft	VG.cv_pst.sim.dcl
VR+bhej_pft	VG.cv_pst.sim.dcl
VR+mar_pft	VG.cv_pst.sim.dcl
VR+maar_pft	VG.cv_pst.sim.dcl
VR+mit_pft	VG.cv_pst.sim.dcl
VR+mita_pft	VG.cv_pst.sim.dcl
VR+rakh_pft	VG.cv_pst.sim.dcl
VR+lag_pft	VG.cv_pst.sim.dcl
VR+laga_pft	VG.cv_pst.sim.dcl
VR+suna_pft	VG.cv_pst.sim.dcl
VR+ho_pft	VG.cv_pst.sim.dcl

VR+a_impf	VG.cv_pst.impf.cfct
VR+uth_impf	VG.cv_pst.impf.cfct
VR+kar_impf	VG.cv_pst.impf.cfct
VR+kha_impf	VG.cv_pst.impf.cfct
VR+gir_impf	VG.cv_pst.impf.cfct
VR+giraa_impf	VG.cv_pst.impf.cfct
VR+gujar_impf	VG.cv_pst.impf.cfct
VR+ghasit_impf	VG.cv_pst.impf.cfct
VR+cal_impf	VG.cv_pst.impf.cfct
VR+cah_impf	VG.cv_pst.impf.cfct
VR+cuk_impf	VG.cv_pst.impf.cfct
VR+chod_impf	VG.cv_pst.impf.cfct
VR+tapak_impf	VG.cv_pst.impf.cfct
VR+dikh_impf	VG.cv_pst.impf.cfct
VR+dikha_impf	VG.cv_pst.impf.cfct
VR+de_impf	VG.cv_pst.impf.cfct
VR+dekh_impf	VG.cv_pst.impf.cfct
VR+daud_impf	VG.cv_pst.impf.cfct
VR+dhamak_impf	VG.cv_pst.impf.cfct
VR+dhar_impf	VG.cv_pst.impf.cfct

VR+nikal_impf	VG.cv_pst.impf.cfct
VR+nikaal_impf	VG.cv_pst.impf.cfct
VR+patak_impf	VG.cv_pst.impf.cfct
VR+pahunch_impf	VG.cv_pst.impf.cfct
VR+phenk_impf	VG.cv_pst.impf.cfct
VR+bach_impf	VG.cv_pst.impf.cfct
VR+ban_impf	VG.cv_pst.impf.cfct
VR+bas_impf	VG.cv_pst.impf.cfct
VR+bit_impf	VG.cv_pst.impf.cfct
VR+baith_impf	VG.cv_pst.impf.cfct
VR+baitha_impf	VG.cv_pst.impf.cfct
VR+bhag_impf	VG.cv_pst.impf.cfct
VR+bhej_impf	VG.cv_pst.impf.cfct
VR+mar_impf	VG.cv_pst.impf.cfct
VR+maar_impf	VG.cv_pst.impf.cfct
VR+mit_impf	VG.cv_pst.impf.cfct
VR+mita_impf	VG.cv_pst.impf.cfct
VR+rakh_impf	VG.cv_pst.impf.cfct
VR+rah_impf	VG.cv_pst.impf.cfct
VR+lag_impf	VG.cv_pst.impf.cfct

VR+laga_impf	VG.cv_pst.impf.cfct
VR+laa_impf	VG.cv_pst.impf.cfct
VR+le_impf	VG.cv_pst.impf.cfct
VR+suna_impf	VG.cv_pst.impf.cfct
VR+ho_impf	VG.cv_pst.impf.cfct
VR+phans_impf	VG.cv_pst.impf.cfct
VR+chadh_impf	VG.cv_pst.impf.cfct
VR+a_imp	VG.cv_0.0.imp
VR+uth_imp	VG.cv_0.0.imp
VR+kar_imp	VG.cv_0.0.imp
VR+kha_imp	VG.cv_0.0.imp
VR+gir_imp	VG.cv_0.0.imp
VR+giraa_imp	VG.cv_0.0.imp
VR+gujar_imp	VG.cv_0.0.imp
VR+ghasit_imp	VG.cv_0.0.imp
VR+cal_imp	VG.cv_0.0.imp
VR+cahiye	VG.cv_0.0.imp
VR+cuk_imp	VG.cv_0.0.imp
VR+chod_imp	VG.cv_0.0.imp
VR+ja_imp	VG.cv_0.0.imp

VR+tapak_imp	VG.cv_0.0.imp
VR+dikh_imp	VG.cv_0.0.imp
VR+dikha_imp	VG.cv_0.0.imp
VR+dekh_imp	VG.cv_0.0.imp
VR+daud_imp	VG.cv_0.0.imp
VR+dhamak_imp	VG.cv_0.0.imp
VR+dhar_imp	VG.cv_0.0.imp
VR+nikal_imp	VG.cv_0.0.imp
VR+nikaal_imp	VG.cv_0.0.imp
VR+patak_imp	VG.cv_0.0.imp
VR+par_imp	VG.cv_0.0.imp
VR+pahunch_imp	VG.cv_0.0.imp
VR+pa_imp	VG.cv_0.0.imp
VR+phenk_imp	VG.cv_0.0.imp
VR+bach_imp	VG.cv_0.0.imp
VR+ban_imp	VG.cv_0.0.imp
VR+bas_imp	VG.cv_0.0.imp
VR+bit_imp	VG.cv_0.0.imp
VR+baith_imp	VG.cv_0.0.imp
VR+baitha_imp	VG.cv_0.0.imp

VR+bhag_imp	VG.cv_0.0.imp
VR+bhej_imp	VG.cv_0.0.imp
VR+mar_imp	VG.cv_0.0.imp
VR+maar_imp	VG.cv_0.0.imp
VR+mit_imp	VG.cv_0.0.imp
VR+mita_imp	VG.cv_0.0.imp
VR+rah_imp	VG.cv_0.0.imp
VR+lag_imp	VG.cv_0.0.imp
VR+laga_imp	VG.cv_0.0.imp
VR+laa_imp	VG.cv_0.0.imp
VR+suna_imp	VG.cv_0.0.imp
VR+ho_imp	VG.cv_0.0.imp
VR+phans_imp	VG.cv_0.0.imp
VR+chadh_imp	VG.cv_0.0.imp
VR+a_opt	VR.cv_0.0.opt
VR+uth_opt	VR.cv_0.0.opt
VR+kar_opt	VR.cv_0.0.opt
VR+kha_opt	VR.cv_0.0.opt
VR+gir_opt	VR.cv_0.0.opt
VR+giraa_opt	VR.cv_0.0.opt

VR+gujar_opt	VR.cv_0.0.opt
VR+ghasit_opt	VR.cv_0.0.opt
VR+cal_opt	VR.cv_0.0.opt
VR+cah_opt	VR.cv_0.0.opt
VR+cuk_opt	VR.cv_0.0.opt
VR+chod_opt	VR.cv_0.0.opt
VR+tapak_opt	VR.cv_0.0.opt
VR+dikh_opt	VR.cv_0.0.opt
VR+dikha_opt	VR.cv_0.0.opt
VR+de_opt	VR.cv_0.0.opt
VR+dekh_opt	VR.cv_0.0.opt
VR+daud_opt	VR.cv_0.0.opt
VR+dhamak_opt	VR.cv_0.0.opt
VR+dhar_opt	VR.cv_0.0.opt
VR+nikal_opt	VR.cv_0.0.opt
VR+nikaal_opt	VR.cv_0.0.opt
VR+patak_opt	VR.cv_0.0.opt
VR+par_opt	VR.cv_0.0.opt
VR+pahunch_opt	VR.cv_0.0.opt
VR+pa_opt	VR.cv_0.0.opt

VR+phenk_opt	VR.cv_0.0.opt
VR+bach_opt	VR.cv_0.0.opt
VR+ban_opt	VR.cv_0.0.opt
VR+bas_opt	VR.cv_0.0.opt
VR+bit_opt	VR.cv_0.0.opt
VR+baith_opt	VR.cv_0.0.opt
VR+baitha_opt	VR.cv_0.0.opt
VR+bhag_opt	VR.cv_0.0.opt
VR+bhej_opt	VR.cv_0.0.opt
VR+mar_opt	VR.cv_0.0.opt
VR+maar_opt	VR.cv_0.0.opt
VR+mit_opt	VR.cv_0.0.opt
VR+mita_opt	VR.cv_0.0.opt
VR+rakh_opt	VR.cv_0.0.opt
VR+rah_opt	VR.cv_0.0.opt
VR+lag_opt	VR.cv_0.0.opt
VR+laga_opt	VR.cv_0.0.opt
VR+laa_opt	VR.cv_0.0.opt
VR+suna_opt	VR.cv_0.0.opt
VR+ho_opt	VR.cv_0.0.opt

VR+phans_opt	VR.cv_0.0.opt
VR+chadh_opt	VR.cv_0.0.opt
VR+a_fut	VG.cv_fut.sim.dcl
VR+uth_fut	VG.cv_fut.sim.dcl
VR+kar_fut	VG.cv_fut.sim.dcl
VR+kha_fut	VG.cv_fut.sim.dcl
VR+gir_fut	VG.cv_fut.sim.dcl
VR+giraa_fut	VG.cv_fut.sim.dcl
VR+gujar_fut	VG.cv_fut.sim.dcl
VR+ghasit_fut	VG.cv_fut.sim.dcl
VR+cal_fut	VG.cv_fut.sim.dcl
VR+cah_fut	VG.cv_fut.sim.dcl
VR+cuk_fut	VG.cv_fut.sim.dcl
VR+chod_fut	VG.cv_fut.sim.dcl
VR+tapak_fut	VG.cv_fut.sim.dcl
VR+dikh_fut	VG.cv_fut.sim.dcl
VR+dikha_fut	VG.cv_fut.sim.dcl
VR+dekh_fut	VG.cv_fut.sim.dcl
VR+daud_fut	VG.cv_fut.sim.dcl
VR+dhamak_fut	VG.cv_fut.sim.dcl

VR+dhar_fut	VG.cv_fut.sim.dcl
VR+nikal_fut	VG.cv_fut.sim.dcl
VR+nikaal_fut	VG.cv_fut.sim.dcl
VR+patak_fut	VG.cv_fut.sim.dcl
VR+par_fut	VG.cv_fut.sim.dcl
VR+pahunch_fut	VG.cv_fut.sim.dcl
VR+phenk_fut	VG.cv_fut.sim.dcl
VR+bach_fut	VG.cv_fut.sim.dcl
VR+ban_fut	VG.cv_fut.sim.dcl
VR+bas_fut	VG.cv_fut.sim.dcl
VR+bit_fut	VG.cv_fut.sim.dcl
VR+baith_fut	VG.cv_fut.sim.dcl
VR+baitha_fut	VG.cv_fut.sim.dcl
VR+bhag_fut	VG.cv_fut.sim.dcl
VR+bhej_fut	VG.cv_fut.sim.dcl
VR+mar_fut	VG.cv_fut.sim.dcl
VR+maar_fut	VG.cv_fut.sim.dcl
VR+mit_fut	VG.cv_fut.sim.dcl
VR+mita_fut	VG.cv_fut.sim.dcl
VR+rakh_fut	VG.cv_fut.sim.dcl

VR+rah_fut	VG.cv_fut.sim.dcl
VR+lag_fut	VG.cv_fut.sim.dcl
VR+laga_fut	VG.cv_fut.sim.dcl
VR+laa_fut	VG.cv_fut.sim.dcl
VR+suna_fut	VG.cv_fut.sim.dcl
VR+ho_fut	VG.cv_fut.sim.dcl
VR+phans_fut	VG.cv_fut.sim.dcl
VR+chadh_fut	VG.cv_fut.sim.dcl
VR+dal_impf+prs_aux	VG.cv_prs.impf.dcl
VR+dal_impf+pst_aux	VG.cv_pst.impf.dcl
VR+dal_pft+prs_aux	VG.cv_prs.pft.dcl
VR+dal_pft+pst_aux	VG.cv_pst.pft.dcl
VR+dal_pft+ho_fut	VG.cv_fut.pft.dcl
VR+VINF+lag_pft	VG.cv_pst.sim_ince.dcl
VR+VINF+lag_fut	VG.cv_fut.sim_ince.dcl
VR+dal_impf+rah_pft	VG.cv_pst.impf_sim_cnt.dcl
VR+dal_impf+rah_fut	VG.cv_fut.cnt.dcl
VR+dal_impf+ja_pft	VG.cv_pst.sim_dur.dcl
VR+dal_impf+ja_fut	VG.cv_fut.dur.dcl
VR+dal_pft+kar_impf	VG.cv_pst.pft_frq.dcl

VR+dal_pft+kar_fut	VG.cv_fut.impf_frq.dcl
VR+dal_impf+ho_fut	VG.cv_fut.impf.prsm
VR+dal_impf+ho_impf	VG.cv_pst.impf.cfct
VR+dal_impf+ho_prsm_aux	VG.cv_fut.impf.cgnt
VR+V2+sak_fut	VG.cv_fut.sim.abil
VR+V2+sak_pft	VG.cv_pst.pft.abil
VR+V2+pa_fut	VG.cv_fut.sim.abil
VR+V2+pa_pft	VG.cv_pst.pft.abil
VR+VINF+de_fut	VG.cv_fut.sim.perm
VR+VINF+de_pft	VG.cv_pst.sim.perm
VR+VINF+par_fut	VG.cv_fut.sim.oblg
VR+VINF+par_pft	VG.cv_pst.sim.oblg
VR+VINF+ho_fut	VG.cv_fut.sim.oblg
VR+VINF+ho_impf	VG.cv_pst.impf.oblg
VR+VINF+prs_aux	VG.cv_prs.sim.oblg
VR+VINF+pst_aux	VG.cv_pst.sim.oblg
VR+dal_pft+ja_fut	VG.cv_pas_fut.sim.dcl
VR+dal_pft+ja_pft	VG.cv_pas_pst.pft.dcl
VR+dal_pft+ja_opt	VG.cv_pas_0.0.opt
VR+de_pft+prs_aux	VG.cv_prs.pft.dcl

VR+de_pft+pst_aux	VG.cv_pst.pft.dcl
VR+cal_pft+prs_aux	VG.cv_prs.pft.dcl
VR+de_pft+ja_opt	VG.cv_pas_pst.pft.opt
VR+rakh_pft+prs_aux	VG.cv_prs.pft.dcl
VR+le_impf+prs_aux	VG.cv_prs.impf.dcl
VR+ja_pft+prs_aux	VG.cv_prs.pft.dcl
VR+le_impf+pst_aux	VG.cv_pst.pft.dcl
VR+ja_pft+pst_aux	VG.cv_pst.pft.dcl
VR+de_pft+ja_fut	VG.cv_pas_fut.pft.dcl
VR+le_pft+prs_aux	VG.cv_prs.pft.dcl
VR+le_pft+pst_aux	VG.cv_pst.pft.dcl
VR+le+ja_impf	VG.cv_pst.impf.cfct
VR+rakh_pft+pst_aux	VG.cv_pst.pft.dcl
VR+de_pft+ja_pft	VG.cv_pas_pst.pft.dcl
VR+kar+le_pft	VG.cv_pst.pft.dcl
VR+uth_impf+prs_aux	VG.cv_prs.impf.dcl
VR+de_impf+prs_aux	VG.cv_prs.impf.dcl
VR+par_pft+prs_aux	VG.cv_prs.pft.dcl
VR+cuk_pft+ho_impf	VG.cv_pst.pft.cgnt
VR+le_pft+ja_pft	VG.cv_pst.pft.dcl

VR+le_pft+ja_fut	VG.cv_pas_fut.sim.dcl
VR+baith_pft+pst_aux	VG.cv_pst.pft.dcl
VR+de_pft_ja_pft	VG.cv_pst.pft.dcl
VR+ja_inf+cahiye	VG.cv_prs.sim.sugg
VR+le_pft+ho_impf	VG.cv_prs.pft.cfct
VR+a_pft+ho_opt	VG.cv_0.0.cfct
VR+par_pft+pst_aux	VG.cv_pst.pft.dcl
VR+a_pft+prs_aux	VG.cv_prs.pft.dcl
VR+nikaal_pft+prs_aux	VG.cv_prs.pft.dcl
VR+par_impf+prs_aux	VG.cv_prs.impf.dcl
VR+dal+rah_pft+prs_aux	VG.cv_prs.prog.dcl
VR+dal+rah_pft+pst_aux	VG.cv_pst.prog.dcl
VR+dal+rah_pft+ho_fut	VG.cv_fut.prog.prsm
VR+VINF+lag_impf+prs_aux	VG.cv_prs.impf_ince.dcl
VR+VINF+lag_pft+prs_aux	VG.cv_prs.pft_ince.dcl
VR+VINF+lag_impf+pst_aux	VG.cv_pst.impf_ince.dcl
VR+VINF+lag_pft+pst_aux	VG.cv_pst.pft_ince.dcl
VR+dal_impf+rah_impf+prs_aux	VG.cv_prs.impf_cnt.dcl
VR+dal_impf+rah_pft+prs_aux	VG.cv_prs.impf_pft_cnt.dcl
VR+dal_impf+rah_impf+pst_aux	VG.cv_pst.impf_cnt.dcl

VR+dal_impf+rah_pft+pst_aux	VG.cv_pst.impf_pft_cnt.dcl
VR+dal_impf+ja_impf+prs_aux	VG.cv_prs.impf_dur.dcl
VR+dal_impf+ja_pft+prs_aux	VG.cv_prs.impf_pft_dur.dcl
VR+dal_impf+ja_impf+pst_aux	VG.cv_pst.impf_dur.dcl
VR+dal_impf+ja_pft+pst_aux	VG.cv_prs.impf_pft_dur.dcl
VR+dal_pft+kar_impf+prs_aux	VG.cv_prs.impf_frq.dcl
VR+dal_pft+kar_impf+pst_aux	VG.cv_pst.impf_frq.dcl
VR+dal_pft+kar_impf+ho_fut	VG.cv_fut.pft_impf_dur.dcl
VR+dal_pft+kar_pft+kar_fut	VG.cv_fut.impf_frq.dcl
VR+V2+sak_impf+prs_aux	VG.cv_prs.impf.abil
VR+V2+sak_impf+pst_aux	VG.cv_pst.impf.abil
VR+V2+sak_pft+prs_aux	VG.cv_prs.pft.abil
VR+V2+sak_pft+pst_aux	VG.cv_pst.pft.abil
VR+V2+sak_pft+ho_fut	VG.cv_fut.pft.abil
VR+V2+pa_impf+prs_aux	VG.cv_prs.impf.abil
VR+V2+pa_impf+pst_aux	VG.cv_pst.impf.abil
VR+V2+pa_pft+prs_aux	VG.cv_prs.pft.abil
VR+V2+pa_pft+pst_aux	VG.cv_pst.pft.abil
VR+V2+pa_pft+ho_fut	VG.cv_fut.pft.abil
VR+VINF+de_impf+prs_aux	VG.cv_prs.impf.perm

VR+VINF+de_impf+pst_aux	VG.cv_pst.impf.perm
VR+VINF+de_pft+prs_aux	VG.cv_prs.pft.perm
VR+VINF+de_pft+pst_aux	VG.cv_pst.pft.perm
VR+VINF+de_pft+ho_fut	VG.cv_fut.pft.perm_cfct
VR+VINF+par_impf+prs_aux	VG.cv_prs.impf.oblg
VR+VINF+par_impf+pst_aux	VG.cv_pst.impf.oblg
VR+VINF+par_pft+prs_aux	VG.cv_prs.pft.oblg
VR+VINF+par_pft+pst_aux	VG.cv_pst.pft.oblg
VR+VINF+par_pft+ho_fut	VG.cv_fut.pft.oblg
VR+VINF+ho_impf+prs_aux	VG.cv_prs.impf.oblg
VR+VINF+ho_impf+pst_aux	VG.cv_pst.impf.oblg
VR+VINF+ho_impf+ho_fut	VG.cv_fut.pft.oblg
VR+VINF+rah_pft+ho_fut	VG.cv_fut.prog.oblg_prsm
VR+dal_pft+ja_impf+prs_aux	VG.cv_pas_prs.impf.dcl
VR+dal_pft+ja_impf+pst_aux	VG.cv_pas_pst.impf.dcl
VR+dal_pft+ja_pft+prs_aux	VG.cv_pas_prs.pft.dcl
VR+dal_pft+ja_pft+pst_aux	VG.cv_pas_pst.pft.dcl
VR+dal_pft+ja_pft+ho_fut	VG.cv_pas_fut.pft.dcl
VR+dal_pft+ja_inf+lag_pft	VG.cv_pas_pst.sim_ince.dcl
VR+dal_pft+ja_inf+lag_fut	VG.cv_pas_pst.pft_ince.dcl

VR+dal_pft+ja_impf+rah_pft	VG.cv_pas_pst.pft_cnt.dcl
VR+dal_pft+ja_impf+rah_fut	VG.cv_pas_pst.pft_cnt.dcl
VR+dal_pft+ja_impf+ho_fut	VG.cv_pas_fut.impf.prsm
VR+dal_pft+ja_pft+ho_impf	VG.cv_pas_pst.impf.cfct
VR+dal_pft+ja+sak_fut	VG.cv_pas_fut.sim.abil
VR+dal_pft+ja+sak_pft	VG.cv_pas_pst.sim.abil
VR+de_impf+ja_pft+prs_aux	VG.cv_prs.pft.dcl
VR+de_pft+ja_pft+prs_aux	VG.cv_pas_prs.pft.dcl
VR+de_impf+ja_pft+pst_aux	VG.cv_pst.pft.dcl
VR+de_pft+ja_pft+pst_aux	VG.cv_pas_pst.pft.dcl
VR+ja+rah_pft+prs_aux	VG.cv_prs.prog.dcl
VR+ja+rah_pft+pst_aux	VG.cv_pst.prog.dcl
VR+ja+rah_pft+ho_fut	VG.cv_fut.prog.dcl
VR+de_pft+ja_impf+prs_aux	VG.cv_pas_prs.impf.dcl
VR+de_pft+ja_impf+pst_aux	VG.cv_pst.impf.dcl
VR+cuk_pft+ho_impf+prs_aux	VG.cv_prs.pft.dcl
VR+ja_pft+ho_impf+prs_aux	VG.cv_prs.pft.dcl
VR+ja_inf+de_pft+ho_impf	VG.cv_pst.impf.cfct_perm
VR+de_pft+ja_inf+cahiye	VG.cv_pas_prs.sim.sugg
VR+ja_inf+cahiye+pst_aux	VG.cv_pst.sim.sugg

VR+le_pft+ja_pft+prs_aux	VG.cv_prs.pft.dcl
VR+le_pft+ja_impf+prs_aux	VG.cv_prs.impf.dcl
VR+dal_impf+ja+rah_pft+prs_aux	VG.cv_prs.prog_dur.dcl
VR+dal_impf+ja+rah_pft+pst_aux	VG.cv_pst.prog_dur.dcl
VR+dal_impf+ja+rah_pft+ho_fut	VG.cv_fut.impf_dur.prsm
VR+V2+sak+rah_pft+prs_aux	VG.cv_prs.prog.abil
VR+V2+sak+rah_pft+pst_aux	VG.cv_pst.prog.abil
VR+V2+sak+rah_pft+ho_fut	VG.cv_fut.prog.abil_prsm
VR+V2+pa+rah_pft+prs_aux	VG.cv_prs.prog.abil
VR+V2+pa+rah_pft+pst_aux	VG.cv_pst.prog.abil
VR+V2+pa+rah_pft+ho_fut	VG.cv_fut.prog.abil_prsm
VR+VINF+de+rah_pft+prs_aux	VG.cv_prs.prog.perm
VR+VINF+de+rah_pft+pst_aux	VG.cv_pst.prog.perm
VR+VINF+de+rah_pft+ho_fut	VG.cv_fut.prog.perm_prsm
VR+VINF+par+rah_pft+prs_aux	VG.cv_prs.prog.oblg
VR+VINF+par+rah_pft+pst_aux	VG.cv_pst.prog.oblg
VR+VINF+par+rah_pft+ho_fut	VG.cv_fut.prog.oblg_prsm
VR+dal_pft+ja+rah_pft+prs_aux	VG.cv_pas_prs.prog.dcl
VR+dal_pft+ja+rah_pft+pst_aux	VG.cv_pas_pst.prog.dcl
VR+dal_pft+ja+rah_pft+ho_fut	VG.cv_pas_fut.prog.prsm

VR+dal_pft+ja_inf+lag_impf+prs_aux	VG.cv_pas_prs.impf_ince.dcl
VR+dal_pft+ja_inf+lag_pft+prs_aux	VG.cv_pas_prs.pft_ince.dcl
VR+dal_pft+ja_inf+lag_impf+pst_aux	VG.cv_pas_pst.pft_ince.dcl
VR+dal_pft+ja_inf+lag_pft+pst_aux	VG.cv_pas_pst.impf_ince.dcl
VR+dal_pft+ja_impf+rah_impf+prs_aux	VG.cv_pas_prs.pft_impf_cnt.dcl
VR+dal_pft+ja_impf+rah_pft+prs_aux	VG.cv_pas_prs.impf_cnt.dcl
VR+dal_pft+ja_impf+rah_impf+pst_aux	VG.cv_pas_pst.pft_cnt.dcl
VR+dal_pft+ja_impf+rah_pft+pst_aux	VG.cv_pas_pst.impf_cnt.dcl
VR+dal_impf+ja_impf+rah_pft+prs_aux	VG.cv_pas_prs.pft_dur.dcl
VR+dal_pft+ja+sak_impf+prs_aux	VG.cv_pas_prs.impf.abil
VR+dal_pft+ja+sak_impf+pst_aux	VG.cv_pas_pst.impf.abil
VR+dal_pft+ja+sak_pft+prs_aux	VG.cv_pas_prs.pft.abil
VR+dal_pft+ja+sak_pft+pst_aux	VG.cv_pas_pst.pft.abil
VR+dal_pft+ja+sak_pft+ho_fut	VG.cv_pas_fut.pft.abil
VR+dal_impf+cal_pft+ja+rah_pft+prs_aux	VG.cv_prs.prog_dur.dcl
VR+dal_impf+cal_pft+ja+rah_pft+pst_aux	VG.cv_pst.prog_dur.dcl
VR+dal_impf+cal_pft+ja+rah_pft+ho_fut	VG.cv_fut.impf.cfct
VR+dal_impf+cal_pft+a+rah_pft+prs_aux	VG.cv_prs.prog_dur.dcl
VR+dal_impf+cal_pft+a+rah_pft+pst_aux	VG.cv_pst.prog_dur.dcl
VR+dal_impf+cal_pft+a+rah_pft+ho_fut	VG.cv_pst.impf.cfct

VR+VINF+cahiye	VG.cv_0.0.sugg
VR+VINF+cahiye+pst_aux	VG.cv_pst.0.sugg
VR_pft+a+rah_pft	VG.ip.pft_dur
VR+ja_pft+ho_opt	VG.cv_0.0.cfct
cal_pft+a_pft	VG.cv_pst.pft.dcl
laa+khada+kar+de_pft	VG.cv_pst.pft.dcl
VR+ke	VG.cp.pft
khada_pft+VR_kar	VG.cv_cp.pft

APPENDIX 2: WORD VARIABLES

ke	के
wala	वाला/वाली/वाले
khada	खड़ा/खड़ा
khada_pft	खड़ा/खड़ी/खड़ीं/खड़े/खड़ा/खड़ी/खड़ीं/खड
prs_aux	है/हैं/हो/हूं/हूँ
pst_aux	था/थी/थीं/थे
ho_cfct	हो/हों
cahiye	चाहिए/चाहिये
jane	जाने
particles	नहीं/ना/ही/तो/भी/सा/सी/से/क्या/न/वाला/वाली/वाले
a	आ
a_fut	आएगा/आयेगा/आएगी/आयेगी/आएंगी/आयेंगी/आईएगा/आईयेगा/आओगे/आओगी/ आउंगा/आऊंगा/आउंगी/आऊंगी
a_imp	आ/आओ/आईए

a_impf	आता/आती/आतीं/आते
a_inf	आना/आनी/आने
a_opt	आऊं/आऊँ/आएं/आयें
a_pft	आया/आई/आईं/आए/आये
bach	बच
bach_fut	बचेगा/बचेगी/बचेंगी/बचेंगे/बचिएगा/बचियेगा/बचोगे/बचोगी/बचूंगा/बचूंगी
bach_imp	बच/बचो/बचें/बचिए/बचिये
bach_impf	बचता/बचती/बचतीं/बचते
bach_inf	बचना/बचनी/बचने
bach_opt	बचूं/बचूँ/बचे/बचें
bach_pft	बचा/बची/बचीं/बचे
baith	बैठ
baith_fut	बैठेगा/बैठेगी/बैठेंगी/बैठेंगे/बैठिएगा/बैठियेगा/बैठोगे/बैठोगी/बैठूंगा/बैठूंगी
baith_imp	बैठ/बैठो/बैठें/बैठिए/बैठिये
baith_impf	बैठता/बैठती/बैठतीं/बैठते

baith_inf	बैठना/बैठनी/बैठने
baith_opt	बैठूं/बैठूँ/बैठे/बैठें
baith_pft	बैठा/बैठी/बैठीं/बैठे
baitha	बैठा
baitha_fut	बैठाएगा/बैठायेगा/बैठाएगी/बैठायेगी/बैठाएंगी/बैठाएँगी/बैठायेंगी/बैठाईएगा/बैठाईयेगा/ बैठाओगे/बैठाओगी/बैठाउंगा/बैठाऊंगा/बैठाउंगी/बैठाऊंगी/बिठाएगा/बिठायेगा/बिठाएगी/ बिठायेगी/बिठाएंगी/बिठाएँगी/बिठायेंगी/बिठाइएगा/बिठाइयेगा/बिठाओगे/बिठाओगी/ बिठाउंगा/बिठाऊंगा/बिठाउंगी/बिठाऊंगी
baitha_imp	बैठा/बैठाओ/बैठाईए/बिठा/बिठाओ/बिठाईए
baitha_impf	बैठाता/बैठाती/बैठातीं/बैठाते/बिठाता/बिठाती/बिठातीं/बिठाते
baitha_inf	बैठाना/बिठाना/बिठानी/बैठानी/बिठाने/बैठाने
baitha_opt	बैठाऊं/बैठाऊँ/बैठाएं/बैठाएँ/बैठाएँ/बैठायें/बिठाऊं/बिठाऊँ/बिठाएं/बिठाएँ/बिठायें
baitha_pft	बैठाया/बैठाई/बैठाईं/बैठाए/बैठाये/बिठाया/बिठाई/बिठाईं/बिठाए/बिठाये
ban	बन
ban_fut	बनेगा/बनेगी/बनेंगी/बनेंगे/बनिएगा/बनियेगा/बनोगे/बनोगी/बनूंगा/बनूंगी
ban_imp	बन/बनो/बनें/बनिए/बनिये

ban_impf	बनता/बनती/बनतीं/बनते
ban_inf	बनना/बननी/बनने
ban_opt	बनूं/बनूँ/बने/बनें
ban_pft	बना/बनी/बनीं/बने
bas	बस
bas_fut	बसेगा/बसेगी/बसेंगी/बसेंगे/बसिएगा/बसियेगा/बसोगे/बसोगी/बसूंगा/बसूंगी
bas_imp	बस/बसो/बसें/बसिए/बसिये
bas_impf	बसता/बसती/बसतीं/बसते
bas_inf	बसना/बसनी/बसने
bas_opt	बसूं/बसूँ/बसे/बसें
bas_pft	बसा/बसी/बसीं/बसे
bhag	भाग
bhag_fut	भागेगा/भागेगी/भागेंगी/भागेंगे/भागिएगा/भागियेगा/भागोगे/भागोगी/भागूंगा/भागूंगी
bhag_imp	भाग/भागो/भागें/भागिए/भागिये
bhag_impf	भागता/भागती/भागतीं/भागते

bhag_inf	भागना/भागनी/भागने
bhag_opt	भागूं/भागूँ/भागे/भागें
bhag_pft	भागा/भागी/भागीं/भागे
bhej	भेज
bhej_fut	भेजेगा/भेजेगी/भेजेंगी/भेजेंगे/भेजिएगा/भेजियेगा/भेजोगे/भेजोगी/भेजूंगा/भेजूंगी
bhej_imp	भेज/भेजो/भेजें/भेजिए/भेजिये
bhej_impf	भेजता/भेजती/भेजतीं/भेजते
bhej_inf	भेजना/भेजनी/भेजने
bhej_opt	भेजूं/भेजूँ/भेजे/भेजें
bhej_pft	भेजा/भेजी/भेजीं/भेजे
bit	बीत
bit_fut	बीतेगा/बीतेंगी/बीतेगी/बीतेंगे/बीतिएगा/बीतियेगा/बीतोगे/बीतोगी/बीतूंगा/बीतूंगी
bit_imp	बीत/बीतो/बीतें/बीतिए/बीतिये
bit_impf	बीतता/बीतती/बीततीं/बीतते
bit_inf	बीतना/बीतने/बीतनी

bit_opt	बीतूं/बीतूँ/बीते/बीतें
bit_pft	बीता/बीती/बीतीं/बीते
cah	चाह
cah_fut	चाहेगा/चाहेगी/चाहेंगी/चाहेंगे/चाहिएगा/चाहियेगा/चाहोगे/चाहोगी/चाहूंगा/चाहूंगी
cah_imp	चाह/चाहो/चाहें/चाहिए/चाहिये
cah_impf	चाहता/चाहती/चाहतीं/चाहते
cah_inf	चाहना/चाहनी/चाहने
cah_opt	चाहूं/चाहूँ/चाहें/चाहे
cah_pft	चाहा/चाही/चाहीं/चाहे
cal	चल
cal_fut	चलेगा/चलेगी/चलेंगी/चलेंगे/चलिएगा/चलियेगा/चलोगे/चलोगी/चलूंगा/चलूंगी
cal_imp	चल/चलो/चलें/चलिए/चलिये
cal_impf	चलता/चलती/चलतीं/चलते
cal_inf	चलना/चलनी/चलने
cal_opt	चलूं/चलूँ/चलें/चले

cal_pft	चला/चली/चलीं/चले
chadh	चढ़/चढ़
chadh_fut	चढ़ेगा/चढ़ेगी/चढ़ेंगी/चढ़ेंगे/चढ़िएगा/चढ़ियेगा/चढ़ोगे/चढ़ोगी/चढ़ूंगी/चढ़ूंगा/चढ़ेगा/चढ़ेगी/चढ़ेंगी/चढ़ेंगे/चढ़िएगा/चढ़ियेगा/चढ़ोगे/चढ़ोगी/चढ़ूंगी/चढ़ूंगा
chadh_imp	चढ़/चढ़ो/चढ़ें/चढ़िए/चढ़िये/चढ़/चढ़ो/चढ़ें/चढ़िए/चढ़िये
chadh_impf	चढ़ता/चढ़ती/चढ़तीं/चढ़ते/चढ़ता/चढ़ती/चढ़तीं/चढ़ते
chadh_inf	चढ़ना/चढ़ना/चढ़नी/चढ़नी/चढ़ने/चढ़ने
chadh_opt	चढ़ूं/चढ़ूँ/चढ़ें/चढ़े/चढ़ूं/चढ़ें/चढ़े
chadh_pft	चढ़ा/चढ़ी/चढ़ीं/चढ़े/चढ़ा/चढ़ी/चढ़ीं/चढ़े
chod	छोड़/छोड़
chod_fut	छोड़ेगा/छोड़ेगी/छोड़ेंगी/छोड़ेंगे/छोड़िएगा/छोड़ियेगा/छोड़ोगे/छोड़ोगी/छोड़ूंगा/छोड़ूंगी/छोड़ेगा/छोड़ेगी/छोड़ेंगी/छोड़ेंगे/छोड़िएगा/छोड़ियेगा/छोड़ोगे/छोड़ोगी/छोड़ूंगा/छोड़ूंगी
chod_imp	छोड़/छोड़ो/छोड़ें/छोड़िए/छोड़िये/छोड़ो/छोड़ें/छोड़िए/छोड़िये
chod_impf	छोड़ता/छोड़ती/छोड़तीं/छोड़ते/छोड़ता/छोड़ती/छोड़तीं/छोड़ते
chod_inf	छोड़ना/छोड़नी/छोड़ने/छोड़ना/छोड़नी/छोड़ने
chod_opt	छोड़ूं/छोड़ूँ/छोड़ें/छोड़े/छोड़ूं/छोड़ें/छोड़े

chod_pft	छोड़ा/छोड़ी/छोड़ीं/छोड़े/छोड़ा/छोड़ी/छोड़ीं/छोड़े
cuk	चुक
cuk_fut	चुकेगा/चुकेगी/चुकेंगी/चुकेंगे/चुकिएगा/चुकियेगा/चुकोगे/चुकोगी/चुकूंगा/चुकूंगी
cuk_imp	चुक/चुको/चुकें/चुकिए/चुकिये
cuk_impf	चुकता/चकती/चुकतीं/चुकते
cuk_inf	चुकना/चुकनी/चुकने
cuk_opt	चुकूं/चुकूँ/चुकें/चुके
cuk_pft	चुका/चुकी/चुकीं/चुके
dal	डाल
dal_fut	डालेगा/डालेगी/डालेंगी/डालेंगे/डालिएगा/डालियेगा/डालोगे/डालोगी/डालूंगा/डालूंगी
dal_imp	डाल/डालो/डालें/डालिए/डालिये
dal_impf	डालता/डालती/डालतीं/डालते
dal_inf	डालना/डालनी/डालने
dal_opt	डालूं/डालूँ/डालो/डालें
dal_pft	डाला/डाली/डालीं/डाले

daud	दौड़/दौड़
daud_fut	दौड़ेगा/दौड़ेगी/दौड़ेंगी/दौड़ेंगे/दौड़िएगा/दौड़ियेगा/दौड़ोगे/दौड़ोगी/दौड़ूंगा/दौड़ूंगी/दौड़ेगा/दौड़ेगी/दौड़ेंगी/दौड़ेंगे/दौड़िएगा/दौड़ियेगा/दौड़ोगे/दौड़ोगी/दौड़ूंगा/दौड़ूंगी
daud_imp	दौड़/दौड़ो/दौड़ें/दौड़िए/दौड़िये/दौड़/दौड़ो/दौड़े/दौड़िए/दौड़िये
daud_impf	दौड़ता/दौड़ती/दौड़तीं/दौड़ते/दौड़ता/दौड़ती/दौड़तीं/दौड़ते
daud_inf	दौड़ना/दौड़नी/दौड़ने/दौड़ना/दौड़नी/दौड़ने
daud_opt	दौड़ूं/दौड़ूँ/दौड़े/दौड़ें/दौड़ूं/दौड़ूँ/दौड़ें/दौड़े
daud_pft	दौड़ा/दौड़ी/दौड़ीं/दौडे/दौड़ा/दौड़ी/दौड़ीं/दौड़े
de	दे
de_fut	देगा/देगी/देंगी//देंगे/दोगे/दोगी/दीजिएगा/दीजियेगा/दूंगा/दूंगी/दूँगा/दूँगी
de_imp	दे/दो/दें/दीजिए/दीजिये
de_impf	देता/देती/देतीं/देते
de_inf	देना/देनी/देने
de_opt	दूं/दूँ/दें/दो
de_pft	दिया/दिए/दिये/दी/दीं
dekh	देख

dekh_fut	देखेगा/देखेगी/देखेंगी/देखेंगे/देखोगे/देखोगी/देखिएगा/देखियेगा/देखूंगा/देखूंगी
dekh_imp	देख/देखो/देखें/देखिए/देखिये
dekh_impf	देखता/देखती/देखतीं/देखते
dekh_inf	देखना/देखनी/देखने
dekh_opt	देखूं/देखूँ/देखें/देखे
dekh_pft	देखा/देखी/देखे/देखीं
dhamak	धमक
dhamak_fut	धमकेगा/धमकेगी/धमकेंगी/धमकेंगे/धमकिएगा/धमकियेगा/धमकोगे/धमकोगी/धमकूंगा/धमकूंगी
dhamak_imp	धमक/धमको/धमकें/धमकिए/धमकिये
dhamak_impf	धमकता/धमकती/धमकतीं/धमकते
dhamak_inf	धमकना/धमकनी/धमकने
dhamak_opt	धमकूं/धमकूँ/धमके/धमकें
dhamak_pft	धमका/धमकी/धमकीं/धमके
dhar	धर

dhar_fut	धरेगा/धरेगी/धरेंगी/धरेंगे/धरिएगा/धरियेगा/धरोगे/धरोगी/धरूंगा/धरूंगी
dhar_imp	धर/धरो/धरें/धरिए/धरिये
dhar_impf	धरता/धरती/धरतीं/धरते
dhar_inf	धरना/धरने/धरनी
dhar_opt	धरूं/धरूँ/धरें/धरे
dhar_pft	धरा/धरी/धरीं/धरे
dikh	दिख
dikh_fut	दिखेगा/दिखेगी/दिखेंगी/दिखेंगे/दिखोगे/दिखोगी/दिखिएगा/दिखियेगा/दिखूंगा/दिखूंगी
dikh_imp	दिख/दिखो/दिखें/दिखिए/दिखिये
dikh_impf	दिखता/दिखती/दिखतीं/दिखते
dikh_inf	दिखना/दिखनी/दिखने
dikh_opt	दिखूं/दिखूँ/दिखो/दिखें
dikh_pft	दिखा/दिखी/दिखीं/दिखे
dikha	दिखा
dikha_fut	दिखाएगा/दिखायेगा/दिखाएगी/दिखायेगी/दिखाएंगी/दिखायेंगी/दिखाईएगा/दिखाईयेगा /दिखाओगे/दिखाओगी/दिखाउंगा/दिखाऊंगा/दिखाऊँगा/दिखाउंगी/दिखाऊँगी/दिखाऊंगी

dikha_imp	दिखा/दिखाओ/दिखाईए
dikha_impf	दिखाता/दिखाती/दिखातीं/दिखाते
dikha_inf	दिखाना/दिखानी/दिखाने
dikha_opt	दिखाऊँ/दिखाऊं/दिखाएं/दिखायें
dikha_pft	दिखाया/दिखाई/दिखाईं/दिखाए/दिखाये
ghasit	घसीट
ghasit_fut	घसीटेगा/घसीटेगी/घसीटेंगी/घसीटेंगे/घसीटिएगा/घसीटियेगा/घसीटोगे/घसीटोगी/घसीटेंगी/घसीटूंगा/घसीटूंगी
ghasit_imp	घसीट/घसीटें/घसीटो/घसीटिए
ghasit_impf	घसीटता/घसीटती/घसीटतीं/घसीटते
ghasit_inf	घसीटना/घसीटनी/घसीटने
ghasit_opt	घसीटूं/घसीटूँ/घसीटें/घसीटे
ghasit_pft	घसीटा/घसीटी/घसीटीं/घसीटे
gir	गिर
gir_fut	गिरेगा/गिरेगी/गिरेंगी/गिरेंगे/गिरोगे/गिरोगी/गिरिएगा/गिरियेगा/गिरूंगा/गिरूंगा/गिरूंगी/गिरूंगी

gir_imp	गिर/गिरें/गिरो/गिरिए/गिरिये
gir_impf	गिरता/गिरती/गिरतीं/गिरते
gir_inf	गिरना/गिरनी/गिरने
gir_opt	गिरूं/गिरूँ/गिरें/गिरे
gir_pft	गिरा/गिरी/गिरीं/गिरे
giraa	गिरा
giraa_fut	गिराएगा/गिरायेगा/गिराएगी/गिरायेगी/गिराएंगी/गिरायेंगी/गिराओगी/गिराओगे/गिराएंगे/गिरायेंगे/गिराईएगा/गिराईयेगा/गिराऊंगा/गिराऊंगी
giraa_imp	गिरा/गिराएं/गिराओ/गिराइए/गिराइये
giraa_impf	गिराता/गिराती/गिरातीं/गिराते
giraa_inf	गिराना/गिरानी/गिराने
giraa_opt	गिराऊं/गिराऊँ/गिराएं/गिराए
giraa_pft	गिराया/गिरायी/गिराई/गिरायीं/गिराईं/गिराए/गिराये
gujar	गुजर
gujar_fut	गुजरेगा/गुजरेगी/गुजरेंगी/गुजरेंगे/गुजरिएगा/गुजरियेगा/गुजरोगे/गुजरेगी/गुजरूंगा/गुजरूंगी

gujar_imp	गुजर/गुजरें/गुजरो/गुजरिए
gujar_impf	गुजरता/गुजरती/गुजरतीं/गुजरते
gujar_inf	गुजरना/गुजरनी/गुजरने
gujar_opt	गुजरूं/गुजरूँ/गुजरें/गुजरे
gujar_pft	गुजरा/गुजरी/गुजरीं/गुजरे
ho	हो
ho_fut	होगा/होगी/होंगीं/होंगी/होंगे/होइएगा/होइयेगा/होगे/होगी/हूंगा/होऊंगा/हूंगी/होऊंगी
ho_imp	हो/होओ/होईए/होईये/होइए/होइये
ho_impf	होता/होती/होतीं/होते
ho_inf	होना/होनी/होने
ho_opt	होऊं/होऊँ/हों/होएं/होए/होवे/होवें
ho_pft	हुआ/हुई/हुईं/हुए/हुये
ja	जा
ja_fut	जाएगा/जायेगा/जाएगी/जाएँगी/जायेगी/जाएंगी/जायेंगी/जाएंगे/जाएँगे/जायेंगे/जाईएगा/ जाओगे/जाओगी/जाऊंगा/जाऊंगी
ja_imp	जा/जाओ/जाएं/जायें/जाईए/जाईये

ja_impf	जाता/जाती/जातीं/जाते
ja_inf	जाना/जानी/जाने
ja_opt	जाऊं/जाऊँ/जाए/जाय/जाएं/जाएँ/जाये/जायें
ja_pft	गया/गयी/गयीं/गई/गईं/गये/गए
Kar	कर
kar_fut	करेगा/करेगी/करेंगी/करेंगे/करियेगा/करिएगा/करोगे/करोगी/करूंगी/करूंगा
kar_imp	कर/करें/करो/करिए/करिये
kar_impf	करता/करती/करतीं/करते
kar_inf	करना/करनी/करने
kar_opt	करूं/करूँ/करें/करे
kar_pft	किया/किए/की/कीं
Kha	खा
kha_fut	खाएगा/खाएगी/खाएंगी/खायेगी/खायेंगी/खाएंगे/खायेंगे/खाओगे/खाओगी/खाईएगा/खाईयेगा/खाउंगा/खाऊंगा/खाउंगी/खाऊंगी
kha_imp	खा/खाएं/खाओ/खाईए/खाईये
kha_impf	खाता/खाती/खातीं/खाते

kha_inf	खाना/खानी/खाने
kha_opt	खाऊं/खाऊँ/खाएं/खायें/खाए/खाएँ
kha_pft	खाया/खायी/खाई//खायीं/खाईं/खाए
Laa	ला
laa_fut	लाएगा/लायेगा/लाएगी/लायेगी/लाएंगी/लायेंगी/लाईएगा/लाईयेगा/लाओगे/लाओगी/लाउंगा/लाऊंगा/लाउंगी/लाऊंगी
laa_imp	ला/लाओ/लाईए
laa_impf	लाता/लाती/लातीं/लाते
laa_inf	लाना/लानी/लाने
laa_opt	लाऊं/लाऊँ/लाएं/लाएँ/लायें
laa_pft	लाया/लाई/लाईं/लाए/लाये
Lag	लग
lag_fut	लगेगा/लगेगी/लगेंगी/लगेंगे/लगिएगा/लगियेगा/लगोगे/लगोगी/लगूंगा/लगूंगी
lag_imp	लग/लगो/लगें/लगिए/लगिये
lag_impf	लगता/लगती/लगतीं/लगते
lag_inf	लगना/लगनी/लगने

lag_opt	लगूं/लगूँ/लगे/लगें
lag_pft	लगा/लगी/लगीं/लगे
Laga	लगा
laga_fut	लगाएगा/लगायेगा/लगाएगी/लगायेगी/लगाएंगी/लगायेंगी/लगाईएगा/लगाईयेगा/लगाओगे/लगाओगी/लगाउंगा/लगाऊंगा/लगाऊँगा/लगाउंगी/लगाउँगी/लगाऊंगी/लगाऊँगी
laga_imp	लगा/लगाओ/लगाईए
laga_impf	लगाता/लगाती/लगातीं/लगाते
laga_inf	लगाना/लगानी/लगाने
laga_opt	लगाऊं/लगाऊँ/लगाएं/लगाएँ/लगायें
laga_pft	लगाया/लगाई/लगाईं/लगाए/लगाये
Le	ले
le_fut	लेगा/लेगी/लेंगी//लेंगे/लोगे/लोगी/लीजिएगा/लीजियेगा/लूंगा/लूंगी
le_imp	ले/लो/लें/लीजिए/लीजिये
le_impf	लेता/लेती/लेतीं/लेते
le_inf	लेना/लेनी/लेने
le_opt	लूं/लूँ/लें/लो

le_pft	लिया/लिए/लिये/ली/लीं
maar	मार
maar_fut	मारेगा/मारेगी/मारेंगी/मारेंगे/मारिएगा/मारियेगा/मारोगे/मारोगी/मारूंगा/मारूंगी
maar_imp	मार/मारो/मारें/मारिए/मारिये
maar_impf	मारता/मारती/मारतीं/मारते
maar_inf	मारना/मारनी/मारने
maar_opt	मारूं/मारूँ/मारे/मारें
maar_pft	मारा/मारी/मारीं/मारे
mar	मर
mar_fut	मरेगा/मरेगी/मरेंगी/मरेंगे/मरिएगा/मरियेगा/मरोगे/मरोगी/मरूंगा/मरूंगी
mar_imp	मर/मरो/मरें/मरिए/मरिये
mar_impf	मरता/मरती/मरतीं/मरते
mar_inf	मरना/मरनी/मरने
mar_opt	मरूं/मरूँ/मरे/मरें
mar_pft	मरा/मरी/मरीं/मरे

mit	मिट
mit_fut	मिटेगा/मिटेगी/मिटेंगी/मिटेंगे/मिटिएगा/मिटियेगा/मिटोगे/मिटोगी/मिटूंगा/मिटूंगी
mit_imp	मिट/मिटो/मिटें/मिटिए/मिटिये
mit_impf	मिटता/मिटती/मिटतीं/मिटते
mit_inf	मिटना/मिटनी/मिटने
mit_opt	मिटूं/मिटूँ/मिटे/मिटें
mit_pft	मिटा/मिटी/मिटीं/मिटे
mita	मिटा
mita_fut	मिटाएगा/मिटायेगा/मिटाएंगी/मिटायेगी/मिटाएंगी/मिटायेंगी/मिटाईएगा/मिटाईयेगा/ मिटाओगे/मिटाओगी/मिटाउंगा/मिटाऊंगा/मिटाउंगी/मिटाऊंगी
mita_imp	मिटा/मिटाओ/मिटाईए
mita_impf	मिटाता/मिटाती/मिटातीं/मिटाते
mita_inf	मिटाना/मिटानी/मिटाने
mita_opt	मिटाऊं/मिटाऊँ/मिटाएं/मिटाएँ/मिटायें
mita_pft	मिटाया/मिटाई/मिटाईं/मिटाए/मिटाये
nikaal	निकाल

nikaal_fut	निकालेगा/निकालेगी/निकालेंगी/निकालेंगे/निकालोगे/निकालोगी/निकालिएगा/निकालियेगा/निकालूंगा/निकालूंगी
nikaal_imp	निकाल/निकालो/निकालें/निकालिए/निकालिये
nikaal_impf	निकालता/निकालती/निकालतीं/निकालते
nikaal_inf	निकालना/निकालनी/निकालने
nikaal_opt	निकालूं/निकालूँ/निकालें/निकाले
nikaal_pft	निकाला/निकाली/निकालीं/निकाले
nikal	निकल
nikal_fut	निकलेगा/निकलेगी/निकलेंगी/निकलेंगे/निकलोगे/निकलोगी/निकलिएगा/निकलियेगा/निकलूंगा/निकलूंगी
nikal_imp	निकल/निकलो/निकलें/निकलिए/निकलिये
nikal_impf	निकलता/निकलती/निकलतीं/निकलते
nikal_inf	निकलना/निकलनी/निकलने
nikal_opt	निकलूं/निकलूँ/निकलें/निकले
nikal_pft	निकला/निकली/निकलीं/निकले
pa	पा

pa_fut	पाएगा/पायेगा/पाएगी/पायेगी/पाएंगी/पाएँगी/पायेंगी/पाएंगे/पाएँगे/पायेंगे/पाईएगा/पाईयेगा/पाओगे/पाओगी/पाउंगा/पाऊंगा/पाउंगी/पाऊंगी
pa_imp	पा/पाओ/पाईए
pa_impf	पाता/पाती/पातीं/पाते
pa_inf	पाना/पानी/पाने
pa_opt	पाऊं/पाऊँ/पाएं/पाएँ/पायें
pa_pft	पाया/पाई/पायी/पाईं/पायीं/पाए/पाये
pahunch	पहुँच/पहुंच
pahunch_fut	पहुंचेगा/पहुंचेगी/पहुंचेंगी/पहुंचेंगे/पहुंचोगे/पहुंचोगी/पहुंचिएगा/पहुंचियेगा/पहुंचूंगा/पहुंचूंगी/पहुँचेगा/पहुँचेगी/पहुँचेंगी/पहुँचेंगे/पहुँचोगे/पहुँचोगी/पहुँचिएगा/पहुँचियेगा/पहुँचूंगा/पहुँचूंगी
pahunch_imp	पहुंच/पहुंचो/पहुंचें/पहुंचिए/पहुंचिये/पहुँच/पहुँचो/पहुँचें/पहुँचिए/पहुँचिये
pahunch_impf	पहुंचता/पहुंचती/पहुंचतीं/पहुंचते/पहुँचता/पहुँचती/पहुँचतीं/पहुँचते
pahunch_inf	पहुंचना/पहुंचनी/पहुंचने/पहुँचना/पहुँचनी/पहुँचने
pahunch_opt	पहुंचूं/पहुंचें/पहुंचे/पहुँचूं/पहुँचें/पहुँचे
pahunch_pft	पहुंचा/पहुंची/पहुंचीं/पहुंचे/पहुँचा/पहुँची/पहुँचीं/पहुँचे

par	पड़/पड़
par_fut	पड़ेगा/पड़ेगी/पड़ेंगी/पड़ेंगी/पड़ेंगे/पड़िएगा/पड़ियेगा/पड़ोगे/पड़ोगी/पड़ूंगी/पड़ूंगा/पड़ेगा/पड़ेगी/पड़ेंगी/पड़ेंगी/पड़ेंगे/पड़िएगा/पड़ियेगा/पड़ोगे/पड़ोगी/पड़ूंगी/पड़ूंगा
par_imp	पड़/पड़ो/पड़ें/पड़िए/पड़िये/पड़/पड़ो/पड़ें/पड़िए/पड़िये
par_impf	पड़ता/पड़ती/पड़तीं/पड़ते/पड़ता/पड़ती/पड़तीं/पड़ते
par_inf	पड़ना/पड़नी/पड़ने/पड़ना/पड़नी/पड़ने
par_opt	पड़ूं/पड़ूँ/पड़ें/पड़े/पड़ूं/पड़ूँ/पड़ें/पड़े
par_pft	पड़ा/पड़ी/पड़ीं/पड़े/पड़ा/पड़ी/पड़ीं/पड़े
patak	पटक
patak_fut	पटकेगा/पटकेगी/पटकेंगी/पटकेंगे/पटकोगे/पटकोगी/पटकिएगा/पटकियेगा/पटकूंगा/पटकूंगी
patak_imp	पटक/पटको/पटकें/पटकिए/पटकिये
patak_impf	पटकता/पटकती/पटकतीं/पटकते
patak_inf	पटकना/पटकनी/पटकने
patak_opt	पटकूं/पटकूँ/पटकें/पटके
patak_pft	पटका/पटकी/पटकीं/पटके

phans	फँस/फंस
phans_fut	फंसेगा/फंसेगी/फंसेगी/फंसेंगे/फंसिएगा/फंसियेगा/फंसोगे/फंसोगी/फंसूंगा/फंसूंगी/फँसेगा/फँसेगी/फँसेगी/फँसेंगे/फँसिएगा/फँसियेगा/फँसोगे/फँसोगी/फँसूँगा/फँसूँगी
phans_imp	फंस/फंसो/फंसें/फंसिए/फंसिये/फँस/फँसो/फँसें/फँसिए/फँसिये
phans_impf	फंसता/फंसती/फंसतीं/फंसते/फँसता/फँसती/फँसतीं/फँसते
phans_inf	फंसना/फंसनी/फंसने/फँसना/फँसनी/फँसने
phans_opt	फंसूं/फंसूँ/फंसे/फंसें/फँसूँ/फँसूँ/फँसे/फँसें
phans_pft	फंसा/फंसी/फंसीं/फंसे/फँसा/फँसी/फँसीं/फँसे
phenk	फेंक
phenk_fut	फेंकेगा/फेंकेगी/फेंकेंगी/फेंकेंगे/फेंकिएगा/फेंकियेगा/फेंकोगे/फेंकोगी/फेंकूंगा/फेंकूंगी
phenk_imp	फेंक/फेंको/फेंकें/फेंकिए/फेंकिये
phenk_impf	फेंकता/फेंकती/फेंकतीं/फेंकते
phenk_inf	फेंकना/फेंकनी/फेंकने
phenk_opt	फेंकूं/फेंकूँ/फेंके/फेंकें
phenk_pft	फेंका/फेंकी/फेंकीं/फेंके
rah	रह

rah_fut	रहेगा/रहेगी/रहेंगी/रहेंगे/रहिएगा/रहियेगा/रहोगे/रहोगी/रहूंगा/रहूंगी
rah_imp	रह/रहो/रहें/रहिए/रहिये
rah_impf	रहता/रहती/रहतीं/रहते
rah_inf	रहना/रहनी/रहने
rah_opt	रहूं/रहूँ/रहे/रहें
rah_pft	रहा/रही/रहीं/रहे
rakh	रख
rakh_fut	रखेगा/रखेगी/रखेंगी/रखेंगे/रखिएगा/रखियेगा/रखोगे/रखोगी/रखूंगा/रखूंगी
rakh_imp	रख/रखो/रखें/रखिए/रखिये
rakh_impf	रखता/रखती/रखतीं/रखते
rakh_inf	रखना/रखनी/रखने
rakh_opt	रखूं/रखूँ/रखे/रखें
rakh_pft	रखा/रखी/रखीं/रखे
sak	सक
sak_fut	सकेगा/सकेगी/सकेंगी/सकेंगे/सकिएगा/सकियेगा/सकोगे/सकोगी/सकूंगा/सकूँगा/सकूँगी/सकूंगी

sak_imp	सक/सको/सकें/सकिए/सकिये
sak_impf	सकता/सकती/सकतीं/सकते
sak_inf	सकना/सकनी/सकने
sak_opt	सकूं/सकूँ/सके/सकें
sak_pft	सका/सकी/सकीं/सके
suna	सुना
suna_fut	सुनाएगा/सुनायेगा/सुनाएंगी/सुनायेगी/सुनाएंगी/सुनायेंगी/सुनाईएगा/सुनाईयेगा/सुनाओगे/सुनाओगी/सुनाउंगा/सुनाऊंगा/सुनाउंगी/सुनाऊंगी
suna_imp	सुना/सुनाओ/सुनाईए
suna_impf	सुनाता/सुनाती/सुनातीं/सुनाते
suna_inf	सुनाना/सुलानी/सुलाने
suna_opt	सुनाऊं/सुनाऊँ/सुनाएं/सुनाएँ/सुनायें
suna_pft	सुनाया/सुनाई/सुनाईं/सुनाए/सुनाये
tapak	टपक
tapak_fut	टपकेगा/टपकेगी/टपकेंगी/टपकेंगे/टपकोगे/टपकोगी/टपकिएगा/टपकियेगा/टपकूंगी/टपकूंगा

tapak_imp	टपक/टपको/टपकें/टपकिए/टपकिये
tapak_impf	टपकता/टपकती/टपकतीं/टपकते
tapak_inf	टपकना/टपकनी/टपकने
tapak_opt	टपकूं/टपकूँ/टपके/टपकें
tapak_pft	टपका/टपकी/टपकीं/टपके
uth	उठ
uth_fut	उठेगा/उठेगी/उठेंगी/उठेंगे/उठोगे/उठोगी/उठिएगा/उठियेगा/उठूंगा/उठूंगी
uth_imp	उठ/उठें/उठो/उठिए/उठिये
uth_impf	उठता/उठती/उठतीं/उठते
uth_inf	उठना/उठनी/उठने
uth_opt	उठूं/उठूँ/उठें/उठे
uth_pft	उठा/उठी/उठे/उठें/उठीं

APPENDIX 3: TAG VARIABLES

Variable	Value(s) to found in the tags
VM	VM
VAUX	VAUX
NV	NV
V1	nfn
VINF	ifn
VINF_imp	ifn,imp
VR	nfn
VR_fut	fut
VR_impf	impf,fin
VR_opt	opt
VR_pft	pft
VR_imp	imp
VR_kar	nfn,pft
VR_pcpl	impf,nfn
VR_pcpl_grnd	impf,fin
VR_opt1sg	opt,1,sg

APPENDIX 4: EXAMPLE SENTENCES AND WITH THEIR GLOSS

E1	वह	घर	गया
IPA	wəh	gʰər	gəja:
Gl	3SG	home	go.PFT.MSG
EGl	He went home.		

E2	वह	घर	गया	है
IPA	wəh	gʰər	gəja:	hɛ:
Gl	3SG	home	go.PFT.MSG	be.PRS.SG
EGl	He has gone home.			

E3	वह	गया
IPA	wəh	gəja:
Gl	3SG	go.PFT.MSG
EGl	He went.	

E4	वह	जाएगा
MB	wəh	ja:-e:ga:

Gl	3SG	go-FUT.3MSG
EGl	He will go.	

E5	वह	जाता	है
IPA	wəh	ɟa:-ta:	hɛ:
Gl.	3SG	go.IMPF.MSG	be.PRS.SG
EGl	He goes.		

E6	वह	घर	जा	रहा	है
IPA	wəh	gʰər	ɟa:	rəh-a:	hɛ:
Gl.	3SG	home	go.VR	live-PFT.MSG	be.PRS-SG
EGl	He is going home.				

E7	वह	घर	जा	ही	रहा	है
IPA	wəh	gʰər	ɟa:	hi:	rəh-a:	hɛ:
Gl.	3SG	home	go.VR	PRT	live-PFT.MSG	be.PRS.SG
EGl	He is just going home.					

E8	वह	घर	जा	भी	रहा	है

IPA	wəh	gʰər	ɟa:	bʰi:	rəh-a:	hɛ:
Gl.	3SG	home	go.VR	PRT	live-PFT.MSG	be.PRS.SG
EGl	He is even *going* home.					

E9	वह	घर	जा	ही	तो	रहा	है
IPA	wəh	gʰər	ɟa:	hi:	to	rəh-a:	hɛ:
Gl.	3SG	home	go.VR	PRT	PRT	live-PFT.MSG	be.PRS.SG
EGl	He is just going home only.						

E10	वह	घर	चला	गया	है
IPA	wəh	gʰər	cəl-a:	gəɟa:	hɛ:
Gl.	3SG	home	walk-PFT.MSG	go.PFT.MSG	be.PRS.SG
EGl	He has gone home.				

E11	उसने	किताब	फाड़	दी	है
IPA	ʊs-ne:	kɪta:b	pʰa:ɖ	di:	hɛ:
Gl.	3SG.OBL-ERG	book	tear.VR	give.PFT	be.PRS.SG
EGl	He has torn the book.				

E12	वह	गुस्से	से	लाल	हो	गया
IPA	wəh	gʊsse:	se:	lal	ho:	gəja:
Gl.	3SG	anger.N	with.ASS	red	become.VR	go.PFT.SG
EGl	He reddened with anger.					

E13	उसने	तुम्हें	मूर्ख	बना	दिया
IPA	ʊs-ne:	tumhẽ	mu:rk^{h}	bəna:	dɪja:
Gl.	3SG.OBL.ERG	2SG.ACC	fool	make.VR	give.PFT.SG
EGl	He fooled you.				

E14	वह	खाकर	बोला
IPA	wəh	k^{h}a:-kər	bol-a:
Gl.	3SG	eat-PFT.PCPL	speak-PFT.SG
EGl	He spoke after eating.		

E15	तुम	मन	लगा	कर	काम	करो
IPA	tʊm	mən	ləga:	kər	ka:m	kər-o:
Gl.	2SG	heart	apply.VR	PFT.PCPL	work	do-IMP
EGl	Work heartedly!					

E16	वह	महीनो	से	पढ़ता	चला	आ	रहा	है
IPA	wəh	məhiːnoː	seː	pəɖʰ-taː	cəl-aː	a	rəh-aː	hɛː
Gl.	3SG	months.OBL	PP.ASS	read-IMPF.MSG	walk-PFT.MSG	come.VR	live-PFT.MSG	be.PRS.SG
Egl.	He has gone on reading for months.							

E17	मुझे	रात	में	खाना	अच्छा	लगता	है
IPA	mʊ̈jʰeː	raːt	mẽː	kʰaːnaː	ɘccʰaː	ləg-taː	hɛː
Gl.	1SG.DAT	night	PP.LOC	eat.INF	good.ADJ	feel-IMPF.MSG	be.PRS.SG
EGl	I like eating at night.						

E18	मैं	खाकर	सो	गया
IPA	mɛ̃ː	kʰaː-kər	so	gəjaː
Gl.	I	eat-PFT.PCPL	sleep.VR	go.PFT.SG
EGl.	I slept after eating.			

E19	वह	खाते-खाते	सो	गया
IPA	wəh	kʰaːteː-kʰaːteː	so	gɔjaː
Gl.	3SG	eat.IMPF.PCPL	sleep.VR	go.PFT.SG

EGl.	He slept while eating.

E20	जरा	मुझे	पानी	देना
IPA	ɟəra:	mʊɟʰe:	pa:ni	de:na:
Gl.	little-ADV	1SG.ACC	water	give.INF.IMP
EGl	Give me some water.			

E21	पानी	बहना	रुक	जाएगा
IPA	pa:ni	bəhna:	rʊk	ɟa:-e:ga:
Gl.	water	flow.INF	stop.VR	go-FUT.MSG
EGl	The water will stop flowing.			

E22	रुक	मैं	भी	साथ	आऊंगा
IPA	rʊk	mɛ̃:	bʰi:	sa:tʰ	a:-ũ:ga:
Gl.	stop	1SG	also.PRT	with.ASS	come-FUT.1MSG
EGl	Stop, I will also come along.				

E23	मेरी	इच्छा	है	कि	वह	खाना	खाए
IPA	me:ri:	ɪccʰa:	hɛ:	kɪ	wəh	kʰa:na:	kʰa:-e:

Gl.	1SG.GEN	wish.N	be.PRS.SG	that.CNJ	3SG	food	eat.OPT
EGl	I wish that he eats food.						

E24	उन्होंने	फल	खाए
IPA	unhõ:ne:	pʰəl	kʰaːeː
Gl.	he/PL.ERG	fruit.N	eat-PFT.PL
EGl	They ate fruits.		

E25	वह	आज	प्रधानमंत्री	होतीं
IPA	wəh	aːj̰	prədʰaːnməntriː	hoː-tĩː
Gl.	she	today	prime minister	be.CFCT
EGl	She would have been prime minister today.			

E26	उसे	किताब	चाहिए
IPA	ʊseː	kɪtaːb	caːhɪjeː
Gl.	3SG.OBL.ACC	book.N	want-DES
EGl	He wants book.		

E27	वह	खाकर	किताब	पढ़ा

IPA	wəh	kʰa:kər	kıta:b	pəɖʰa:
Gl.	3SG	eat.VR.PCPL	book	read-PFT
EGl.	He read book after eating.			

E28	उसको	खाते	देख	वह	बोला
IPA	ʊs-ko:	kʰa:-te:	de:kʰ	wəh	bol-a:
Gl.	3SG.OBL-ACC	eat-IMPF.PL	see.VR	3SG	speak-PFT.MSG
EGl.	Seeing him eat, he spoke.				

E29	सुरक्षा	बहाल	करना	सरकार	की	जिम्मेदारी	है
IPA	surəkʃa:	bəha:l	kərna:	sərka:r	ki:	jımme:da:ri:	hɛ:
Gl.	security	maintain	do.INF	government	GEN	responsibility	be.PRS.3SG
EGl.	Maintaining security is the government’s responsibility.						

E30	वह	डॉक्टर	है
IPA	wəh	dɔ:kʈər	hɛ:
Gl.	3SG	doctor	be.PRS.SG
EGl.	He is a doctor.		

E31	तुम	कैसे	थे
IPA	tʊm	kɛ:se:	tʰe:
Gl.	2SG	how	be.PST.PL
EGl.	How were you?		

E32	वह	किताब	पढ़ेगा
IPA	wəh	kɪta:b	pəɖʰ-e:ga:
Gl.	3SG	book	read-FUT.3MSG
EGl.	He will read book.		

E33	उसने	किताब	पढ़ी
IPA	ʊs-ne:	kɪta:b	pəɖʰ-i:
Gl.	3SG.OBL-ERG	book	read.PFT.FSG
EGl.	He read book.		

E34	किताब	पढ़
IPA	kɪta:b	pəɖʰ
Gl.	book	read.IMP
EGl.	Read book.	

E35	वह	किताब	पढ़े
IPA	wəh	kɪta:b	pəɖʰe:
Gl.	3SG	book	read.OPT
EGl.	(I wish) He should read book.		

E36	वह	किताब	पढ़ता
IPA	wəh	kɪta:b	pəɖʰ-ta:
Gl.	3SG	book	read-IMPF.MSG
EGl.	Had would read book.		

E37	जरा	किताब	पढ़ना
IPA	ɟəra:	kɪta:b	pəɖʰna:
Gl.	little.ADV	book	read.INF
EGl.	Please read the book.		

E38	सोना	स्वास्थ्य	के	लिए	जरूरी	है
IPA	sona:	swa:stʰjə	ke:	lɪe:	ɟəru:ri:	hɛ:
Gl.	sleep.INF	health	GEN	for.BEN	necessary	be.PRS.SG
EGl.	Sleeping is necessary for health.					

E39	वह	घर	गया	था
IPA	wəh	gʰər	gəja:	tʰa:
Gl	3SG	home	go.PFT.MSG	be.PST.MSG
EGl	He had gone home.			

E40	वह	किताब	पढ़	रही	है
IPA	wəh	kɪta:b	pəɖʰ	rəh-i:	hɛ:
Gl.	3SG	book.N	read.VR	live-PFT.3FSG	be.PRS.SG
EGl	She is reading a book.				

E41	वह	किताब	पढ़	रहा	था
IPA	wəh	kɪta:b	pəɖʰ	rəh-a:	tʰa:
Gl.	3SG	book.N	read.VR	live-PFT.MSG	be.PST.MSG
EGl	He was reading a book.				

E42	तुम	किताब	पढ़	रहे	होगे
IPA	tʊm	kɪta:b	pəɖʰ	rəh-e:	ho:-ge:
Gl.	2SG	book.N	read.VR	live-PFT.PL	be-FUT.2SG
EGl	You will be reading a book.				

E43	वह	किताब	पढ़ने	लगा
IPA	wəh	kɪta:b	pəɖʰne:	ləg-a:
Gl.	3SG	book.N	read.INF	begin-PFT.MSG
EGl	He started reading a book.			

E44	वह	किताब	पढ़ता	रहा
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-a:
Gl.	3SG	book.N	read-IMPF.MSG	live-PFT.3MSG
EGl	He kept on reading a book.			

E45	वह	किताब	पढ़ता	गया
IPA	wəh	kɪta:b	pəɖʰ-ta:	gəja:
Gl.	3SG	book.N	read-IMPF.MSG	go.PFT-MSG
EGl	He went on reading a book.			

E46	वह	किताब	पढ़ा	करता
IPA	wəh	kɪta:b	pəɖʰ-a:	kər-ta:
Gl.	3SG	book.N	read-PFT.MSG	do-IMPF.MSG
EGl	He used to read a book.			

E47	उसने	किताब	पढ़ी	होगी
IPA	ʊs-ne:	kɪta:b	pəɖʰ-i:	ho:-gi:
Gl.	3SG.OBL-ERG	book.N	read-PFT.FSG	be-FUT.3FSG
EGl	He would have read the book.			

E48	वह	किताब	पढ़ा	हो
IPA	wəh	kɪta:b	pəɖʰ-a:	ho:
Gl.	3SG	book.N	read-PFT.MSG	be.OPT
EGl	He should have read the book.			

E49	उसने	किताब	पढ़ी	होती
IPA	ʊs-ne:	kɪta:b	pəɖʰ-i:	ho:-ti:
Gl.	3SG.OBL-ERG	book.N	read-PFT.FSG	be-IMPF.FSG
EGl	He could have read the book.			

E50	वह	किताब	पढ़ने	लगता	था
IPA	wəh	kɪta:b	pəɖʰne:	ləg-ta:	tʰa:
Gl.	3SG	book.N	read.INF	begin-IMPF.MSG	be.PST.MSG
EGl	He used to start reading the book.				

E51	वह	किताब	पढ़ता	जाता	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	ɟa:-ta:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	go-IMPF.MSG	be.PST.SG
EGl	He keeps on reading the book.				

E52	वह	बैठा	रहा
IPA	wəh	bɛ:ʈʰ-a:	rəh-a:
Gl.	3SG	sit-PFT.MSG	live-PFT.MSG
EGl	He kept on sitting.		

E53	वह	किताब	पढ़ता	रहा	था
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-a:	tʰa:
Gl.	3SG	book.N	read-IMPF.MSG	live-PFT.MSG	be.PST.MSG
EGl	He had kept on reading the book.				

E54	वह	बैठा	रहता	है
IPA	wəh	bɛ:ʈʰ-a:	rəh-ta:	hɛ:
Gl.	3SG	sit-PFT	live-IMPF.MSG	be.PRS.SG
EGl	He keeps on sitting.			

E55 मुझे किताब चाहिए

E55	मुझे	किताब	चाहिए
IPA	mʊjʰe:	kɪta:b	ca:hɪje:
Gl.	1SG.DAT	book.N	want.DES
EGl	I want a book.		

E56 उसे किताब चाहिए थी

E56	उसे	किताब	चाहिए	थी
IPA	ʊse:	kɪta:b	ca:hɪje:	tʰi:
Gl.	3SG.OBL.DAT	book.N	want.DES	be.PST.FSG
EGl	He wanted a book.			

E57 उसे जाना चाहिए

E57	उसे	जाना	चाहिए
IPA	ʊse:	ɟa:na:	ca:hɪje:
Gl.	3SG.OBL.DAT	go.INF	want.DES
EGl	He should go.		

E58 काश तुमने उसे जाने दिया होता

E58	काश	तुमने	उसे	जाने	दिया	होता
IPA	ka:ʃ	tʊmne:	ʊse:	ɟa:ne:	dɪja:	ho:-ta:
Gl.	wish.INJ	2SG.ERG	3SG.OBL.ACC	go.INF	de.PFT.MSG	be-IMPF.MSG
EGl	(I wish) you had let him go.					

E59 उसे तुम्हें यह काम करने देना चाहिए

E59	तुम्हें	यह	काम	उसे	करने	देना	चाहिए
IPA	tʊmhẽ:	jəh	ka:m	ʊse:	kərne:	de:na:	ca:hɪje
Gl.	2SG.ACC	this	job	3SG.OBL.ACC	do.INF	give.INF	should
EGl	You should let him do this job.						

E60 उसे खाने दो

E60	उसे	खाने	दो
IPA	ʊse:	k^ha:ne:	do:
Gl.	3SG.OBL.ACC	eat.INF	give.IMP
EGl	Let him eat.		

E61 यदि उसे भूख लगी तो खाने देना

E61	यदि	उसे	भूख	लगी	तो	खाने	देना
IPA	jədɪ	ʊse:	$b^hu:k^h$	ləg-i:	to	$k^ha:ne:$	de:na:
Gl.	if	3SG.OBL.DAT	hunger.N	begin-PFT.FSG	then.CONJ	eat.INF	give.INF.IMP
EGl	If he feels hungry then (you) let him eat.						

E62 ताकि मैं उसे दिल्ली ला सकूं

E62	ताकि	मैं	उसे	दिल्ली	ला	सकूं
IPA	ta:kɪ	mɛ̃:	ʊse:	dɪlli:	la:	səkũ:
Gl.	so that.CONJ	1SG	3SG.OBL.ACC	Delhi	bring.VR	can.OPT
EGl	So that I could bring him/her to Delhi.					

E63 फिर तो वह पढ़ चुका

E63	फिर	तो	वह	पढ़	चुका
IPA	p^hir	to:	wəh	$pəɖ^h$	cʊk-a:
Gl.	then.CONJ	PRT	3SG	read.VR	finish-PFT.MSG
EGl	It’s unlikely that he would have read then.				

E64 वह पढ़ चुका

*E64	वह	पढ़	चुका
IPA	wəh	pəɖʰ	cʊk-a:
Gl.	3SG	read.VR	finish-PFT.MSG
EGl	He had read.		

E65 कहीं उसने किताब पढ़ी हो तो

E65	कहीं	उसने	किताब	पढ़ी	हो	तो
IPA	kəhĩ:	ʊs-ne:	kɪta:b	pəɖʰ-i:	ho:	to:
Gl.	if-PRT	3SG.OBL-ERG	book.N	read-PFT.FSG	be.CFCT	then
EGl	If he had the book read then?					

E66 वह अब तक किताब पढ़ चुका हो

E66	वह	अब	तक	किताब	पढ़	चुका	हो
IPA	wəh	əb	tək	kɪta:b	pəɖʰ	cʊk-a:	ho:
Gl.	3SG	now	till	book.N	read.VR	finish-PFT.MSG	be.OPT
EGl	He should have read the book by now.						

E67 काश वह पढ़ा होता

E67	काश	वह	पढ़ा	होता
IPA	ka:ʃ	wəh	pəɖʰ-a:	ho:-ta:
Gl.	wish.INJ	3SG	read-PFT.MSG	be-IMPF.MSG
EGl	I wish he had read (it).			

E68 अब तक तो तुम किताब पढ़ चुके होते

E68	अब	तक	तो	तुम	किताब	पढ़	चुके	होते
IPA	əb	tək	to:	tʊm	kɪta:b	pəɖʰ	cʊk-e:	ho:-te:
Gl.	now	till	PRT	2SG	book.N	read.VR	finish-PFT.PL	be-IMPF.PL
EGl	By now you would have read the book.							

E69 उसने किताब पढ़ डाली है

E69	उसने	किताब	पढ़	डाली	है
IPA	ʊs-ne:	kɪta:b	pəɖʰ	ɖa:l-i:	hɛ:
Gl.	3SG.OBL-ERG	book.N	read.VR	put-PFT.FSG	be.PRS.SG
EGl	I wish he had read (it).				

E70 वह किताब पढ़ता है

E70	वह	किताब	पढ़ता	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	be.PRS.SG
EGl	He reads a book.			

E71 उसने किताब पढ़ी है

E71	उसने	किताब	पढ़ी	है
IPA	ʊs-ne:	kɪta:b	pəɖʰ-i:	hɛ:
Gl.	3SG.OBL.ERG	book.N	read-PFT.FSG	be.PRS.SG
EGl	He has read a book.			

E72 उसे किताब पढनी है

E72	उसे	किताब	पढ़नी	है
IPA	ʊs-e	kɪta:b	pəɖʰni:	hɛ:
Gl.	3SG.OBL-ACC	book.N	read.INF	be.PRS.SG
EGl	He has to read a book.			

E73 वह पढ़ने वाला है

E73	वह	पढ़ने	वाला	है
IPA	wəh	pəɖʰne:	wa:la:	hɛ:
Gl.	3SG	read.INF	PRT	be.PRS.SG
EGl	He is about to read.			

E74 वह किताब पढ़ती थी

E74	वह	किताब	पढ़ती	है
IPA	wəh	kɪta:b	pəɖʰ-ti:	hɛ:
Gl.	3.FSG	book.N	read-IMPF.FSG	be.PRS.SG
EGl	She reads a book.			

E75 उसने किताब पढ़ी थी

E75	उसने	किताब	पढ़ी	थी
IPA	ʊsne:	kɪta:b	pəɖʰ-i:	tʰi:
Gl.	3SG.OBL.ERG	book.N	read-PFT.FSG	be.PST.FSG
EGl	He/She had read a book.			

E76 वह किताब पढ़ता होता

E76	वह	किताब	पढ़ता	होता
IPA	wəh	kɪta:b	pəɖʰ-ta:	ho:-ta:
Gl.	3SG	book.N	read-IMPF.MSG	be-IMPF.MSG
EGl	He would have been reading the book.			

E77 वह किताब पढ़ सकी

E77	वह	किताब	पढ़	सकी
IPA	wəh	kɪta:b	pəɖʰ	səki
Gl.	3SG	book.N	read.VR	can-PFT.FSG
EGl	She could read the book.			

E78 वह किताब पढ़ पाई

E78	वह	किताब	पढ़	पाई
IPA	wəh	kɪta:b	pəɖʰ	pa:i:
Gl.	3SG	book.N	read.VR	get-PFT.FSG
EGl	She could read the book.			

E79 उसने किताब पढने दी

E79	उसने	किताब	पढ़ने	दी
IPA	ʊsəne:	kɪta:b	pəɖʰne:	di:
Gl.	3SG.OBL.ERG	book.N	read.INF	give.PFT.FSG
EGl	He/She let (somone) read the book.			

E80 उसे किताब पढनी पड़ी

E80	उसे	किताब	पढ़नी	पड़ी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ-i:
Gl.	3SG.OBL.DAT	book.N	read.INF	lay-PFT.FSG
EGl	He/She had to read the book.			

E81 उसे किताब पढनी होती

E81	उसे	किताब	पढ़नी	होती
IPA	ʊse:	kɪta:b	pəɖʰni:	ho:-ti:
Gl.	3SG.OBL.DAT	book.N	read.INF	be-IMPF.FSG
EGl	He/She had to read the book.			

E82 उसे किताब पढनी थी

E82	उसे	किताब	पढ़नी	थी
IPA	ʊse:	kɪta:b	pəɖʰni:	tʰi:
Gl.	3SG.OBL.DAT	book.N	read.INF	be.PST.FSG
EGl	He/She had to read the book.			

E83 उससे खाते नहीं बना

E83	उससे	खाते	नहीं	बना
IPA	ʊs-se:	kʰa:-te:	nəhĩ:	bən-a:
Gl.	3SG.OBL-INS	eat-IMPF.PL	NEG	make-PFT.MSG
EGl	He could not eat.			

E84 वह पढ़ने वाला था

E84	वह	पढ़ने	वाला	था
IPA	wəh	pəɖʰne:	wa:la:	tʰa:
Gl.	3SG	read.INF	PRT	be.PST.MSG
EGl	He was about to read.			

E85 करने देता

E85	करने	देता
IPA	kərne:	de:ta:
Gl.	do.INF	give.IMPF.MSG
EGl	Would let (someone) do.	

E86 न रहने पाए

E86	न	रहने	पाए
IPA	nə	rəhne:	pa:-e:
Gl.	NEG	read.INF	get-PFT.PL
EGl	Should not remain.		

E87 करता चला

E87	करता	चला
IPA	kər-ta:	cəl-a:
Gl.	do-IMPF.MSG	walk-PFT.MSG
EGl	... went on doing.	

E88 करता रहता

E88	करता	रहता
IPA	kər-ta:	rəh-ta:
Gl.	do-IMPF.MSG	live-IMPF.MSG
EGl	... would keep on doing.	

E89 आए होते

E89	आए	होते
IPA	a:-e:	ho:-te:
Gl.	come-PFT.PL	be-IMPF.PL
EGl	... would have come.	

E90 बनाए रखी

E90	बनाए	रखी
IPA	bəna:-e	rək^{h}-i:
Gl.	mantain-PFT.PL	keep-PFT.FSG
EGl	... kept mantained.	

E91 उसने किताब पढ़ा होगा

E91	उसने	किताब	पढ़ा	होगा
IPA	ʊsne:	kɪta:b	pəɖʰ-a:	ho:-ga:
Gl.	3SG.OBL.ERG	book.N	read-PFT.MSG	be-FUT.3MSG
EGl	He would have read the book.			

E92 वह किताब पढ़ने लगेगा

E92	वह	किताब	पढ़ने	लगेगा
IPA	wəh	kɪta:b	pəɖʰne:	ləg-e:ga:
Gl.	3SG	book.N	read-INF	begin-FUT.3MSG
EGl	He would start reading book.			

E93 वह किताब पढ़ता रहेगा

E93	वह	किताब	पढ़ता	रहेगा
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-e:ga:
Gl.	3SG	book.N	read-IMPF.MSG	live-FUT.3MSG
EGl	He would keep reading the book.			

E94 वह किताब पढ़ता जाएगा

E94	वह	किताब	पढ़ता	जाएगा
IPA	wəh	kɪta:b	pəɖʰ-ta:	ja:-e:ga:
Gl.	3SG	book.N	read-IMPF.MSG	go-FUT.3MSG
EGl	He would go on reading the book.			

E95 वह किताब पढ़ा करेगा

E95	वह	किताब	पढ़ा	करेगा
IPA	wəh	kɪta:b	pəɖʰ-a:	kər-e:ga:
Gl.	3SG	book.N	read-PFT.MSG	go-FUT.3MSG
EGl	He would keep on reading the book.			

E96 वह किताब पढ़ता होगा

E96	वह	किताब	पढ़ता	होगा
IPA	wəh	kɪta:b	pəɖʰ-ta:	ho:-ga:
Gl.	3SG	book.N	read-IMPF.MSG	be-FUT.3MSG
EGl	He would be reading the book.			

E97 वह किताब पढ़ता होवे/होए

E97	वह	किताब	पढ़ता	होवे
IPA	wəh	kɪta:b	pəɖʰ-ta:	ho:-ve
Gl.	3SG	book.N	read-IMPF.MSG	be.OPT
EGl	He should be reading the book.			

E98 वह किताब पढ़ सकेगा

E98	वह	किताब	पढ़	सकेगा
IPA	wəh	kɪta:b	pəɖʰ	sək-e:ga:
Gl.	3SG	book.N	read.VR	can-FUT.3MSG
EGl	He would be able to read the book.			

E99 वह किताब पढ़ पाएगा

E99	वह	किताब	पढ़	पाएगा
IPA	wəh	kɪta:b	pəɖʰ	pa:-e:ga:
Gl.	3SG	book.N	read.VR	get-FUT.3MSG
EGl	He would be able to read the book.			

E100 वह किताब पढने देगा

E100	वह	किताब	पढ़ने	देगा
IPA	wəh	kɪta:b	pəɖʰne:	de:-ga:
Gl.	3SG	book.N	read.INF	give-FUT.3MSG
EGl	He would let (someone) read the book.			

E101 उसे किताब पढनी पड़ेगी

E101	उसे	किताब	पढ़नी	पड़ेगी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ-e:gi:
Gl.	3SG.OBL.DAT	book.N	read.INF	lay-FUT.3FSG
EGl	He will *have* to read the book.			

E102 उसे किताब पढनी होगी

E102	उसे	किताब	पढ़नी	होगी
IPA	wəh	kɪta:b	pəɖʰni:	ho:-gi:
Gl.	3SG.OBL.DAT	book.N	read.INF	be-FUT.3FSG
EGl	He might have to read the book.			

E103 वह बैठा रहेगा

E103	वह	बैठा	रहेगा
IPA	wəh	bɛ:ʈha:	rəh-e:ga:
Gl.	3SG	sit.PFT.MSG	live-FUT.MSG
EGl	He will be sitting.		

E104 वह पढ़ने वाला होगा

E104	वह	पढ़ने	वाला	होगा
IPA	wəh	pəɖʰne:	wa:la:	ho:-ga:
Gl.	3SG	read.INF	PRT	be-FUT.3MSG
EGl	He will be about to read.			

E105 वे लिखते रहें

E105	वे	लिखते	रहें
IPA	we	lɪkʰ-te:	rəh-ẽ:
Gl.	they	write-IMPF.PL	live-FUT.3PL
EGl	They should keep writing.		

E106 मैं बैठा रहूं

E106	मैं	बैठा	रहूं
IPA	mɛ̃:	bɛːʈh-aː	rəh-ũː
Gl.	1SG	sit-PFT.MSG	live.OPT
EGl	I should keep sitting.		

E107 तुम बैठे रहो

E107	तुम	बैठे	रहो
IPA	tʊm	bɛːʈh-eː	rəh-oː
Gl.	2SG	sit-PFT.PL	live-IMP
EGl	(You) keep sitting.		

E108 ... अपने को भी जान सकें

E108	अपने	को	भी	जान	सकें
IPA	əpneː	koː	bʰiː	ɟaːn	səkɛ̃ː
Gl.	REFL	PP.ACC	also.PRT	know.VR	can-OPT
EGl	... so that you should know yourself too.				

E109 तू किताब पढ़ा कर

E109	तू	किताब	पढा	कर
IPA	tu:	kɪta:b	pəɖʰ-a:	kər
Gl.	2SG	book.N	read-PFT.MSG	do-IMP
EGl	You keep reading book.			

E110 करने दो

E110	करने	दो
IPA	kərne:	do:
Gl.	do.INF	give-IMP
EGl	Let (him/her) do.	

E111 करने दूं

E111	करने	दूं
IPA	kərne:	dũ:
Gl.	do.INF	give.OPT.1SG
EGl	Should I let (him/her) do.	

E112 उसे आना चाहिए

E112	उसे	आना	चाहिए
IPA	ʊse:	a:na:	ca:hɪje:
Gl.	3SG.OBL.ACC	come.INF	should
EGl	S/He should come.		

E113 करने दे

E113	करने	दे
IPA	kərne:	de:
Gl.	do.INF	give-IMP
EGl	Let (him/her) do.	

E114 बटोरते चलें

E114	बटोरते	चलें
IPA	bəʈo:r-te:	cəl-ẽ:
Gl.	pick-IMPF.PL	walk.OPT
EGl	Keep on picking.	

E115 करता चलूं

E115	करता	चलूं
IPA	kər-ta:	cəl-ũ:
Gl.	do-IMPF.MSG	walk.OPT.1SG
EGl	I should keep on doing.	

E116 अगर वह पढ़ती हों तो...

E116	अगर	वह	पढ़ती	हों	तो
IPA	əgər	wəh	pəɖʱ-ti:	hõ:	to:
Gl.	if	she	read-IMPF.3FSG	be.OPT	then-PRT
EGl	If she is reading then ...				

E117 आप चलते रहें

E117	आप	चलते	रहें
IPA	a:p	cəl-te:	rəh-ẽ:
Gl.	2SG.HON	walk-IMPF	live.OPT
EGl	You keep walking.		

E118 आप कुछ भी रहे हों

E118	आप	कुछ	भी	रहे	हों
IPA	a:p	kʊcʰ	bʰi:	rəhe:	hõ:
Gl.	2SG.HON	something	also. PRT	live-PFT	be. CFCT
EGl	Whatever you were!				

E119 जरा मेरे घर भी चले जाना

E119	जरा	मेरे	घर	भी	चले	जाना
IPA	ɟəra:	me:re:	gʰər	bʰi:	cəle:	ɟa:na:
Gl.	little	1SG.GEN	house.N	also.PRT	go.PFT	go.INF
EGl	Please be at my house also for a while.					

E120 पढ़ने वाले लड़के

E120	पढ़ने	वाले	लड़के
IPA	pəɖʰne:	wa:le:	ləɖke:
Gl.	read-NF	PRT	boy.OBL.PL
EGl	The sutudius boys.		

E121 मुरझा कर सूख गया

E121	मुरझा	कर	सूख	गया
IPA	mʊrjʰa:	kər	su:kʰ	gəja:
Gl.	fade.VR	do-PCPL	dry.VR	go.PFT.3MSG
EGl	Having faded it dried away.			

E122 वह खाते हुए बोला

E122	वह	खाते	हुए	बोला
IPA	wəh	kʰa:te:	hʊe	bola:
Gl.	3SG	eat-IMPF	be.PFT	speak-PFT.3MSG
EGl	He spoke while eating.			

E123

E123	चोट	खाया	हुआ	सांप	गिर	पड़ा
IPA	co:ʈ	kʰa:ɟa:	hʊa:	sã:p	gɪr	pəɖa:
Gl.	blow.N	eat-PFT.3MSG	be.PFT.3MSG	snake.N	fall.VR	lay-PFT.3MSG
EGl	The hurt snake fell down.					

E124 भगवान तुम्हें पापों से बचे रहने की शक्ति दें

E124	भगवान	तुम्हें	पापों	से	बचे	रहने	की	शक्ति	दें
IPA	bʰəgwa:n	tʊmhẽ:	pa:põ:	se:	bəce:	rəhne:	ki:	ʃəktɪ	dẽ
Gl.	god.N	2SG.ACC	sins-OBL	ABL	save-PFT	live.INF	GEN	strength-N	give.OPT
EGl	May God give you the strength to stay away from sins!								

E125 भूख हड़ताल कर रही सामाजिक कार्यकर्ता

E125	भूख	हड़ताल	कर	रही	सामाजिक	कार्यकर्ता
IPA	bʰu:kʰ	həɖta:l	kər	rəhi:	sa:ma:jɪk	ka:rjkərta:
Gl.	hunger.N	strike.N	do.VR	live-PFT.3FSG	social-ADJ	activist.N
EGl	The social activist on hunger strike ...					

E126 सिमटती जाती दुनिया...

E126	सिमटती	जाती	दुनिया
IPA	sɪməʈti:	ɟa:-ti:	dʊnɪja
Gl.	shrink-IMPF.3FSG	go-IMPF.3FSG	world.N
EGl	The shrinking world ...		

E127 वहां का नजारा देखते ही बनता है

E127	वहां	का	नजारा	देखते	ही	बनता	है
IPA	wəhã:	ka:	nəɟa:ra:	de:kʰ-te:	hi:	bənta:	hɛ:
Gl.	there-DEM	GEN	view.N	see-IMPF	only.PRT	become-IMPF	be. PRS
EGl	The view from there is worth watching.						

E128 वह किताब पढ़ चुका है

E128	वह	किताब	पढ़	चुका	है
IPA	wəh	kɪta:b	pəɖʱ	cʊk-a:	hɛ:
Gl.	3SG	book.N	read.VR	finish-PFT.3MSG	be. PRS
EGl	He has read the book.				

E129 वह किताब पढ़ रहा है

E129	वह	किताब	पढ़	रहा	है
IPA	wəh	kɪta:b	pəɖʱ	rəh-a:	hɛ:
Gl.	3SG	book.N	read.VR	live-PFT.3MSG	be. PRS
EGl	He is reading the book.				

E130 वह किताब पढ़ने लगता है

E130	वह	किताब	पढ़ने	लगता	है
IPA	wəh	kɪta:b	pəɖʰne:	ləg-ta:	hɛ:
Gl.	3SG	book.N	read.INF	begin-IMPF.3MSG	be. PRS
EGl	He starts reading the book.				

E131 वह किताब पढ़ने लगा है

E131	वह	किताब	पढ़ने	लगा	है
IPA	wəh	kɪta:b	pəɖʰne	ləg-a:	hɛ:
Gl.	3SG	book.N	read.INF	begin-PFT.3MSG	be. PRS
EGl	He has started reading the book.				

E132 वह किताब पढ़ता रहता है

E132	वह	किताब	पढ़ता	रहता	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	ləg-ta:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	apply-IMPF.3MSG	be. PRS
EGl	He keeps on reading the book.				

E133 वह किताब पढ़ता रहा है

E133	वह	किताब	पढ़ता	रहा	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-a:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	live-PFT.3MSG	be. PRS
EGl	He has kept on reading the book.				

E134 वह बैठा रहता है

E134	वह	बैठा	रहता	है
IPA	wəh	bɛ:ʈh-a:	rəh-ta:	hɛ:
Gl.	3SG	sit-PFT.MSG	live-IMPF.MSG	be. PRS
EGl	He keeps on sitting.			

E135 वह किताब पढ़ता गया है

E135	वह	किताब	पढ़ता	गया	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	gəja:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	go.PFT.MSG	be. PRS
EGl	He has kept on reading the book.				

E136 वह किताब पढ़ा करता है

E136	वह	किताब	पढ़ा	करता	है
IPA	wəh	kɪta:b	pəɖʰ-a:	kərta:	hɛ:
Gl.	3SG	book.N	read-PFT.MSG	do-IMPF.MSG	be. PRS
EGl	He has keeps on reading the book.				

E137 वह किताब पढ़ सकती है

E137	वह	किताब	पढ़	सकती	है
IPA	wəh	kɪta:b	pəɖʰ	səkti:	hɛ:
Gl.	3SG	book.N	read.VR	can-IMPF.FSG	be. PRS
EGl	She can read the book.				

E138 वह किताब पढ़ सकी है

E138	वह	किताब	पढ़	सकी	है
IPA	wəh	kɪta:b	pəɖʰ	səki:	hɛ:
Gl.	3SG	book.N	read.VR	can-PFT.FSG	be. PRS
EGl	She has been able to read the book.				

E139 वह किताब पढ़ पाती है

E139	वह	किताब	पढ़	पाती	है
IPA	wəh	kɪta:b	pəɖʰ	pa:ti:	hɛ:
Gl.	3SG	book.N	read.VR	get-IMPF.FSG	be. PRS
EGl	She is able to read the book.				

E140 वह किताब पढ़ पाई है

E140	वह	किताब	पढ़	पाई	है
IPA	wəh	kɪta:b	pəɖʰ	pa:i:	hɛ:
Gl.	3SG	book.N	read.VR	get-PFT.FSG	be. PRS
EGl	She has been able to read the book.				

E141 वह किताब पढने देती है

E141	वह	किताब	पढने	देती	है
IPA	wəh	kɪta:b	pəɖʰne:	de:-ti:	hɛ:
Gl.	3SG	book.N	read.INF	give-IMPF.FSG	be. PRS
EGl	She lets (someone) read the book.				

E142 उसने किताब पढने दी है

E142	उसने	किताब	पढने	दी	है
IPA	ʊsne:	kɪta:b	pəɖʰne:	di:	hɛ:
Gl.	3SG.OBL.ERG	book.N	read.INF	give-PFT.FSG	be. PRS
EGl	She has let (someone) read the book.				

E143 उसे किताब पढनी पड़ती है

E143	उसे	किताब	पढनी	पड़ती	है
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖti:	hɛ:
Gl.	3SG.OBL.ACC	book.N	read.INF	lay-IMPF.FSG	be. PRS
EGl	She has to read the book.				

E144 उसे किताब पढनी पड़ी है

E144	उसे	किताब	पढनी	पड़ी	है
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖi:	hɛ:
Gl.	3SG.OBL.ACC	book.N	read.INF	lay-PFT.FSG	be. PRS
EGl	She has had to read the book.				

E145 उसे किताब पढनी होती है

E145	उसे	किताब	पढनी	होती	है
IPA	ʊse:	kɪta:b	pəɖʰni:	ho:-ti:	hɛ:
Gl.	3SG.OBL.ACC	book.N	read.INF	be-IMPF.FSG	be. PRS
EGl	She has got to read the book.				

E146 वे यह काम सालों से करते आए हैं

E146	वे	यह	काम	सालों	से	करते	आए	हैं
IPA	we	jəh	ka:m	salõ:	se:	kərte:	a:e:	hɛ̃:
Gl.	3PL	this-DEM	work.N	years.N	since-PP	do-IMPF.PL	come-PFT.PL	be.PRS.PL
EGl	They have been doing this work for years.							

E147 अब तक उसे आता होना चाहिए

E147	अब	तक	उसे	आता	होना	चाहिए
IPA	əb	tək	ʊse:	a:-ta:	ho:na:	ca:hɪje
Gl.	now	till-PP	3SG.OBL.ACC	come-IMPF.MSG	be.INF	should
EGl	By now he should be coming.					

E148 टिकट लेनेवालों का तांता लगा होता है

E148	टिकट	लेनेवालों	का	तांता	लगा	होता	है
IPA	ʈɪkəʈ	le:ne:-walõ:	ka:	tã:ta:	ləg-a:	ho:-ta:	hɛ:
Gl.	ticket	take.INF-PRT.PL	GEN.MSG	queue.N	begin-PFT.MSG	be-IMPF.MSG	be.PRS
EGl	A queue of the ticket takers keeps formed.						

E149 हर जगह फैले हुए हैं

E149	हर	जगह	फैले	हुए	हैं
IPA	hər	ɟəgəh	pʰɛ:le:	hʊe:	hɛ̃:
Gl.	each	space.N	spread-PFT.PL	be.PFT.PL	be.PRS.PL
EGl	(They are) spread everywhere.				

E150 यह नाव वर्षों से डूबी पड़ी है

E150	यह	नाव	वर्षों	से	डूबी	पड़ी	है
IPA	jəh	na:w	bərʂõ:	se:	ɖu:bi:	pəɖi:	hɛ:
Gl.	this	boat.N	years.N.PL.OBL	since-PP	sink-PFT.FSG	lay-PFT.FSG	be.PRS.SG
EGl	This boat is lying sunk for years.						

E151 बच्चे शांति बनाए रखते हैं

E151	बच्चे	शांति	बनाए	रखते	हैं
IPA	bəcce:	ʃa:ntɪ	bəna:e:	rəkʰte:	hẽ:
Gl.	children	peace.N	maintain-PFT.PL	keep-IMPF.PL	be.PRS.PL
EGl	Children keep the peace mantained.				

E152 वहां का नजारा देखते ही बनता था

E151	वहां	का	नजारा	देखते	ही	बनता	था
IPA	wəhã:	ka:	nəja:ra:	de:kʰ-te:	hi:	bən-ta:	tʰa:
Gl.	there	GEN	view.N	see-IMPF.PL	only.PRT	form-IMPF.MSG	be.PST.SG
EGl	The view there was worth watching.						

E153 वह किताब पढ़ चुका था

E153	वह	किताब	पढ़	चुका	था
IPA	wəh	kɪta:b	pəɖʰ	cʊk-a:	tʰa:
Gl.	3SG	book.N	read.VR	finish-PFT.MSG	be.PST.SG
EGl	He had read the book.				

E154 वह किताब पढ़ने लगा था

E154	वह	किताब	पढ़ने	लगा	था
IPA	wəh	kɪta:b	pəɖʰne:	ləg-a:	tʰa:
Gl.	3SG	book.N	read.INF	begin-PFT.MSG	be.PST.SG
EGl	He had started reading the book.				

E155 वह किताब पढ़ता रहता था

E155	वह	किताब	पढ़ता	रहता	था
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-ta:	tʰa:
Gl.	3SG	book.N	read-IMPF.MSG	live-IMPF.MSG	be.PST.SG
EGl	He used to keep reading the book.				

E156 वह बैठा रहता था

E156	वह	बैठा	रहता	था
IPA	wəh	bɛ:ʈʰa	rəh-ta:	tʰa:
Gl.	3SG	sit-PFT.MSG	live-IMPF.MSG	be.PST.SG
EGl	He used to keep sitting.			

E157 वह किताब पढ़ता जाता था

E157	वह	किताब	पढ़ता	जाता	था
IPA	wəh	kɪta:b	pəɖʰ-ta:	ja:-ta:	tʰa:
Gl.	3SG	book.N	read-PFT.MSG	go-IMPF.MSG	be.PST.SG
EGl	He used to go on reading the book.				

E158 वह किताब पढ़ता गया था

E158	वह	किताब	पढ़ता	गया	था
IPA	wəh	kɪta:b	pəɖʰ-ta:	gəja:	tʰa:
Gl.	3SG	book.N	read-IMPF.MSG	go.PFT.MSG	be.PST.SG
EGl	He had gone on reading the book.				

E159 वह किताब पढ़ा करता था

E159	वह	किताब	पढ़ा	करता	था
IPA	wəh	kɪta:b	pəɖʰ-a:	kər-ta:	tʰa:
Gl.	3SG	book.N	read-PFT.MSG	do-IMPF.MSG	be.PST.MSG
EGl	He used to read the book.				

E160 वह किताब पढ़ सकती थी

E160	वह	किताब	पढ़	सकती	थी
IPA	wəh	kɪta:b	pəɖʰ	sək-ti:	tʰi:
Gl.	she	book.N	read.VR	can-IMPF.FSG	be.PST.FSG
EGl	She could read the book.				

E161 वह किताब पढ़ सकी थी

E161	वह	किताब	पढ़	सकी	थी
IPA	wəh	kɪta:b	pəɖʰ	sək-i:	tʰi:
Gl.	she	book.N	read.VR	can-PFT.FSG	be.PST.FSG
EGl	She was able to read the book.				

E162 वह किताब पढ़ पाती थी

E162	वह	किताब	पढ़	पाती	थी
IPA	wəh	kɪta:b	pəɖʰ	pa:-ti:	tʰi:
Gl.	she	book.N	read.VR	get-IMPF.FSG	be.PST.FSG
EGl	She could read the book.				

E163 वह किताब पढ़ पाई थी

E163	वह	किताब	पढ़	पाई	थी
IPA	wəh	kɪta:b	pəɖʰ	pa:-i:	tʰi:
Gl.	she	book.N	read.VR	get-PFT.FSG	be.PST.FSG
EGl	She was able to read the book.				

E164 वह किताब पढने देती थी

E164	वह	किताब	पढने	देती	थी
IPA	wəh	kɪta:b	pəɖʰne:	de:-ti:	tʰi:
Gl.	she	book.N	read.INF	give-IMPF.FSG	be.PST.FSG
EGl	She would let (someone) read the book.				

E165 उसने किताब पढने दी थी

E165	उसने	किताब	पढने	दी	थी
IPA	ʊsne:	kɪta:b	pəɖʰne:	di:	tʰi:
Gl.	he-OBL.ERG	book.N	read.INF	give-PFT.FSG	be.PST.FSG
EGl	She had let (someone) read the book.				

E166 उसे किताब पढनी पड़ती थी

E166	उसे	किताब	पढनी	पड़ती	थी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ-ti:	tʰi:
Gl.	he-OBL.DAT	book.N	read.INF	lay-IMPF.FSG	be.PST.FSG
EGl	She *had* to read the book.				

E167 उसे किताब पढनी पड़ी थी

E167	उसे	किताब	पढनी	पड़ी	थी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ-i:	tʰi:
Gl.	he-OBL.DAT	book.N	read.INF	lay-PFT.FSG	be.PST.FSG
EGl	She had *had* to read the book.				

E168 उसे किताब पढनी होती थी

E168	उसे	किताब	पढनी	होती	थी
IPA	ʊse:	kɪta:b	pəɖʰni:	ho:-ti:	tʰi:
Gl.	he-OBL.DAT	book.N	read.INF	be-IMPF.FSG	be.PST.FSG
EGl	She had to read the book.				

E169 तुमने उसे यह करने दिया होता

E169	तुमने	उसे	यह	करने	दिया	होता
IPA	tʊmne:	ʊse:	jəh	kərne:	dɪja:	ho:-ta:
Gl.	2SG.ERG	he-OBL	this	do.INF	give.PFT.MSG	be-IMPF.MSG
EGl	You should have let him do it.					

E170 उसे आना चाहिए था

E170	उसे	आना	चाहिए	था
IPA	ʊse:	a:na:	ca:hɪje:	tʰa:
Gl.	he-OBL.DAT	come.INF	should	be.PST.MSG
EGl	He should have come.			

E171 वह वहीं खड़ा रह गया

E171	वह	वहीं	खड़ा	रह	गया
IPA	wəh	wəhĩ:	kʰəɖa:	rəh	gəja:
Gl.	3SG	there-EMPH	stand.VR	live.VR	go.PFT.MSG
EGl	He went on standing there itself.				

E172 जुड़ती चली गयीं

E172	जुड़ती	चली	गयीं
IPA	jʊɖ-ti:	cəl-i:	gəji:
Gl.	join-IMPF.FSG	walk-PFT.FSG	go.PFT.FSG
EGl	... went on joining.		

E173 देखती रह गई

E173	देखती	रह	गई
IPA	de:kʰ-ti:	rəh	gəi:
Gl.	see-IMPF.FSG	live.VR	go.PFT.FSG
EGl	... went on watching.		

E174 लगे हुए थे

E174	लगे	हुए	थे
IPA	ləg-e:	hʊe:	tʰe:
Gl.	apply-PFT.PL	be-PFT.PL	be.PST.PL
EGl	... were indulged.		

E175 ढंके पड़े थे

E175	ढंके	पड़े	थे
IPA	ɖʰə̃k-e:	pəɖ-e:	tʰe:
Gl.	cover-PFT.PL	lay-PFT.PL	be.PST.PL
EGl	... were covered.		

E176 वहां का नजारा देखते ही बनेगा

E176	वहां	का	नजारा	देखते	ही	बनेगा
IPA	wəhã:	ka:	nəɟa:ra:	de:kʰ-te:	hi:	bən-e:ga:
Gl.	there	GEN	view	see-IMPF.PL	only.PRT	form-FUT.3MSG
EGl	The view there would be worth watching.					

E177 वह किताब पढ़ चुका होगा

E177	वह	किताब	पढ़	चुका	होगा
IPA	wəh	kɪta:b	pəɖʰ	cʊk-a:	ho:-ga:
Gl.	3SG	book.N	read.VR	finish-PFT.MSG	be-FUT.3MSG
EGl	He would have read the book.				

E178 वह किताब पढ़ रहा होगा

E178	वह	किताब	पढ़	रहा	होगा
IPA	wəh	kɪta:b	pəɖʰ	rəh-a:	ho:-ga:
Gl.	3SG	book.N	read.VR	live-PFT.MSG	be-FUT.3MSG
EGl	He would be reading the book.				

E179 वह किताब पढ़ा करता होगा

E179	वह	किताब	पढ़ा	करता	होगा
IPA	wəh	kɪta:b	pəɖʰ-a:	kər-ta:	ho:-ga:
Gl.	3SG	book.N	read-PFT.MSG	do-IMPF.MSG	be-FUT.3MSG
EGl	He would keep on reading the book.				

E180 वह किताब पढ़ सका होगा

E180	वह	किताब	पढ़	सका	होगा
IPA	wəh	kɪta:b	pəɖʰ	sək-a:	ho:-ga:
Gl.	3SG	book.N	read.VR	can-PFT.MSG	be-FUT.3MSG
EGl	He could have read the book.				

E181 वह किताब पढ़ पाया होगा

E181	वह	किताब	पढ़	पाया	होगा
IPA	wəh	kɪta:b	pəɖʰ	pa:-ɟa:	ho:-ga:
Gl.	3SG	book.N	read.VR	get-PFT.MSG	be-FUT.3MSG
EGl	He could have been able to read the book.				

E182 उसने किताब पढने दिया होगा

E182	उसने	किताब	पढने	दिया	होगा
IPA	ʊs-ne:	kɪta:b	pəɖʰne:	dɪja:	ho:-ga:
Gl.	3SG.OBL-ERG	book.N	read.INF	de.PFT.MSG	be-FUT.3MSG
EGl	He would have let (someone) read the book.				

E183 उसे किताब पढनी पड़ी होगी

E183	उसे	किताब	पढनी	पड़ी	होगी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ-i:	ho:-gi:
Gl.	3SG.OBL.DAT	book.N	read.INF	lay-PFT.FSG	be-FUT.3FSG
EGl	He would have had to read the book.				

E184 उसे किताब पढनी होती होगी

E184	उसे	किताब	पढनी	होती	होगी
IPA	ʊse:	kɪta:b	pəɖʰni:	ho:-ti:	ho:-gi:
Gl.	he-OBL.DAT	book.N	read.INF	be-IMPF.FSG	be-FUT.3FSG
EGl	He might have had to read the book.				

E185 उसे किताब पढनी रही होगी

E185	उसे	किताब	पढनी	रही	होगी
IPA	ʊse:	kɪta:b	pəɖʰni:	rəh-i:	ho:-gi:
Gl.	he-OBL.DAT	book.N	read.INF	live-PFT.FSG	be-FUT.3FSG
EGl	He might have had to read the book.				

E186 करने देता होगा

E186	करने	देता	होगा
IPA	kərne:	de:-ta:	ho:-ga:
Gl.	do.INF	give-IMPF.MSG	be-FUT.3MSG
EGl	… would have let (someone) do.		

E187 वह किताब पढ़ने लगा होगा

E187	वह	किताब	पढ़ने	लगा	होगा
IPA	wəh	kɪta:b	pəɖʰne:	ləg-a:	ho:-ga:
Gl.	3SG	book.N	read.INF	begin-PFT.MSG	be-FUT.3MSG
EGl	He might have started reading the book.				

E188 वह किताब पढ़ने लगता होगा

E188	वह	किताब	पढ़ने	लगता	होगा
IPA	wəh	kɪta:b	pəɖʰne:	ləg-ta:	ho:-ga:
Gl.	3SG	book.N	read.INF	begin-IMPF.MSG	be-FUT.3MSG
EGl	He would have started reading the book.				

E189 वह किताब पढ़ता रहा होगा

E189	वह	किताब	पढ़ता	रहा	होगा
IPA	wəh	kɪta:b	pəɖʰ-ta:	rəh-a:	ho:-ga:
Gl.	3SG	book.N	read.IMPF.MSG	live-PFT.MSG	be-FUT.3MSG
EGl	He might have been reading the book.				

E190 अगर वह काम कर रहा हो तो

E190	अगर	वह	काम	कर	रहा	हो	तो
IPA	əgər	wəh	ka:m	kər	rəh-a:	ho:	to
Gl.	if	3SG	work.N	do.VR	live-PFT.MSG	be.CFCT	then-PRT
EGl	If he is working then …						

E191 करने देना चाहिए

E191	करने	देना	चाहिए
IPA	kərne:	de:na:	ca:hɪje:
Gl.	do.INF	give.INF	should
EGl	… should let (someone) do (it).		

E192 सिमटती जा रही दुनिया...

E192	सिमटती	जा	रही	दुनिया
IPA	sɪmət̪-ti:	ja:	rəhi:	dʊnɪja:
Gl.	shrink-IMPF.FSG	go.VR	live-PFT.FSG	world.N
EGl	… shrinking world …			

E193 बच्चे ले जा रहे लोग...

E193	बच्चे	ले	जा	रहे	लोग
IPA	bəcce:	le:	ɟa:	rəh-e:	lo:g
Gl.	children	take.VR	go.VR	live-PFT.FSG	people.N
EGl	… the people taking children …				

E194 वह किताब पढ़ता जा रहा है

E194	वह	किताब	पढ़ता	जा	रहा	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	ɟa:	rəh-a:	hɛ:
Gl.	3SG	book	read-IMPF.MSG	go.VR	live-PFT.MSG	be.PRS.SG
EGl	He keeps on reading the book.					

E195 वह किताब पढ़ सक रहा है

E195	वह	किताब	पढ़	सक	रहा	है
IPA	wəh	kɪta:b	pəɖʰ	sək	rəh-a:	hɛ:
Gl.	3SG	book	read.VR	can.VR	live-PFT.MSG	be.PRS.SG
EGl	He is able keep on reading the book.					

E196 वह किताब पढ़ पा रहा है

E196	वह	किताब	पढ़	पा	रहा	है
IPA	wəh	kɪta:b	pəɖʰ	pa:	rəh-a:	hɛ:
Gl.	3SG	book.N	read.VR	get.VR	live-PFT.MSG	be.PRS.SG
EGl	He is able keep on reading the book.					

E197 वह किताब पढने दे रहा है

E197	वह	किताब	पढने	दे	रहा	है
IPA	wəh	kɪta:b	pəɖʰne:	de:	rəh-a:	hɛ:
Gl.	3SG	book	read.INF	give.VR	live-PFT.MSG	be.PRS.SG
EGl	He letting (someone) read the book.					

E198 उसे किताब पढनी पड़ रही है

E198	उसे	किताब	पढनी	पड़	रही	है
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ	rəh-i:	hɛ:
Gl.	3SG.OBL.DAT	book	read.INF	lay.VR	live-PFT.FSG	be.PRS.SG
EGl	He has to keep on reading the book.					

E199 देनी पड़ सकती है

E199	देनी	पड़	सकती	है
IPA	de:ni:	pəɖ	sək-ti:	hɛ:
Gl.	de.INF	lay.VR	can-IMPF.FSG	be.PRS.SG
EGl	… might have to give away.			

E200 सुनते आ रहे हैं

E200	सुनते	आ	रहे	हैं
IPA	sʊn-te:	a:	rəh-e:	hɛ̃:
Gl.	listen-IMPF.PL	come.VR	live-PFT.PL	be.PRS.PL
EGl	… have gone on listening.			

E201 चली आ रही है

E201	चली	आ	रही	है
IPA	cəl-i:	a:	rəh-i:	hɛ:
Gl.	listen-PFT.FSG	come.VR	live-PFT.FSG	be.PRS.SG
EGl	… have been going on.			

E202 बना रह सकता है

E202	बना	रह	सकता	है
IPA	bən-a:	rəh	sək-ta:	hɛ:
Gl.	form-PFT.MSG	live.VR	can-IMPF.MSG	be.PRS.SG
EGl	… may keep on formed.			

E203 वह किताब पढ़ता जा रहा था

E203	वह	किताब	पढ़ता	जा	रहा	था
IPA	wəh	kɪta:b	pəɖʱ-ta:	ɟa:	rəh-a:	tʰa:
Gl.	3SG	book.N	form-IMPF.MSG	go.VR	can-PFT.MSG	be.PST.MSG
EGl	He had kept on reading the book.					

E204 वह किताब पढ़ सक रहा था

E204	वह	किताब	पढ़	सक	रहा	था
IPA	wəh	kɪta:b	pəɖʱ	sək	rəh-a:	tʰa:
Gl.	3SG	book.N	form.VR	can.VR	can-PFT.MSG	be.PST.MSG
EGl	He has been able to read the book.					

E205 वह किताब पढ़ पा रहा था

E205	वह	किताब	पढ़	पा	रहा	था
IPA	wəh	kɪta:b	pəɖʰ	pa:	rəh-a:	tʰa:
Gl.	3SG	book.N	form.VR	get.VR	live-PFT.MSG	be.PST.MSG
EGl	He has been able to read the book.					

E206 वह किताब पढने दे रहा था

E206	वह	किताब	पढने	दे	रहा	था
IPA	wəh	kɪta:b	pəɖʰne:	de:	rəh-a:	tʰa:
Gl.	3SG	book.N	read.INF	give.VR	live-PFT.MSG	be.PST.MSG
EGl	He had been letting (someone) read the book.					

E207 उसे किताब पढनी पड़ रही थी

E207	उसे	किताब	पढनी	पड़	रही	थी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ	rəh-i:	tʰi:
Gl.	he-OBL	book.N	read.INF	lay.VR	live-PFT.FSG	be.PST.FSG
EGl	He had been having to read the book.					

E208 देनी पड़ सकती थी

E208	देनी	पड़	सकती	थी
IPA	de:ni:	pəɖ	sək-ti:	t^{h}i:
Gl.	give.INF	lay.VR	can-IMPF.FSG	be.PST.FSG
EGl	…might have had to give away.			

E209 सुनते आ रहे थे

E209	सुनते	आ	रहे	थे
IPA	sʊn-te:	a:	rəh-e:	t^{h}e:
Gl.	listen-IMPF-PL	come.VR	live-PFT.PL	be.PST.PL
EGl	…had ben listening (for a long time).			

E210 चली आ रही थी

E210	चली	आ	रही	थी
IPA	cəl-i:	a:	rəh-i:	t^{h}i:
Gl.	listen-PFT-FSG	come.VR	live-PFT.FSG	be.PST.FSG
EGl	…had ben going on (for a long time).			

E211 बना रह सकता था

E211	बना	रह	सकता	था
IPA	bən-a:	rəh	sək-ta:	t^ha:
Gl.	form-PFT-MSG	live.VR	can-IMPF.MSG	be.PST.MSG
EGl	…might have gone on staying formed.			

E212 वह किताब पढ़ता जा रहा होगा

E212	वह	किताब	पढ़ता	जा	रहा	होगा
IPA	wəh	kɪta:b	pəɖ^h-ta:	ɟa:	rəh-a:	ho:-ga:
Gl.	3SG	book.N	read-IMPF.MSG	go.VR	live-PFT.MSG	be-FUT.3MSG
EGl	He might have gone on reading the book.					

E213 वह किताब पढ़ सक रहा होगा

E213	वह	किताब	पढ़	सक	रहा	होगा
IPA	wəh	kɪta:b	pəɖ^h	sək	rəh-a:	ho:-ga:
Gl.	3SG	book.N	read.VR	go.VR	live-PFT.MSG	be-FUT.3MSG
EGl	He might have been able in reading the book.					

E214 वह किताब पढ़ पा रहा होगा

E214	वह	किताब	पढ़	पा	रहा	होगा
IPA	wəh	kɪta:b	pəɖʰ	pa:	rəh-a:	ho:-ga:
Gl.	3SG	book.N	read.VR	get.VR	live-PFT.MSG	be-FUT.3MSG
EGl	He might have been able in reading the book.					

E215 वह किताब पढने दे रहा होगा

E215	वह	किताब	पढने	दे	रहा	होगा
IPA	wəh	kɪta:b	pəɖʰne:	de:	rəh-a:	ho:-ga:
Gl.	3SG	book.N	read.INF	give.VR	live-PFT.MSG	be-FUT.3MSG
EGl	He might have been letting (someone) read the book.					

E216 उसे किताब पढनी पड़ रही होगी

E216	उसे	किताब	पढनी	पड़	रही	होगी
IPA	ʊse:	kɪta:b	pəɖʰni:	pəɖ	rəh-i:	ho:-gi:
Gl.	he-OBL	book.N	read.INF	lay.VR	live-PFT.FSG	be-FUT.3FSG
EGl	He would have had to read to the book.					

E217 सुनते आ रहे होंगे

E217	सुनते	आ	रहे	होंगे
IPA	sʊn-te:	a:	rəh-e:	hõ:-ge
Gl.	listen-IMPF.PL	come.VR	read-PFT.PL	be-FUT.3PL
EGl	…might have gone on listening (for a long time).			

E218 चली आ रही होगी

E218	चली	आ	रही	होगी
IPA	cəl-i:	a:	rəh-i:	ho:-gi:
Gl.	walk-PFT.FSG	come.VR	read-PFT.FSG	be-FUT.FSG
EGl	…might have been going on (for a long time).			

E219 वह किताब पढ़ता चला जा रहा है

E219	वह	किताब	पढ़ता	चला	जा	रहा	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	cəl-a:	ja:	rəha:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	walk-PFT.MSG	go.VR	live-PFT.MSG	be.PRS.SG
EGl	He has gone on reading a book (for a long time).						

E220 वह किताब पढ़ता चला जा रहा था

E220	वह	किताब	पढ़ता	चला	जा	रहा	था
IPA	wəh	kɪta:b	pəɖʰ-ta:	cəl-a:	ja:	rəha:	tʰa:
Gl.	3SG	book.N	read-IMPF.MSG	walk-PFT.MSG	go.VR	live-PFT.MSG	be.PST.MSG
EGl	He had gone on reading a book (for a long time).						

E221 वह किताब पढ़ता चला आ रहा है

E221	वह	किताब	पढ़ता	चला	आ	रहा	है
IPA	wəh	kɪta:b	pəɖʰ-ta:	cəl-a:	a	rəha:	hɛ:
Gl.	3SG	book.N	read-IMPF.MSG	walk-PFT.MSG	come.VR	live-PFT.MSG	be.PRS.MSG
EGl	He has gone on reading a book (for a long time).						

E222 वह किताब पढ़ता चला आ रहा था

E222	वह	किताब	पढ़ता	चला	आ	रहा	था
IPA	wəh	kɪta:b	pəɖʰ-ta:	cəl-a:	a	rəha:	tʰa:
Gl.	3SG	book.N	read-IMPF.MSG	walk-PFT.MSG	come.VR	live-PFT.MSG	be.PST.MSG
EGl	He had gone on reading a book (for a long time).						

E223 वह किताब पढ़ता चला जा रहा होगा

E223	वह	किताब	पढ़ता	चला	जा	रहा	होगा
IPA	wəh	kɪta:b	pəɖʰ-ta:	cəl-a:	ja:	rəha:	ho:-ga:
Gl.	3SG	book.N	read-IMPF.MSG	walk-PFT.MSG	go.VR	live-PFT.MSG	be-FUT.MSG
EGl	He would have gone on reading a book (for a long time).						

E224 वह किताब पढ़ता चला आ रहा होगा

E224	वह	किताब	पढ़ता	चला	आ	रहा	होगा
IPA	wəh	kɪta:b	pəɖʰ-ta:	cəl-a:	a	rəha:	ho:-ga:
Gl.	3SG	book.N	read-IMPF.MSG	walk-PFT.MSG	come.VR	live-PFT.MSG	be-FUT.MSG
EGl	He would have gone on reading a book (for a long time).						

E225 उसके द्वारा किताब पढ़ी गयी

E225	उसके	द्वारा	किताब	पढ़ी	गयी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.PFT.FSG
EGl	The book was read by him/her.				

E226 उससे किताब पढ़ी जाती

E226	उससे	किताब	पढ़ी	जाती
IPA	ʊs-se:	kɪta:b	pəɖʰ-i:	ɟa:-ti:
Gl.	he-INS	book.N	read-PFT.FSG	go-IMPF.FSG
EGl	The book would be read by him/her.			

E227 उसके द्वारा किताब पढ़ी जाएगी

E227	उसके	द्वारा	किताब	पढ़ी	जाएगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-e:gi
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go-FUT.3FSG
EGl	The book will be read by him/her.				

E228 उसके द्वारा किताब पढ़ी जाए

E228	उसके	द्वारा	किताब	पढ़ी	जाए
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-e:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.OPT
EGl	The book should be read by him/her.				

E229 मुझे इस किताब के पढ़े जाने की आशा काफी है

E229	मुझे	इस	किताब	के	पढ़े	जाने	की	आशा	काफी	है
IPA	mʊjʰe:	ɪs	kɪta:b	ke:	pəɖʰe:	ɟa:ne:	ki:	a:ʃa:	ka:pʰi	hɛ:
Gl.	1SG.DAT	this-OBL	book.N	GEN	read-PFT.PL	go.INF	GEN	hope.N	very	be.PRS.SG
EGl	I am very hopeful of this book being read.									

E230 उसके द्वारा पढ़ी जाने वाली किताब

E230	उसके	द्वारा	पढ़ी	जाने	वाली	किताब
IPA	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ne:	wa:li:	kɪta:b
Gl.	3SG.OBL-GEN	by.INS	read-PFT.FSG	go.INF	PRT	book.N
EGl	... the book being read by him/her ...					

E231 उसके द्वारा किताब पढ़ी जाती है

E231	उसके	द्वारा	किताब	पढ़ी	जाती	है
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ti:	hɛ:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go-IMPF.FSG	be.PRS.SG
EGl	The book gets read by him/her.					

E232 उसके द्वारा किताब पढ़ी गयी है

E232	उसके	द्वारा	किताब	पढ़ी	गयी	है
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəji:	hɛ:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.PFT.FSG	be.PRS.SG
EGl	The book has been read by him/her.					

E233 उसके द्वारा किताब पढ़ी जा रही है

E233	उसके	द्वारा	किताब	पढ़ी	जा	रही	है
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	rəh-i:	hɛ:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.VR	live-PFT.FSG	be.PRS.SG
EGl	The book is being read by him/her.						

E234 उसके द्वारा किताब पढ़ी जानी है

E234	उसके	द्वारा	किताब	पढ़ी	जानी	है
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ni:	hɛ:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.INF	be.PRS.SG
EGl	The book is to be read by him/her.					

E235 उसके द्वारा किताब पढ़ी जाने वाली है

E234	उसके	द्वारा	किताब	पढ़ी	जाने	वाली	है
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ne:	wa:li:	hɛ:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.INF	PRT	be.PRS.SG
EGl	The book is about to be read by him/her.						

E236 उसके द्वारा किताब पढ़ी जाती थी

E236	उसके	द्वारा	किताब	पढ़ी	जाती	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	tʰi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go-IMPF.FSG	be.PST.FSG
EGl	The book used to get read by him/her.					

E237 उसके द्वारा किताब पढ़ी गयी थी

E237	उसके	द्वारा	किताब	पढ़ी	गयी	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəji:	tʰi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.PFT.FSG	be.PST.FSG
EGl	The book had been read by him/her.					

E238 उसके द्वारा किताब पढ़ी जा रही थी

E238	उसके	द्वारा	किताब	पढ़ी	जा	रही	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	rəhi:	tʰi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.VR	live-PFT.FSG	be.PST.FSG
EGl	The book was being read by him/her.						

E239 किताब उसके द्वारा पढ़ी जाने लगी

E239	किताब	उसके	द्वारा	पढ़ी	जाने	लगी
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ne:	ləgi
Gl.	book.N	3SG.OBL-GEN	by.INS	read-PFT.FSG	go.INF	begin-PFT.FSG
EGl	The book was started to be read by him/her.					

E240 किताब उसके द्वारा पढ़ी जाने लगेगी

E240	किताब	उसके	द्वारा	पढ़ी	जाने	लगेगी
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:	ləge:gi:
Gl.	book.N	3SG.OBL-GEN	by.INS	read-PFT.FSG	go.INF	begin-FUT.3FSG
EGl	The book will start being read by him/her.					

E241 उसके द्वारा किताब पढ़ी जाती रही

E241	उसके	द्वारा	किताब	पढ़ी	जाती	रही
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	rəhi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go-IMPF.FSG	live-PFT.FSG
EGl	The book was kept on being read by him/her.					

E242 उसके द्वारा किताब पढ़ी जाती रहेगी

E242	उसके	द्वारा	किताब	पढ़ी	जाती	रहेगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	rəh-e:gi
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go-IMPF.FSG	live-FUT.3FSG
EGl	The book will keep on being read by him/her.					

E243 उसके द्वारा किताब पढ़ी गई होती

E243	उसके	द्वारा	किताब	पढ़ी	गई	होती
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəi:	ho:-ti:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.PFT.FSG	be-IMPF.FSG
EGl	The book might have been read by him/her.					

E244 उससे किताब पढी जा सकी

E244	उससे	किताब	पढ़ी	जा	सकी
IPA	ʊs-se:	kɪta:b	pəɖʰ-i:	ɟa:	sək-i:
Gl.	3SG.OBL-INS	book.N	read-PFT.FSG	go.VR	can-PFT.FSG
EGl	The book could get read by him/her.				

E245 उसके द्वारा किताब पढ़ी जानी होती

E245	उसके	द्वारा	किताब	पढ़ी	जानी	होती
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ni:	ho:-ti:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.INF	be-IMPF.FSG
EGl	The book had had to be read by him/her.					

E246 उसके द्वारा किताब पढ़ी जानी थी

E246	उसके	द्वारा	किताब	पढ़ी	जानी	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ni:	tʰi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.INF	be.PST.FSG
EGl	The book was to be read by him/her.					

E247 किताब उसकी पढ़ी होती थी

E247	किताब	उसकी	पढ़ी	होती	थी
IPA	kɪta:b	us-ki:	pəɖʰ-i:	ho:-ti:	tʰi:
Gl.	book.N	3SG.OBL-GEN	read-PFT.FSG	be-IMPF.FSG	be.PST.FSG
EGl	He had the book read.				

E248 उसके द्वारा किताब पढ़ी जाने वाली थी

E248	उसके	द्वारा	किताब	पढ़ी	जाने	वाली	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ne:	wa:li:	tʰi:
Gl.	3SG.OBL-GEN	by.INS	book.N	read-PFT.FSG	go.INF	PRT	be.PST.FSG
EGl	The book was about to be read by him/her.						

E249 इसे भुलाया नहीं जा सकता

E249	इसे	भुलाया	नहीं	जा	सकता
IPA	ɪse:	bʰʊla:ja:	nəhĩ:	ɟa:	sək-ta:
Gl.	it.OBL.ACC	forget-PFT.MSG	NEG	go.VR	can-IMPF.MSG
EGl	It cannot be forgotten.				

E250 उसके द्वारा किताब पढ़ी गयी होगी

E250	उसके	द्वारा	किताब	पढ़ी	गयी	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəji:	ho:-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go.PFT.FSG	be-FUT.FSG
EGl	The book would have been read by him.					

E251 उसके द्वारा किताब पढ़ी जा रही होगी

E251	उसके	द्वारा	किताब	पढ़ी	जा	रही	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	rəh-i:	ho:-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go.VR	live-PFT.FSG	be-FUT.FSG
EGl	The book would have been getting read by him.						

E252 उसके द्वारा किताब पढ़ी जाती होगी

E252	उसके	द्वारा	किताब	पढ़ी	जाती	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	ho:-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go-IMPF.FSG	be-FUT.FSG
EGl	The book would be getting read by him.					

E253 उसके द्वारा किताब पढी जा सकेगी

E253	उसके	द्वारा	किताब	पढ़ी	जा	सकेगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	sək-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go.VR	can-FUT.FSG
EGl	?The book would be able to get read by him.					

E254 उसके द्वारा किताब पढ़ी गई हो

E254	उसके	द्वारा	किताब	पढ़ी	गई	हो
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəi:	ho:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go.PFT.FSG	be.OPT
EGl	The book might have gotten read by him.					

E255 उसके द्वारा किताब पढ़ी जानी होगी

E255	उसके	द्वारा	किताब	पढ़ी	जानी	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ni:	ho:-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go.INF	be-FUT.FSG
EGl	The book was to be read by him.					

E256 उसके द्वारा किताब पढ़ी होती होगी

E256	उसके	द्वारा	किताब	पढ़ी	होती	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ho:-ti:	ho:-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go-IMPF.FSG	be-FUT.FSG
EGl	The book might have been read by him.					

E257 उसके द्वारा किताब पढ़ी जाने वाली होगी

E257	उसके	द्वारा	किताब	पढ़ी	जाने	वाली	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ne:	wa:li:	ho:-gi:
Gl.	3SG.OBL-GEN	GEN	book.N	read-PFT.FSG	go.INF	PRT	be-FUT.FSG
EGl	The book would have been about to be read by him.						

E258 उसके द्वारा यह काम होने दी जायेगी

E258	उसके	द्वारा	यह	काम	होने	दी	जायेगी
IPA	ʊs-ke:	dwa:ra:	jəh	ka:m	ho:ne:	di:	ɟa:-e:gi
Gl.	3SG.OBL-GEN	INS	this	job.N	be.INF	give-PFT.FSG	go-FUT.FSG
EGl	This job will let be done by him.						

E259 उसके द्वारा किताब पढ़ी जाती रहे

E259	उसके	द्वारा	किताब	पढ़ी	जाती	रहे
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	rəh-e:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go-IMPF.FSG	live-PFT.PL
EGl	The book should get read read by him.					

E260 उसके द्वारा किताब पढ़ी गई हो

E260	उसके	द्वारा	किताब	पढ़ी	गई	हो
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəi:	ho:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.PFT.FSG	be.VR
EGl	The book should have been read by him.					

E261 किताब उसके द्वारा पढ़ी जानी चाहिए

E261	किताब	उसके	द्वारा	पढ़ी	जानी	चाहिए
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ni:	ca:hɪje
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.INF	should
EGl	The book should be read by him.					

E262 नियुक्ति उसके द्वारा की जानी हो तो

E262	नियुक्ति	उसके	द्वारा	की	जानी	हो	तो
IPA	nɪju:kti:	ʊs-ke:	dwa:ra:	ki:	ɟa:ni:	ho:	to:
Gl.	appointment.N	3SG.OBL.GEN	INS	GEN	go.INF	be.CFCT	then
EGl	If the appointment is to be done by him.						

E263 उसके द्वारा किए जा सकें

E263	उसके	द्वारा	किए	जा	सकें
IPA	ʊs-ke:	dwa:ra:	kɪje:	ɟa:	səkẽ:
Gl.	3SG.OBL.GEN	INS	do-PFT.PL	ɟa:.VR	can.OPT
EGl	… could be done by him.				

E264 भुनाए जा सकने वाले चेक

E264	भुनाए	जा	सकने	वाले	चेक
IPA	bʰunaː-e:	ɟa:	səkne:	wa:le:	ce:k
Gl.	encash.PFT.PL	go.VR	can.INF	PRT	cheque.N
EGl	… cheques that can be encashed …				

E265 उसके द्वारा पढ़ी जा रही किताब

E265	उसके	द्वारा	पढ़ी	जा	रही	किताब
IPA	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:	rəh-i:	kıta:b
Gl.	3SG.OBL.GEN	INS	read-PFT.FSG	go.VR	live-PFT.FSG	book.N
EGl	… the book being read by him …					

E266 किताब उसके द्वारा पढ़ी जा चुकी है

E266	किताब	उसके	द्वारा	पढ़ी	जा	चुकी	है
IPA	kıta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:	cʊk-i:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.VR	finish-PFT.FSG	be.PRS.SG
EGl	The book has been read by him.						

E267 किताब उसके द्वारा पढ़ी जाने लगती है

E267	किताब	उसके	द्वारा	पढ़ी	जाने	लगती	है
IPA	kıta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ne:	ləg-ti:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.INF	begin-IMPF.FSG	be.PRS.SG
EGl	The book starts being read by him.						

E268 किताब उसके द्वारा पढ़ी जाने लगी है

E268	किताब	उसके	द्वारा	पढ़ी	जाने	लगी	है
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ne:	ləg-i:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.INF	begin-PFT.FSG	be.PRS.SG
EGl	The book has started being read by him.						

E269 किताब उसके द्वारा पढ़ी जाती रही है

E269	किताब	उसके	द्वारा	पढ़ी	जाती	रही	है
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:-ti:	rəh-i:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go-IMPF.FSG	live-PFT.FSG	be.PRS.SG
EGl	? The book has been being read by him.						

E270 ?? उसके द्वारा किताब पढ़ती जाती रही है

??E270	उसके	द्वारा	किताब	पढ़ती	जाती	रही	है
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-ti:	ɟa:-ti:	rəh-i:	hɛ:
Gl.	3SG.OBL.GEN	INS	book.N	read-IMPF.FSG	go-IMPF.FSG	live-PFT.FSG	be.PRS.SG
EGl	? The book has gone on being read by him.						

E271 किताब उसके द्वारा पढी जा सकती है

E271	किताब	उसके	द्वारा	पढ़ी	जा	सकती	है
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:	sək-ti:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.VR	can-IMPF.FSG	be.PRS.SG
EGl	The book can be read by him.						

E272 किताब उसके द्वारा पढी जा सकी है

E272	किताब	उसके	द्वारा	पढ़ी	जा	सकी	है
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:	sək-i:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.VR	can-PFT.FSG	be.PRS.SG
EGl	The book could have been read by him.						

E273 ??किताब उसके द्वारा पढ़ी गई होती है

??E273	किताब	उसके	द्वारा	पढ़ी	गई	होती	है
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	gəi:	ho:-ti:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.PFT.FSG	be-IMPF.FSG	be.PRS.SG
EGl	The book has got read by him.						

E274 ??किताब उसके द्वारा पढ़ी जाती रहती है

??E274	किताब	उसके	द्वारा	पढ़ी	जाती	रहती	है
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:-ti:	rəh-ti:	hɛ:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go-IMPF.FSG	live-IMPF.FSG	be.PRS.SG
EGl	The book has been getting read by him.						

E275 ??किताब उसके द्वारा पढ़ी जाती रहती थी

??E275	किताब	उसके	द्वारा	पढ़ी	जाती	रहती	थी
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:-ti:	rəh-ti:	tʰi:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go-IMPF.FSG	live-IMPF.FSG	be.PST.FSG
EGl	The book had been getting read by him.						

E276 उसके द्वारा किताब पढ़ी जा चुकी थी

E276	उसके	द्वारा	किताब	पढ़ी	जा	चुकी	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	cʊk-i:	tʰi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.VR	finish-PFT.FSG	be.PST.FSG
EGl	The book had got read by him.						

E277 किताब उसके द्वारा पढ़ी जाने लगती थी

E277	किताब	उसके	द्वारा	पढ़ी	जाने	लगती	थी
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ne:	ləg-ti:	tʰi:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.INF	begin-IMPF.FSG	be.PST.FSG
EGl	The book started getting read by him.						

E278 किताब उसके द्वारा पढ़ी जाने लगी थी

E278	किताब	उसके	द्वारा	पढ़ी	जाने	लगी	थी
IPA	kɪta:b	ʊs-ke:	dwa:ra:	pəɖʰ-i:	ɟa:ne:	ləg-i:	tʰi:
Gl.	book.N	3SG.OBL.GEN	INS	read-PFT.FSG	go.INF	begin-PFT.FSG	be.PST.FSG
EGl	The book had started getting read by him.						

E279 उसके द्वारा किताब पढ़ी जाती रही थी

E279	उसके	द्वारा	किताब	पढ़ी	जाती	रही	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	rəh-i:	tʰi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go-IMPF.FSG	live-PFT.FSG	be.PST.FSG
EGl	The book has been getting read by him.						

E280 उसके द्वारा किताब पढी जा सकती थी

E280	उसके	द्वारा	किताब	पढ़ी	जा	सकती	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	sək-ti:	tʰi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.VR	can-IMPF.FSG	be.PST.FSG
EGl	The book could have been read by him.						

E281 उसके द्वारा किताब पढी जा सकी थी

E281	उसके	द्वारा	किताब	पढ़ी	जा	सकी	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	sək-i:	tʰi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.VR	can-PFT.FSG	be.PST.FSG
EGl	The book could get read by him.						

E282 उसके द्वारा किताब पढ़ी जानी चाहिए थी

E282	उसके	द्वारा	किताब	पढ़ी	जानी	चाहिए	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ni:	ca:hɪje:	tʰi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.INF	should	be.PST.FSG
EGl	The book should have been read by him.						

E283 ?? उसके द्वारा किताब पढ़ी गई होती थी

??E283	उसके	द्वारा	किताब	पढ़ी	गई	होती	थी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəi:	ho:-ti:	tʰi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.PFT.FSG	be-IMPF.FSG	be.PST.FSG
EGl	The book used to get read by him.						

E284 उसके द्वारा किताब पढ़ी जा चुकी होगी

E284	उसके	द्वारा	किताब	पढ़ी	जा	चुकी	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	cʊk-i:	ho:-gi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.VR	finish-PFT.FSG	be-FUT.FSG
EGl	The book would have been read by him.						

E285 उसके द्वारा किताब पढी जा सकी होगी

E285	उसके	द्वारा	किताब	पढ़ी	जा	सकी	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	sək-i:	ho:-gi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.VR	can-PFT.FSG	be-FUT.FSG
EGl	The book would have been read by him.						

E286 उसके द्वारा किताब पढ़ी जाने लगी होगी

E286	उसके	द्वारा	किताब	पढ़ी	जाने	लगी	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ne:	ləg-i:	ho:-gi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.INF	begin-PFT.FSG	be-FUT.FSG
EGl	The book would have started getting read by him.						

E287 उसके द्वारा किताब पढ़ी जाने लगती होगी

E287	उसके	द्वारा	किताब	पढ़ी	जाने	लगती	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:ne:	ləg-ti:	ho:-gi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.INF	begin-IMPF.FSG	be-FUT.FSG
EGl	The book would have started getting read by him.						

E288 उसके द्वारा किताब पढ़ी जाती रही होगी

E288	उसके	द्वारा	किताब	पढ़ी	जाती	रही	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:-ti:	rəh-i:	ho:-gi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go-IMPF.FSG	live-PFT.FSG	be-FUT.FSG
EGl	The book would have been getting read by him.						

E289 ?? उसके द्वारा किताब पढ़ी गई होती होगी

??E289	उसके	द्वारा	किताब	पढ़ी	गई	होती	होगी
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	gəi:	ho:-ti:	ho:-gi:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.PFT.FSG	be-IMPF.FSG	be-FUT.FSG
EGl	The book would have got being read by him.						

E290 उसके द्वारा किताब पढ़ी जा रही हो

E290	उसके	द्वारा	किताब	पढ़ी	जा	रही	हो
IPA	ʊs-ke:	dwa:ra:	kɪta:b	pəɖʰ-i:	ɟa:	rəh-i:	ho:
Gl.	3SG.OBL.GEN	INS	book.N	read-PFT.FSG	go.VR	live-IMPF.FSG	be.OPT
EGl	The book should have been being read by him.						

E291 काम उसके द्वारा किए जा सके हों

E291	काम	उसके	द्वारा	किए	जा	सके	हों
IPA	ka:m	ʊs-ke:	dwa:ra:	kɪe:	ɟa:	sək-e:	hõ:
Gl.	work.N	3SG.OBL.GEN	INS	do-PFT.FSG	go.VR	can-PFT.PL	be.OPT
EGl	The work should have been done by him.						

E292 वह दौड़ सका

E292	वह	दौड़	सका
IPA	wəh	dɔːɖ	səkˑaː
Gl.	3SG	run.VR	can-PFT.MSG
EGl	He could run.		

E293 वह चल नहीं पाया

E293	वह	चल	नहीं	पाया
IPA	wəh	cəl	nəhĩː	paːjaː
Gl.	3SG	walk.VR	NEG	get-PFT.MSG
EGl	He could not walk.			

E294 ?*वह किताब पढ़ डालकर लिखने बैठा

?*E294	वह	किताब	पढ़	डालकर	लिखने	बैठा
IPA	wəh	kɪtaːb	pəɖʰ	ɖaːlˑkər	lɪkʰneː	bɛːʈʰˑaː
Gl.	3SG	book.N	read.VR	put-PFT.PCPL	write.INF	sit-PFT.MSG
EGl	Having read the book he sat to write.					

E295 ?*उपाय सोच निकालकर वह बोला

?*E295	उपाय	सोच	निकालकर	वह	बोला
IPA	ʊpa:j	so:c	nɪka:l-kər	wəh	bol-a:
Gl.	solution.N	thought.N	take out-PFT.PCPL	3SG	speak-PFT.MSG
EGl	Having found the solution he spoke.				

E296 ?*कमरा छान मारकर वह थक गया

?*E296	कमरा	छान	मारकर	वह	थक	गया
IPA	kəmra:	cʰa:n	ma:r-kər	wəh	tʰək	gəja:
Gl.	room.N	sift.N	kill-PFT.PCPL	3SG	tire.VR	go.PFT.MSG
EGl	Having sifted through the room he tired away.					

E297 जो मार खाकर खड़ा रहे, वही आगे चलता है

E297	जो	मार	खाकर	खड़ा	रहे	वही	आगे	चलता	है
IPA	ɟo:	ma:r	kʰa:-kər	kʰəɖ-a:	rəh-e:	wəh-i:	a:ge:	cəl-ta:	hɛ:
Gl.	who	beat.VR	eat-PFT.PCPL	stand-PFT.MSG	live-PFT.MSG	he-EMPH	ahead	walk-PFT.MSG	be.PRS.SG
EGl	Those who keep standing even after being beaten, move ahead.								

E298 चकाचक कपड़े पर जा बैठकर सुगंध पर दुर्गन्ध का मुलम्मा चढ़ा डालता है

E298	चकाचक	कपड़े	पर	जा	बैठकर
IPA	cəka:cək	kəpɖe:	pər	ɟa:	bɛ:ʈʰ-kər
Gl.	shining	clothes.N	LOC	go.VR	sit-PFT.PCPL
EGl	… having gone, sitting on the shining clothes …				

E299 तू उतर आ

E299	तू	उतर	आ
IPA	tu:	ʊtər	a:
Gl.	2SG	climb down.VR	come-VR
EGl	You climb down!		

E300 दुश्मन को चारों ओर से जा घेरो

E300	दुश्मन	को	चारों	ओर	से	जा	घेरो
IPA	dʊʃmən	ko:	carõ:	o:r	se:	ɟa:	gʰe:r-o:
Gl.	enemy	ACC	four-PL.OBL	side	ASS	ɟa:.VR	surround-IMP
EGl	… Surround the enemies from all four sides …						

E301 जिसको चाहे ला घसीटो

E301	जिसको	चाहे	ला	घसीटो
IPA	ɟɪs-ko:	ca:he:	la:	gʰəsi:ʈ-o:
Gl.	who-ACC	want-PFT.PL	bring.VR	pull-IMP
EGl	… Pull any one you want …			

E302 मुझे भी दिलवा दो

E302	मुझे	भी	दिलवा	दो
IPA	mʊjʰe:	bʰi:	dɪlwa:	do:
Gl.	1SG.ACC	also.PRT	cause to give.VR	give-IMP
EGl	Get me also given.			

E303 कभी भी आ धमको

E303	कभी	भी	आ	धमको
IPA	kəbʰi:	bʰi:	a:	dʰəmk-o:
Gl.	any time	also	come.VR	thump-IMP
EGl	Get here any time.			

E304 बसों के सड़कों पर उतर आने से जाम के नजारे आम होने लगे

E304	बसों	के	सड़कों	पर	उतर	आने	से
IPA	bəsõ:	ke:	səɖkõ:	pər	ʊtər	a:ne:	se:
Gl.	buses	GEN	roads	LOC	come down.VR	come.INF	INS
EGl	… due to buses coming out on the roads …						

E305 अपनी लड़ाई को किसी राजनीतिक पार्टी द्वारा ले उड़ने की इजाजत न दें

E305	ले	उड़ने	की	इजाजत	न	दें
IPA	le:	uɖne:	ki:	ɪja:jət	nə	dẽ:
Gl.	take.INF	fly.INF	GEN	permission.N	NEG	give-IMP
EGl	… do not permit to fly away with …					

E306 हजार दीपक के एक साथ जल उठने से पूरा मंदिर परिसर रात में ही दिन का एहसास देने लगा

E306	हजार	दीपक	के	एक	साथ	जल	उठने	से
IPA	həɟa:r	di:pək	ke:	e:k	sa:t^{h}	jəl	ʊʈ^{h}ne:	se:
Gl.	thousand.N	lamp.N	GEN	one	with	burn.VR	rise.INF	ASS
EGl	… with a thousand lamps lighting up …							

E307 पुरस्कार कैटरीना की झोली में आ गिरने को बेकरार है

E307	पुरस्कार	कैटरीना	की	झोली	में	आ	गिरने	को	बेकरार	है
IPA	pʊrəskaːr	kɛːʈriːnaː	kiː	ʄʰoːliː	mɛ̃ː	aː	gɪrneː	koː	beːkraːr	hɛː
Gl.	prize.N	Katrina	GEN	bag	LOC	come.VR	fall.INF	ACC	eager	hɛː.PRS.SG
EGl	The prize is eager to fall into Katrina’s bag.									

E308 सूरज को मार गिराने की कहानी

E308	सूरज	को	मार	गिराने	की	कहानी
IPA	suːrəj	koː	maːr	gɪraːneː	kiː	kəhaːniː
Gl.	Suraj.N	ACC	mar.VR	fall.INF	GEN	story.N
EGl	The story of killing down Suraj.					

E309 वह किताब पढ़ डालेगा

E309	वह	किताब	पढ़	डालेगा
IPA	wəh	kɪtaːb	pəɖʰ	ɖaːl-eːgaː
Gl.	3SG	book.N	read.VR	put-FUT.3MSG
EGl	He will complete reading the book.			

E310 उसने किताब पढ़ डाली

E310	उसने	किताब	पढ़	डाली
IPA	ʊs-ne:	kɪta:b	pəɖʱ	ɖa:l-i:
Gl.	3SG.OBL-ERG	book.N	read.VR	put-PFT.3FSG
EGl	He completed reading the book.			

E311 किताब पढ़ डाल

E311	किताब	पढ़	डाल
IPA	kɪta:b	pəɖʱ	ɖa:l
Gl.	book.N	read.VR	put-IMP
EGl	He completed reading the book.		

E312 वह किताब पढ़ डाले

E311	वह	किताब	पढ़	डाले
IPA	wəh	kɪta:b	pəɖʱ	ɖa:l-e
Gl.	3SG	book.N	read.VR	put-IMP/OPT
EGl	He should complete reading book.			

E313 वह किताब पढ़ डालता

E313	वह	किताब	पढ़	डालता
IPA	wəh	kɪta:b	pəɖʰ	ɖa:l-ta:
Gl.	3SG	book.N	read.VR	put-IMPF.MSG
EGl	He would complete reading book.			

E314 जरा किताब पढ़ देना!

E314	जरा	किताब	पढ़	देना
IPA	ɟəra:	kɪta:b	pəɖʰ	de:na:
Gl.	little	book.N	read.VR	give.INF
EGl	Could (you) read the book a bit?			

E315 सुबह आ जाना तुम्हारा काम है

E315	सुबह	आ	जाना	तुम्हारा	काम	है
IPA	sʊbəh	a:	ɟa:na:	tʊmha:ra:	ka:m	hɛ:
Gl.	morning	come.VR	go.INF	2SG.POSS	job.N	be.PRS.SG
EGl	It’s your job to come in the morning?					

E316 वह किताब पढ़ डालता है

E316	वह	किताब	पढ़	डालता	है
IPA	wəh	kɪta:b	pəɖʰ	ɖa:l-ta:	hɛ:
Gl.	3SG	book.N	read.VR	put-IMPF.MSG	be.PRS.SG
EGl	He completes reading the book.				

E317 वह किताब पढ़ डालती थी

E317	वह	किताब	पढ़	डालती	थी
IPA	wəh	kɪta:b	pəɖʰ	ɖa:l-ti:	tʰi:
Gl.	3SG	book.N	read.VR	put-IMPF.FSG	be.PST.FSG
EGl	She used to complete reading the book.				

E318 उसने किताब पढ़ डाली थी

E318	उसने	किताब	पढ़	डाली	थी
IPA	ʊsne:	kɪta:b	pəɖʰ	ɖa:l-i:	tʰi:
Gl.	3SG.OBL.ERG	book.N	read.VR	put-PFT.FSG	be.PST.FSG
EGl	She had completed reading the book.				

E319 देखना वह मार खाने न पाए

E319	देखना	वह	मार	खाने	न	पाए
IPA	de:kʰna	wəh	ma:r	kʰa:ne:	nə	pa:-e:
Gl.	see.INF	3SG	kill.INF	eat.INF	NEG	get-PFT.PL
EGl	Watch that he should not get beaten.					

E320 काश तुम उसे मार भगाए होते

E320	काश	तुम	उसे	मार	भगाए	होते
IPA	ka:ʃ	tʊm	ʊse:	ma:r	bʰəga:-e:	ho:-te:
Gl.	wish.INJ	2SG	he-OBL.ACC	beat.VR	throw-PFT.PL	be-IMPF.PL
EGl	I wish you had beaten him away.					

E321 उसने किताब पढ़ डाली होगी

E321	उसने	किताब	पढ़	डाली	होगी
IPA	ʊs-ne:	kɪta:b	pəɖʰ	ɖa:l-i:	ho:-gi:
Gl.	3SG.OBL.ERG	book.N	read-VR	put.PFT.FSG	be-FUT.FSG
EGl	He would have had the book read.				

E322 वह सुबह सुबह आ जाया करेगा

E322	वह	सुबह-सुबह	आ	जाया	करेगा
IPA	wəh	sʊbəh-sʊbəh	a:	ɟa:ja:	kər-e:ga:
Gl.	3SG	morning-morning	come.VR	go.PFT.MSG	do-FUT.MSG
EGl	He would keep on coming early in the morning.				

E323 वह किताब पढ़ डालता होगा

E323	वह	किताब	पढ़	डालता	होगा
IPA	wəh	kɪta:b	pəɖʰ	ɖa:l-ta:	ho:-ga:
Gl.	3SG	book	read.VR	put-IMPF.MSG	do-FUT.MSG
EGl	He would keep reading the book completed.				

E324 वह किताब पढ़ डालता होवे

E324	वह	किताब	पढ़	डालता	होवे
IPA	wəh	kɪta:b	pəɖʰ	ɖa:l-ta:	ho:-we
Gl.	3SG	book	read.VR	put-IMPF.MSG	be.OPT
EGl	He should have reading the book completed.				

E325 मुझे जल्द से जल्द यह काम कर लेनी पड़ेगी

E325	मुझे	जल्द	से	जल्द	यह	काम	कर	लेनी	पड़ेगी
IPA	mʊjʰe:	ɟəld	se:	ɟəld	jəh	ka:m	kər	le:ni:	pəɖe:-e:gi:
Gl.	1SG.ACC	soon	ASS	soon	this	job	do.VR	take.INF	lay-FUT.3FSG
EGl	I will have to complete this job as soon as possible.								

E326 शेर सियार को मार डालने वाला होगा

E326	शेर	सियार	को	मार	डालने	वाला	होगा
IPA	ʃe:r	sɪja:r	ko:	ma:r	ɖa:lne:	wa:la:	ho:ga:
Gl.	lion	jackal	ACC	kill.VR	put.INF	PRT	be-FUT.MSG
EGl	The lion will be about to kill the jackal.						

E327 आम इसी तरह आ लगते रहें तो क्या कहने

E327	आम	इसी	तरह	आ	लगते	रहें	तो	क्या	कहने
IPA	a:m	ɪsi:	tərəh	a:	ləgte:	rəhẽ:	to:	kja:	kəhne:
Gl.	mango	this-EMPH	like	come.VR	begin-IMPF.PL	live.OPT	then	what	say.INF
EGl	If the mangoes keep on fruiting like this, then it’s wow!								

E328 तू यहां हर रोज आ जाया कर

E328	तू	यहां	हर	रोज	आ	जाया	कर
IPA	tu:	jəhã:	hər	ro:j	a:	ja:ja:	kər
Gl.	2SG	here	every	day	come.VR	go.PFT.MSG	do.IMP
EGl	You keep on coming here everyday.						

E329 कर लेने दो

E329	कर	लेने	दो
IPA	kər	le:ne:	do:
Gl.	do.VR	take.INF	give-IMP
EGl	Let (someone) do it.		

E330 कर लेने दूं

E330	कर	लेने	दूं
IPA	kər	le:ne:	dũ:
Gl.	do.VR	take.INF	give.OPT.1SG
EGl	(I should) let (someone) do it.		

E331 उसे आ जाना चाहिए

E331	उसे	आ	जाना	चाहिए
IPA	ʊs-e:	a:	ɟa:na:	ca:hɪje:
Gl.	he-OBL-ACC	come.VR	go.INF	should
EGl	He should have come.			

E332 कर लेने देना

E332	कर	लेने	देना
IPA	kər	le:ne:	de:na:
Gl.	do.VR	take.INF	go.INF.IMP
EGl	(You) let (someone) get done.		

E333 मार करते चलें

E333	मार	करते	चलें
IPA	ma:r	kər-te:	cəlẽ:
Gl.	kill.VR	do-IMPF.PL	walk.OPT/IMP
EGl	Go on fighting.		

E334 मार करता चलूं

E334	मार	करता	चलूं
IPA	ma:r	kər-ta:	cəlũ:
Gl.	kill.VR	do-IMPF.MSG	walk.OPT.1SG
EGl	(should I) go on fighting.		

E335 अगर वह छूट जाती हों

E335	अगर	वह	छूट	जाती	हों
IPA	əgər	wəh	$c^{h}u:ʈ$	ɟa:-ti:	hõ:
Gl.	if	she	leave.VR	go-IMPF.FSG	be.CFCT
EGl	If she is left out…				

E336 लिख डालते रहें

E336	लिख	डालते	रहें
IPA	$lɪk^{h}$	ɖa:l-te:	rəh-ẽ:
Gl.	write.VR	put-IMPF-PL	live.OPT
EGl	… should keep on completing the writing.		

E337 काट खाए हों

E337	काट	खाए	हों
IPA	ka:ʈ	kʰa:-e:	hõ:
Gl.	cut.VR	eat-PFT-PL	be.OPT
EGl	(if it) is eaten out.		

E338 काट खाने वाले कुत्ते

E338	काट	खाने	वाले	कुत्ते
IPA	ka:ʈ	kʰa:-ne:	wa:le:	kʊtte:
Gl.	cut.VR	eat-INF	PRT	dogs
EGl	… the biting dogs …			

E339 उसे आ जाना चाहिए था

E339	उसे	आ	जाना	चाहिए	था
IPA	ʊs-e:	a:	ɟa:na:	ca:hɪje	tʰa:
Gl.	3SG.OBL-ACC	come.VR	go.INF	should	be.PST.MSG
EGl	He should have had gone.				

E340 उसने चोर को आ घुसने दिया होगा

E340	उसने	चोर	को	आ	घुसने	दिया	होगा
IPA	ʊs-ne:	co:r	ko:	a:	gʰʊsne:	dɪja:	ho:-ga:
Gl.	3SG.OBL-ERG	thief.N	ACC	come.VR	enter.INF	give-PFT.MSG	be-FUT.3MSG
EGl	He might have had let the thief enter.						

E341 उसे मार खानी पड़ी होगी

E341	उसे	मार	खानी	पड़ी	होगी
IPA	ʊs-e:	ma:r	kʰa:ni	pəɖ-i:	ho:-gi:
Gl.	3SG.OBL-ACC	beat.VR	eat.INF	lay.PFT.FSG	be-FUT.3FSG
EGl	He might have had to get the beatings.				

E342 ले जाने देना चाहिए

E342	ले	जाने	देना	चाहिए
IPA	le:	ɟa:ne:	de:na:	ca:hɪje:
Gl.	take.VR	go.INF	give.INF	should
EGl	(one) should let (someone) take away.			

E343 पढ़ डाली जाएगी

E343	पढ़	डाली	जाएगी
IPA	pəɖʰ	ɖa:li	ja:-e:gi
Gl.	read.VR	put-PFT.FSG	go-FUT.3FSG
EGl	… will get read.		

E344 शिकारी शिकार को मार गिरा रहा है

E344	शिकारी	शिकार	को	मार	गिरा	रहा	है
IPA	ʃɪka:ri:	ʃɪka:r	ko:	ma:r	gɪra:	rəh-a:	hɛ:
Gl.	hunter	hunt	ACC	kill.VR	make fall.VR	live-PFT.MSG	be.PRS.SG
EGl	The hunter has been getting the hunt killed.						

E345 उसे उठ बैठना पड़ता है

E345	उसे	उठ	बैठना	पड़ता	है
IPA	ʊs-e:	ʊʈʰ	bɛ:ʈʰ-na:	pəɖ-ta:	hɛ:
Gl.	3SG.OBL.ACC	rise.VR	sit.INF	lay-IMPF.MSG	be.PRS.SG
EGl	He has to get up and sit down.				

E346 *वह लिख चुकता रहा है

*E346	वह	लिख	चुकता	रहा	है
IPA	wəh	lɪkʰ	cʊk-ta:	rəh-a:	hɛ:
Gl.	3SG	write.VR	finish-IMPF.MSG	live-PFT.MSG	be.PRS.SG
EGl	He has kept on completing the writing.				

E347 *उसे लिख चुकना पड़ता है

*E347	उसे	लिख	चुकना	पड़ता	है
IPA	ʊse:	lɪkʰ	cʊkna:	pəɖ-ta:	hɛ:
Gl.	3SG.OBL.ACC	write.VR	finish.INF	lay-IMPF.MSG	be.PRS.SG
EGl	He has to complete the writing.				

E348 *वह लिख चुकने लगा है

*E348	वह	लिख	चुकने	लगा	है
IPA	wəh	lɪkʰ	cʊkne:	ləg-a:	hɛ:
Gl.	3SG	write.VR	finish.INF	begin-PFT.MSG	be.PRS.SG
EGl	He has started completing the writing.				

E349 जंगली घास उखाड़ फेंकने पड़ते हैं

E349	जंगली	घास	उखाड़	फेंकने	पड़ते	हैं
IPA	jə̃gli:	gʰa:s	ʊkʰa:ɖ	pʰẽ:kne:	pəɖ-te:	hɛ̃:
Gl.	wild	grass	uproot.VR	throw.INF	lay-IMPF.PL	be.PRS.PL
EGl	The wild grass has to be uprooted.					

E350 वह काट कर खाने लगा है

E350	वह	काट	कर	खाने	लगा	है
IPA	wəh	ka:ʈ	kər	kʰa:ne:	ləga:	hɛ:
Gl.	3SG	cut	do.VR	eat.INF	begin-PFT.PL	be.PRS.SG
EGl	He has started cutting and eating.					

E350a वह काट खाने लगा है

E350a	वह	काट	खाने	लगा	है
IPA	wəh	ka:ʈ	kʰa:ne:	ləg-a:	hɛ:
Gl.	3SG	cut	eat.INF	begin-PFT.PL	be.PRS.SG
EGl	He has started cutting and eating.				

E351 जंगली घास उखाड़ कर फेंकने पड़ते हैं

E351	जंगली	घास	उखाड़	कर	फेंकने	पड़ते	हैं
IPA	jə̃gli:	gʰa:s	ʊkʰa:ɖ	kər	pʰẽ:kne:	pəɖ-te:	hẽ:
Gl.	wild	grass	uproot.VR	do.VR	throw.INF	lay-IMPF.PL	be.PRS.PL
EGl	The wild grass has to be thrown away.						

E351a जंगली घास उखाड़ फेंकने पड़ते हैं

E351a	जंगली	घास	उखाड़	फेंकने	पड़ते	हैं
IPA	jə̃gli:	gʰa:s	ʊkʰa:ɖ	pʰẽ:kne:	pəɖ-te:	hɛ̃:
Gl.	wild	grass	uproot.VR	throw.INF	lay-IMPF.PL	be.PRS.PL
EGl	The wild grass has to be thrown away.					

E352 ?*वह किताब पढ़ सकता रहा

?*E352	वह	किताब	पढ़	सकता	रहा
IPA	wəh	kıta:b	pəɖʰ	sək-ta:	rəh-a:
Gl.	3SG	book	read.VR	can-IMPF.MSG	live-PFT.MSG
EGl	He went on being able to read the book.				

E353 वह खाते ही बोल पड़ा “वाह”

E353	वह	खाते	ही	बोल	पड़ा	“वाह”
IPA	wəh	kʰaː-teː	hiː	boːl	pəɖ-aː	waːh
Gl.	3SG	eat-IMPF.PL	only-PRT	speak.VR	lay-PFT.MSG	wow
EGl	Just as he ate, he said ‘wow’!					

E354 वह रोने लगा

E354	वह	रोने	लगा
IPA	wəh	roːneː	ləg-aː
Gl.	3SG	cry.INF	begin-PFT.MSG
EGl	He started crying.		

E355 मैं यह काम करना चाहता हूं

E355	मैं	यह	काम	करना	चाहता	हूं
IPA	mɛ̃ː	jəh	kaːm	kərnaː	caːh-taː	hũ
Gl.	I	this	job	do.INF	want-IMPF.MSG	be.PRS.1SG
EGl	I want to do this job.					

E356 उसने उसे टोपी पहना कर ठग लिया

E356	उसने	उसे	टोपी	पहना	कर	ठग	लिया
IPA	ʊs-ne:	ʊs-e:	ʈopi:	pəhna:	kər	ʈʰəg	lɪja:
Gl.	he.OBL-ERG	he.OBL-ACC	cap	wear.INF	do.VR	cheat.VR	take-PFT.MSG
EGl	Having capped him, he cheated him.						

E357 तू रोज टोपी पहना कर

E357	तू	रोज	टोपी	पहना	कर
IPA	tu:	ro:ɟ	ʈopi:	pəhna:	kər
Gl.	2SG	everyday	cap	wear.INF	do.VR
EGl	You wear a cap everyday!				

E358 वे लोग किताब पढ़े

E358	वे	लोग	किताब	पढ़े
IPA	we	lo:g	kɪta:b	pəɖʰ-e:
Gl.	3PL	people	book	read.PFT.PL
EGl	They read a book!			

E359 उसने उसे अपने पास बैठ रखवाली करने को कहा

E359	उसने	उसे	अपने	पास	बैठ	रखवाली	करने	को	कहा
IPA	ʊs-ne:	ʊse:	əpne:	pa:s	bɛ:ʈʰ	rəkʰwa:li:	kərne:	ko:	kəh-a:
Gl.	3SG.OBL-ERG	3SG.OBL.ACC	own.REFL	near	sit.VR.PCPL	watch.N	do.INF	ACC	say-PFT.MSG
EGl	He asked him to watch while sitting near him.								

E360 वह खाते-खाते बोला

E360	वह	खाते-खाते	बोला
IPA	wəh	kʰa:te:-kʰa:te:	bol-a:
Gl.	3SG	eat-IMPF.PCPL	speak-PFT.MSG
EGl	He spoke while eating.		

E361 अपने सिर पर लगाते नजर आए

E361	अपने	सिर	पर	लगाते	नजर	आए
IPA	əpne:	sɪr	pər	ləga:-te:	nəʄər	a:-e:
Gl.	own	head	on-LOC	apply-IMPF.PL	view	come-PFT.MSG
EGl	was seen applying on his/her/their head.					

E362 अपने बिगड़ते स्वास्थ्य की वजह से

E362	अपने	बिगड़ते	स्वास्थ्य	की	वजह	से
IPA	əpne:	bɪgəɖ-te:	swa:stʰjə	ki:	vəɟəh	se:
Gl.	own	worsen-IMPF.PL	health	GEN	cause	ASS
EGl	… due to his/her/their worsening health …					

E363 राम का कहना है कि वह वहां नहीं था

E363	राम	का	कहना	है	कि	वह	वहां	नहीं	था
IPA	ra:m	ka:	kəhna:	hɛ:	kɪ	wəh	wəhã:	nəhĩ:	tʰa:
Gl.	Ram	GEN	say.INF	be.PRS.SG	that	3SG	there	NEG	be.PST.MSG
EGl	… Ram has to say that he was not there.								

E364 राम को कहना है वह वहां नहीं था

E364	राम	को	कहना	है	कि	वह	वहां	नहीं	था
IPA	ra:m	ko:	kəhna:	hɛ:	kɪ	wəh	wəhã:	nəhĩ:	tʰa:
Gl.	Ram	ACC	say.INF	be.PRS.SG	that	3SG	there	NEG	be.PST.MSG
EGl	… Ram has to say that he was not there.								

E365 गाड़ी खुलने ही वाली है

E365	गाड़ी	खुलने	ही	वाली	है
IPA	ga:ɖi:	kʰʊlne:	hi:	wa:li:	hɛ:
Gl.	vehicle	open.INF	only.PRT	PRT	be.PRS.SG
EGl	The train is just about to start.				

E366 पढ़ने वाले लड़के को बुलाओ

E366	पढ़ने	वाले	लड़के	को	बुलाओ
IPA	pəɖʰne:	wa:le:	ləɖke:	ko:	bʊlao
Gl.	read.INF	PRT	boy.OBL	ACC	call-IMP
EGl	Call the reading boy.				

E367 पढ़ी जा रही हो

E367	पढ़ी	जा	रही	हो
IPA	pəɖʰ-i:	ja:	rəhi:	ho:
Gl.	read-PFT.FSG	go.VR	live.PFT.FSG	be.CFCT
EGl	… should be getting read.			

E368 खा रहा हो

E368	खा	रहा	हो
IPA	ɟa:	rəh-i:	ho:
Gl.	go.VR	live.PFT-FSG	be.CFCT
EGl	… should be getting eaten.		

E369 नियुक्ति की जानी हो तो

E369	नियुक्ति	की	जानी	हो	तो
IPA	nɪjʊktɪ	ki:	ɟa:ni:	ho:	to:
Gl.	appointment	do.PFT.FSG	go.INF	be.CFCT	then
EGl	… if the appointment has to be done then …				

E370 *सुलाई दे रहा है

*E370	सुलाई	दे	रहा	है
IPA	sʊlai:	de:	rəha:	hɛ:
Gl.	sleep-PFT.FSG	give.VR	live.INF	be.PRS.SG
EGl	is letting (someone) sleep.			

E371 *जम्हाई दे रहा है

*E371	जम्हाई	दे	रहा	है
IPA	jəmha:i:	de:	rəh-a:	hɛ:
Gl.	doze-PFT.FSG	give.VR	live.INF	be.PRS.SG
EGl	… is letting (someone) doze off.			

E372 *पढ़ाई दे रहा है

*E372	पढ़ाई	दे	रहा	है
IPA	pəɖʰa:i:	de:	rəh-a:	hɛ:
Gl.	read-PFT.FSG	give.VR	live.PFT.MSG	be.PRS.SG
EGl	… is letting (someone) read.			

E373 वह किताब पढ़ डालता रहा

E373	वह	किताब	पढ़	डालता	रहा
IPA	wəh	kɪta:b	pəɖʰ	ɖa:l-ta:	rəh-a:
Gl.	3SG	book	read.VR	put-IMPF.MSG	live.PFT.MSG
EGl	He kept on completing reading the book.				

E374 वह उठ बैठता है

E374	वह	उठ	बैठता	है
IPA	wəh	uʈʰ	bɛːʈʰ-ta	hɛː
Gl.	3SG	rise.VR	sit.IMPF.MSG	be.PRS.SG
EGl	He gets up.			

E375 वह चल देता है

E375	वह	चल	देता	है
IPA	wəh	cəl	de-ta:	hɛː
Gl.	3SG	walk.VR	give.IMPF.MSG	be.PRS.SG
EGl	He starts moving ahead.			

E376 वह उसे ले भागा है

E376	वह	उसे	ले	भागा	है
IPA	wəh	ʊse:	le:	bʰaːg-a:	hɛː
Gl.	3SG	he-OBL.ACC	take.VR	run away.PFT.MSG	be.PRS.SG
EGl	He has taken it away.				

E377 उसने इसका चटनी बना डाला है

E377	उसने	इसका	चटनी	बना	डाला	है
IPA	ʊs-ne:	ɪs-ka:	cəʈni:	bən-a:	ɖa:l-a:	hɛ:
Gl.	3SG.OBL-ERG	this.OBL-GEN	sauce	make.VR	put-PFT.MSG	be.PRS.SG
EGl	He has smashed it.					

E378 पटना में उसने एक अच्छा घर देखा ।

E378	पटना	में	उसने	एक	अच्छा	घर	देखा
IPA	pəʈna	me:	ʊs-ne:	e:k	əcc^{h}a:	g^{h}ər	de:k^{h}-a:
Gl.	Patna	LOC	3SG.OBL.ERG	one	good	house	see-PFT.MSG
EGl	He has smashed it.						

APPENDIX 5: EXAMPLE SENTENCES WITH THEIR ANNOTATIONS IN THE IL-POST FRAMEWORK

The example annotated sentences given here are after the corrections made by this researcher. The actual corpus identification (CID) number of these sentences are not the same as provided in the actual LDC corpus. This ID has changed because I merged the separate text files into one single corpus file and deleted a few undesired line breaks that appeared in the single corpus file.

CID	Annotated Sentence
46	उत्तरप्रदेश\NP.mas.sg.obl.gen के\PP.0.0.gen इलाहाबाद\NP.mas.sg.obl.loc में\PP.0.0.loc गंगा\NP.fem.sg.obl.gen ,\PU यमुना\NP.fem.sg.obl.gen एवं\CCD.n अदृश्य\JJ.0.0.dir सरस्वती\NP.fem.sg.obl.gen के\PP.0.0.gen संगम\NC.mas.sg.obl.loc पर\PP.0.0.loc चल\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.prs.pft.dcl.fin.n अर्द्धकुंभ\NP.mas.sg.obl.loc मेले\NC.mas.sg.obl.loc में\PP.0.0.loc फर्जी\JJ.0.0.dir पत्रकारों\NC.0.pl.obl.gen की\PP.fem.sg.gen भी\CX.n भरमार\NC.fem.sg.dir.0 हैं\VM.0.sg.3.prs.sim.dcl.fin.n और\CCD.n वे\PPR.0.sg.3.dir.0.n.n.dst.n पत्रकाि◌रता\NC.fem.sg.obl.gen की\PP.fem.sg.gen आड़\NC.fem.sg.obl.loc में\PP.0.0.loc साधुसंतों-\NC.mas.pl.obl.acc एवं\CCD.n जनसामान्य-\NC.mas.sg.obl.acc को\PP.0.0.acc ◌िदगभ्रिमत\JJ.0.0.dir कर\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.prs.pft.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n ।\PU
53	इलाहाबाद\NP.mas.sg.obl.loc में\PP.0.0.loc गंगा\NP.fem.sg.obl.gen की\PP.fem.sg.gen रेती\NC.fem.sg.obl.loc पर\PP.0.0.loc बसें\VM.mas.pl.3.pst.pft.dcl.fin.n

	तंबुओं\NC.mas.pl.obl.gen की\PP.fem.sg.gen नगरी\NC.fem.sg.obl.loc में\PP.0.0.loc सोमवार\NP.mas.sg.dir.0 रात\NC.fem.sg.dir.0 एक\JQ.0.sg.dir.n.crd पंडाल\NC.mas.sg.obl.loc में\PP.0.0.loc आग\NC.fem.sg.dir.0 लग\VM.0.0.0.0.0.0.nfn.0 गई\VAUX.fem.sg.3.pst.pft.dcl.fin.n ि◌जसमें\PRL.sg.obl.loc.n.n.n एक\JQ.0.sg.dir.n.crd बच्ची\NC.fem.sg.dir.0 समेत\JQ.0.pl.dir.n.nnm दो\JQ.0.pl.dir.n.crd लोग\NC.mas.sg.dir.0 झुलस\VM.0.0.0.0.0.0.nfn.0 गए\VAUX.mas.pl.3.pst.pft.dcl.fin.n तथा\CCD.n पंडाल\NC.mas.sg.obl.loc में\PP.0.0.loc रखा\VM.mas.sg.3.pst.pft.dcl.fin.n सारा\JJ.mas.sg.dir सामान\NC.mas.sg.dir.0 जलकर\VM.0.0.0.0.pft.0.nfn.0 राख\NC.mas.sg.dir.0 हो\VM.0.0.0.0.pft.0.nfn.0 गया\VAUX.mas.sg.3.pst.pft.dcl.fin.n ।\PU
96	यहाँ\ALC.dir.0.n चल\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.prs.pft.dcl.fin.n अर्द्धकुंभ\NP.mas.sg.obl.loc में\PP.0.0.loc गंगा\NP.fem.sg.obl.gen के\PP.0.0.gen प्रद◌ूषित\JJ.0.0.dir जल\NC.mas.sg.obl.abl से\PP.0.0.abl कई\JQ.0.pl.dir.n.nnm लोग\NC.0.sg.dir.0 बीमार\JJ.0.0.dir हो\VM.0.0.0.0.pft.0.nfn.0 गए\VAUX.mas.pl.3.prs.pft.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n ।\PU
113	धार्◌िमक\JJ.0.0.dir नगरी\NC.fem.sg.dir.0 इलाहाबाद\NP.mas.sg.obl.loc में\PP.0.0.loc चल\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.sg.3.prs.pft.dcl.fin.n अर्द्धकुंभ\NP.mas.sg.obl.loc मेले\NC.mas.sg.obl.loc में\PP.0.0.loc देश\NC.mas.sg.obl.gen के\PP.0.0.gen कोने-कोने\NC.mas.sg.obl.abl से\PP.0.0.abl आए\VM.mas.pl.3.pst.pft.dcl.fin.n साधु-संतों\NC.mas.pl.obl.gen और\CCD.n श्रद्धालुओं\NC.sg.n.0.n के\PP.0.0.gen अलावा\CX.n

	देशभर\AMN.obl.gen.n के\PP.0.0.gen ठगों\NC.mas.pl.obl.erg ने\PP.0.0.erg भी\CX.n डेरा\NC.mas.sg.dir.0 डाल\VM.0.0.0.0.0.0.nfn.0 रखा\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
125	तीर्थ\NC.mas.sg.dir.0 नगरी\NC.fem.sg.dir.0 प्रयाग\NP.mas.sg.obl.loc (\RDS इलाहाबाद\NP.mas.sg.dir.0)\RDS में\PP.0.0.loc चल\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.sg.3.prs.pft.dcl.fin.n अर्द्धकुंभ\NP.mas.sg.obl.loc में\PP.0.0.loc आने\NV.dir.0 वाले\CX.n ि◌वदेशी\JJ.0.0.dir पर्यटकों\NC.0.pl.obl.abl में\PP.0.0.loc से\PP.0.0.abl कुछ\JQ.0.pl.dir.n.nnm के\PP.0.0.gen लिए\PP.0.0.x यह\DAB.sg.dir.n.prx महज\JJ.0.0.dir एक\JQ.0.sg.dir.n.crd मेला\NC.mas.sg.dir.0 हो\VM.0.0.0.0.pft.0.nfn.0 सकता\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ले◌िकन\CCD.n इनमें\PPR.0.pl.1.obl.abl.n.n.prx.n से\PP.0.0.abl कुछ\JQ.0.pl.dir.n.nnm पर्यटक\NC.0.sg.dir.0 ऐसे\AMN.dir.0.n भी\CX.n हैं\VM.0.pl.3.prs.sim.dcl.fin.n जो\PRL.sg.dir.0.n.n.n यहाँ\ALC.dir.0.n आकर\VM.0.0.0.0.pft.0.nfn.0 अपनी\PRF.fem.sg.obl.gen.n वर्षो\NC.0.pl.obl.gen की\PP.fem.sg.gen ख्वा◌िहश\NC.fem.sg.dir.0 पूरी\JJ.fem.sg.dir कर\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.sg.3.prs.pft.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n या\CCD.n अपने\PRF.mas.sg.obl.gen.n ि◌कसी\PPR.0.sg.3.dir.0.n.n.dst.n अजीज\NC.0.sg.obl.acc को\PP.0.0.acc श्रद्धांजिल\NC.fem.sg.dir.0 अर्पित\JJ.0.0.dir कर\VM.0.0.0.0.0.0.nfn.0
154	लोस्वेट\NP.fem.sg.dir.0 लोगो\NC.0.pl.obl.acc को\PP.0.0.acc संगम\NP.mas.sg.obl.loc में\PP.0.0.loc स्नान\NC.mas.sg.dir.0 करते\VM.mas.sg.3.pst.pft.dcl.fin.n

	देख\VM.0.0.0.0.pft.0.nfn.0 इस\DAB.sg.dir.n.prx कदर\CX.n प्रभाि‍वत\JJ.0.00.dir हुई\VM.fem.sg.3.pst.pft.dcl.fin.n की\CSB.n वह\PPR.0.sg.3.dir.0.n.n.dst.n संगम\NP.mas.sg.obl.loc में\PP.0.0.loc डुबकी\NC.fem.sg.dir.0 लगाने\NV.obl.abl से\PP.0.0.abl खुद\NC.0.sg.obl.acc को\PP.0.0.acc रोक\VM.0.0.0.0.0.0.nfn.n नहीं\CX.y पाई\VAUX.fem.sg.3.pst.pft.dcl.fin.n ।\PU
172	*जुलूस\NC.mas.sg.obl.acc को\PP.0.0.acc देखने\NV.obl.x के\PP.0.0.gen लिए\PP.0.0.x लोग\NC.mas.sg.dir.0 अखाड़ा\NC.mas.sg.dir.0 मार्ग\NC.mas.sg.obl.loc पर\PP.0.0.loc लगाई\VM.fem.sg.3.pst.pft.dcl.fin.n गई\VAUX.fem.sg.3.pst.pft.dcl.fin.n बिलल्यो\NC.fem.pl.obl.loc पर\PP.0.0.loc भी\CX.n चढ़ने\NV.obl.abl से\PP.0.0.abl बाज\CX.n नहीं\CX.y आए\VM.mas.pl.3.pst.pft.dcl.fin.n ।\PU
175	*ि‍नर्मोही\JJ.0.0.dir अखाडा\NC.mas.sg.obl.gen के\PP.0.0.gen जुलूस\NC.mas.sg.obl.loc में\PP.0.0.loc सबसे\AMN.dir.0.n पीछे\NST.dir.0 चल\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.pst.pft.dcl.fin.n ि‍वदेशी\JJ.0.0.dir श्रद्धालुओ\NC.mas.pl.obl.gen के\PP.0.0.gen एक\JQ.0.sg.dir.n.crd समूह\NC.mas.sg.obl.erg ने\PP.0.0.erg सबका\JQ.mas.pl.dir.n.nnm ध्यान\NC.mas.sg.dir.0 आकर्ि‍षत\JJ.0.0.dir ि‍कया\VM.mas.sg.3.pst.pft.dcl.fin.n ।\PU
185	सुश्री\NP.fem.sg.obl.erg मायावती\NP.fem.sg.obl.erg ने\PP.0.0.erg कहा\VM.mas.sg.3.pst.pft.dcl.fin.n ि‍क\CSB.n बसपा\NP.fem.sg.obl.abl से\PP.0.0.abl अलग\JJ.0.0.dir हुए\VM.mas.sg.3.pst.pft.dcl.fin.n ३८\RDX ि‍वधायकों\NC.0.pl.obl.gen के\PP.0.0.gen मामले\NC.mas.sg.obl.loc में\PP.0.0.loc उच्चतम\JJ.0.0.dir

	न्यायालय\NC.mas.sg.obl.gen के\PP.0.0.gen सम्भाित\JJ.0.0.dir आदेश\NC.fem.sg.obl.loc के\PP.0.0.gen बाद\PP.0.0.loc मुलायम\NP.mas.sg.dir.0 िसंह\NP.mas.sg.dir.0 यादव\NP.mas.sg.dir.0 सरकार\NC.fem.sg.dir.0 खुद\NC.0.sg.dir.0 ब\CX.n खुद\NC.0.sg.dir.0 िगर\VM.0.0.0.0.0.0.nfn.0 जाएगी\VAUX.fem.sg.3.fut.0.dcl.fin.n ।\PU
196	जो\PRL.sg.dir.0.n.n.n १६\RDX फरवरी\NP.fem.sg.dir.0 महाशिवरात्रि\NP.fem.sg.obl.loc तक\PP.0.0.loc चलेगा\VM.mas.sg.3.fut.0.dcl.fin.n ।\PU
199	संगम\NP.mas.sg.obl.acc को\PP.0.0.acc मुख्य\JJ.0.0.dir केन्द्र\NC.mas.sg.dir.0 मानकर\VM.0.0.0.0.pft.0.nfn.0 पूरे\JJ.mas.sg.dir मेला\NC.mas.sg.dir.0 क्षेत्र\NC.mas.sg.obl.loc में\PP.0.0.loc लगे\VM.mas.pl.3.pst.pft.dcl.fin.n पत्थरों\NC.mas.pl.obl.loc पर\PP.0.0.loc िकलोमीटर\CCL.n दर्शाया\VM.mas.sg.3.prs.pft.dcl.fin.n गया\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n िजससे\DRL.sg.dir.n.dst लोगों\NC.0.pl.obl.acc को\PP.0.0.acc दूरी\JJ.fem.sg.dir मालूम\JJ.0.0.dir हो\VM.0.0.0.0.pft.0.nfn.0 सकेगी\VAUX.fem.sg.3.fut.0.dcl.fin.n ।\PU
203	राज्य\NC.mas.sg.dir.0 सरकार\NC.fem.sg.dir.0 मंगलवार\NP.mas.sg.obl.abl से\PP.0.0.abl प्रतापगढ़\NP.mas.sg.dir.0 िजले\NC.mas.sg.obl.gen के\PP.0.0.gen बदरी\NC.fem.sg.obl.loc के\PP.0.0.gen िनकट\PP.0.0.loc शारदा\NP.fem.sg.dir.0 नहर\NC.fem.sg.obl.abl से\PP.0.0.abl प्रितिदन\AMN.dir.0.n ३००\RDX क्यूसेक\CCL.n

	पानी\NC.mas.sg.dir.0 गंगा\NP.fem.sg.obl.loc में\PP.0.0.loc छोड़ेगी\VM.fem.sg.3.fut.0.dcl.fin.n ताि\CSB.n इलाहाबाद\NP.mas.sg.obl.loc में\PP.0.0.loc उसका\PPR.mas.sg.3.obl.gen.n.n.dst.n जलस्तर\NC.mas.sg.dir.0 बना\VM.mas.sg.3.0.pft.dcl.fin.n रह\VAUX.mas.sg.3.pst.pft.dcl.fin.n ।\PU
218	यह\DAB.sg.dir.n.prx क्रम\NC.mas.sg.dir.0 शाम\ALC.obl.loc.n तक\PP.0.0.loc चलेगा\VM.mas.sg.3.fut.0.dcl.fin.n ।\PU
229	बसंत\NP.0.sg.dir.0 पंचमी\NP.fem.sg.obl.gen का\PP.mas.sg.gen मुहरत\NC.mas.sg.dir.0 रात\NC.fem.sg.obl.abl से\PP.0.0.abl शुरू\JJ.0.0.dir हो\VM.0.0.0.0.pft.0.nfn.0 जाने\NV.obl.gen के\PP.0.0.gen कारण\NC.0.sg.dir.0 श्रद्धालुओं\NC.mas.pl.obl.erg ने\PP.0.0.erg तड़के\NC.fem.sg.dir.0 ही\CX.n स्नान\NC.mas.sg.dir.0 करना\NV.dir.0 शुरू\JJ.0.0.dir कर\VM.0.0.0.0.0.0.nfn.0 दिया\VAUX.mas.sg.3.pst.pft.dcl.fin.n और\CCD.n श्रद्धालुओं\NC.mas.pl.obl.gen के\PP.0.0.gen आने\NV.obl.gen का\PP.mas.sg.gen क्रम\NC.mas.sg.dir.0 जारी\NC.0.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ।\PU
238	यही\DAB.sg.dir.y.prx कारण\NC.0.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n िंक\CSB.n दो\JQ.0.pl.dir.n.crd के\PP.0.0.gen िंविभन्न\JJ.0.0.dir िंहस्सो\NC.mas.pl.obl.loc में\PP.0.0.loc तीर्थ\NC.mas.sg.dir.0 बनाए\VM.mas.pl.3.pst.pft.dcl.fin.n और\CCD.n धर्म\NC.mas.sg.obl.acc को\PP.0.0.acc यात्रा\NC.fem.sg.obl.soc के\PP.0.0.gen साथ\PP.0.0.soc जोड़ा\VM.mas.sg.3.pst.pft.dcl.fin.n गया\VAUX.mas.sg.3.pst.pft.dcl.fin.n ताि\CSB.n अिंधक\JJ.0.0.dir से\PP.0.0.abl अिंधक\JJ.0.0.dir संख्या\NC.fem.sg.obl.loc में\PP.0.0.loc यात्री\NC.mas.sg.dir.0 तीर्थ\NC.mas.sg.dir.0 यात्राओ\NC.fem.pl.obl.loc

	पर\PP.0.0.loc ि◌ंनकलें\VM.mas.pl.3.0.0.imp.fin.n ।\PU
323	-\PU श्री\NP.mas.sg.obl.acc सलीम\NP.mas.sg.obl.acc खान\NP.mas.sg.obl.acc को\PP.0.0.acc कांग्रेस\NP.fem.sg.obl.gen नशा\NP.fem.sg.obl.gen विरोधी\NP.fem.sg.obl.gen प्रकोष्ठ\NP.fem.sg.obl.gen का\PP.mas.sg.gen कार्यकारिणी\NC.mas.sg.dir.0 सदस्य\NC.mas.sg.dir.0
361	इसके\PPR.mas.sg.3.obl.gen.n.n.dst.n लिए\PP.0.0.x शिक्षकों\NC.mas.pl.obl.acc को\PP.0.0.acc उचित\JJ.0.0.dir प्रशिक्षण\NC.mas.sg.dir.0 एवं\CCD.n मार्गदर्शन\NC.mas.sg.dir.0 दिया\VM.mas.sg.3.prs.pft.dcl.fin.n जाना\NV.dir.0 भी\CX.n परमावश्यक\JJ.0.0.dir है\VM.0.sg.3.prs.sim.dcl.fin.n ।\PU
369	मूर्ख\NP.mas.sg.dir.0 दिवस\NP.mas.sg.dir.0 मनाया\VM.mas.sg.3.prs.pft.dcl.fin.n जाना\NV.dir.0 एक\JQ.0.sg.dir.n.crd अति\JJ.0.0.dir प्राचीन\JJ.0.0.dir प्रथा\NC.fem.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ,\PU लेकिन\CSB.n इसकी\DAB.sg.obl.n.prx प्रारंभिक\JJ.0.0.dir शुरुआत\NC.fem.sg.obl.gen के\PP.0.0.gen संदर्भ\NC.mas.sg.obl.loc में\PP.0.0.loc अलगअलग-\JJ.0.0.dir धारणाएँ\NC.fem.pl.dir.0 हैं\VM.0.pl.3.prs.sim.dcl.fin.n ।\PU
382	काफी\JJ.0.0.dir समय\NC.mas.sg.dir.0 गुजर\VM.0.0.0.0.0.0.nfn.0 जाने\NV.obl.gen के\PP.0.0.gen बावजूद\NC.mas.sg.dir.0 देवता\NC.mas.sg.dir.0 प्रकट\JJ.0.0.dir न\CX.y होने\NV.obl.gen के\PP.0.0.gen कारण\NC.mas.sg.dir.0 मायूस\JJ.0.0.dir व\CCD.n हताश\JJ.0.0.dir होकर\VM.0.0.0.0.pft.0.nfn.0 वह\DAB.sg.dir.n.dst व्यक्ति\NC.mas.sg.dir.0 अपने\PRF.0.sg.dir.0.n साथियों\NC.mas.pl.obl.soc के\PP.0.0.gen साथ\PP.0.0.soc

	वापस\JJ.0.0.dir लौट\VM.0.0.0.0.0.0.nfn.0 आया\VAUX.mas.sg.3.pst.pft.dcl.fin.n और\CCD.n बीच\NST.dir.0 राह\NC.mas.sg.obl.loc में\PP.0.0.loc ही\CX.n उसे\PPR.0.sg.3.obl.acc.n.n.dst.n अपने\PRF.0.sg.dir.0.n वे\PPR.mas.pl.3.dir.0.n.n.dst.n साथी\NC.mas.sg.dir.0 मिल\VM.0.0.0.0.0.0.nfn.0 गए\VAUX.mas.pl.3.pst.pft.dcl.fin.n जिन्होंने\PRL.mas.pl.3.obl.erg.n.n.0.n उसे\PPR.mas.sg.3.obl.acc.n.n.dst.n देवता\NC.mas.sg.dir.0 प्रकट\JJ.0.0.dir होने\NV.obl.gen की\PP.fem.sg.gen बात\NC.fem.sg.dir.0 कही\VM.fem.sg.3.pst.pft.dcl.fin.n थी\VAUX.fem.sg.3.pst.sim.dcl.fin.n ।\PU
388	नौकर\NC.mas.sg.dir.0 अपने\PRF.0.sg.dir.0.n मालिक\NC.mas.sg.obl.gen के\PP.0.0.gen शानदार\JJ.0.0.dir कपड़े\NC.mas.sg.dir.0 पहनकर\VM.0.0.0.0.pft.0.nfn.0 आराम\NC.mas.sg.dir.0 कुर्सी\NC.mas.sg.obl.loc पर\PP.0.0.loc बैठकर\VM.0.0.0.0.pft.0.nfn.0 हुक्मनामे\NC.mas.sg.dir.0 जारी\JJ.0.0.dir करता\VM.mas.sg.3.pst.impf.dcl.fin.n था\VAUX.mas.sg.3.pst.sim.dcl.fin.n और\CCD.n नौकर\NC.mas.sg.dir.0 बना\VM.mas.sg.3.prs.pft.dcl.fin.n मालिक\NC.mas.sg.dir.0 उसके\PPR.mas.sg.3.obl.gen.n.n.dst.n हर\JJ.0.0.dir हुक्म\NC.mas.sg.obl.gen का\PP.0.0.gen सादर\JJ.0.0.dir पालन\NC.mas.sg.dir.0 करता\VM.mas.sg.3.pst.impf.dcl.fin.n था\VAUX.mas.sg.3.pst.sim.dcl.fin.n ।\PU
390	यहाँ\PPR.mas.sg.3.dir.0.n.n.prx.n तक\PP.0.0.loc कि\CSB.n मालिक\NC.mas.sg.dir.0 बने\VM.mas.sg.3.pst.pft.dcl.fin.n नौकर\NC.mas.sg.obl.acc को\PP.0.0.acc अपने\PRF.0.sg.dir.0.n नौकर\NC.mas.sg.dir.0 बने\VM.0.sg.3.pst.pft.dcl.fin.n

	मालिक\NC.mas.sg.obl.acc को\PP.0.0.acc दंड\NC.mas.sg.dir.0 देने\NV.obl.gen का\PP.mas.sg.gen भी\CX.n पूरा\JJ.mas.sg.dir अधिकार\NC.mas.sg.dir.0 व\CCD.n छूट\NC.mas.sg.dir.0 थी\VM.fem.sg.3.pst.sim.dcl.fin.n ।\PU
407	यह\DAB.sg.dir.n.prx दो\JQ.0.pl.dir.n.crd शताब्दी\NC.fem.sg.dir.0 ईसा\NP.mas.sg.dir.0 पूर्व\JJ.0.0.dir बना\VM.mas.sg.3.pst.pft.dcl.fin.n बौद्ध\JJ.0.0.dir कला\NC.fem.sg.obl.gen का\PP.mas.sg.gen नमूना\NC.mas.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ।\PU
430	इसकी\DAB.sg.dir.n.prx दर्शनीयता\JJ.0.0.dir देखते\VM.mas.sg.3.prs.impf.dcl.fin.n ही\CX.n बनती\VM.fem.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
445	िडजाइनरों\NC.mas.pl.obl.acc को\PP.0.0.acc हमारी\PPR.fem.sg.1.obl.gen.n.n.prx.n व\CCD.n अगली\NST.dir.0 पीढ़ी\NC.fem.sg.obl.gen के\PP.0.0.gen बीच\NST.dir.0 सेतु\NC.mas.sg.dir.0 बनाना\NV.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ।\PU
481	जब\ALC.dir.0.n हम\PPR.0.pl.1.dir.0.n.n.prx.n सातवीं\JQ.fem.pl.dir.n.ord से\PP.0.0.abl तेरहवी\JQ.fem.pl.dir.n.ord शताब्दी\NC.fem.sg.obl.gen के\PP.0.0.gen स्थापत्य\NC.mas.sg.obl.gen का\PP.mas.sg.gen पर्यवेक्षण\NC.mas.sg.dir.0 करते\VM.mas.pl.3.prs.impf.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n तो\CX.n हम\PPR.0.sg.1.dir.0.n.n.prx.n पाते\VM.mas.pl.3.prs.impf.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n िक\CSB.n वह\PPR.0.sg.3.dir.0.n.n.dst.n एक\JQ.0.sg.dir.n.crd नये\JJ.mas.pl.dir राजनीितक\JJ.0.0.dir पिरस्िथित\NC.fem.sg.obl.gen के\PP.0.0.gen अनुरुप\NC.0.sg.dir.0 ढल\VM.0.0.0.0.0.0.nfn.0 जाता\VAUX.mas.sg.3.prs.impf.dcl.fin.n

	है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
521	पहाड़ी\JJ.0.0.dir ढ़लान\NC.mas.sg.obl.acc ,\PU चढ़ाव\NC.mas.sg.obl.acc और\CCD.n उतार\NC.mas.sg.obl.acc को\PP.0.0.acc ध्यान\NC.mas.sg.obl.loc में\PP.0.0.loc रखते\VM.mas.sg.3.pst.impf.dcl.fin.n हुए\VAUX.mas.sg.3.pst.pft.dcl.fin.n सारे\JQ.0.pl.dir.n.nnm नगर\NC.mas.sg.obl.acc को\PP.0.0.acc टेढ़ामेढ़ा-\JJ.mas.sg.dir इस\DAB.sg.dir.n.prx तरह\AMN.dir.0.n बसाया\VM.mas.sg.3.pst.pft.dcl.fin.n गया\VAUX.mas.sg.3.pst.pft.dcl.fin.n िक\CSB.n मध्यकालीन\JJ.0.0.dir उदयपुर\NP.mas.sg.obl.gen की\PP.fem.sg.gen योजना\NC.fem.sg.obl.loc में\PP.0.0.loc कहीं\ALC.dir.0.n रास्ते\NC.mas.sg.obl.gen या\CCD.n चौपड़\NC.fem.sg.obl.gen की\PP.fem.sg.gen व्यवस्था\NC.fem.sg.dir.0 नहीं\CX.y दीख\VM.0.0.0.0.0.0.nfn.0 पड़ती\VAUX.fem.sg.3.pst.impf.dcl.fin.n ।\PU
548	रावल\NP.mas.sg.obl.erg जैसल\NP.mas.sg.obl.erg ने\PP.0.0.erg अपनी\PRF.fem.sg.obl.gen.n स्वतंत्र\JJ.0.0.dir राजधानी\NC.fem.sg.dir.0 स्थापित\JJ.0.0.dir की\VM.fem.sg.3.pst.pft.dcl.fin.n थी\VAUX.fem.sg.3.pst.sim.dcl.fin.n ।\PU
696	जैसे\AMN.dir.0.n किसी\PPR.0.sg.3.dir.0.n.n.dst.n बच्ची\NC.fem.sg.obl.acc को\PP.0.0.acc पुचकार\VM.0.0.0.0.0.0.nfn.n ,\PU दुलार\VM.0.0.0.0.0.0.nfn.n रहे\VAUX.mas.pl.3.prs.pft.dcl.fin.n हों\VAUX.0.pl.3.0.pft.0.cfct.n -\PU ऐसे\AMN.dir.0.n शब्द\NC.mas.sg.dir.0 और\CCD.n आवाजें\NC.fem.pl.dir.0 मेरे\PPR.fem.sg.1.dir.gen.n.n.prx.n कानों\NC.mas.pl.obl.loc में\PP.0.0.loc पड़\VM.0.0.0.0.0.0.nfn.n रही\VAUX.fem.sg.1.pst.ppft.dcl.fin.n

	थीं\VAUX.fem.sg.1.pst.sim.dcl.fin.n और\CCD.n मेरा\PPR.mas.sg.1.dir.0.n.n.prx.n पूरा\JJ.mas.sg.dir शरीर\NC.mas.sg.dir.0 पिघल\VM.0.0.0.0.pft.0.nfn.n रहा\VAUX.mas.sg.1.pst.pft.dcl.fin.n था\VAUX.mas.sg.1.pst.sim.dcl.fin.n ।\PU
706	सालोंसाल-\NC.mas.pl.dir.0 वंचित\JJ.0.0.dir रही\VM.fem.sg.1.prs.pft.dcl.fin.n हूं\VAUX.fem.sg.1.prs.sim.dcl.fin.n मैं\PPR.fem.sg.1.dir.0.n.n.prx.n ।\PU
789	'\PU अचानक\JJ.0.0.dir गीती\NP.fem.sg.obl.gen भाभी\NC.fem.sg.obl.gen का\PP.mas.sg.gen धीमा\JJ.mas.sg.dir और\CCD.n दृढ़\JJ.fem.sg.dir स्वर\NC.mas.sg.dir.0 सुनाई\NV.dir.0 दिया\VM.mas.sg.3.pst.pft.dcl.fin.n ,\PU '\PU हां\CX.n हां\CX.n कहो\VM.0.sg.2.0.0.imp.fin.n न\CX.n ,\PU तब\CX.n से\PP.0.0.abl चुपचाप\JJ.0.0.dir ,\PU स्क्रिप्ट\NC.fem.sg.obl.loc पर\PP.0.0.loc नजर\NC.fem.sg.dir.0 गड़ाए\VM.0.sg.2.0.pft.dcl.fin.n बैठी\VM.fem.sg.2.prs.pft.dcl.fin.n हो\VAUX.0.0.0.0.pft.0.nfn.n ।\PU
876	'\PU हां\CX.n क्यों\PWH.sg.dir.0.n.n.n नहीं\CX.y इन\DAB.pl.dir.n.prx तीन\JQ.0.pl.dir.n.crd दिनों\NC.0.pl.obl.loc में\PP.0.n आपको\PPR.0.sg.3.obl.acc.n.n.dst.y चाय\NC.fem.sg.obl.soc के\PP.0.0.gen साथ\PP.0.0.soc बिस्किट\NC.mas.sg.dir.0 ऑफर\RDF किए\VM.0.pl.3.0.pft.dcl.fin.n तो\CX.n आपने\PPR.0.sg.3.obl.gen.n.n.dst.y मना\VM.0.0.0.0.pft.0.nfn.0 कर\VAUX.0.0.0.0.0.0.nfn.0 दिया\VAUX.mas.sg.3.pst.pft.dcl.fin.n हमेशा\AMN.dir.0.n ।\PU
920	क्षणभर\NC.0.sg.obl.x के\PP.0.0.gen लिए\PP.0.0.x मेरे\PPR.0.sg.1.obl.gen.n.n.prx.n मन\NC.mas.sg.obl.loc में\PP.0.0.loc जैसे\AMN.dir.0.n गुस्से\NC.mas.sg.obl.gen

	का\PP.mas.sg.gen उबाल\NC.mas.sg.dir.0 आ\VM.0.0.0.0.0.0.nfn.0 गया\VAUX.mas.sg.3.pst.pft.dcl.fin.n ,\PU लेकिन\CSB.n मैंने\PPR.0.sg.1.obl.gen.n.n.prx.n तुरंत\AMN.0.0.n उसे\PPR.0.sg.3.dir.0.n.n.dst.n दबा\VM.mas.sg.3.0.0.0.nfn.n दिया\VAUX.mas.sg.3.pst.pft.dcl.fin.n ।\PU
999	बांध\NC.mas.sg.obl.loc में\PP.0.0.loc एक\JQ.0.sg.dir.n.crd दरार\NC.fem.sg.dir.0 हो\VM.0.0.0.0.0.0.nfn.0 जाने\NV.obl.gen की\PP.fem.sg.gen देर\JJ.fem.sg.dir थी\VM.fem.sg.3.pst.sim.dcl.fin.n ,\PU फिर\CSB.n तो\CX.n मन\NC.mas.sg.obl.gen की\PP.fem.sg.gen उमंगो\NC.fem.pl.obl.erg ने\PP.0.0.erg खुदखुद-ब-\JJ.0.0.dir काम\NC.mas.sg.dir.0 करना\NV.dir.0 शुरु\JJ.0.0.dir किया\VM.mas.sg.3.pst.pft.dcl.fin.n ।\PU
1074	मेरे\PPR.0.sg.1.obl.x.n.n.prx.n लिए\PP.0.0.x यह\DAB.sg.dir.n.prx डूब\VM.0.0.0.0.0.0.nfn.0 मरने\NV.obl.gen की\PP.fem.sg.gen जगह\NC.fem.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n कि\CSB.n मैं\PPR.0.sg.1.dir.0.n.n.prx.n अपनी\PRF.fem.sg.obl.gen.n बीवी\NC.fem.sg.obl.abl से\PP.0.0.abl इस\DAB.sg.dir.n.prx तरह\NC.0.sg.obl.gen की\PP.fem.sg.gen बातें\NC.fem.pl.dir.0 करूं\VM.0.sg.1.prs.sim.opt.fin.n ,\PU उसकी\PPR.fem.sg.3.obl.gen.n.n.dst.n ओर\NC.0.sg.obl.abl से\PP.0.0.abl मेरे\PPR.0.sg.1.obl.gen.n.n.prx.n दिल\NC.mas.sg.obl.loc मे\PP.0.0.loc संदेह\NC.mas.sg.dir.0 पैदा\VM.0.0.0.0.0.0.nfn.0 हो\VAUX.0.0.0.0.pft.0.nfn.0 ।\PU
1097	मैं\PPR.0.sg.1.dir.0.n.n.prx.n तो\CX.n अपनी\PRF.fem.sg.obl.gen.n

	किस्मत\NC.fem.sg.obl.gen का\PP.mas.sg.gen फैसला\NC.mas.sg.dir.0 पहले\NST.obl.abl से\PP.0.0.abl ही\CX.n किये\VM.mas.pl.1.pst.pft.dcl.fin.n बैठा\VM.mas.sg.1.pst.pft.dcl.fin.n था\VAUX.mas.sg.1.pst.sim.dcl.fin.n ।\PU
1456	भारतवर्ष\NP.0.sg.obl.gen की\PP.fem.sg.gen गौरवमयी\JJ.0.sg.dir ,\PU महिमामंडित\JJ.0.sg.dir संत\NC.mas.sg.dir.0 परंपरा\NC.fem.sg.obl.loc में\PP.0.0.loc एक\JQ.0.sg.dir.n.crd श्लाघनीय\JJ.0.sg.dir नाम\NC.0.sg.dir.0 स्वर्णाक्षरों\NC.mas.pl.obl.loc में\PP.0.0.loc उल्लिखित\JJ.0.sg.dir है\VM.0.sg.3.prs.sim.dcl.fin.n -\PU भारत\NP.0.sg.dir.0 ज्योति\NC.0.sg.dir.0 ,\PU अणुव्रत\NP.0.sg.dir.0 अनुशास्ता\NC.0.sg.dir.0 आचार्य\NC.mas.sg.dir.0 श्री\NP.mas.sg.dir.0 तुलसी\NP.mas.sg.dir.0 ।\PU
1505	सच\NC.mas.sg.dir.0 यह\DAB.sg.dir.n.prx है\VM.0.sg.3.prs.sim.dcl.fin.n कि\CSB.n शस्त्र\NC.mas.sg.dir.0 विराम\NC.mas.sg.obl.erg ने\PP.0.0.erg कश्मीर\NP.0.sg.obl.acc को\PP.0.0.acc ऐसे\PPR.0.sg.3.dir.0.n.n.prx.n रास्ते\NC.mas.sg.obl.loc पर\PP.0.0.loc ला\VM.0.0.0.0.0.0.nfn.0 खड़ा\VAUX.mas.sg.3.prs.pft.dcl.fin.n कर\VAUX.0.0.0.0.0.0.nfn.0 दिया\VAUX.mas.sg.3.pst.pft.dcl.fin.n ,\PU जहाँ\ALC.dir.0.n सत्ताधारी\JJ.0.sg.dir नेशनल\NP.mas.sg.obl.gen कॉन्फ्रेंस\NP.mas.sg.obl.gen का\PP.mas.sg.gen भविष्य\NC.0.sg.dir.0 तो\CX.n दाँव\NC.mas.sg.obl.loc पर\PP.0.0.loc लग\VM.0.0.0.0.0.0.nfn.0 ही\CX.n गया\VAUX.mas.sg.3.pst.pft.dcl.fin.n ,\PU कश्मीर\NP.0.sg.obl.loc में\PP.0.0.loc बढ़े\JJ.0.sg.dir मौत\NC.fem.sg.obl.gen के\PP.0.0.gen आँकड़ों\NC.mas.pl.obl.ins से\PP.0.0.ins सुरक्षाबल\NC.mas.sg.dir.0 भी\CX.n

	हतोत्साहित\JJ.0.sg.dir हुए\VM.mas.sg.3.pst.pft.dcl.fin.n ,\PU क्योंकि\CSB.n हाथ\NC.mas.sg.dir.0 बाँध\VM.0.sg.3.pst.pft.dcl.fin.n उन्हें\PPR.mas.pl.3.obl.acc.n.n.dst.n आतंकवादियों\NC.mas.pl.obl.gen के\PP.mas.pl.gen शिकार\NC.mas.sg.obl.x के\PP.mas.sg.gen लिए\PP.0.sg.x छोड़\VM.0.0.0.0.0.0.nfn.0 गया\VAUX.mas.sg.3.prs.sim.dcl.fin.n ।\PU
1554	सुरक्षाबलों\NC.mas.pl.obl.acc को\PP.0.0.acc चोट\NC.0.sg.dir.0 पहुँची\VM.fem.sg.3.pst.pft.dcl.fin.n आतंकवादियों\NC.mas.pl.obl.gen की\PP.fem.sg.gen गतिविधियों\NC.fem.pl.obl.gen के\PP.0.0.gen कारण\NC.0.sg.dir.0 ,\PU जिन्हें\PRL.pl.obl.acc.n.n.n वे\PPR.0.pl.3.dir.0.n.n.dst.n रोक\VM.0.0.0.0.0.0.nfn.0 पाने\NV.obl.loc में\PP.0.0.loc सक्षम\JJ.0.sg.dir तो\CX.n हैं\VM.0.pl.3.prs.sim.dcl.fin.n मगर\CCD.n शस्त्र\NC.mas.sg.dir.0 विराम\NC.mas.sg.obl.gen की\PP.fem.sg.gen बेड़ियाँ\NC.fem.pl.dir.0 पाँव\NC.mas.sg.obl.loc में\PP.0.0.loc होने\NV.obl.ins से\PP.0.0.ins वे\PPR.0.pl.3.dir.0.n.n.dst.n ऐसा\PPR.mas.sg.3.obl.0.n.n.prx.n कोई\PPR.0.sg.3.dir.0.n.n.0.n कदम\NC.mas.sg.dir.0 न\CX.y उठाने\NV.obl.acc को\PP.0.0.acc मजबूर\JJ.0.sg.dir हैं\VM.0.pl.3.prs.sim.dcl.fin.n ,\PU जिससे\PRL.sg.obl.ins.n.n.n शस्त्र\NC.0.sg.dir.0 विराम\NC.mas.sg.dir.0 '\PU बदनाम\JJ.0.0.dir '\PU होता\VM.mas.sg.3.prs.impf.dcl.fin.n हो\VAUX.0.0.0.0.pft.0.nfn.0 ।\PU
1592	शहरों\NC.mas.pl.obl.gen के\PP.mas.pl.gen बागबगीचों-\NC.mas.pl.obl.gen के\PP.0.0.gen बाहर\NST.dir.0 और\CCD.n खुली\JJ.fem.sg.dir जगह\NC.fem.sg.obl.loc में\PP.0.0.loc

	पैदा\VM.0.0.0.0.0.0.nfn.0 होने\NV.dir.0 वाले\CX.n पौधों\NC.mas.sg.obl.loc में\PP.0.0.loc एक\JQ.0.sg.dir.n.crd पौधा\NC.mas.sg.dir.0 होता\VM.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n अपामार्ग\NC.mas.sg.obl.gen का\PP.mas.sg.gen जिसे\PRL.sg.obl.acc.n.n.n बोलचाल\NC.mas.sg.obl.gen की\PP.fem.sg.gen भाषा\NC.fem.sg.obl.loc में\PP.0.0.loc आँधीझाड़ा\NC.mas.sg.dir.0 भी\CX.n कहते\VM.0.pl.3.prs.impf.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n ।\PU
1595	भावप्रकाश\NP.fem.sg.obl.loc निघण्टु\NP.fem.sg.obl.loc में\PP.0.0.loc लिखा\VM.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n -\PU
1681	परंतु\CCD.n हाँ\CX.n ,\PU मेरी\PPR.fem.sg.1.obl.gen.n.n.0.n आम\NC.fem.sg.obl.gen की\PP.fem.sg.gen गुठली\NC.fem.sg.obl.loc में\PP.0.0.loc बना\JJ.mas.sg.dir चित्र\NC.mas.sg.dir.0 देखने\NV.obl.loc के\PP.0.0.gen बाद\NST.dir.0 ,\PU कोई\JJ.0.0.dir व्यक्ति\NC.fem.sg.dir.0 आम\NC.mas.sg.dir.0 खाने\NV.obl.loc के\PP.0.0.gen बाद\NST.dir.0 गुठली\NC.fem.sg.dir.0 फेंक\VM.0.0.0.0.0.0.nfn.0 देने\NV.obl.gen के\PP.0.0.gen पूर्व\NST.dir.0 रुककर\VM.0.0.0.0.pft.0.nfn.0 यदि\CSB.n एक\JQ.0.sg.dir.n.crd क्षण\NC.mas.sg.obl.x के\PP.0.0.gen लिए\PP.0.0.x भी\CX.n उसे\PPR.0.sg.3.obl.acc.n.n.dst.n गौर\NC.mas.sg.obl.soc से\PP.0.0.soc निहारे\VM.mas.pl.3.pst.pft.dcl.fin.n तो\CX.n मैं\PPR.0.sg.1.dir.0.n.n.0.n मेरी\PPR.0.sg.1.obl.gen.n.n.0.n आंशिक\JJ.0.0.dir सफलता\NC.fem.sg.dir.0 मान\VM.0.0.0.0.0.0.nfn.0 सकता\VAUX.mas.sg.3.prs.impf.dcl.fin.n हूँ\VAUX.0.sg.1.prs.sim.dcl.fin.n ।\PU

1682	परंतु\CCD.n मेरी\PPR.fem.sg.1.obl.gen.n.n.0.n टूटी\JJ.fem.sg.dir हुई\VM.fem.sg.3.pst.pft.dcl.fin.n चप्पल\NC.fem.sg.obl.gen की\PP.fem.sg.gen '\PU मोनालिसा\NC.mas.sg.dir.0 '\PU देखने\NV.obl.gen के\PP.0.0.gen बाद\NST.dir.0 अपनी\PPR.fem.sg.3.dir.0.n.n.0.n टूटी\JJ.fem.sg.dir चप्पल\NC.fem.sg.dir.0 यदि\CSB.n कोई\PPR.0.sg.3.dir.0.n.n.0.n मेरे\PPR.0.sg.1.obl.gen.n.n.0.n ही\CX.n पास\NST.dir.0 भेज\VM.0.0.0.0.0.0.0.nfn.0 दे\VAUX.0.0.0.0.0.0.0.nfn.0 तो\CSB.n फिर\CCD.n सिवाय\AMN.dir.0.n माथा\NC.fem.sg.dir.0 ठोक\VM.0.0.0.0.0.0.0.nfn.0 लेने\NV.obl.gen के\PP.0.0.gen मैं\PPR.0.sg.1.obl.0.n.n.0.n कर\VM.0.0.0.0.0.0.0.nfn.0 भी\CX.n क्या\PWH.sg.dir.0.n.n.n सकता\VAUX.mas.sg.3.prs.impf.dcl.fin.n हूँ\VAUX.0.sg.1.prs.sim.dcl.fin.n ?\PU
1746	वैज्ञानिक\JJ.0.0.dir अपने\PRF.0.pl.dir.0.n तरीके\NC.mas.pl.obl.ins से\PP.0.0.ins ,\PU चित्रकार\NC.mas.sg.dir.0 अपने\PRF.mas.pl.dir.0.n तरीके\NC.mas.sg.obl.ins से\PP.0.0.ins तो\CX.n अन्य\JJ.0.0.dir विधा\NC.fem.sg.obl.gen के\PP.0.0.gen धनी\JJ.0.0.dir अपने-अपने\PRF.0.pl.dir.0.n तरीके\NC.mas.sg.obl.ins से\PP.0.0.ins प्रकृति\NC.fem.sg.obl.acc को\PP.0.0.acc समझने\NV.obl.gen का\PP.0.0.gen प्रयास\NC.mas.sg.dir.0 कर\VM.0.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.prs.pft.dcl.fin.y हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n और\CCD.n प्रकृति\NC.fem.sg.dir.0 भी\CX.n अपना\PRF.mas.sg.dir.0.n पूरा\JJ.mas.sg.dir सहयोग\NC.mas.sg.dir.0 दे\VM.0.0.0.0.0.0.0.nfn.0 रही\VAUX.fem.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU

1852	इस\DAB.sg.obl.n.prx ऊर्जाहीनता\NC.fem.sg.obl.gen की\PP.fem.sg.gen स्थिति\NC.0.sg.obl.acc को\PP.0.sg.acc विरक्ति\NC.fem.sg.obl.gen का\PP.0.sg.gen नाम\NC.0.sg.dir.0 देना\NV.dir.0 सिर्फ़\JQ.0.sg.dir.n.nnm अपने\PRF.0.sg.dir.0.n अस्तित्व\NC.0.sg.obl.acc को\PP.0.sg.acc बनाए\VM.0.sg.3.prs.sim.dcl.fin.n रखने\NV.obl.gen की\PP.0.sg.gen छटपटाहट\NC.0.sg.obl.abl से\PP.0.sg.abl अधिक\JJ.0.sg.dir कुछ\JQ.0.sg.dir.n.nnm नहीं\CX.y कहा\VM.0.sg.3.prs.pft.dcl.fin.n जा\VAUX.0.0.0.0.0.0.0.nfn.0 सकता\VAUX.0.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
1863	कुछकुछ-\JJ.0.sg.dir घबराहट\NC.0.sg.dir.0 दिखाई\NV.dir.0 दे\VM.0.0.0.0.0.0.0.nfn.0 रही\VAUX है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
1871	इसी\DAB.sg.obl.y.prx प्रकार\NC.0.sg.dir.0 मानवीय\JJ.0.sg.dir संवेदनाओं\NC.0.pl.obl.gen के\PP.0.sg.gen पाखंड\NC.0.sg.obl.gen और\CCD.n लेखकीय\JJ.0.sg.dir उत्तरदायित्व\NC.0.sg.obl.gen का\PP.0.sg.gen दिखावा\NC.0.sg.dir.0 करने\NV.dir.0 वाले\CX.n समाज\NC.0.sg.obl.loc में\PP.0.sg.loc किसी\PWH.sg.obl.0.n.n.n सजग\JJ.0.sg.dir और\CCD.n चेतनशील\JJ.0.sg.dir व्यक्ति\NC.0.sg.obl.ins द्वारा\PP.0.sg.ins लेखन\NC.0.sg.obl.acc को\PP.0.sg.acc मजबूरीवश\NC.0.sg.dir.0 चलाए\VM.0.sg.3.pst.pft.dcl.fin.n रखना\NV.dir.0 क्या\PWH.sg.dir.0.n.n.n विरक्ति\NC.0.sg.obl.abl से\PP.0.sg.abl भी\CX.n अच्छा\JJ.0.sg.dir हो\VM.0.0.0.0.0.0.0.nfn.0 सकता\VAUX.0.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ?\PU
1876	कुछ\JQ.0.pl.dir.n.nnm के\PP.0.0.gen मन\NC.mas.sg.obl.loc में\PP.0.0.loc

	तत्कालीनता\NC.fem.sg.obl.loc में\PP.0.0.loc मंदिर\NC.mas.sg.dir.0 तोड़कर\VM.0.0.0.0.pft.0.nfn.n बाबर\NC.mas.sg.obl.ins द्वारा\PP.0.0.ins मस्जिद\NC.fem.sg.dir.0 बनाए\VM.0.0.0.prs.pft.dcl.fin.n जाने\NV.obl.gen की\PP.fem.sg.gen कुंठा\NC.fem.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ,\PU तो\CSB.n किसी\PPR.0.sg.3.obl.gen.n.n.dst.n के\PP.0.0.gen मन\NC.mas.sg.obl.loc में\PP.0.sg.loc मस्जिद\NC.fem.sg.dir.0 तोड़कर\VM.0.0.0.0.pft.0.nfn.n धर्मनिरपेक्षता\NC.fem.sg.obl.gen के\PP.0.sg.gen स्वरूप\NC.mas.sg.obl.acc को\PP.0.0.acc ढहाए\VM.0.0.0.prs.sim.dcl.fin.n जाने\NV.obl.gen की\PP.0.sg.gen कुंठा\NC.fem.sg.dir.0 घर\NC.mas.sg.dir.0 कर\VM.0.0.0.0.0.0.nfn.n रही\VAUX.fem.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
1916	वे\PPR.mas.sg.3.dir.0.n.y.dst.y फिर\AMN.dir.0.n कहते\VM.0.sg.3.prs.impf.dcl.fin.y हुए\VAUX.0.sg.3.prs.sim.dcl.fin.n दिखाई\NV.dir.0 देते\VM.0.pl.0.0.impf.dcl.fin.n हैं\VAUX.mas.sg.3.prs.sim.dcl.fin.y -\RDS बेटी\NC.fem.sg.dir.0 मारोगे\VM.mas.pl.2.fut.sim.dcl.fin.0 तो\LC.y.0.0 अनर्थ\NC.mas.sg.dir.0 हो\VM.0.0.0.0.0.0.nfn.n जाएगा\VAUX.mas.sg.3.fut.sim.dcl.fin.n ।\PU
2271	उपभोक्ताओं\NC.mas.pl.obl.acc को\PP.0.0.acc जल\NC.mas.sg.obl.gen की\PP.fem.sg.gen आपूर्ति\NC.fem.sg.dir.0 सामान्य\JJ.0.0.dir रूप\NC.mas.sg.obl.abl से\PP.0.0.abl प्रभावकारी\JJ.fem.sg.dir मांग\NC.fem.sg.dir.0 सिद्धान्त\NC.mas.sg.obl.loc पर\PP.0.0.loc आधारित\JJ.0.0.dir होनी\VM.fem.sg.0.0.0.0.ifn.n चाहिए\VAUX.0.0.0.prs.sim.imp.fin.n और\CCD.n यह\DAB.sg.dir.y.prx मांग\NC.fem.sg.dir.0 मोटे\JJ.0.sg.obl तौर\NC.0.0.dir.0

	पर\PP.0.0.loc उस\DRL.sg.obl.y.dst सेवा\NC.fem.sg.obl.gen के\PP.0.0.gen स्तर\NC.mas.sg.obl.gen के\PP.0.0.gen अनुरूप\NC.0.sg.dir.0 होनी\VM.fem.sg.3.0.0.0.ifn.0 चाहिए\VAUX.0.0.0.prs.sim.imp.fin.n जिसे\PRL.sg.obl.0.n.y.n उपभोक्ता\NC.mas.sg.dir.0 एक\JQ.0.sg.0.y.crd समुदाय\NC.mas.sg.obl.gen के\PP.0.0.gen रूप\NC.mas.sg.obl.abl में\PP.0.0.loc चलाने\NV.dir.0 ,\PU बनाए\VM.0.0.0.prs.sim.dcl.fin.n रखने\NV.dir.0 और\CCD.n वित्त-पोषित\NC.mas.sg.dir.0 करने\NV.obl.x के\PP.0.0.gen लिए\PP.0.0.x तैयार\JJ.0.0.dir हों\VM.0.pl.3.fut.pft.0.fin.n ।\PU
2392	कुछ\JQ.0.pl.dir.n.nnm विशेषज्ञों\NC.mas.pl.obl.gen का\PP.0.0.gen मानना\NV.dir.0 था\VM.mas.sg.3.pst.sim.dcl.fin.n कि\CSB.n पश्चिमी\JJ.0.0.dir तट\NC.mas.sg.obl.abl से\PP.0.0.abl हटाए\VM.mas.pl.3.pst.pft.dcl.fin.n गए\VAUX.mas.pl.3.pst.pft.dcl.fin.n यहूदियों\NC.mas.pl.obl.gen के\PP.0.0.gen कारण\NC.mas.sg.dir.0 इतना\JQ.mas.sg.dir.n.nnm क्रोध\NC.mas.sg.dir.0 फैलेगा\VM.mas.sg.3.fut.sim.dcl.fin.n कि\CSB.n इसराइली\JJ.0.0.dir मतदाता\NC.mas.sg.dir.0 शेरोन\NP.mas.sg.obl.gen की\PP.fem.sg.gen पार्टी\NC.fem.sg.obl.acc को\PP.0.0.acc हरा\VM.mas.sg.3.0.0.0.nfn.n देंगे\VAUX.mas.pl.3.fut.sim.dcl.fin.n लेकिन\CCD.n कदीमा\NP.fem.sg.obl.gen का\PP.mas.sg.gen सबसे\JJ.0.0.dir बड़ी\JJ.fem.sg.dir पार्टी\NC.fem.sg.dir.0 बनकर\VM.0.0.0.0.pft.0.nfn.0 उभरना\NV.dir.0 यह\PPR.0.sg.3.dir.0.n.n.prx.n सिद्ध\JJ.0.0.dir करता\VM.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n कि\CSB.n इसराइल\NP.mas.sg.obl.gen की\PP.fem.sg.gen जनता\NC.fem.sg.dir.0 खुद\PRF.0.sg.dir.0.n यह\PPR.0.sg.3.dir.0.n.n.prx.n मानने\VM.0.0.0.0.0.0.ifn.0

	लगी\VAUX.fem.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n कि\CSB.n फिलिस्तीनियों\NC.mas.pl.obl.gen की\PP.fem.sg.gen ज़मीन\NC.fem.sg.obl.loc पर\PP.0.0.loc ३८\RDX साल\NC.mas.sg.obl.abl से\PP.0.0.abl चला\VM.mas.sg.3.0.pft.0.fin.n आ\VAUX.0.0.0.0.0.0.nfn.0 रहा\VAUX.mas.sg.3.prs.pft.dcl.fin.n उसका\PPR.mas.sg.3.obl.gen.n.n.dst.n कब्जा\NC.mas.sg.dir.0 अवैध\JJ.0.0.dir और\CCD.n अनैतिक\JJ.0.0.dir है\VM.0.sg.3.prs.sim.dcl.fin.n तथा\CCD.n यह\PPR.0.sg.3.dir.0.n.n.prx.n अशांति\NC.fem.sg.dir.0 और\CCD.n युद्ध\NC.mas.sg.obl.gen का\PP.mas.sg.gen सबसे\JQ.0.sg.dir.n.nnm बड़ा\JJ.mas.sg.dir कारण\NC.mas.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ।\PU
2456	उन्होंने\PPR.0.sg.3.obl.erg.n.n.dst.y इन\DAB.pl.obl.n.prx युवाओं\NC.mas.pl.obl.gen का\PP.mas.sg.gen आह्वान\NC.mas.sg.dir.0 करते\VM.mas.sg.3.prs.impf.dcl.fin.n हुए\VAUX.0.sg.3.pst.pft.dcl.fin.n कहा\VM.mas.sg.3.pst.pft.dcl.fin.n क\CSB.n भारत\NP.mas.sg.obl.gen के\PP.0.0.gen चकत्सकों\NC.mas.sg.obl.erg ने\PP.0.0.erg पूरे\JJ.0.0.dir वव\NC.mas.sg.obl.loc में\PP.0.0.loc अपनी\PRF.fem.sg.dir.0.n वोषज्ञता\NC.fem.sg.obl.gen और\CCD.n सेवा\NC.mas.sg.obl.gen भावना\NC.fem.sg.obl.gen की\PP.fem.sg.gen धाक\VM.0.0.0.0.0.0.nfn.n जमाई\VAUX.fem.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
2524	ऐसी\DAB.sg.obl.n.prx एक\JQ.0.sg.dir.n.crd या\CCD.n अधिक\JJ.0.0.dir विधियों\NC.mas.pl.obl.gen का\PP.mas.sg.gen संयोग\NC.mas.sg.dir.0 भी\CX.n

	हो\VM.0.x.0.0.0.0.nfn.0 सकता\VAUX.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n .\PU
2551	मैंने\PPR.0.sg.1.obl.gen.n.n.prx.n अपने\PRF.0.sg.obl.gen.n मकान\NC.mas.sg.obl.acc को\PP.0.0.acc रात\NC.fem.sg.obl.gen के\PP.0.0.gen समय\NC.mas.sg.dir.0 लालसा-भरी\JJ.fem.sg.dir आंखों\NC.fem.pl.obl.abl से\PP.0.0.abl देखा\VM.mas.sg.3.pst.pft.dcl.fin.n -\RDS दरवाज़े\NC.mas.sg.obl.loc पर\PP.0.0.loc दो\JQ.0.pl.dir.n.crd लालटेनें\NC.fem.pl.dir.0 जल\VM.0.0.0.0.0.0.nfn.0 रही\VAUX.fem.sg.3.pst.pft.dcl.fin.n थीं\VAUX.fem.sg.3.pst.sim.dcl.fin.n और\CCD.n बच्चे\NC.mas.pl.dir.0 इधरउधर-\ALC.dir.0.n खेल\VM.0.0.0.0.0.0.nfn.0 रहे\VAUX.mas.pl.3.pst.pft.dcl.fin.n थे\VAUX.mas.pl.3.pst.sim.dcl.fin.n ,\PU हर\JQ.0.sg.dir.n.nnm सफ़ाई\JJ.fem.sg.dir और\CCD.n सुथरापन\NC.mas.sg.dir.0 दिखायी\NV.dir.0 देता\VM.mas.sg.3.pst.impf.dcl.fin.n था\VAUX.mas.sg.3.pst.sim.dcl.fin.n I\PU
3034	उन्होंने\PPR.0.sg.3.obl.erg.n.n.dst.y कहा\VM.mas.sg.3.pst.pft.dcl.fin.n -\PU इस\DAB.sg.obl.n.prx स्थिति\NC.fem.sg.obl.loc में\PP.0.0.loc स्थगन\NC.mas.sg.dir.0 उचित\JJ.0.0.dir नहीं\CX.y होगा\VM.mas.sg.3.fut.sim.dcl.fin.n ,\PU लेकिन\CCD.n हम\PPR.0.pl.1.dir.0.n.n.0.n याचिका\NC.fem.sg.obl.loc में\PP.0.0.loc यह\DAB.sg.dir.n.0 अनुरोध\NC.mas.sg.dir.0 जोड़\VM.0.0.0.0.0.0.nfn.0 सकते\VAUX.mas.pl.3.prs.impf.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n कि\CSB.n अर्जुन\NP.mas.sg.dir.0 तथा\CCD.n अन्य\JJ.0.0.dir राष्ट्रीय\JJ.0.0.dir

	पुरस्कारों\NC.mas.pl.obl.gen के\PP.0.pl.gen चयन\NC.mas.sg.obl.x हेतु\PP.0.0.x उचित\JJ.0.0.dir दिशानिर्देश-\NC.mas.sg.dir.0 तय\VM.0.0.0.0.0.0.nfn.0 किए\VAUX.mas.sg.3.pst.pft.dcl.fin.n जाएँ\VAUX.mas.pl.3.prs.pft.dcl.fin.n ।\PU
3040	आयोग\NC.mas.sg.obl.gen के\PP.0.0.gen वचारार्थ\JJ.0.0.dir वषय\NC.mas.sg.dir.0 इस\DAB.sg.obl.n.prx प्रकार\NC.mas.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ः-\PU
3056	आयोग\NC.mas.sg.obl.gen के\PP.0.0.gen वचारार्थ\JJ.0.0.dir वषय\NC.mas.sg.dir.0 इस\DAB.sg.obl.n.prx प्रकार\NC.mas.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ः-\PU
3131	श्री\NP.mas.sg.obl.erg फातमी\NP.mas.sg.obl.erg ने\PP.0.0.erg कहा\VM.mas.sg.3.pst.pft.dcl.fin.n क\CSB.n जब\ALC.dir.0.n तक\PP.0.0.loc वकसत\JJ.0.0.dir वव\CCD.n ,\PU वकासाील\JJ.0.0.dir दाोेंं\NC.mas.pl.obl.acc को\PP.0.0.acc सर्वाक्षा\NC.mas.sg.obl.loc में\PP.0.0.loc अधक\JJ.0.0.dir सहायता\NC.fem.sg.dir.0 प्रदान\NC.mas.sg.dir.0 करने\NV.obl.gen के\PP.0.0.gen प्रत\JJ.0.0.dir वचनबद्ध\JJ.0.0.dir नहीं\CX.y होता\VM.mas.sg.3.prs.impf.dcl.fin.n ,\PU डाकर\NP.mas.sg.obl.gen समझौते\NP.mas.sg.obl.gen के\PP.0.0.gen लक्ष्यों\NC.mas.pl.obl.acc को\PP.0.0.acc पूरी\JJ.fem.sg.dir तरह\AMN.dir.0.n प्राप्त\JJ.0.0.dir नहीं\CX.y कया\VM.mas.sg.3.pst.pft.dcl.fin.n जा\VAUX.0.0.0.0.0
3133	केन्द्रीय\NP.mas.sg.obl.erg मानव\NP.mas.sg.obl.erg संसाधन\NP.mas.sg.obl.erg वकास\NP.mas.sg.obl.erg मंत्री\NP.mas.sg.obl.erg श्री\NP.mas.sg.obl.erg अर्जुन\NP.mas.sg.obl.erg सिंह\NP.mas.sg.obl.erg ने\PP.0.0.erg भारतीय\NP.mas.sg.obl.gen राष्ट्रीय\NP.mas.sg.obl.gen सहयोग\NP.mas.sg.obl.gen

	आयोग\NP.mas.sg.obl.gen के\PP.0.0.gen अध्यक्ष\NC.mas.sg.obl.gen के\PP.0.0.gen रूप\NC.mas.sg.obl.loc में\PP.0.0.loc कार्यभार\NC.mas.sg.dir.0 ग्रहण\NC.mas.sg.dir.0 कर\VM.0.0.0.0.0.0.nfn.0 लिया\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ,\PU जबक\CSB.n श्री\NP.mas.sg.obl.acc बीबासवान.एस.\NP.mas.sg.obl.acc को\PP.0.0.acc आयोग\NP.mas.sg.obl.gen का\PP.mas.sg.gen नया\JJ.mas.sg.dir महासचव\NP.mas.sg.dir.0 नयुक्त\JJ.0.0.dir कया\VM.mas.sg.3.prs.pft.dcl.fin.n गया\VAUX.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
3196	वे\PPR.0.pl.3.dir.0.n.n.dst.y कहते\VM.0.sg.3.prs.impf.dcl.fin.y हैं\VAUX.0.sg.3.prs.sim.dcl.fin.y -\PU अभिनय\NC.mas.sg.dir.0 रचनात्मकता\NC.fem.sg.obl.soc से\PP.0.0.soc जुड़ी\JJ.fem.sg.dir विधा\NC.fem.sg.dir.0 है\VM.0.sg.3.prs.sim.dcl.fin.n ,\PU नकल\NC.mas.sg.obl.soc से\PP.0.0.soc नहीं\CX.y ।\PU
3211	इस\DAB.sg.obl.n.prx जोशीले\JJ.0.0.dir अभिनेता\NC.mas.sg.obl.bnf के\PP.0.0.gen लिए\PP.0.0.bnf शेर\NC.mas.sg.dir.0 मुलहिजा\JJ.0.0.dir है\VM.0.sg.3.prs.sim.dcl.fin.n -\PU
3236	भाव\NP.mas.sg.obl.loc प्रकाश\NP.mas.sg.obl.loc निघण्टु\NP.mas.sg.obl.loc में\PP.0.0.loc लिखा\VM.mas.sg.3.prs.pft.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n -\PU
3321	ये\PPR.0.pl.3.dir.0.n.n.prx.n यहाँ\DAB.sg.dir.n.prx के\PP.0.sg.gen बरसाती\JJ.0.sg.dir जंगल\NC.0.sg.obl.loc में\PP.0.sg.loc जमीन\NC.0.sg.obl.loc पर\PP.0.sg.loc पड़ी\JJ.0.sg.dir

	पत्तियों\NC.0.pl.obl.acc को\PP.0.sg.acc चट\VM.0.0.0.0.0.0.nfn.0 करते\VAUX.0.sg.3.prs.impf.dcl.fin.n हैं\VAUX.0.pl.3.prs.sim.dcl.fin.n ।\PU
3950	अब\ALC.dir.0.n अपने\PRF.0.sg.obl.gen.n बेटे\NC.mas.sg.dir.0 अमित\NP.mas.sg.obl.acc को\PP.0.0.acc लाँच\VM.0.0.0.0.pft.0.nfn.0 कर\VAUX.0.0.0.0.pft.0.nfn.0 रहे\VAUX.mas.sg.3.prs.pft.dcl.fin.y हैं\VAUX.0.sg.3.prs.sim.dcl.fin.y ।\PU
4002	पर्यटन\NC.0.sg.dir.0 मंत्री\NC.0.sg.dir.0 श्री\NP.mas.sg.obl.gen अनूप\NP.mas.sg.obl.gen मिश्रा\NP.mas.sg.obl.gen के\PP.0.0.gen बंगले\NC.mas.sg.obl.loc पर\PP.0.0.loc तैनात\JJ.0.0.dir सुरक्षाकर्मी\NC.mas.sg.dir.0 तब\CX.n हतप्रभ\JJ.0.sg.dir रह\VM.0.0.0.0.0.0.nfn.n गए\VAUX.mas.pl.3.pst.pft.dcl.fin.n,\PU जब\PRL.sg.dir.0.n.n.n विवाह\NC.mas.sg.obl.loc में\PP.0.0.loc शामिल\NC.0.sg.dir.0 होने\NV.obl.x के\PP.0.0.gen लिए\PP.0.0.x प्रधानमंत्री\NC.mas.sg.dir.0 श्री\NP.mas.sg.dir.0 अटलबिहारी\NP.mas.sg.obl.abl वाजपेयी\NP.mas.sg.obl.abl बंगले\NC.mas.sg.obl.abl से\PP.0.0.abl सड़क\NC.fem.sg.obl.loc तक\PP.0.0.loc पैदल\JJ.0.sg.dir ही\CX.n चले\VM.mas.sg.3.pst.pft.dcl.fin.y आए\VAUX.mas.sg.3.pst.pft.dcl.fin.y ।\PU
4096	जिनके\PRL.sg.obl.gen.n.n.y हाथों\NC.mas.pl.obl.loc में\PP.0.0.loc क्रमशः\AMN.dir.0.n धनुष\NC.mas.sg.dir.0 ,\PU गदा\NC.mas.sg.dir.0 ,\PU भाला\NC.mas.sg.dir.0 एवं\CCD.n लड्डू\NC.mas.sg.dir.0 दिखाई\NV.dir.0 देता\VM.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n ।\PU
4347	4.3\RDF सामान्य\JJ.0.0.dir परम्परा\NC.mas.sg.dir.0 यह\PPR.0.sg.3.dir.0.n.n.prx.n है\VM.0.sg.3.prs.sim.dcl.fin.n कि\CSB.n प्रोन्नति\NC.fem.sg.obl.loc पर\PP.0.0.loc

	स्थानांतरण\NC.mas.sg.dir.0 होता\VM.mas.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n .\PU कुछ\JQ.0.pl.dir.n.nnm मामलों\NC.mas.pl.obl.loc में\PP.0.0.loc इससे\PPR.mas.sg.3.obl.soc.n.n.dst.n संबंधित\JJ.0.0.dir व्यक्तियों\NC.mas.pl.obl.acc को\PP.0.0.acc कठिनाई\NC.fem.sg.dir.0 हो\VM.0.0.0.0.0.0.nfn.0 सकती\VAUX.fem.sg.3.prs.impf.dcl.fin.n है\VAUX.0.sg.3.prs.sim.dcl.fin.n .\PU अत:\CCD.n इस\DAB.sg.obl.n.prx मामले\NC.mas.sg.obl.acc को\PP.0.0.acc नियुक्ति\NC.fem.sg.dir.0 समिति\NC.fem.sg.obl.gen के\PP.0.0.gen निर्णय\NC.mas.sg.obl.loc पर\PP.0.0.loc छोड़\VM.0.0.0.0.0.0.nfn.0 देना\VAUX.0.0.0.0.0.0.ifn.0 चाहिए\VAUX.mas.sg.3.prs.sim.imp.fin.n .\PU इस\DAB.sg.obl.n.prx संदर्भ\NC.mas.sg.obl.loc में\PP.0.0.loc 'जेएजी\NC.mas.sg.dir.0 (\RDS JAG\RDF)\RDS '\PU प्रदान\NC.mas.sg.dir.0 जाने\NV.obl.acc को\PP.0.0.acc प्रोन्नति\NC.fem.sg.dir.0NV.n.0.n नहीं\CX.y माना\VM.mas.sg.3.prs.pft.dcl.fin.n जाएगा\VAUX.mas.sg.3.fut.sim.dcl.fin.n .\PU
4515	वह\PPR.sg.3.0.n.y.n.n छूटते\AMN.0.n.n ही\CX.n बोली\VM.3.prs.pft.dcl.fin.n.n.n ,\PU मम्मी\NC.n.0.sg.n ,\PU आप\PPR.sg.3.0.n.y.n.y हर\JQ.y.n.nnm काम\NC.n.0.sg.n अपने\PRF.sg.gen.n मन\NC.n.x.sg.n से\PP.x.n क्यों\PWH.n.0.n.n.n.n करती\VM.3.prs.impf.dcl.fin.n.n.n हैं\VAUX.3.prs.sim.dcl.fin.n.n.y ,\PU कम\JJ.n.n से\PP.x.n कम\JJ.n.n फोन\NC.n.0.sg.n कर\VM.0.0.0.0.nfn.n.n.n के\VAUX.0.0.pft.0.nfn.n.n.n पूछ\VM.0.0.0.0.nfn.n.n.n तो\CX.n

	लिया\VAUX.3.0.pft.dcl.fin.n.n.n होता\VAUX.3.pst.impf.dcl.fin.n.n.n ।\PU

REFERENCES

Abbi, Anvia. 1992. *The Explicator Compound Verbs: Some Definitional Issues and Criteria for Identification.* In: Indian Linguistics. Vol. 51. No.1.

Abbi, Anvita & Devi Gopalakrishnan. 1991. *Semantics of Explicator Compound Verbs in South Asian Languages.* In: Language Sciences. Tokyo. Vol. XIII NO. 2 pp. 161-180

Abbi, Anvita. 1994. *Why Explicators are not Auxiliaries? An Area between Morphology and Syntax*. In: Compound Verbs in Indian Languages. ed. Alice Davison. Iowa Univ. Press. Pp. 1-10

Abbi, Anvita. 1997. *Reduplication in South Asian Languages: An Areal, Typological and Historical Study*. Allied Publishers: Delhi.

Abbi, Anvita. 2004. *Typology of 'Manner' in Verb Sequences in South Asian Languages*. In Indian Linguistics. Vol. 65:1-4.

Abney, Steven 1991. *Parsing by Chunks*. In: Robert Berwick, Steven Abney and Carol Tenny (eds.), Principle-Based Parsing. Kluwer Academic Publishers, Dordrecht.

Abney, Steven. 1996. *Part-of-Speech Tagging and Partial Parsing*. In: Steve Young, Gerrit Bloothooft (eds). Corpus-Based Methods in Language and Speech, Kluwer Academic Publishers. Pages. 118-136.

Abney, Steven. 2008. *Semi-Supervised Learning for Computational Linguistics.* Chapman & Hall/CRC.

Adam L. Berger, Stephen Della Pietra, and Vincent J. Della Pietra. 1996. *A maximum entropy approach to natural language processing*. In: Computational Linguistics, 22(1):39–71.

Agnihotri, Rama Kant. 2006. *Hindi: An Essential Grammar*. Routledge.

Agrawal, Himanshu. 2007. *PoS Tagging and Chunking for Indian Languages*. In: Proceedings of IJCAI Workshop on Shallow Parsing for SOuth Asian Languages, Hyderabad.

Agrawal, Himanshu and Anirudh Mani. 2006. *Part of Speech Tagging and Chunk-ing with Conditional Random Fields*. In: The Proceedings of NWAI Workshop.

Allen, James. 2004. *Natural Language Understanding*. Pearson Education Pvt. Ltd.: Singapore-New Delhi.

Alsina, Alex, Joan Bresnan and Peter Sells. 1997. *Complex Predicates*. CSLI Publications:Stanford.

Ashby, William J. 1977. *Clitic Inflection in French, an historical perspective*. Rodopi:Amsterdam

Atserias, Jordi, Bernardino Casas, Elisabet Comelles, Meritxell González, Lluís Padró and Muntsa Padró. 2006. *FreeLing 1.3: Syntactic and Semantic Services in an Open- source NLP Library*. In: Proceedings of the fifth International Conference on Language Resources and Evaluation (LREC 2006), ELRA.

Avinesh PVS and Karthik Gali. 2007. *Part of Speech Tagging and Chunking using Conditional Random Fields and Transformation Based Learning*. In: Proceedings of the SPSAL workshop.Proceedings of the SPSAL workshop, IJCAI '07.

Babko-Malaya, Olga. 2005. *PropBank Annotation Guidelines.* source: <http://verbs.colorado.edu/~mpalmer/projects/ace/PBguidelines.pdf>

accessed: 1 January, 2010

Bahl, Kalicharan.1967. *A Reference Grammar of Hindi.* Chicago: South Asia Language and Area Center, Univ. of Chicago.

Bailey, T. Grahame. 1950. *Teach Yourself Hindustani.* English Universities Press: London.

Bali, Kalika, Monojit Choudhury, Priyanka Biswas, Girish Nath Jha, Narayan Choudhary & Maansi Sharma. (2010). *Indian Language Part-of-Speech Tagset: Hindi*. Linguistic Data Consortium, Philadelphia. ISBN: 1-58563-571-5

Barker, Abd-al-Rahman Muhammad. 1967. *A Course in Urdu.* Vols. I, II and III. Montreal:

Baskaran, Sankaran, Kalika Bali, Monojit Choudhury, Tanmoy Bhattacharya, Pushpak Bhattacharyya, Girish Nath Jha, S. Rajendran, K. Saravanan, L. Sobha and K.V. Subbarao. 2008. *A Common Parts-of-Speech Tagset Framework for Indian Languages*. In: Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odjik, Stelios Piperidis, Daniel Tapias (Eds.) Proceedings of the Sixth International Language Resources and Evaluation (LREC'08), Marrakech, Morocco.

Begum, Rafiya, Karan Jindal, Ashish Jain, Samar Husain & Dipti M. Sharma. 2011. *Identification of conjunct verbs in hindi and its effect on parsing accuracy*. In: Proceedings of the 12th international conference on Computational linguistics and intelligent text processing - Volume Part I, Springer-Verlag Berlin, Heidelberg

Bharati, Akshar, Dipti Misra Sharma, Lakshmi Bai, Rajeev Sangal. 2006. Anncorra: Annotating Corpora. Language Technologies Research Centre, IIIT, Hyderabad

Source: <http://ltrc.iiit.ac.in/tr031/posguidelines.pdf> accessed: 2 February, 2010

Bharati, Akshar, Vineet Chaitanya, Amba P. Kulkarni, Rajeev Sangal. 2001. Anusaaraka: Overcoming the Language Barrier in India. In: Rukmini Bhaya Nair, (ed.) Anuvad: Approaches to Translation, Sage, New Delhi.

Bharati, Akshar, Vineet Chaitanya, Rajeev Sangal. 1995. *Natural Language Processing: A Paninian Perspective*. Prentice-Hall of India, New Delhi.

Bhatia, Tej K. 1987. *A History of Hindi Grammatical Tradition: Hindi Hindustani Grammar, Grammarians and Problems*. E.J. Brill: New York

Bhatt, Rajesh, Bhuvana Narasimhan, Martha Palmer, Owen Rambow, Dipti Misra Sharma, Fei Xia. 2009. *A multi-representational and multi-layered treebank for Hindi/Urdu.* In: Proceedings of the Third Linguistic Annotation Workshop. pp.186-189. Association for Computational Linguistics Morristown, NJ, USA

Bijalwan, C.D. 1897. *Indian Theory of Knowledge*. Heritage Publishers: New Delhi.

Bresnan, Joan (ed.). 1982. The Mental Representation of Grammatical Relations. MIT Press: Cambridge, MA.

Brill, Eric. 1995. *Transformation-based error driven learning and natural language processing: A case study in part-of-speech tagging.* In: Computational Linguistics 21, no. 4 (1995): 543-566

Brill, Eric, David Magerman, Mitchel Marcus, Beatrice Santorini. 1990. *Deducing linguistic structure from the statistics of large corpora.* In Proceedings of the workshop on Speech and Natural Language . Pages 275-282

Brill, Eric. 1992. *A Simple Rule Based Part of Speech Tagger*. In: Proceedings of the Third Conference on Applied Natural Language Processing (1992), pp. 152-155.

Brill, Eric. 1995. *Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part of Speech Tagging.* In Computational Linguistics, December. Vol. 21. pp. 543-565

Brown, P. et al. 1988. *A statistical approach to language translation.* In: Coling 88, 71-76. Also: Computational Linguistics 16 (1990), 79-85.

Burnard, Lou and Syd Bauman (eds). 2008. *TEI P5: Guidelines for Electronic Text Encoding and Exchange by the TIE Consortium*. The TEI Consortium. Oxford — Providence — Charlottesville — Nancy.

Burton-Page, John. 1957. *Compound and Conjunct Verbs in Hindi*. BSOAS 19, 469-78.

Butt, Miriam. 2003. *The Light Verb Jungle*. In: G. Aygen, C. Bowern, and C. Quinn (eds.) Harvard Working Papers in Linguistics. Volume 9: Papers from the GSAS/Dudley House Workshop on Light Verbs, 1–49.

Bybee, Joan & Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar – Tense, Aspect, and Modality in the Languages of the World.* The University of Chicago Press: Chicago.

Calder, J. 1989. *Paradigmatic Morphology*. In: Proceedings of the 4th European Chapter of the Association for Computational Linguistic. (ACL 1989) Manchester pp.58-65.

Carrol, John. 2003. *Parsing*. In Ruslan Mitkov (ed). The Oxford Handbook of Computaional Linguistics. Oxford: New York pp.233-248.

Catell, R. 1984. *Complex Predicates in English*. Academic Press Australia, North Ryde, New South Wales.

Chakrabarty, Debasri, Vaijayanthi Sarma and Pushpak Bhattacharyya. 2007. *Complex Predicates in Indian Language Wordnets*. In Lexical Resources and Evaluation Journal. Forthcoming. < http://www.cse.iitb.ac.in/~pb/papers/Debasri-Complex-Predicates-26mar07.pdf>

Chakrabarti, D., Mandalia, H., Priya, R., Sarma, V., Bhattacharyya, P.2008. *Hindi Compound Verbs and their Automatic Extraction*. In: International Conference on Computational Linguistics, pp. 27-30

Chanod, Jean-Pierre and Pasi Tapanainen. 1995. *Tagging French: Comparing a Statistical and a Constraint-based Method*. In Proceedings of the Seventh Cconference on European Chapter of the Association for Computational Linguistics, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.

Chattopadhyay, Santanu. 2005. *Compiler Design*. Prentice-Hall of India: New Delhi.

Chomsky, N. 1956. *Three Models for the Description of Language*. I.R.E. Transactions on Information Theory, vol. IT-2, no. 3: 113-24.

Chomsky, N. 1981. *Lectures on Government and Binding*. Foris, Dordrecht. Reprint. 7th Edition. Berlin and New York: Mouton de Gruyter, 1993.

Chomsky, N. and H. Lasnik. 1993. *The Theory of Principles and Parameters*. In J. Jacobs et al. (eds.) Syntax: An International Handbook of Contemporary Research, Vol. 1. Walter de Gruyter, pp. 506-569. (Reprinted in N. Chomsky, The Minimalist Program. MIT Press, 1995)

Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English.* Harper and Row: New York.

Chomsky, Noam. 1957. *Syntactic Structures.* Mouton: Hague.

Church, Kenneth Ward. 1988. A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text. In: *Proceedings of the second conference on Applied Natural Language Processing*. Austin, Texas . Pages: 136 - 143

Clark, P. and T. Niblett 1989. *The CN2 Induction Algorithm*. In: Journal of Machine Learning, vol(3), pages 261-283, 1989

Cutting, D., Kupiec J, Pedersen J, Sibun P. 1992. *A Practical Part-of-Speech Tagger*. In: Cutting, D. (eds.) Proceedings of the Third Conference on Applied Natural Language Processing.

Dahl, Östen. 1985. *Tense and Aspect Systems*. Basil Blackwell Ltd: Oxford, UK.

Dalal, Aniket, Kumar Nagaraj, Uma Sawant, Sandeep Shelke. 2006. *Hindi Part-of-Speech Tagging and Chunking : A Maximum Entropy Approach*. In: Proceedings of the NLPAI Machine Learning Contest 2006, NLPAI, 2006.

DeRose, Steven J. 1988. *Grammatical Category Disambiguation by Statistical Optimization.* In: Computational Linguistics. Volume 14 , Issue 1. Pages: 31 - 39

Dorr, B. 1992. *Parameterization of the Interlingua in Machine Translation*. In: Coling 92 (2), 624- 630.

Dorr, B. J. 1990. *A Cross-Linguistic Approach to Translation.* In: TMI-90, 13-32.

Dymsic, Z.M. 1966. *Hindi Vyakaran ki Ruprekhaa*. Delhi: Rajkamal Prakashan.

Emeneau, M. B. 1956. *India as a Linguistic Area*. Language vol.32 1:3-16

Farwell, D. & Wilks, Y. 1991. *ULTRA: a multilingual machine translator*. In: MT Summit 3, 19-24.

Feldman, Anna and Jirka Hana. 2010. *A Resource Light Approach to Morpho-Syntactic Tagging*. Rodopi, Amsterdam-New York.

Fellbaum, Christiane (ed.). 1998. *WordNet: An Electronic Lexical Database*. MIT Press: Cambridge, USA.

Fraser, B. 1976. *The Verb-Particle Combination in English*. Mouton: Hague.

Furuse, O. & Iida, H. 1992. *An Example-based Method for Transfer-driven Machine Translation*. In: TMI-92, 139-150.

Furuse, O. & Iida, H. 1992a. *Transfer-driven Machine Translation*. In: FGNLP, 95-111.

Garnham, Allan. 1988. *Artificial Intelligence: An Introduction.* Routledge and Kegan Paul: London, NY.

Garside, Roger, Geoffrey Leech and Geoffrey Sampson (eds.). 1987. *The Computational Analysis of English: A Corpus-Based Approach*. Longman.

Gazdar, G., Klein, E., Pullum, G. K., and Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford.

Greene, B. and Rubin, G. M. 1971. *Automatic Grammatical Tagging of English*, Providence, R. I.: Department of Linguistics, Brown University.

Grinberg, Dennis, John Lafferty and Daniel Sleator. 1995. *A Robust Parsing Algorithm for Link Grammars*. Carnegie Mellon University Computer Science technical report CMU-CS-95-125, and Proceedings of the Fourth International Workshop on Parsing Technologies, Prague, September, 1995

Grishman, Ralph. 1986. *Computational Linguistics: An Introduction*. Cambridge University Press: New York.

Gross, Maurice. 1972. *Mathemetical Models in Linguistics*. Prentice Hall, Inc.: New Jersey.

Grosz, Barbara J. et al. (eds.). 1990. *Readings in Natural Language Processing*. Morgan- -Kaufmann: Los Altos, CA.

Gupta, Deepa. 2005. *Contributions to English to Hindi Machine Translation Using Example-Based Approach*. Unpublished Ph.D. Thesis submitted at Indian Institute of Technology, Delhi.

Guru, Kamataprasad. 1978. *Hindi Vyaakaran*. Nagari Pracharini Sabha: Kashi.

Hacker, Paul. 1958. *Zur Franktion eineger Hilfsverben im Modern Hindi*. Mainz: Academie der Wissenschaften und der Literatur.

Hardie, Andrew. 2003. *Developing a tag-set for automated part-of-speech tagging in Urdu*. In: Archer, D, Rayson, P, Wilson, A, and McEnery, T (eds.) Proceedings of the Corpus Linguistics 2003 conference. UCREL Technical Papers Volume 16. Department of Linguistics, Lancaster University, UK.

Hardie, Andrew. 2004. *The Computational Analysis of Morphosyntactic Categories in Urdu.* PhD thesis submitted to Lancaster University.

Hook, Peter E. 1974. *The Compund Verb in Hindi*. The Michigan Series in South and South East Asian Languages and Lingusitics: The University of Michigan.

Hutchins, John. 1993. *Latest developments in machine translation technology: beginning a new era in MT research*. In: The Fourth Machine Translation Summit: MT Summit IV. Proceedings: International cooperation for global communication, July 20-22, 1993, Kobe, Japan. [Tokyo: AAMT, 1993], p.11-34.

Johnsbaugh, Richard and Marcus Schaefer. 2004. *Algorithms*. Pearson Education Pvt. Ltd: Singapore-New Delhi.

Joshi, Aravind K., Yves Schabes 1992. *Tree-adjoining Grammars and Lexicalized Grammars*. Tree Automata and Languages pp.409-432

Joshi, Arvind K. 1985. *Tree Adjoining Grammars: how much context sensitivity is required to reasonable structural descriptions*?. In Dowty, D. R. Kartunnen, L., and Zwicky, A. (eds.) Natural Language Parsing. pp. 206-250. CUP: Cambridge.

Jurafsky, Daniel and James H. Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing*, Computational Linguistics and Speech Recognition. Englewood Cliffs, Prentice Hall: NJ.

Kachru, Yamuna. 1965. *Transformational Treatment of Hindi Verbal Syntax*. London: Univ. of London Dissertation.

Kachru, Yamuna. 1966. *An Introduction to Hindi Syntax*. Urbana, Illinois: Deptt. Of Linguistics, Univ. of Illinois.

Kachru, Yamuna. 2006. *Hindi*. John Benjamins: Amsterdam/Philadelphia.

Kaji, H. et al. 1992. *Learning Translation Templates from Bilingual Text*. In: Coling 92 (2), 672- 678.

Kaplan, R. M. and Bresnan, J. (1982). *Lexical-functional grammar: A formal system for grammatical representation*. In Bresnan, J. (Ed.), The Mental Representation of Grammatical Relations, pp. 173-281. MIT Press, Cambridge, MA.

Karlsson, Fred, Atro Voutilainen, Juha Heikkilä and Arto Anttila, (eds.). 1995. *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Berlin, Mouton de Gruyter.

Karlsson, Fred. 1990. *Constraint Grammar as a Framework for Parsing Running Text*. In: Proceedings of the 13th Conference on Computational Linguistics - Volume 3. ACL, Morristown, NJ, USA.

Kay, M. (1984). *Functional Unification Grammar: A formalism for machine translation*. In: COLING-84, Stanford, CA, pp. 75-78.

Kay, Martin. 2003. *Introduction*. In: The Oxford Handbook of Computaional Linguistics. Oxford: New York pp: xvii-xx.

Kearns, K. 1988. *Light Verbs in English*. MIT: Mass.

Kelogg, S.H. 1938. *A Grammar of the Hindi Language*. (3rd ed.). London: Routledge and Kegan Paul.

Kim, Su Nam and Timothy Baldwin. 2006. *Automatic Identification of English Verb Particle Constructions using Linguistic Features*. In: Proceedings of the Third ACL-SIGSEM Workshop on Prepositions. ACL: Trento, Italy. pp: 65-72.

Kingsbury, P., Palmer, M., and Marcus, M. 2002. *Adding Semantic Annotation to the Penn TreeBank*. In: Proceedings of the Human Language Technology ConferenceHLT-2002 (San Diego, California).

Kingsbury, Paul and Martha Palmer. 2003. *PropBank: the Next Level of TreeBank*. In: Proceedings of Tree-banks and Lexical Theories.

Klein, Sheldon & Robert F. Simmons. 1963. *A Computational Approach to Grammatical Coding of English Words*. In Journal of the ACM (JACM). Volume 10 , Issue 3. pages: 334 - 347

Koskenniemi, Kimmo, Pasi Tapanainen, Atro Voutilainen. 1992. *Compiling and Using Finite-State Syntactic Rules*. COLING. Pp. 156-162

Leech, G. and Wilson, A. 1996. *Recommendations for the Morphosyntactic Annotation of Corpora*. EAGLES Report EAG-TCWG-MAC/R.

Leech, G. and Wilson, A (1999) *Standards for Tagsets*. (Edited version of EAGLES Recommendations for the Morphosyntactic Annotation of Corpora (1996): available on the internet at http://www.ilc.cnr.it/EAGLES96/browse.html .) In: van Halteren, H (ed.) (1999)

Leech, G and Wilson, A. 1999. *Standards for Tag-sets*. In: Hans van Halte-ren (eds.) Syntactic Wordclass Tagging, Dordrecht: Kluwer Academic.

Leech, G. 1997. *Grammatical Tagging*. In: Corpus An-notation: Linguistic Information from Computer Text Corpora, ed: Garside, Leech and McEnery, London: Longman

Levin, B.1993. *English Verb Classes and Alternation: A Preliminary Investigation*. The University of Chicago Press.

Malmkjaer, Kirsten and James M. Anderson (eds.). 1991. The Linguistics Encyclopaedia. Routledge: London and New York.

Manning, C. D. and H. Schutze. 2002. *Foundation of statistical Natural Language Processing.* MIT Press.

Marcus, M., et.al. 1993. *Building a Large Annotated Corpus of English: The Penn Treebank.* In: Computational Linguistics. Vol. 19. Issue 2. pp. 313-330.

Masica (mimeo). *Verb Sequences and Verbal System in Hindi-Urdu*. Chicago: South Asia Language and Area Centre, Univ. Chicago. Undated

Masica, Colin P. 1976. *Defining a Linguistic Area: South Asia*. The Univ. of Chicago Press:Chicago and London.

McCarthy, Diana, Bill Keller, and John Carroll. 2003. *Detecting a continuum of compositionality in phrasal verbs*. In Proc. of the ACL-2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment. Sapporo, Japan.

McEnery, A. M., P. Baker, R. Gaizauskas, and H. Cunningham. 2000. *EMILLE: Building a Corpus of South Asian Languages*. In: Vivek: A Quarterly in Artiificial Intelligence. 13(3): pp: 23–32.

McGregor, R.S. 1972. *Outline of Hindi Grammar*. Oxford: Clarendon Press.

Melcuk, Igor A. 1987. *Dependency Syntax : Theory and Practice,* SUNY Series in Linguistics. State University of New York Press.

Merialdo, Bernard. 1994. *Tagging English Text with a Probabilistic Model*. In: Computational Linguistics. Volume 20 , Issue 2 . MIT Press Cambridge. Pages: 155 - 171

Miller, G. 1995. *Wordnet: A lexical database.* Communication of the ACM 38, 11, 39–41.

Mitchell, Tom M. 1997. *Machine Learning*. MacGraw-Hill

Montague, R. 1970. *Universal Grammar*. Theoria. 36: 373-398.

Montague, R. 1970a. *English as a Formal Language*. In: B Visentini et al. (ed.), Linguaggi nella Società e nella Tecnica. Edizioni di Comunità, Milan, pp. 189-224.

Montague, R. 1973. *The proper treatment of quantification in ordinary English.* In: K J J Hintikka, J M E Moravcsik, P Suppes (eds.), Approaches to Natural Language. Reidel, Dordrecht, pp. 221-242.

Montague, R. 1974. *Formal Philosophy*. In: Selected Papers of Richard Montague. Yale University Press, Edited and with an introduction by Richmond H. Thomason. New Haven/London.

Mukerjee, Amitabha, Ankit Soni & Achla M. Raina. 2006. *Detecting Complex Predicates in Hindi using POS Projection across Parallel Corpora*. In Proceedings of the 5th Workshop on Important Unresolved Matters. Sydney, July 2006. pp:. 11–18.

Nagao, M. 1984. *A framework of a mechanical translation between Japanese and English by analogy principle*. In: Elithorn, A. & Banerji, R.(eds.) Artificial and human intelligence (Amsterdam: North-Holland, 1984)

Nirenburg, Sergei (ed.). 1987. *Machine Translation: Theoretical and Methodological Issues*. CUP: Cambridge.

Okumura, A. et al. 1991. *Multi-lingual sentence generation from the PIVOT interlingua*. In: MT Summit 3, 67-71.

Pahwa, Munshi Thakardass. 1936. *The Pucca Munshi*. Peshawar Cantonment: Faqir Chand Marwah.

Pammi, Sathish Chandra and Kishore Prahallad. 2007. *POS Tagging and Chunking using Decision Forests*. In: Proceedings of Workshop on Shallow Parsing in South Asian Languages at IJCAI, Hyderabad, India.

Paul, Soma. 2003. *Composition of Compund Verbs in Bangla*. In: Proceedings of the Workshop on Multi-Verb Constructions. Trondheim Summer School.

Paul, Soma. 2006. *An HPSG Account of Bangla Compound Verbs with LKB Implementation*. Unpublished PhD thesis submitted to University of Hyderabad.

Pollard, C. and I.A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. CSLI Publications, Stanford, CA.

Porizka. 1963. *Hindistina: Hindi Language Course*. Prague: State Pedagogical Publishing House.

Pray, Bruce R. 1970. *Topics in Hindi-Urdu Grammar*. Berkeley, California: Centre for South and South East Asian Studies, Univ. of California.

Rabiner. L. R. 1989. *A tutorial on hidden Markov models and selected applications in speech recognition*. Proceedings of the IEEE, 77(2):257–286

Ratnaparkhi, Adwait. 1996. *A maximum entropy model for part-of-speech tagging*. In Eric Brill and Kenneth Church, editors, Proceedings of the Conference on Empirical Methods in Natural Language Processing, pages 133–142. Association for Computational Linguistics, Somerset, New Jersey.

Ratnaparkhi, Adwait. 1997. *A simple introduction to maximum entropy models for natural language processing*. Technical Report 97-08, Institute for Research in Cognitive Science, University of Pennsylvania, May.

Ray, Pradipta Ranjan, Harish V. Sudeshna Sarkar and Anupam Basu., 2003. *Part of Speech Tagging and Local Word Grouping Techniques for Natural Language Parsing in Hindi.* In: Proceedings of (ICON) 2003.

Ringland, G.A. and Duce, D.A. 1988. *Approaches to Knowledge Representation: An Introduction*. Research Studies Press Ltd: Letchworth, Hertfordshire, England.

Ritchie, G. D., G. J. Russel Black and S. G. Pulman. 1992. *Computational Morphology*. Cambridge, Mass.:MIT Press.

Rocher, Ludo. 1958. *Leerboek van het Moderne Hindi*. Gent: Oosters-Afrikaans Instituut bij de Rijksuniversiteit te Gent.

Roland, Douglas and Daniel Jurafsky. 1998. *Verb Sense and Verb Subcategorisation Probabilities*. In: Stevension, Suzanne and Paola Merlo (eds.) In Proceedings of the CUNY Sentence Processing Conference. Benjamins.

Ruppenhofer, J. et. al. 2006. *FrameNet II: Extended Theory and Practice*. http://framenet.icsi.berkeley.edu/book/book.pdf accessed: 15 September, 2010

Sadler, L. & Arnold, D. 1992. *Unification and Machine Translation*. Meta 37(4), 657-680.

Sadler, V. 1989. *Working with Analogical Semantics: Disambiguation Techniques in DLT*. Dordrecht: Foris.

Sag, Ivan A., Thomas Wasow, Emily Bender. 2003. *Syntactic Theory: A formal Introduction*. 2nd Edition. Chicago: University of Chicago Press.

Sag, Ivan, Timothy Baldwin, Francis Bond, Ann Copestake and Dan Flickinger. 2002. *Multiword expressions: A pain in the neck for NLP*. In: Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics. (CICLing-2002). p1- 15.

Samuelsson, Christer and Atro Voutilainen. 1997. *Comparing a Linguistic and a Stochastic Tagger*. In: Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics. ACL. Pages 246-253.

Samuelsson, Christer, Pasi Tapanainen, Atro Voutilainen. 1996. *Inducing Constraint Grammars*. ICGI pp.146-155

Santorini, Beatrice. 1990. *Part-of-speech Tagging Guidelines for the Penn Treebank Project*. Technical report MS-CIS-90-47, Department of Computer and Information Science, University of Pennsylvania.

Sharma, Aryendra. 1958. *A Basic Grammar of Modern Hindi*. Delhi: Ministry of Education and Scientific Research, Govt. Of India.

Sharma, Dipti M. et.al. 2002. *AnnCorra : TreeBanks for Indian Languages (version – 1.9)*. In: Proceedings of the 3rd workshop on Asian language resources and international standardization - Volume 12 . Pp. 1-8

Sharp, R. & Streiter, O. 1992. *Simplifying the complexity of machine translation*. Meta 37(4), 681-692

Shrivastava, Manish, R. Melz, S. Singh, K. Gupta and P. Bhattacharyya. 2006. *Conditional Random Field Based POS Tagger for Hindi*. In: Proceedings of the MSPIL, pp. 63-68.

Shrivastava, Manish, Nitin Agrawal, Bibhuti Mohapatra, Smriti Singh & Pushpak Bhattacharya. 2005. *Morphology Based Natural Language Processing Tools for Indian Languages*. Paper presented at 4th Annual Inter Research Institute Student Seminar in Computer Science, Indian Institute of Technology, Kanpur. April 1-2, 2005. <http://www.cse.iitk.ac.in/users/iriss05/presentation_m_shrivastava.pdf>

Simon, H. F. 1958. *Some Remarks on the Structure of the Verb Complex in Standard Chinese*. In Bulletin of the School of Oriental and African Studies. University of London, Vol. 21, No. 1/3. pp. 553-577.

Singh, Akshay. S M Bendre, Rajeev Sangal. 2006. *HMM Based Chunker for Hindi*. In: Proceedings of IJCNLP-05: The Second International Joint Conference on Natural Language Processing, 11-13 October, 2005, Jeju Island, Republic of Korea.

Singh, Rajendra and R. K. Agnihotri. 1997. *Hindi Morphology: A Word Based Description*. MLBD: New Delhi.

Singh, Smriti, Kuhoo Gupta, Manish Shrivastava, and Pushpak Bhattacharyya. 2006. *Morphological richness offsets resource demand – experiences in constructing a pos tagger for hindi*. In: Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions, pages 779–786, Sydney, Australia, July. Association for Computational Linguistics.

Sinha, R. M. K. & Anil Thakur. 2005. *Dealing with replicative words in Hindi for machine translation to English*. In: The Proceedings of the Tenth Machinde Translation Summit. Phuket, Thailand, September 13-15, 2005. pp.157-164

Southworth, Franklin C. 1971. *The Students Hindi-Urdu Reference Manual*. Tucson, Arizona: Univ. of Arizona Press.

Sowa, J. F. 1984. *Conceptual Structures: Information Processing in Mind and Machine*. Addision-Wesley: Reading, Massachusetts, California, London.

Spencer, A. and A. Zwicky (eds.). 1998. *The Handbook of Morphology*. Blackwell: Oxford.

Sproat, R.W. 1992. *Morphology and Computation*. MIT Press: Cambridge, Mass.

Stella Markantonatou, Sokratis Sofianopoulos, Vassiliki Spilioti, Yiorgos Tambouratzis, Marina Vassiliou, Olga Yannoutsou, & Nikos Ioannou. 2005. *Monolingual corpus- based MT using chunks*. MT Summit X, Phuket, Thailand, September 16, 2005, Proceedings of Second Workshop on Example-Based Machine Translation. pp.91-98.

Stevenson, Suzane and Afsaney Fazly and Ryan North. 2004. *Statistical Measures of the Semi-Productivity of Light Verb Constructions*. In: Second ACL Workshop on Multiword Expressions: Integrating Processing.

Sumita, E. et al. 1990. *Translating with examples: a new approach to machine translation*. In: TMI-90, 203-212.

Tapanainen, Pasi and Atro Voutilainen. 1994. *Tagging accurately - Don't Guess if You Know*. In: Proceedings of ANLP pp. 47-52

Tapanainen, Pasi and Timo Järvinen. 1997. *A Non-Projective Dependency Parser*. In: Proceedings of the 5th Conference on Applied Natural Language Processing . Pages 64--71

Timothy Baldwin and Aline Villavicencio. 2002. *Extracting the unextractable: A case study on verb-particles*. In: Proceedings of the 6th Conference on Natural Language Learning (CoNLL-2002). Taipei, Taiwan. Pp.: 98–104,

Timothy Baldwin, Colin Bannard, Takaaki Tanaka, and Dominic Widdows. 2003. *An empirical model of multiword expression decomposability*. In: MWE '03 Proceedings of the ACL 2003 workshop on Multiword expressions: analysis, acquisition and treatment - Volume 18

Timothy Baldwin. 2005. *The deep lexical acquisition of English verb-particle constructions*. In: Computer Speech and Language, Special Issue on Multiword Expressions. 19(4):398–414.

Toutanova, Kristina and Christopher D. Manning. 2000. *Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger*. In: Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000), pp. 63-70. Hong Kong.

<http://nlp.stanford.edu/~manning/papers/emnlp2000.pdf> accessed: 12 June, 2010

Toutanova, Kristina and Christopher D. Manning. 2000. *Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger*. In: Proceedings of the Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora (EMNLP/VLC-2000), pp. 63-70.

Trost, Harald. 2003. *Morphology*. In: The Oxford Handbook of Computaional Linguistics. Oxford: New York pp.25-47.

Utsuro, T., Yuji Matsumoto and Makoto Nagao. 1992. *Lexical knowledge acquisition from bilingual corpora*. In: Coling 92 (2), 581-587.

Vajpeyi, Kishoridas. 1979. *Hindi Shabdanushasan*. Nagari Pracharini Sabha: Kashi. 5th edition.

Vale, Ramchandra Narayan. 1948. *Verbal Composition in Indo-Aryan*. Poona: Deccan College.

van Haltener, Hans (ed.) 1999. *Syntactic Wordclass Tagging*. Springer/Kluwer Academic Publishers, Netherlands.

Van Olphen, Hermen Hendrik. 1970. *The Structure of the Hindi Verb Phrase.* Austin, Texas: Univ. of Texas Dissertation.

Verma, M. K. 1993. *Complex Predicates and Light Verbs in Hindi*. In M. K. Verma ed. (1993) Complex Predicates in South Asian Languages. Manohar Publishers and Distributors, New Delhi.

Voutilainen, Atro and Juha Heikkilä. 1993. *An English Constraint Grammar (ENGCG) a surface-syntactic parser of English*. In Udo Fries and Gunnel Tottie and Peter Schneider (eds). Creating and Using English Language Corpora. Rodopi: Amsterdam and Atlanta. Pages 189-199.

Voutilainen, Atro and Lluís Padró. 1997. *Developing a Hybrid NP Parser*. ANLP pp.80-87

Voutilainen, Atro. 1999. *An Experiment on the Upper Bound of Interjudge Agreement: the Case of Tagging*. EACL pp. 204-208

Voutilainen, Atro. 2003. *Part-of-Speech Tagging*. In: Ruslan Mitkov (ed). The Oxford Handbook of Computaional Linguistics. Oxford: New York pp.219-232.

Watanabe, Taro, Eiichiro Sumita, & Hiroshi G. Okuno. 2003. *Chunk-based statistical translation*. In: ACL-2003: 41st Annual meeting of the Association for Computational Linguistics. Sapporo, Japan. July 7-12, 2003.

Web Resources

Anna University.

<http://www.au-kbc.org/frameresearch.html> accessed 10 Sept, 2007

Anubharati at IIT Kanpur.

<http://www.cse.iitk.ac.in/users/langtech/anubharti.htm> accessed: 11 Sept, 2010

Anusaaraka at IIIT Hyderabad.

<http://www.iiit.net/ltrc/Anusaaraka/anu_home.html> (accessed: June 10, 2007).

Comrie, Bernard. In an article entry in Encarta Encyclopaedia. <http://www2.ignatius.edu/faculty/turner/languages.htm> accessed: 26 September, 2007

LTRC, IIIT Hyderabad. Morphological Analysers.

<http://ltrc.iiit.net/showfile.php?filename=onlineServices/morph/index.htm> accessed: 29 September, 2007

IIIT-tagset. A Parts-of-Speech tagset for Indian languages.

< http://shiva.iiit.ac.in/SPSAL2007/iiit_tagset_guidelines.pdf>

accessed: 28 September, 2007

IIT Bombay. Search Hindi Corpora.

<www.cfilt.iitb.ac.in/~corpus/hindi> accessed: 4 October, 2007

RCILTS, IIT Guwahati.

<http://www.iitg.ernet.in/rcilts> 15 September, 2007

Jain, Ajai. 1999. Directions. IIT Kanpur. Vol.2. No.6. November, 1999.

<http://www.iitk.ac.in/infocell/Archive/dirnov2/techno_machine.html>

accessed 12, September, 2007.

Machine Translation Archive.

<www.mt-archive.info> Source of several papers in this area.

Media Lab Asia, Kharagpur.

<http://www.mla.iitkgp.ernet.in/Resource/index.html> accessed: 1 September, 2007

Media Lab Asia, Kharagpur.

<http://www.mla.iitkgp.ernet.in/technology.html> accessed: 1 September, 2007

Punjabi University.

<http://www.advancedcentrepunjabi.org/punjabi_mor_ana.asp>

accessed:1 September, 2007

Rajendran, S. 2006. Parsing in Tamil: Present State of Art

<http://www.languageinindia.com/aug2006/parsingrajendran.pdf> accessed 23 July, 2007.

RCILTS, IIT Bombay. <http://www.cfilt.iitb.ac.in/>accessed: 17 August, 2010

Renganathan, Vasu. Univ. of Pennsyslvania.

<http://ccat.sas.upenn.edu/plc/tamilweb/hindi.html> Accessed 17 July, 2007

TDIL, Desika.

<http://tdil.mit.gov.in/download/Desika.htm> (accessed: June 02, 2007)

TDIL, RCILTS, Anna University. <http://ns.annauniv.edu> (accessed 15 September, 2006)

TDIL, RCILTS, JNU. <http://rcilts.jnu.ac.in> (Accessed: 17 April 2006).

TDIL. <http://tdil.mit.gov.in/mat/ach-mat.htm> accessed: 10, November 2005

The Language Group at IIT Kanpur. <http://www.cse.iitk.ac.in/users/language> 22 Sept, 2007

Utkal University. <http://www.ilts-utkal.org/oriyamorphological.htm> 22 Sept, 2007