

**A COMPUTATIONAL ANALYSIS OF NEPALI MORPHOLOGY:
A MODEL FOR NATURAL LANGUAGE PROCESSING**

A Dissertation

Submitted to the Faculty of Humanities and Social Sciences of
Tribhuvan University in Fulfillment of the Requirements for the Degree
of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

By

BALARAM PRASAIN

**Ph.D. Reg. No.: 19-2058 Magh
TU Reg. No.: 288-83**

March 2011

Recommendation Letter

We hereby recommend that this dissertation, entitled **A COMPUTATIONAL ANALYSIS OF NEPALI MORPHOLOGY: A MODEL FOR NATURAL LANGUAGE PROCESSING** prepared by Mr. **Balaram Prasain** under our supervision and guidance be accepted by the research committee for the final examination in fulfillment of the requirement for the degree of doctor of philosophy in linguistics.



Professor Dr. Yogendra Prasad Yadava
Supervisor
Central Department of Linguistics
Tribhuvan University

DISSERTATION COMMITTEE



Professor Dr. Miriam Butt
Expert
Department of Linguistics
Konstanz University, Germany



TRIBHUVAN UNIVERSITY
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
OFFICE OF THE DEAN

Ref. No.

APPROVAL LETTER

This dissertation entitled *A Computational Analysis of Nepali Morphology: A Model For Natural Language Processing* was submitted by Mr. *Balaram Prasain* for final examination by the Research Committee of the Faculty of Humanities and Social Sciences, Tribhuvan University, in fulfillment of the requirements for the Degree of Doctor of Philosophy in *Linguistics*. I hereby certify that the Research Committee of the Faculty has found this dissertation satisfactory in scope and quality and has therefore accepted it for the degree.

Nav R. Kanel

Prof. Nav Raj Kanel, Ph.D.
Dean and Chairman
Research Committee

Date: *March 21, 2012*



DECLARATION

I hereby declare that this dissertation entitled " A COMPUTATIONAL ANALYSIS OF NEPALI MORPHOLOGY: A MODEL FOR NATURAL LANGUAGE PROCESSING" submitted to the office of the Dean, Faculty of Humanities and Social Sciences, Tribhuvan University, is an entirely original work prepared under the supervision of my supervisors. I have made due acknowledgements to all ideas and information borrowed from different sources in the course of writing this dissertation. The result presented in this dissertation have not been presented or submitted anywhere else for the award of any degree or for any other purposes. No part of the contents of this dissertation has ever been published in any form before. I shall be solely responsible if any evidence is found against my dissertation.



Balaram Prasain
Tribhuvan University

Date: 2012/03/18

ACKNOWLEDGEMENTS

My profound indebtedness is due to my supervisor prof. Dr. Yogendra Prasad Yadava, the former head, Central Department of Linguistics, Tribhuvan University, Nepal for his insistent encouragement, continuous guidance, valuable suggestions and insightful comments in accomplishing this dissertation.

I would like to express my sincere gratitude to Professor Dr. Miriam Butt, Department of Linguistics, Konstanz University, Germany for her constructive suggestions, proper guidance, insightful comments to improve this dissertation.

I owe a great deal to Dr. Andrew Hardie, Lancaster University, for his valuable suggestions and providing his articles which help me understand the basic concepts and also for helping in using online NNC corpus.

I would like to extend my thanks to Dr. Dan Raj Regmi, head of the Central Department of Linguistics, Tribhuvan University, for his encouragement, useful suggestions and comments provided to my study.

I would like to extend my sincere gratitude to Prof. Madhav Prasad Pokharel, Central Department of Linguistics, Tribhuvan University, for his prompt answer to any queries regarding Nepali morphology and its structure.

I would like to express my gratitude to Prof. Dr. Chuda Mani Bandhu and Prof. Dr. Tej R. Kansakar, Former heads, Central Department of Linguistics, for their inspiration and encouragement given to this study.

I extend my thanks to Krishna Prasad Chalise, Central Department of Linguistics, Dubi Nanda Dhakal, Central Department of Linguistics, Krishna Poudel, Central Department of Linguistics, for their valuable comments and active participation in discussion whenever the problem was raised.

I am equally thankful to Ram Raj Lohani, Central Department of Linguistics, Bhim Narayan Regmi, Central Department of Linguistics, Karnakhar Khatiwada, Central Department of Linguistics, Bhim Lal Gautam, Central Department of Linguistics, Krishna Prasad Parajuli, Central Department of Linguistics for their support and encouragement.

I am extremely thankful to Santa Bahadur Basnet for his tokenizer and a number of computational concepts.

I would like to thank Madan Puraskar Library and its staff for their help in different point of time.

My sincere thanks go to Dr. Tika Ram Poudel, Tina Bögel, Sebastian Sulger, Kanstanz Univeristy for their help.

I would like to thank Tribhuvan University for providing me study leave for two years, SIL International for providing me travel grants while attending the workshops and institute in University of California, Bangkok, Thailand and IIT Hyderabad, India; Bhashasanchar project for supporting me financially while attending the training on text-to-speech in Gotenberg University, Sweden; and Department of Linguistics, University of Kanstanz for supporting me financially to attend the school of computational and natural language processing in Konstanz University, Germany.

I would like to take this opportunity to express my sincere appreciation to my spouse Mrs. Nirmala Prasain, son Aryan Prasain and daughter Sanskriti Prasain for their tolerance.

Finally, I express my thanks to all the Central Department's non-teaching staff for their help whenever required.

BALARAM PRASAIN

ABSTRACT

The main goal of this study is to present a computational analysis of morphology in Nepali for developing a model for natural language processing by applying the finite state approach. The morphological categories have been analyzed according to the principle of Two-level morphology (Koskeniemmi 1983), and these categories have been implemented using Xerox finite state tool (Beesley and Karttunen 2003) to create the morphological analyzer. A version of finite state automaton called finite state transducer is used in this study which handles relation between two languages, namely upper language and lower language. Upper language is equivalent to lexical level and lower language is equivalent to surface level. The finite state transducer is bidirectional, i.e., moving from surface level to lexical level is analysis and from lexical level to surface level is generation.

This study is organized into eight chapters. Chapter 1 presents the general morphological concepts, the objectives, methodology, the significance and limitations of the study. Chapter 2 presents the theoretical framework that is adopted for the study. Chapter 3 analyzes nouns, pronouns, adjectives, numerals and classifiers in Nepali. Chapter 4 analyzes the verbs in Nepali from computational approach in the first part and verbal inflections in the second part. Chapter 5 deals with indeclinable words in Nepali. Chapter 6 analyzes the derivational process. Chapter 7 implements the outcome of analysis in previous chapters into a finite state transducer using Xerox Finite State Tool. Chapter 8 summarizes the findings of the study.

This study has identified fourteen groups of nouns, eight groups of pronouns, four groups of adjectives, one group of cardinal numerals, two groups of ordinal numerals, three groups of classifiers, ten groups of verbs, seven groups of adverbs, two groups of conjunctions, three groups of postpositions, one group of particles and fifteen groups of derivations in Nepali. The phonological rules for each group have also been identified. The finite state transducer for each group with corresponding morphological tags and phonological rules have been created; and all of them have been put together into a single transducer which can be used as a morphological analyzer for Nepali.

TABLE OF CONTENTS

Recommendation Letter	i
Approval Letter	ii
Declaration	iii
Acknowledgments	iv
Abstract	vi
List of tables	xiii
List of figures	xx
List of abbreviations	xxv
Chapter 1: Introduction	1
1.1 Background	1
1.2 Statement of the problem	6
1.3 Objectives of the study	7
1.3 Literature review	7
1.5 Significance of the study	16
1.6 Research methodology	16
1.7 Limitations	17
1.8 Organization of the study	17
Chapter 2: Theoretical framework	19
2.0 Outline	19
2.1 Computational concept	19
2.2 Regular expression	20
2.3 Finite state technology	22
2.4 Regular language	23
2.5. Finite state machine	24
2.5.1 Finite state automata (FSA)	24
2.5.2 Finite state transducer (FST)	25
2.5.3 Some important operations on FSTs	26

2.6 FST in computational morphology	30
2.7 Xerox finite state tool syntax (XFST)	31
2.7.1 LEXC grammar	33
2.7.2 XFST interface	38
2.8 Summary	41
Chapter 3: Nominal morphology	42
3.0 Outline	42
3.1 Nouns in Nepali	42
3.1.1 Characteristics of nouns in Nepali	43
3.2 Classification of nouns in Nepali	55
3.2.1 <i>O</i> -ending nouns	55
3.2.2 Non- <i>o</i> -ending nouns	60
3.3 Pronouns	69
3.3.1 Characteristics of pronouns in Nepali	69
3.3.2 Grouping of pronouns	71
3.4 Adjectives	91
3.4.1 Characteristics of adjectives in Nepali	91
3.4.2 Classification of adjectives	95
3.5 Numerals	100
3.5.1 Cardinal numbers	100
3.5.2 Ordinal number	101
3.5.2 Other numerals	105
3.6 Classifiers in Nepali	107
3.6.1 Numeral classifiers	107
3.6.2 Quasi classifiers	108
3.7 Summary	110
Chapter 4: Verbal morphology	111
4.0 Outline	111
4.1 Characteristics of verb in Nepali	111

4.1.1 Significant verb stem finals	111
4.1.2 Transitivity	118
4.1.3 Syllabicity	120
4.1.4 Sound आ <i>a</i>	122
4.2 Morphological processes	122
4.2.1 Causativization/transitivization	122
4.2.2 Passivization	127
4.2.3 Negativization	129
4.3 Stem formation	130
4.4 Grouping of verb stems	131
4.4.1 Intransitive verb stems	131
4.4.2 Transitive verb stem	138
4.4.3 Irregular verb stems	144
4.4.4 Suppletive verb stems	145
4.5 Verbal inflections	147
4.5.1 Auxiliary verbs in Nepali	147
4.5.2 Tense	155
4.5.3 Aspects	163
4.5.4 Moods	173
4.5.5 Participial forms	179
4.6 Summary	187
Chapter 5: Adverbs, conjunctions, postpositions and particles	188
5.0 Outline	188
5.1 Adverbs in Nepali	188
5.1.1 Temporal adverbs	188
5.1.2 Spatial adverbs	189
5.1.3 Amount adverbs	190
5.1.4 Manner adverbs	191
5.1.5 Frequency adverbs	191
5.1.6 Reason adverbs	192

5.1.7 Sentential adverbs	193
5.2 Conjunctions in Nepali	194
5.2.1 Coordinate conjunctions	194
5.2.2 Subordinate conjunctions	195
5.3 Postpositions in Nepali	196
5.3.1 Plural/collective marker	196
5.3.2 Case markers in Nepali	196
5.3.3 Adverbial postpositions	198
5.4 Particles and interjections in Nepali	201
5.4.1 Particles	201
5.4.2 Emphatic markers	202
5.4.3 Interjections in Nepali	203
5.5 Summary	204
Chapter 6: Derivational morphology	205
6.0 Outline	205
6.1 Prefixation	205
6.1.2 Noun to noun derivation	206
6.1.3 Noun to adjective derivation	207
6.1.4 Noun to adverb derivation	208
6.1.5 Adjective to adjective derivation	209
6.2 Suffixation	209
6.2.1 Noun to noun derivation	209
6.2.2 Noun to adjective derivation	210
6.2.3 Noun to noun/adjective derivation	212
6.2.4 Adjective to noun derivation	213
6.2.5 Adjective/noun to noun derivation	214
6.2.6 Verb to noun derivation	215
6.2.7 Verb to adjective derivation	217
6.2.8 Verb to adverb derivation	219
6.2.9 Adverb to adjective derivation	220

6.2.10 Verb to noun conversion	221
6.2.11 Verb to adjective/noun conversion	222
6.2.12 Verb to noun derivation	223
6.3 Summary	224
Chapter 7: Implementation	225
7.0 Outline	225
7.1 Morphotactics: syntax of morphemes	225
7.1.1 Morphological categories	225
7.1.2 Grammatical categories	226
7.2 Lexc grammar	228
7.2.1 Nouns	228
7.2.2 Pronouns	230
7.2.3 Verbs	235
7.2.4 Adjectives	246
7.2.5 Numerals and classifiers	247
7.2.6 Adverbs	250
7.2.7 Postpositions	251
7.2.8 Conjunctions, particles and interjections	252
7.2.9 Derivations	256
7.3 Realization: rules of alternations	269
7.3.1 Phonological rules for nouns	269
7.3.2 Phonological rules for pronouns	271
7.3.3 Phonological rules for verbs	271
7.3.4 Phonological rules for adjectives	274
7.3.5 Phonological rules for adverbs	275
7.3.6 Phonological rules for postpositions	275
7.3.7 Phonological rules for particles and interjections	276
7.3.8 Phonological rules for numerals and classifiers	276
7.3.9 Phonological rules for derivations	276
7.4 Summary	278

Chapter Eight: Summary and conclusion	279
Annexes	282
Annex-1: Devanagari – IPA	282
Annex-2: Nepali nouns sample	284
Annex-3: Nepali pronouns	305
Annex-4: Adjectives in Nepali	310
Annex-5: Numerals and classifiers in Nepali	311
Annex-6: Adverbs in Nepali	318
Annex-7: Verbs in Nepali	325
Annex-8: Verbal Inflections in Nepali	328
Annex-9: Conjunctions and particles in Nepali	334
Annex-10: Postpositions in Nepali	337
Annex-11: Words and affixes for derivation in Nepali	340
References	348

List of Tables

Table 1.1: Simple, complex, compound and reduplicated words	3
Table 1.2 Free morphemes in Nepali	3
Table 1.3 Bound morphemes in Nepali	4
Table 1.4: Lexical and surface levels representation	5
Table 2.1: The sample regular expressions	20
Table 2.2: Some operators used in regular expressions	21
Table 2.3: Regular expressions and regular language	24
Table 2.4: The transition table for घर and घरहरू	25
Table 3.1: The <i>o</i> -ending and non- <i>o</i> -ending nouns	42
Table 3.2: Number: singular and plural	44
Table 3.3: Lexical gender	46
Table 3.4: Morphological gender	47
Table 3.5: Direct and oblique forms	48
Table 3.6: Honorificity: non-honorific and honorific	49
Table 3.7: Augmentative and diminutive	50
Table 3.8: NounType 1a	54
Table 3.9: NounType 1b	56
Table 3.10: NounType 1c	57
Table 3.11: NounType 1d	58
Table 3.12: NounType 21a	59
Table 3.13: NounType 21b	60
Table 3.14: NounType 21c	62
Table 3.15: NounType 21d	63
Table 3.16: NounType 22a	64
Table 3.17: NounType 22b	64
Table 3.18: NounType 22c	65
Table 3.19: NounType 22d	66
Table 3.20: NounType 22e	66
Table 3.21: NounType 22f	67
Table 3.22: Pronouns with respect to persons	69
Table 3.23: Persons number in number distinctions	70

Table 3.24: Form of pronouns: direct and oblique	70
Table 3.25: Honorific levels in Nepali pronouns	71
Table 3.26: First person singular pronouns	72
Table 3.27: First person plural pronouns	73
Table 3.28: Second person singular non-honorific pronouns	74
Table 3.29: Second person honorific pronouns	75
Table 3.30: Second person high honorific pronouns	75
Table 3.31: Second person royal honorific pronoun	76
Table 3.32: Third person pronoun ऊ <i>u</i> :	77
Table 3.33: Third person pronouns त्यो <i>tjo</i> and ती <i>ti</i> :	78
Table 3.34: Third person pronouns यो <i>jo</i> and यी <i>ji</i> :	79
Table 3.35: The reflexive pronouns	80
Table 3.36: The demonstrative pronouns यो <i>jo</i> and यी <i>ji</i> :	81
Table 3.37: The demonstrative pronouns त्यो <i>tjo</i> and ती <i>ti</i> :	82
Table 3.38: The demonstrative pronouns ऊ <i>u</i> :	83
Table 3.39: The remaining demonstrative pronouns	84
Table 3.40: The relative pronouns	84
Table 3.41a: The interrogative pronouns	85
Table 3.41b: The indefinite pronouns derived from interrogative pronouns	86
Table 3.42: The indefinite pronouns derived from relative pronouns	87
Table 3.43a: The definite pronouns	88
Table 3.43b: The definite pronoun अर्को	88
Table 3.44a: The reciprocal pronouns	89
Table 3.44b: The reciprocal pronouns	90
Table 3.45: <i>O</i> -ending and non- <i>o</i> -ending adjectives	91
Table 3.46: Number: singular and plural	92
Table 3.47: Gender: masculine and feminine	93
Table 3.48: Form: direct and oblique	93
Table 3.49: Honorificity: non-honorific and honorific	94
Table 3.50: Degree: positive, comparative and superlative	95

Table 3.51: <i>O</i> -ending adjectives	96
Table 3.52: Type 1 marked adjectives	97
Table 3.53: Type 2 marked adjectives	99
Table 3.54: Unmarked adjectives	99
Table 3.55: Some cardinal numbers	100
Table 3.56: Some regular ordinal numbers	102
Table 3.57: Irregular ordinal numbers of one	103
Table 3.58: Irregular ordinal numbers of two	103
Table 3.59: Irregular ordinal numbers of three	103
Table 3.60: Irregular ordinal numbers of four	103
Table 3.61: Some ordinal numbers from Sanskrit loan	105
Table 3.62: Frequency numerals (I)	105
Table 3.63: Frequency numerals (II)	105
Table 3.64: Frequency numerals (III)	106
Table 3.65: Frequency numerals (IV)	106
Table 3.66: Some portion numerals	106
Table 3.67: Numeral classifiers	107
Table 3.68: <i>o</i> -ending classifiers	108
Table 3.69: General non- <i>o</i> -ending classifiers	109
Table 4.1: <i>i</i> -ending intransitive verb stems	112
Table 4.2: <i>i</i> -ending transitive verb stems	112
Table 4.2a: <i>i</i> -ending transitive verb stems	113
Table 4.3: Alternative forms of <i>i</i> -ending verb stems	114
Table 4.4: <i>a</i> -ending verb stems (group 1)	115
Table 4.5: <i>a</i> -ending verb stems (group 2)	115
Table 4.6: <i>o</i> -ending verb stems	116
Table 4.7a: Change of <i>o</i> to <i>u</i> in <i>o</i> -ending verb stems	116
Table 4.7b: <i>Ḍ</i> -ending verb stems	116
Table 4.7c: <i>Ḍ</i> -ending verb stems	116
Table 4.7d: Verb stems ending with a voiceless consonant	117

Table 4.8: Alternative forms from stems ending with voiceless consonant	117
Table 4.9: Verb stems ending with voiced consonant	118
Table 4.10: Alternative forms from stems ending with voiced consonant	118
Table 4.11: Intransitive verbs	119
Table 4.12: Some transitive verbs	120
Table 4.13: Some ditransitive verbs	120
Table 4.14: Monosyllabic verb stems	121
Table 4.15: Polysyllabic verb stems	121
Table 4.16: Verb stems with <i>a</i> sound	122
Table 4.17: Verb stems without <i>a</i> sound	122
Table 4.18: Causative verb stems	123
Table 4.19: Verb stems forming causatives with - <i>आ</i> - <i>a</i> and <i>आल्</i> - <i>al</i>	124
Table 4.20: Verb stems forming causatives by changing <i>अ</i> Λ to <i>आ</i> <i>a</i>	125
Table 4.21a: Verb stems forming causatives by changing <i>उ</i> <i>u</i> to <i>ओ</i> <i>o</i>	126
Table 4.21b: Verb stems forming causatives by suffixing - <i>आ</i> - <i>a</i>	126
Table 4.22: Verb stems forming causatives by inserting <i>a</i>	127
Table 4.23: Some passive verb stems	129
Table 4.24: Negation by the prefixation of negative marker <i>न</i> <i>n</i> -	130
Table 4.25: Negation by the suffixation of negative marker - <i>न</i> <i>n</i> -	130
Table 4.26: Pattern of the stem formation	131
Table 4.27: Type 1a verb stems	132
Table 4.28: Type 1b verb stems	133
Table 4.29: Type 1c verb stems	134
Table 4.30: Type 1d verb stems	136
Table 4.31: Type 1e verb stems (i)	137
Table 4.32: Type 1e verb stems (ii)	137
Table 4.33: Type 2a verb stems (i)	139
Table 4.34: Type 2a verb stems (ii)	139
Table 4.35: Type 2b verb stems	140
Table 4.36: Type 2c verb stems	142

Table 4.37: Type2d verb stems	143
Table 4.38: Irregular verb stems	145
Table 4.39: Suppletive verb stems	146
Table 4.39: Inflections for non-past existential verb छे <i>cha</i> 'be' (affirmative)	148
Table 4.40: Inflection for non-past existential verb छे <i>cha</i> 'be' (negative)	149
Table 4.41: Inflections for non-past identificational verb हो <i>ho</i> 'be' (affirmative)	151
Table 4.42: Inflection for non-past identificational verb हो <i>ho</i> 'be' (negative)	152
Table 4.43: Inflections for past existential verb थि <i>thi</i> 'be' (affirmative)	153
Table 4.44: Inflections for past existential verb थि <i>thi</i> 'be' (negative)	155
Table 4.45: Inflections for non-past tense (affirmative)	157
Table 4.46: Inflections for non-past tense negative 1	158
Table 4.47: Inflections for non-past tense negative 2	159
Table 4.48: Inflections for past tense (affirmative)	161
Table 4.49: Inflections for past tense (negative)	162
Table 4.50: Inflections for perfect aspect	164
Table 4.51: Inflections for imperfect aspect	166
Table 4.52: Inflections for past habitual aspect (affirmative)	168
Table 4.53: Inflections for habitual aspect (negative)	169
Table 4.54: Inflections for inferential aspect (affirmative)	171
Table 4.55: Inflections for inferential aspect (negative)	172
Table 4.56: Inflection for imperative mood	174
Table 4.57: Inflections for optative mood (affirmative)	176
Table 4.58: Inflections for potential mood (affirmative)	178
Table 4.59: Inflection for absolute participle	179
Table 4.60: Inflections for infinitive participle	181
Table 4.61: Inflections for purposive participle	182
Table 4.62: Inflection for prospective participle	183
Table 4.63: Inflections for durative participle	183
Table 4.64: Inflections for conjunctive participle	184
Table 4.65: Inflection for conditional participle	185
Table 4.65: Inflection for perfective participle	186

Table 5.1: Temporal adverbs	189
Table 5.2: Spatial adverbs	189
Table 5.3: Amount adverbs	190
Table 5.4: Manner adverbs	191
Table 5.5: Frequency adverbs	192
Table 5.6: Reason adverbs	192
Table 5.7: Sentential adverbs	193
Table 5.8: Coordinate conjunctions	194
Table 5.9: Subordinate conjunctions	195
Table 5.10: Collective/plural marker	196
Table 5.11a: Case marker postpositions (i)	197
Table 5.11b: Case marker postpositions (ii)	197
Table 5.12a: Adverbial postpositions (a)	198
Table 5.12b: Adverbial postpositions (b)	199
Table 5.13: Particles in Nepali	202
Table 5.14: Interjections in Nepali	204
Table 6.1: Noun to noun derivation	206
Table 6.2: Noun to adjective derivation	207
Table 6.3: Noun to adverb derivation	208
Table 6.4: Adjective to adjective derivation	209
Table 6.5: Noun to noun derivation	209
Table 6.6: Noun to adjective derivation	211
Table 6.7: Noun to noun/adjective derivation	212
Table 6.8: Adjective to noun derivation	213
Table 6.9: Adjective/noun to noun derivation	214
Table 6.10: Verb to noun derivation	216
Table 6.11: Verb to adjective derivation	218
Table 6.12: Verb to adverb derivation	219
Table 6.13: Adverb to adjective derivation	220
Table 6.14: Verb to noun conversion	221
Table 6.15: Verb to adjective/noun conversion	222
Table 6.16: Verb to noun (vowel insertion)	223
Table 7.1: The open word classes	225
Table 7.2: The closed word classes	226

Table 7.3: The grammatical categories and features	226
Table 7.3: The arbitrary tags	227

List of Figures

Figure 1.1: Lexical and surface levels of Nepali word केटो <i>keṭo</i> 'boy'	6
Figure 2.1. A finite state automaton that accepts घर 'house' and घरहरू 'houses'	24
Figure 2.2: A finite state transducer that transduces between घर 'house' and घर+NOUN+SG	26
Figure 2.3: FST unioned from three FSTs for nouns, adjectives and adverbs	27
Figure 2.4: A finite state transducer concatenated from two FSTs above	28
Figure 2.5: A finite state transducer from composing two FSTs above	29
Figure 2.6: The interrelation among language, regular expression and finite state network	30
Figure 2.7: The structure of Lexc grammar (Beesley and Karttunen 2003:205)	34
Figure 2.8: xfst interface can compile lexicon and rule and compose them into single FST (Karttunen 2000)	39
Figure 3.1: A finite state transducer for NounsType 1a	56
Figure 3.2: A finite state transducer for NounType 1b	57
Figure 3.3: A finite state transducer for NounType 1c	58
Figure 3.4: A finite state transducer for NounsType 1d	60
Figure 3.5: A finite state transducer for NounType21a	61
Figure 3.6: A finite state transducer for NounType 21b	62
Figure 3.7: A finite state transducer for NounType 21d	64
Figure 3.8: A finite state transducer for NounType 22a	65
Figure 3.9: A finite state transducer for NounsType 22b	66
Figure 3.10: A finite state transducer for NounType 22c	66
Figure 3.11: A finite state transducer for NounType 22d	67
Figure 3.12: A finite state transducer for NounType 22e	68
Figure 3.13: A finite state transducer for NounType 22f	68
Figure 3.14: A finite state transducer for first person singular pronouns	72
Figure 3.15: A finite state transducer for first person plural pronouns	73
Figure 3.16: A finite state transducer for second person singular non-honorific pronouns	74
Figure 3.17: A finite state transducer for second person honorific pronouns	75
Figure 3.18: A finite state transducer for second person higher honorific pronouns	76

Figure 3.19: A finite state transducer for second person highest honorific pronoun	76
Figure 3.20: A finite state transducer for third person <i>u</i> :	77
Figure 3.21: A finite state transducer for third person pronouns त्थो <i>tjo</i> and ती <i>ti</i> :	78
Figure 3.22: A finite state transducer for third person pronouns यो <i>jo</i> and यी <i>ji</i> :	79
Figure 3.23: A finite state transducer for reflexive pronouns	80
Figure 3.24: A finite state transducer for demonstrative pronouns यो <i>jo</i> and यी <i>ji</i> :	81
Figure 3.25: A finite state transducer for demonstrative pronouns त्थो <i>tjo</i> and ती <i>ti</i> :	82
Figure 3.26: A finite state transducer for demonstrative pronouns ऊ <i>u</i> :	83
Figure 3.27: A finite state transducer for remaining demonstrative pronouns	84
Figure 3.28: A finite state transducer for relative pronouns	85
Figure 3.29: A finite state transducer for interrogative pronouns	86
Figure 3.30: A finite state transducer for indefinite pronouns derived from interrogative pronouns	87
Figure 3.31: A finite state transducer for indefinite pronouns derived from relative pronouns	87
Figure 3.32a: A finite state transducer for definite pronouns	88
Figure 3.32b: A finite state transducer for definite pronouns	89
Figure 3.33a: A finite state transducer for reciprocal pronouns	90
Figure 3.33b: A finite state transducer for reciprocal pronouns	90
Figure 3.34: A finite state transducer for <i>o</i> -ending adjectives	96
Figure 3.35: A finite state transducer for Type 1 marked adjectives	98
Figure 3.36: A finite state transducer for Sanskrit loan adjectives	99
Figure 3.37: A finite state transducer for unmarked adjectives	100
Figure 3.38 A finite state transducer for cardinal numbers and regular ordinal numbers	102
Figure 3.39: A finite state transducer for irregular ordinal numerals	104
Figure 3.40: A finite state transducer for ordinal numerals form Sanskrit loan	105
Figure 3.41: A finite state transducer for frequency numerals	106
Figure 3.42: A finite state transducer for portion numerals	107
Figure 3.43: A finite state transducer for numeral classifiers	107
Figure 3.44: A finite state transducer for general classifier type 1	108
Figure 3.45: A finite state transducer for general classifier type 2	109

Figure 4.1: A finite state transducer for Type1a verb stems	132
Figure 4.2: A finite state transducer for Type1b verb stems	133
Figure 4.3: A finite state transducer for Type1c verb stems	135
Figure 4.4 A finite state transducer for Type1d verb stems	136
Figure 4.5: A finite state transducer for Type1e verb stems	138
Figure 4.6: A finite state transducer for Type2a verb stems	139
Figure 4.7: A finite state transducer for Type2b verb stems	141
Figure 4.8: A finite state transducer for Type2c verb stems	142
Figure 4.9: A finite state transducer for Type2d verb stems	144
Figure 4.10: A finite state transducer for inflections of non-past existential verb छे <i>chA</i> 'be' (affirmative)	149
Figure 4.11: A finite state transducer for inflections of non-past existential verb छे <i>chA</i> 'be' (negative)	150
Figure 4.12: A finite state transducer for inflections of non-past identificational verb हो <i>ho</i> 'be' (affirmative)	151
Figure 4.13: A finite state transducer for inflection of non-past identificational verb हो <i>ho</i> 'be' (negative)	153
Figure 4.14: A finite state transducer for inflections of past existential verb थि <i>t^{hi}</i> 'be' (affirmative)	154
Figure 4.15: A finite state transducer for inflections of past existential verb थि <i>t^{hi}</i> 'be' (negative)	155
Figure 4.16: A finite state transducer for inflections of non-past tense	157
Figure 4.17: A finite state transducer for inflections of non-past tense negative 1	159
Figure 4.17a: A finite state transducer for inflections of non-past tense negative 2	160
Figure 4.18: A Finite State Transducer for inflections of past tense (affirmative)	161
Figure 4.19: A finite state transducer for inflections of past tense (negative)	163
Figure 4.20: A finite state transducer for inflections of perfect aspect	165
Figure 4.21 A finite state transducer for Inflections of imperfect aspect	166
Figure 4.22: A finite state transducer for inflections of habitual aspect (affirmative)	168
Figure 4.23: A finite state transducer for inflections of habitual aspect (negative)	170
Figure 4.24: A finite state transducer for inflections of inferential aspect (affirmative)	171

Figure 4.25: A finite state transducer for inflections of inferential aspect (negative)	173
Figure 4.26: A finite state transducer for inflections of imperative mood	175
Figure 4.27: A finite state transducer for inflections of optative mood	176
Figure 4.28: A finite state transducer for inflections of potential mood	178
Figure 4.29: A finite state transducer for inflection of absolutive form	180
Figure 4.30: A finite state transducer for inflections of infinitive participial form	181
Figure 4.31: A finite state transducer for inflections of purposive participial form	182
Figure 4.32: A finite state transducer for inflection of prospective participial form	183
Figure 4.33: A finite state transducer for inflections of durative participial forms	184
Figure 4.34: A finite state transducer for inflections of conjunctive participial form	185
Figure 4.35: A finite state transducer for conditional participial form	186
Figure 4.36: A finite state transducer for inflection of conditional participial form	187
Figure 5.1: A finite state transducer for temporal adverbs	189
Figure 5.2: A finite state transducer for spatial adverbs	190
Figure 5.3: A finite state transducer for amount adverbs	190
Figure 5.4: A finite state transducer for manner adverbs	191
Figure 5.5: A finite state transducer for frequency adverbs	192
Figure 5.6: A finite state transducer for reason adverbs	193
Figure 5.7: A finite state transducer for sentential adverbs	193
Figure 5.8: A finite state transducer for coordinate conjunctions	195
Figure 5.9: A finite state transducer for subordinate conjunctions	195
Figure 5.10: A finite state transducer plural/collective marker	196
Figure 5.12a: A finite state transducer for adverbial postpositions that do not take emphatic marker	199
Figure 5.12b: A finite state transducer for adverbial postpositions that take emphatic marker	201
Figure 5.13: A finite state transducer for particles	202
Figure 5.14: A finite state transducer for interjections	204
Figure 6.1: A finite state transducer for noun to noun derivation	207
Figure 6.2: A finite state transducer for noun to adjective derivation	208
Figure 6.3: A finite state transducer for noun to adverb derivation	208
Figure 6.4: A finite state transducer for adjective to adjective derivation	209
Figure 6.5: A finite state transducer for noun to noun derivation	210
Figure 6.6: A finite state transducer for noun to adjective derivation	211

Figure 6.7: A finite state transducer for noun to noun/adjective derivation	212
Figure 6.8: A finite state transducer for noun to adjective derivation	214
Figure 6.9: A finite state transducer for noun/adjective to noun derivation	215
Figure 6.10: A finite state transducer for verb to noun derivation	217
Figure 6.11: A finite state transducer for verb to adjective derivation	218
Figure 6.12: A finite state transducer for verb to adverb derivation	220
Figure 6.13: A finite state transducer for noun to adjective derivation	221
Figure 6.14: A finite state transducer for verb to adverb derivation	222
Figure 6.15: A finite state transducer for verb to adverb derivation	222
Figure 6.16: A finite state transducer for verb to adverb derivation	223

List of abbreviations

+ABL	=	Ablative
+ABS	=	Absolutive
+ADJ	=	Adjective
+ADV	=	Adverb
+AMOUNT	=	Amount
+AUG	=	Augmentative
+CARD	=	Cardinal
+CAUSE	=	Causative
+CCONJ	=	Coordinate Conjunction
+CLF	=	Classifier
+COM	=	Commitative
+COMP	=	Comparative
+COND	=	Conditional
+CONJUNCT	=	Conjunctive
+DAT	=	Dative
+DEF	=	Definite
+DEM	=	Demonstrative
+DIM	=	Diminutive
+ALL	=	Directional
+DIRT	=	Direct
+DUR	=	Durative
+EMPH	=	Emphatic
+ERG	=	Ergative
+EXIST	=	Existential
+FEM	=	Feminine
+FREQ	=	Frequency
+GEN	=	Genitive
+HAB	=	Habitual
+HHON	=	High Honorific
+HON	=	Honorific
+ID	=	Identificational
+IMP	=	Imperative
+IMPERF	=	Imperfect
+INDEF	=	Indefinite
+INF	=	Infinitive
+INFER	=	Inferential
+INST	=	Instrument
+INTERJ	=	Interjection
+INTERRO	=	Interrogative
IPA	=	International Phonetic Alphabet
+LOC	=	Locative
+MANNER	=	Manner
+MASC	=	Masculine
+NHON	=	Non honorific
+NOUN	=	Noun
NP	=	Noun Phrase
+NPST	=	Non past

+NUM	=	Numeral
+OBL	=	Oblique
+OPT	=	Optative
+ORD	=	Ordinal
+PST	=	Past
+PARTICLE	=	Particle
+PASS	=	Passive
+PERF	=	Perfect
+PERFT	=	Perfective
+PL	=	Plural
+PLACE	=	Place
+PORT	=	Portion
POS	=	Parts of speech
+POSIT	=	Positive
+POSTP	=	Postposition
+POT	=	Potential
+PRON	=	Pronoun
+PROPER	=	Proper
+PROS	=	Prospective
+PROX	=	Proximate
+PURP	=	Purposive
+REASON	=	Reason
+RECIP	=	Reciprocal
+REFL	=	Reflexive
+REL	=	Relative
+RHON	=	Royal honorific
+SCONJ	=	Subordinate conjunction
+SENT	=	Sentential
+SG	=	Singular
+SPAC	=	Spatial
+SUPER	=	Superlative
+TEMP	=	Temporal
+VERB	=	Verb
+VOC	=	Vocative
1	=	First person
2	=	Second person
3	=	Third person

CHAPTER 1

INTRODUCTION

1.1 Background

This study is an attempt to analyze morphology in Nepali and design a computational model for natural language processing within the framework of the finite state technology in general and 'Two-level morphology' developed by Koskeniemmi (1983) in particular. For the implementing of analyzed data and creating a computational model, the Xerox Finite State Tool developed by Beesley and Karttunen (2003) has been employed.

Language as a means of human communication is a tool to express the greater part of human ideas and emotions. Language shapes human thoughts, has a structure and carries meaning. Learning and expressing new concepts and ideas through language are so natural that it is hardly realized how the natural language is processed in our brain. Thus, it can be claimed that there must be some sorts of language representation and a processing module in the human brain (Siddiqui and Tiwari 2008:1). This type of content in the brain also helps to represent the language in real time world. Every time the language activities take place, there is always a very fast and accurate natural language processing that finally performs a successful communicative event. To capture this reality, the computational linguistics attempts to develop computational models of aspects of human language processing. Developing such automated tools for language processing and gaining a better understanding of human communication are the main reasons that have inspired the linguists and the computer scientists in this ever growing field.

In fact, language is the outer form of the content it expresses. Therefore, the language processing means the processing of content it possesses. The language, generally, is manifested either in the written or spoken form. Both forms of the language can be processed with the help of computer. To achieve this goal, a computational model of a particular language based on formal approach is required to be designed and implemented into the computer.

As computers are not able to understand the natural language, the computational models and methods are developed to map its content in a formal language. And such formal languages are extended to account the natural language phenomena at various

levels of the language. The representation of the whole body of the knowledge of language can be an ambitious project. Thus, the language can be graded into various levels such as phonology, morphology, syntax, semantics and pragmatics. Each level is perceived and defined in different ways by the people in different disciplines according to the goals set up. It is not possible to create a mega computational model at a time to cover the entire language. Therefore, the main goal of this study is to represent, design a computational model and process the morphology of written form of the Nepali text.

Nepali is an Indo-Aryan language characterized by agglutinating morphology in general. The verb is predominantly inflectional whereas the noun is heavily agglutinating. There have not yet been made attempts to analyze Nepali morphology for the development of computational model. The basic linguistic concepts in relation to morphology and its representation in order to further clarify the computational aspects of morphology are briefly discussed in the following subsections.¹

I. Morphology

The words are considered to be fundamental building blocks of language (O'Grady, Dobrovolsky and Aronoff 1997:118; Jurafsky and Martin 2000:45). Every human language has words (Mathews 1991:20), they are said to be granted for at least in descriptive linguistics (Katamba 1993:17). Of all the units of linguistic analysis, the word is the central, most familiar and crucial. The smallest free form found in a language are said to be words (Bloomfield 1933); however, a free form that can occur in isolation is not only atomic but also molecular in its structure. A word (i.e. word-form), in real sense, can either be in simple, complex, compound or reduplicated form. Table 1.1 presents simple, complex, compound and reduplicated words in Nepali.

¹ The general concepts of the linguistic categories may be required to be qualified and specified for computational purposes. Therefore, some basic concepts are briefly discussed in sections I and II.

Table 1.1: Simple, complex, compound and reduplicated words

Type	Word	IPA/gloss	Meaning
Simple	घर	g ^h ʌr 'house'	house
Complex	घरबाट	g ^h ʌr-bat 'house-ABL'	from house
Compound	घरपरिवार	g ^h ʌr-pʌriwar 'house-family'	family
Reduplication	घरघरै	g ^h ʌr-g ^h ʌr-ʌi 'house-house-EMPH'	each house

Table 1.1 shows that the word घर g^hʌr 'house' is simple; घरबाट g^hʌr-batʌ 'from house' is complex; घरपरिवार g^hʌr-pʌriwar 'family' is compound; and घरघरै g^hʌr-g^hʌr-ʌi 'each house' is reduplicated.

a. Morpheme: free and bound

The smallest (minimal) unit of grammar that carries information about the meaning and function is said to be a morpheme (Bloomfield 1933; Katmaba 1993; O'Grady, Dobrovolsky, Aronoff 1997; Mathews 1991). The morpheme is an abstract entity that may correspond to various forms at the surface level. The morpheme can be free and bound. A morpheme may be a word by itself or it may not be. The lexical categories such as nouns, verbs, adjective, adverbs, etc. are free morphemes whereas a morpheme that must be attached to another element (normally a free morpheme) is called a bound morpheme. Free morphemes are lexical and bound morphemes are generally grammatical. Table 1.2 lists some free morphemes in Nepali.

Table 1.2: Free morphemes in Nepali

Morpheme	IPA	Gloss
घर	g ^h ʌr	house
जा	dza	go
असल	ʌsʌl	good
आज	adzʌ	today

The morphemes घर g^hʌr 'house', जा dza 'go', असल ʌsʌl 'good' and आज adzʌ 'today' in Table 1.2 can stand as words. The bound morphemes in Nepali cannot stand by themselves. Table 1.3 lists some bound morphemes in Nepali.

Table 1.3: Bound morphemes in Nepali

Morpheme	IPA	Gloss
-नु	-nu	-INF
-एको	-eko	-PERF
-ला	-la	-POT
-आइ	-ai	-NML

The morphemes -नु -nu 'INF', -एको -eko 'PERF', -ला -la 'POT' and -आइ -ai 'NML' in Table 1.3 cannot stand alone. They appear with other free morphemes and express some grammatical functions.

b. Root, stem, base, affixes and word

Root is an ultimate and irreducible constituent element common to all word-forms of the same family. It is not an abstract but a concrete form. The root constitutes the core of the word and carries the major component of its meaning and typically belongs to a lexical category such as noun, verb, adjective or adverb. Therefore, a root corresponds to a free morpheme (Katamba 1993:45; Payne 1997:25). For example, *man*, *book*, *tea*, etc. are the roots. A stem is a part of the word except the last (generally for inflectional purpose) affix. Therefore, the stem may be composed of minimally a root and it may have more elements (Katamba 1993:45). *cat* in *cats* is the stem for instance. A base is the word or part of the word to which an affix can be attached. It may be called stem for the inflectional purpose (Katamba 1993:45). In English, the word *work* can be a root, a stem and a base whereas the word *worker* can only be a base for the word *workers*. The root, stem and base have something in common in their definition from linguistic perspective. A base can be a stem as well as a root and a stem can be a root as well.²

The concept of root, stem and base overlaps with one another from theoretical point of view. But, from computational perspective, it makes no difference among them especially in the computer processing. So, the term 'stem' is used to represent any one of them in this study. That means any sequence of character to which some other sequences of characters can be attached.

² I have not discussed the bound roots, see Katamba (1993) for details.

An affix is a bound morpheme when added to the radical element (root, stem or base) it changes the meaning or function of a word by creating a new word-form. Therefore, the affixes are basically involved in the inflectional and derivational phenomena of the language. The affixes can be of various kinds but only prefixes and suffixes are discussed here for the present purpose. The prefix is an affix that gets attached in front of the stem and suffix at the end (Katamba 1993:44). In English, affixes *re-*, *un-* and *in-* in the words *reunion*, *unhappy* and *intolerable* are prefixes whereas *-ly*, *-ing*, *-er*, *-ed* in the words *slowly*, *working*, *walker* and *walked* are suffixes.

Thus, it is clear that from computational point of view a word minimally consists of a stem and optionally one or more affixes.³ In analysis, words are decomposed into their constituents and represented by following certain formalism whereas in generation, the process is reversed.

II. Levels of representation: lexical and surface

Words in written or spoken texts, in fact, represent the outer form, i.e. surface form. But a word carries various kinds of information which can be represented at least at two levels. The lexical level of the word is its canonical form or lemma word and a set of tags showing its syntactic category and morphological features. They are the possible parts of speech and/or inflectional properties such as gender, number, person, tense, aspect and mood. Thus, the lexical level represents the sequence of morphemes in a certain fashion. The actual arrangement of morphemes is governed by the language specific rules. Table 1.4 presents lexical and surface level representations of words in Nepali.

Table 1.4: Lexical level and surface level representation

Lexical level	Surface level	Gloss of stem
केटो+NOUN+MASC+SG	केटो	boy
खा+VERB+P.3SG	खायो	eat
न्यून+ADJ+SUPER	न्यूनतम	least
र+CC	र	and
र+PART	र	uncertain

³ See Katamba (1993:17-23) for the detail.

Table 1.4 presents the lexical level representation consisting of the sequences of morphemes attached to the stem resulting to certain word forms at the surface level. Thus, the representation at the surface level in Nepali corresponds to actual spelling of the word.

In this study, an attempt has been made to represent a word in Nepali at the two levels. The pair of lexical level and surface level can be taken as a relation between two languages and can be used as morphological analyzer and generator simply by changing the direction of transitions. Figure 1.1 illustrates two levels and the process of analysis and generation of words in Nepali.

LEXICAL LEVEL : केटो+NOUN+MASC+SG
SURFACE LEVEL : केटो
Figure 1.1: Lexical and surface levels of Nepali word केटो <i>keṭo</i> 'boy'

In Figure 1.1, for instance, moving from the surface level केटो to lexical level केटो+NOUN+MASC+SG represents morphological analysis and reverse represents generation, respectively.

1.2 Statement of the problem

Nepali is a morphologically rich language. There exist a number of morphological studies in Nepali. Most of them are descriptive in nature. There also exist some scanty works from computational point of view (see, review of literature in 1.3). Morphology of Nepali has not yet been fully analyzed from computational perspective. The main problem of this study is to analyze morphology in Nepali that can be implemented from computational point of view. The specific problems of this study are as follows:

- a. What are the morphological categories in Nepali?
- b. What are the morphological processes in the language?
- c. What are the rules involved in the morphological processes?
- d. What is the computational model for morphology in Nepali?

1.3 Objectives of the study

The main objective of this study is to analyze the morphology in Nepali from computation perspective. The specific objectives of the study are as follows:

- a. To identify the morphological categories in Nepali;
- b. To identify the morphological processes in the language;
- c. To formulate the phonological/orthographic rules in Nepali; and
- d. To design and develop the computational model for Nepali morphology.

1.3 Review of literature

There are only a few scanty works in Nepali morphology from computational perspective. However, there are a number of works in Nepali morphology from traditional and descriptive perspectives. These works contribute to the understanding of the main problem of the study to some extent. Such works have been thematically reviewed in four groups.⁴

a. Nepali morphology

Pandit (2051VS (1969 VS)) has classified word categories in Nepali from traditional point of view. The categories include noun, pronoun, adjective, verb and indeclinable. Nouns are grouped into common noun, proper noun and abstract noun. Each noun is discussed with respect to gender: masculine, feminine and neuter; number: singular and plural; cases and case markers: subjective, objective, instrumental, dative, ablative and locative. He has also presented a detailed inflectional paradigm of nouns and verbs.

Bandhu (1973) has analyzed clause patterns of Nepali from tagmemic approach. The basic clause patterns are sub-classified and illustrated with examples. Under the

⁴ The literatures available are not directly related to this study. However, they provide knowledge to understand the research problems. Therefore, instead of evaluating them critically as per the review style, their contributions to this study have been mentioned under four themes, they are (i) Nepali morphology, (ii) Nepali computational morphology, (iii) Nepali language and related NLP works, and (iv) NLP works in selected languages.

inflected patterns, he has analyzed the inflectional categories, inflectional system, mood and finite system, aspects and copulas, modals, negation and post verbal particles. The paradigms for each inflectional category are presented alone with the illustrations in the sentence.

Dahal (1974) is an extensive description of colloquial and literary Nepali. He has classified the stem formation process into three classes, namely, derived stems, composite stems and reduplicated stems and has also discussed the derived stems as suffix-derived noun stems, suffix-derived adjective stems, suffix-derived verb stems, suffix-derived adverbs, modification-derived stems and prefix-derived stems. Inflectional categories and their realizations have been described under two headings, namely nominal inflections and verbal inflections.

Adhikari (1980) has examined a set of Nepali verbs ending in $\bar{\text{ch}}$ [ts^h] and the relation of $\bar{\text{ch}}$ with the time of speech. There can be other elements between verb stem and element $\bar{\text{ch}}$ and $\bar{\text{ch}}$ always refers to the non-past tense and it is always followed by concord marker.

Sharma (1980) has presented the verbal structure in a formulation as $V \rightarrow \text{stem} ((+\text{Aspect} + \text{BE}) + \text{Tense}) (+\text{Neg}) + \text{concord}$ from descriptive approach. Among them, only the stem and concord are obligatory and others are controlled by some other constraints. The verbal stems are divided into simple and complex. The morphophonemic changes that occur when the suffixes for tense, aspect, negation and concord appear to the simple and complex stems are discussed in a length.

Wallace (1985) is an attempt to test Nepali data against the Relational Grammar and Government and Binding theory. It is mentioned that Nepali nouns show number and the case relations that are indicated by postpositions. Finite verb forms indicate tense, aspect, affirmation or negation and they agree with subject. Adjectives are discussed as noun phrase modifiers where they agree with the head noun in terms of their number and gender.

Chapagain (2046 VS) has also classified the Nepali words into nouns, pronouns, adjectives, verbs, adverbs, conjunction and from traditional and descriptive point of view. The word formation processes such as prefixation, suffixation, compounding and reduplication are well discussed with illustrations.

Acharya (1991) is a corpus based study. He has classified the form classes into inflected forms: noun, adjective, pronoun and verb; and uninflected forms: adverb, conjunction, postposition, interjection and nuance particle. Nouns are discussed with respect to the number: singular and plural; cases: nominative, accusative, instrumental, dative, ablative, genitive and locative. Adjectives are discussed with reference to their endings. The verbs are discussed according to the inflectional suffixes for present, past and future tenses with their corresponding number, person, gender and honorificity. The verbs can have simple and compound stems. The adverbs are placed under the uninflected form class and described in term of comparative and superlative structures with examples from the corpus. The conjunctions, postpositions, interjection and particles are also discussed.

Adhikari (1993) has classified Nepali words into nouns, pronouns, adjectives, verbs, adverbs, postpositions, conjunctions and interjections from descriptive approach. The agreement system is extensively discussed with respect to gender, number, person and five levels of honorificity. The classification of the verbs and stem formation, inflections and derivation are of a great help in designing the alternation rules.

Adhikari (2052 VS) is a study of Nepali case system using Filmorian framework. He has classified the Nepali cases into two main categories. They are core cases and peripheral cases. The former includes semantic cases such as agent, affected, resultative, neutral, experiencer, recipient, essive and the latter includes cases such as locative, instrumental, cause, ablative, beneficiary, purposive and comitative.

Pokharel (2054 VS) has presented the analysis of morphological and syntactic levels of voice, causativization, tense, aspect and mood in the simple verb, compound verbs and negation from descriptive approach. The classification of the verbs, the agreement, honorificity, various kinds of classifiers, gender, number, case in nouns and the grammar of the pronouns have been described and illustrated.

Lohani (1999) has studied the complex predicates in Nepali using theoretical framework of lexical functional grammar. He has classified complex predicates into nominal, verbal, adjectival and adverbial complex predicates.

Sharma (2056 VS) has discussed nine parts of speech categories, namely nouns, pronouns, adjectives, adverbs, verbs, postpositions, conjunctions, interjections and

particles extensively from traditional and descriptive approach. Each class is further sub-classified and discussed with illustrations.

Acharya (2058 VS) has classified words into various classes on different bases, viz: original and loan words; underived and derived; compound and reduplication; declinable and indeclinable from traditional approach. Nouns, pronouns, adjectives, verbs, adverbs are discussed with illustration.

Dhakal (2058 VS) has analyzed the Nepali numerical words from a historical perspective. He has compared the Nepali word forms with Sanskrit and Prakrit and has also shown the changes that occurred during the evolutionary period. He has also analyzed numerical words into their component parts and has observed the sound changes.

Pokharel (2010a) has analyzed noun class agreement system in Nepali from descriptive point of view. On the basis of analysis, Nepali nouns are grouped into eleven agreement classes. The gender assignment in Nepali is 'strictly semantic'. The use of classifiers for gender distinction is a unique feature not commonly found in languages. Nepali has based nominal agreement on human vs. non-human distinction.

Pokharel (2010b) has presented various strategies to derive the verb root in Nepal. Derivation from citation form, imperative singular form and probalilitative singular form are compared. None of the strategies can derive the entire verb roots, so he has proposed the mathematical strategy of generative phonology to this problem. According to this, if all the verb forms of a root are taken together and calculated the highest common factor, it will generate the root form and this will be the general formula of verb root derivation in Nepali.

b. Nepali computational morphology

Keshari et al. (2005) has discussed the development of a rule based system that guesses the part of speech of words in Nepali in a raw corpus without the use of lexicon. The system uses the linguistic information at morphological level and guesses the POS by looking at the affixes. The system has three modules, namely, lexicon maintainer module, rule maintainer module and POS guesser module. The modules interact with lexicon database, guessing rules and corpus.

Upadhyaya et al. (2005) has developed a morphological analyzer for Nepali language. The finite state technology is used in designing the analyzer but it can handle only the surface forms but not on the lexical level. It also lacks the detailed description of various aspects dealt in the process.

Aryal et al. (2006) has developed a system that produces the parsed text with maximum possible POS tag from computational perspective. The process consists of three phases: tokenizing into syllables, morphological analysis and disambiguation.

Paudel et al. (2006) has developed a morphological analyzer along with a spell checker. Nepali words are categorized into two major types, namely, declinable words and indeclinable words. The declinable and indeclinable words are further grouped into subclasses. The morphological analyzer consists of a root word dictionary and a rule dictionary. In the main engine, the root dictionary and rule dictionary interact with one another at various levels and spell checking has also been done in a coordinated manner.

Bal (2007) is an attempt to analyze the structure of Nepali grammar from the computational perspective. Even though the main focus of the paper is on the morphological and syntactic aspects of the Nepali language, it has also given a space for writing system of Nepali language. Despite of a novel start, it lacks detailed and deeper observation into the morphological structure of the words in Nepali.

Bal and Shrestha (2007c) discuss the design and implementation issues as well as the linguistic aspects of the morphological analyzer and a stemmer for Nepali language. The stemming algorithms and their limitations have also been discussed.

Bal and Shrestha (2007b) has presented a stemmer for Nepali language. This stemmer is especially designed to assist the morphological analysis and parts of speech tagging. Based on paradigm approach the stemmer is capable of splitting the words into meaningful units.

Aryal et al. (2007) has presented the techniques used for the syntactic and semantic disambiguation for Nepali language. The process of parsing has been done with two components, namely tokenizer and morphological analyzer. The work includes both syntactic and semantic levels.

Prasain (2008) is an attempt to analyze the Nepali basic verbs from computational perspective. The basic verbs are classified into two broad categories in terms of the

ending, viz., consonant ending and vowel ending. The former group has two types of ending: voiced and voiceless and latter has five types: *i*-ending, *u*-ending, *a*-ending, *ʌ*-ending and vowel sequence ending. The analysis uses the finite state technology; since it is a preliminary work in nature, it does not cover all the aspects of the verbs.

Shrestha (2008) has developed a system that disambiguates the Nepali word senses from natural language processing perspective. He has used modified Lesk algorithm and the wordNet. The processes in Nepali word sense disambiguation have been completed in four stages, namely, (i) tokenizer, (ii) context selection, (iii) finding the senses of target word and (iv) sense identification.

Hardie (2008) has analyzed Nepali postpositions applying a collocation-based technique to the categorization of postposition in Nepali from corpus linguistic perspective using the Nepali National Corpus. He has examined the most significant collocations of several postpositions for patterns that characterize postposition as a category or categories. The collocation with semantically coherent nouns, and collocation with words for which the postposition functions as a subcategorizer are identified.

Hardie et al. (2009) has described the linguistic rationale underlying the part-of-speech tagset used for tagging the Nepali National Corpus. The implementation of the tagset in an automated tagging system has also been outlined. This work further supports the classification of words into various groups for designing the finite state transducer for each of the groups.

Prasain (2010) is an attempt to analyze Nepali basic nouns and implement them into computer using finite state approach. Various noun characteristics: number, gender, form, honorificity, augmentative/diminutive and significant stem finals are analyzed. On the basis of these features, Nepali basic nouns are grouped into fourteen classes. Each group of nouns are implemented following xfst format (Beesley and Karttunen 2003): lexc grammar to create lexicon finite state transducer and xfst interface to create rule finite state transducer. And finally these finite state transducers are composed into one to create a single finite state transducer which can directly be used to analyze and generate the basic nouns.

c. Nepali language and related NLP works

Bandhu (1971) is a computer concordance of Nepali spoken corpus. The corpus has been morphologically analyzed and forms are segmented according to their functions. Most of the data for the collection were collected from Palpa district (January, 1971), Syangja and Pokhara (January, 1970) and Gorkha district (April-May 1971). Now, this corpus is available at [<http://cqpweb.lancs.ac.uk/bandhu/index.php>] and the information such as title, speakers list, text type and POS tags based on Nelralec tagset are available.

Gurung and Khatiwada (2007) is an attempt to analyze Devanagari script used in Nepali writing system for the collation sequence. The study also discusses the development process of a lexicon for Nepali language. To make the lexicon computer readable, the XML format is used. Each entry is provided with pronunciation, syllable break, parts of speech, meaning, and synonyms. And the framework of Hunspell is used so that it could be used in the spellchecker in OpenOffice.

Bista et al. (2007) has presented the Nepali lexicon development process to be used in spell checking system for Nepali language. This paper also reports about the collection of words and the tools used the problems and issues faced during the process of lexicon development. The architecture consisting of various modules such as Lexicon database, Lexicon maintainer, Rules database, Rule maintainer, Corpus and Rule Interpreter. These modules interact with other concerned modules as required during the process of developing the lexicon.

Bal and Shrestha (2007a) has developed simple spellchecker to be used in Nepali OpenOffice.org. The head words are stored in a file and the affix rules in another file.

Bal et al. (2007) discusses a general overview of the technical and linguistic research and development being carried out for the development of Nepali spellchecker 1.1. OO.org using HunSpell framework with Unicode support. The dictionary is populated with stems of nouns, verbs, pronouns, adjectives, adverbs, conjunctions, interjections, particles, postpositions and compound words. The possible word forms are generated applying the affix rules.

Hyoju and Shrestha (2007) presents an overview of the contemporary Nepali dictionary based on Nepali national corpus. The various components that incorporated in this dictionary are the headword, part of speech, phrase category, guide word,

pragmatics, definition, example, usage note, various form, suppletive form, extra information, phrase, idiom, compound, proverb and cross-reference.

Gurung and Thapa (2007) has described the process of building text-to-speech for the Nepali language from speech processing perspective. The multilingual speech synthesis system known as Festival has been used. A component of Festival called Festvox provides a framework in building synthetic voices. Corpus based rule generation and statistical modeling methodologies are used. The sentences for building speech database are taken from the Nepali National Corpus. The normalized text is fed to module where generation of wave form takes place using letter-to-sound rules, concatenation of diphones and the pitch extraction.

Yadava et al. (2008) describes the construction of the 14-million-word Nepali National Corpus (NNC) (<http://cqpweb.lancs.ac.uk/nncv2/index.php>) which includes spoken corpus, written corpus, parallel corpus and speech corpus. The NNC is encoded as Unicode text and marked up in CES-compatible XML and follows FLOB and Frown frameworks.

d. NLP works in some selected languages

Megerdoomian (2003) provides a detailed description and analysis of Persian inflectional morphology from a computational perspective. The morphological analyzer designed for Persian language uses a unification-based grammar with typed feature structure. The linguistic analysis and implementation to the Samba Grammar for developing the morphological analyzer are main tasks. The surface form is formally represented as a regular expression. The morphological features are specified as a feature structure that contains the lexical and inflectional information provided by the rule. These features describe how the stem and the morphological features of the affixes are combined.

Hussain (2004) has developed a finite-state morphological analyzer for Urdu. She has described the general morphological concepts such as morpheme, roots, bases, affixes, inflection, derivation and causation. The analysis has been done following the two-level morphology formalism using finite-state transducer.

Makedonski (2005) is a finite state approach to the inflectional morphology of Turkish nouns. The finite state transducer is used for analyzing the nouns and the

implementation is done in Xerox Finite State Toolbox in two levels namely the lexicon and rule component.

Ziai (2006) has developed a finite state morphological analyzer for Persian simple verbs from finite state technology approach. The system presented covers the full inflectional paradigm of modern Persian for both regular and a large number of irregular verbs.

Islam (2007) describes the inflectional Bangla verb and noun morphology and also mentions the rules, lexicons and grammar for Bangla morphological analysis. This analysis is based on PC-KIMMO, a two level morphological analyzer.

Dasgupta et al. (2007) has discussed the inflectional behaviors of the compound words in Bangla language from computational perspective. The Bangla compound words may retain the inflectional suffixes on both the constituents and the resultant compound that may further be inflected as a single word.

Khan and Fatima (2007) investigate the inflectional properties of Pashto nouns from finite state perspective. The main focus is on the classification of the Pashto nouns. The finite state transducer is used for analyzing the Pashto nouns.

Bharati and Kulkarni (2007) has discussed the importance of Paninian grammar from the perspective of information coding. The study has applied the finite state technology. The theoretical and practical aspects of computational linguistics concerning Hindi and Sanskrit and application of Paninian approach to English language are highlighted. The complexity of word formation in Sanskrit is captured by a finite state automata, analyzer for Sanskrit has been developed which provides the output with morph analysis.

Bögel et al. (2007) discusses a number of issues, in particular, potential ambiguity and non-concatenative morphology . This approach deals with the treatment of both Urdu and Hindi via a cascade of FSTs that transliterates the very different scripts into a common ASCII transcription system; and the implementation of the analysis is based on the xerox finite state toolkit.

Shrivastava et al. (n.d.) has developed a rule-based part-of-speech tagger for Hindi with stemmer and morphological analyzer. The developed stemmer and morphological analyzer are integrated with Hindi WordNet, Hindi Generation and Question Answering Projects.

1.5 Significance of the study

It has been clear from the literature review (1.3) that the Nepali language is primarily described from two approaches: notional and descriptive. There have been very few and sporadic works done from formal and computational perspective. The numbers of natural language processing works are growing in many languages of the world and in South Asian languages in particular. In this context, the Nepali language is lagging behind. Therefore, there is an urgent need to develop the computational models and implement it in computing processes so that various kinds of computer applications such as spell checker, grammar checker, part of speech tagger and syntactic parser can be developed. When the applications related to Nepali language are developed, specially, end users who are seeking the information stored in Nepali language (i.e. texts) can be benefited. In this regard, this work can be a foundational and very much useful. The computational analysis of morphology in Nepali would be a central and essential component for the development of various kinds of other Nepali language processing applications. Further analysis of linguistic levels such as syntax, semantics, and pragmatics can also be done taking this work as the reference point. Therefore, this study can be of a great importance by itself and much more useful for both academicians and practitioners of natural language processing.

1.6 Research methodology

The methodology consists of data collection, classification, analysis and implementation.

Data collection: The study is primarily based upon secondary data for morphological analysis; however the example sentences illustrated are elicited. The secondary data, especially word-forms have been taken from the Nepali National Corpus developed by Bhashanchar Project (Nelralec) and cross-checked them from 'Brihad Nepali Sabdakosh' (Pokharel et al., 2040 VS). Being a native speaker of Nepali, I have also used my intuition for cross checking and analyzing the data.

Classification: The unique word-forms are classified into different categories such as nouns, verbs, adjective, etc. and further subdivisions have been made according to their morpho-syntactic behaviors.

Analysis: The classified data are analyzed into stems and affixes for each category and the inflectional and derivational processes have been treated separately. Then phonological rules have been identified and formalized.

Implementation: Finite state transducers for each group of words have been created following concept of ‘two-level morphology’. Then, a computational model for Nepali morphology has been implemented by using the tool referred to as Xerox Finite State Tool (XFST) developed by Beesley and Karttunen (2003). See Chapter 2 for detailed description of the theoretical framework.

1.7 Limitations of the study

This study has dealt mainly with the written form of words in Nepali. Only inflectional and derivational aspects of the words have been taken care of in this study. Compounding and reduplication are also morphologically important, but they have not been dealt with in this study. Despite the fact that there are a number of models/approaches for computational analysis in the literature, only finite state approach is employed in this study. Moreover, only the representative words have been considered in the implementation of the analysis.

1.8 Organization of the study

This study has been structured into eight chapters. Chapter 1 introduces the concept of computational analysis of word level categories following the Two-level morphology. This chapter also deals with the statement of the problems, objectives, review of literature, research methodology and justification of the study. In Chapter 2, we present theoretical framework for the study. Chapter 3 looks into the general characteristic features and analyzes nominals in Nepali computationally and presents finite state transducer for each of them. In this chapter, we also deal with the phonological rules in the form of regular expressions which are implemented later on. In Chapter 4, we discuss the various features possessed by the verbs. It also presents finite state transducers and phonological rules involved in the verb morphology. Chapter 5 analyzes adverbs, conjunctions, case markers, postpositions, particles and interjections. Separate finite state transducers for each category have been presented in this chapter. In Chapter 6, we present the analyses of various derivational systems

in Nepali. The finite state transducers, phonological rules have also been presented in this chapter. Chapter 7 implements all those analyses done in the preceding chapters. We have summarized the study in chapter 8. And finally, a number of annexes have been provided.

CHAPTER 2

THEORETICAL FRAMEWORK

2.0 Outline

This chapter presents the theoretical framework employed in this study. It consists of eight sections. Section 2.1 deals with general idea of computational concepts. In section 2.2, we deal with the concept of regular expression. In section 2.3, we briefly discuss the finite state technology. Section 2.4 introduces a brief idea of regular language. Section 2.5 presents the introduction of finite state machine which includes finite state automata, finite state transducer and some relevant and important operations that can be performed on finite state transducer. Section 2.6 deals with the use of finite state transducer in computational morphology in natural language. It also briefly shows the relation among regular expression, regular language and finite state network. In section 2.7, we discuss the basic concepts and application of Xerox finite state tool (xfst) used in this study for the development of computational model of Nepal morphology that can be used as morphological analyzer. In Section 2.8, we present the summary.

2.1 Computational concepts

Finite state morphology has been an important and active field of research and development for a number of decades. The Natural Language Processing (NLP) system remains incomplete without morphological analysis. The words are the units of syntax and meaning of word is the basis of semantics. In fact the input to the syntactic and semantic analysis comes from the morphology. Therefore, the morphological analysis of a natural language has become important and fundamental.

Relating word forms and detecting the structure of word forms are what morphological analysis is all about. The task of relating a given form to a canonical form is called lemmatization. Both lemmatization and the decomposition into parts have their uses, however, they share some common processes. The task of morphological analysis, then, is to take forms and relate them to other word forms, at the same time deriving featural information about the form (Roark and Sproat 2006).

It is customary in discussion of morphology to talk about inflection versus derivational morphology in terms of the kinds of features each of these encodes. This distinction is not relevant here for discussion. Rather we will concentrate purely on the computational mechanisms for performing morphological analysis and the way these mechanisms represent two kinds of linguistic information. The formal properties of morphological operations, viz. the syntagmatic combination of morphological elements and the paradigmatic relation between the forms are the crucial aspects (Roark and Sprout 2006). To realize this objective, one needs to understand some mathematical and computational notions and operations which are introduced in the subsequent sections.

2.2 Regular expression

Regular expressions are the standard notation for characterizing text sequences and it is used for specifying the text strings in searching text (Jurafsky and Martin 2000:48-59). They are highly applied in various natural language processing activities such as information retrieval, word-processing, computation of frequencies from corpora and other such tasks.

A regular expression is a formula and a special language that is used for specifying classes of strings. A string is any sequences of alphanumeric characters or symbols for the purpose of most text-based search techniques. The set of strings in the regular expression has a pattern which is actually a value for the algebraic formula (Siddiqui and Tiwari 2008:54-9). The regular expressions are kept between the slashes to distinguish them from other ordinary set of characters. Table 2.1 lists some of the simplest regular expressions and the matches in the text.

Table 2.1: The sample regular expressions

Regular expressions	Example pattern matched
/a/	There is <u>a</u> dog.
/book/	I have read many <u>books</u> .
/घर/ 'house'	तिमी <u>घर</u> मा बस। 'You stay at home.'
/म/ '1SG'	<u>म</u> लाई भोक लाग्यो। 'I am hungry.'

The simple regular expressions in Table 2.1 are used for searching the text. The regular expression in the left can search the underlined text in right.

Formally, regular expressions are an algebraic notation for characterizing a set of strings. Thus, they can be used to specify search strings as well as to define a language in a formal way. The characters are grouped by putting them between square brackets. For example, the pattern `/[कखगघड]/` will match any one of them. The square brackets specify the disjunction of the characters used within the square brackets. A dash '-', which specifies a range, can be used when the set of the characters within the brackets is very big. For example, `[a-z]` specifies any lowercase letter of Latin alphabet and `[0-9]` specifies any digit from 0 to 9. Some important operators used in the regular expressions, patterns and their meanings are listed in Table 2.2.

Table 2.2: Some operators used in regular expressions

Operators in RE	Pattern	Meaning
[]	[abc]	a or b or c
[-]	[A-Z]	any one of the capital letters
^	[^a-z]	not a lowercase letter
*	ab*c	zero or more bs
.	a.c	any character between a and c
?	ab?c	either zero or b in between a and c
+	ab+	one or more bs
	a b	either a or b
()	appl(y ies)	apply or applies
{n}		n occurrences
{n,m}		from n to m occurrences
\n		a new line
\t		a tab

Source: Jurafsky and Martin 2000

The operators illustrated in Table 2.2 are used for creating the complex regular expressions.

The simple regular expression, i.e, any alphanumeric alphabets and various operators, can be combined together and a very complex regular expression can be constructed according to the requirement.¹

Now, it is clear that a regular expression requires a pattern of the text to be searched and a corpus of texts to search through. And finally, a regular expression search function will search through the corpus returning all the texts of that pattern provided. Thus, a regular expression specifies a language according to its pattern, the complexity of the language that is represented depends on the complexity of the pattern used to specify the language.

However, the regular expression is more than just a convenient metalanguage for text searching. Firstly, a regular expression is one way of describing a finite state automaton. It means the regular expression can be compiled into a finite state automaton. Secondly, it is a way of characterizing a particular kind of formal language called a regular language (see 2.6).

2.3 Finite state technology

In order to understand how to build the linguistic application, we first need to be acquainted with the basics of how a finite-state machine works. A finite-state machine is a network consisting of states indicating one start state and one or more final states. Transitions between states are possible only if the required input is recognized. A path is a sequence of transition over arcs to a particular state. In computational morphology, a path is a set of alphabets equivalent to a word in natural language. So, it can be said that the technology that utilizes the finite-state network in the processing of creating an application is said to be a finite state technology.

Therefore, the basic concept behind the finite state technology is a set of states with different properties and a set of arcs that connect these states. The arcs have a direction and an input symbol. That means there is a set of outgoing arcs with their respective input symbols. The sets of these states and arcs together form a network.

As Chomsky (1957) stated, the finite state devices were limited in generative capacity i.e., the power to accurately describe all natural language phenomena. Therefore,

¹ A detailed description, how a complex regular expression is created, is out of the scope of this study.

finite state technology was considered to be inefficient by the linguists at the earlier stages of its development. One reason was that it is a mathematical formal abstract device, so it was believed that it doesn't have the descriptive power for natural language analysis. The second reason was that in its developing stages, it was not really powerful to account for the linguistic phenomena. But later on, it proved to be quite useful in modeling the parts of languages that could be considered as finite and regular. As far as the natural language is concerned, it shows this quality of being regular and finite at least in its parts if not in whole.

Various tasks such as POS disambiguation, tokenization, shallow parsing, etc. are successfully accomplished using the finite state technology. Morphology is the core component of the natural language and it can be considered more or less finite and regular. Thus, the most significant application of the finite state technology has been the computational morphology in which both analysis and generation of the morphology of the natural language is performed. The computational morphological analysis has been the basis for any further kind of natural language processing.

2.4 Regular language

As discussed in section 2.2, a regular expression denotes or specifies or describes a regular language using a specified pattern. A formal language is a set of strings each of which is composed of symbols from a finite symbol-set called an alphabet (Jurafsky and Martin 2000:48). A regular language is a formal language that is possibly an infinite set of finite sequences of symbols from a finite alphabet that satisfies particular mathematical properties: "A class of languages that are definable by regular expression is exactly the same as the class of languages that are characterized by finite-state automaton is said to be a regular language" (Jurafsky and Martin 2000:75). Table 2.3 illustrates the regular language from Nepali defined by a regular expression.

Table 2.3: Regular expressions and regular language

Regular Expression	Regular Language
$/\text{खा} \cdot * \text{छ} \cdot */$ $'k^h a \cdot * t s^h *'$	खान्छ $k^h a n t s^h \Lambda$, खान्छौ $k^h a n t s^h \Lambda \tilde{u}$, खान्छु $k^h a n t s^h u$, खान्छन् $k^h a n t s^h \Lambda n$, खाइन्छ $k^h a i n t s^h \Lambda$, खाएछ $k^h a e t s^h \Lambda$, खाइरहन्छ $k^h a i r \Lambda f i n t s^h \Lambda$, खानुहुन्छ $k^h a n u f i u t s^h \Lambda$, खाएछन् $k^h a e t s^h \Lambda n$, खान्छ्यौ $k^h a n t s^h \Lambda u$, खाएँछु $k^h a \tilde{e} t s^h u$ खाँदैछस् $k^h \tilde{a} d \Lambda i t s^h \Lambda s$, खान्छेस् $k^h a n t s^h e s$,

The language in right column of Table 2.3 is a regular language denoted by a regular expression in left column. All the strings (words) matched by a regular expression $/\text{खा} \cdot * \text{छ} \cdot */$ $'k^h a \cdot * t s^h *'$ have the same pattern.

2.5. Finite state machine

2.5.1 Finite state automata (FSA)

Finite state automata are a mathematical abstract device. As discussed in (2.3), they consist of states and arcs called transitions. Each FSA has exactly only one initial state and one or more final states. In between initial and final states, there can be any finite number of states called intermediate states. The transitions are the connections between these states and thus responsible for moving from one state to another. Conventionally, the states are represented as circles and the transitions between them are represented as labeled arcs; and an arrow is used to indicate the initial state and double circles are used to indicate the final states. The finite state automata are best understood as recognizers because they accept a finite set of input strings. For illustration, an automaton that accepts a string from the Nepali language घर 'house' and घरहरू 'houses' is visualized in Figure 2.1.

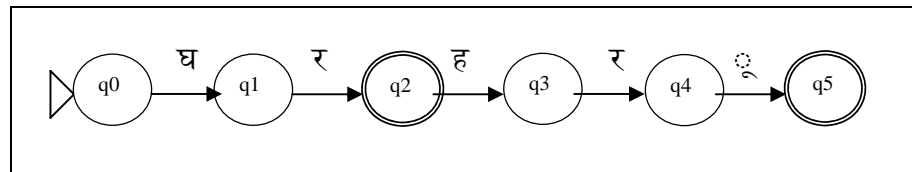


Figure 2.1. A finite state automaton that accepts घर 'house' and घरहरू 'houses'

The finite state automaton shown in Figure 2.1 recognizes the words by reading the input string symbol-by-symbol and matching the symbols to the labels on the arcs. This FSA accepts घर 'house' and घरहरू 'houses' because the inputs lead to final states. No other strings are accepted by this FSA.

Formally, the finite state automaton can be defined by the following five parameters:

- Q: a finite set of N states q_0, q_1, \dots, q_N
- Σ : a finite input alphabet of symbols
- q_0 : the start state
- F: the set of final states $F \subseteq Q$
- $\delta(q,i)$: the transition function or transition matrix between states. Given a state $q \in Q$, δ is thus a relation from $Q \times \Sigma$ to Q (Jurafsky and Martin 2000:62).

For the language automaton in Figure 2.1, $Q = \{q_0, q_1, q_2, q_3, q_4, q_5\}$, $\Sigma = \{घ, र, ह, ्र, ्रु\}$, $F = \{q_2, q_5\}$ and $\delta(q,i)$ is defined by the transition in Table 2.4.

Table 2.4: The transition table for घर and घरहरू

State	Input				
	घ	र	ह	्र	्रु
0	1	\emptyset	\emptyset	\emptyset	\emptyset
1	\emptyset	2	\emptyset	\emptyset	\emptyset
2:	\emptyset	\emptyset	3	\emptyset	\emptyset
3	\emptyset	\emptyset	\emptyset	4	\emptyset
4	\emptyset	\emptyset	\emptyset	\emptyset	5
5:	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset

2.5.2 Finite state transducer (FST)

The finite state automaton discussed in (2.5.1) accomplishes the task of recognizing strings in a regular language by providing a way to systematically explore all the possible paths through a machine (Jurafsky and Martin 2000:97-108). However, this exploration can only address the problem whether the string is present in its language or not. The automaton of this capacity cannot be used to show the relation between

two or more languages. However, this problem can be solved by the use of another version of the FSA called ‘Finite State Transducer’. An FST is similar to an FSA; it consists of states and transitions with labeled arcs. However, in an FST the labels can be in a pair of symbols, i.e., the relation between two languages, instead of simple symbols. Whenever an arc has such a label, it is traversed and the input symbol matches, then it is transduced to the output symbols (Makedonski 2005; Ziai, 2006). Consider the example transducer in Figure 2.2, in the upper side is labeled as घर+NOUN+SG and lower side is labeled as घर. This FST transduces घर+NOUN+SG to घर and vice versa. That means, the input घर is matched and it outputs घर+NOUN+SG.

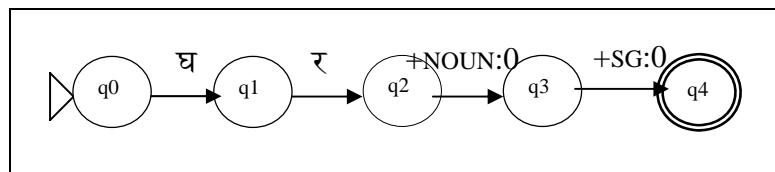


Figure 2.2: A finite state transducer that transduces between घर ‘house’ and घर+NOUN+SG

Therefore, the important property of FSTs is that they are in principle bi-directional, meaning that they can also be applied backwards. Thus, the bi-directionality feature of the FST can be applied to the morphological analysis and generation.

2.5.3 Some important operations on FSTs

a. Union

The union of two or more networks is another set that contains all the elements of constituent networks. There is no ordering of the arcs in the network and it is denoted as $[A|B]$, where A and B are the networks (Beesley and Kartumnen 2003). To illustrate this operation, there are three FSTs for nouns, adjectives and adverbs which are illustrated in upper part of the Figure 2.3. When operation union is performed on these three FSTs, it results into a single FST, as illustrated in the lower part of Figure 2.3.

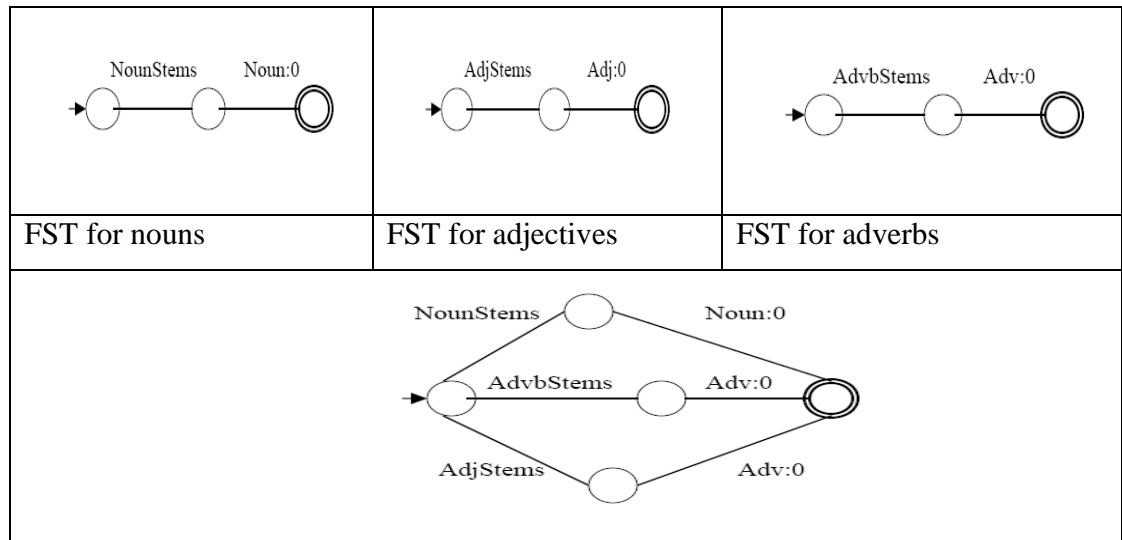


Figure 2.3: FST unioned from three FSTs for nouns, adjectives and adverbs

The Figure 2.3 shows the process of unioning two or more finite state transducers into a single network which contain the elements of its constituent finite state transducers.

This union operation is very much useful and powerful to create the large and complex FST from smaller FSTs of the morphological word classes. Therefore, it allows working in the modular concept.

b. Concatenation

Concatenation is an operation which keeps the networks in sequence. One can also concatenate two existing finite state networks with one another to build up new words productively or dynamically (Beesley and Kartumnen 2003). This is usually denoted as [A B] where A and B are the networks. This phenomenon is illustrated in Figure 2.4.

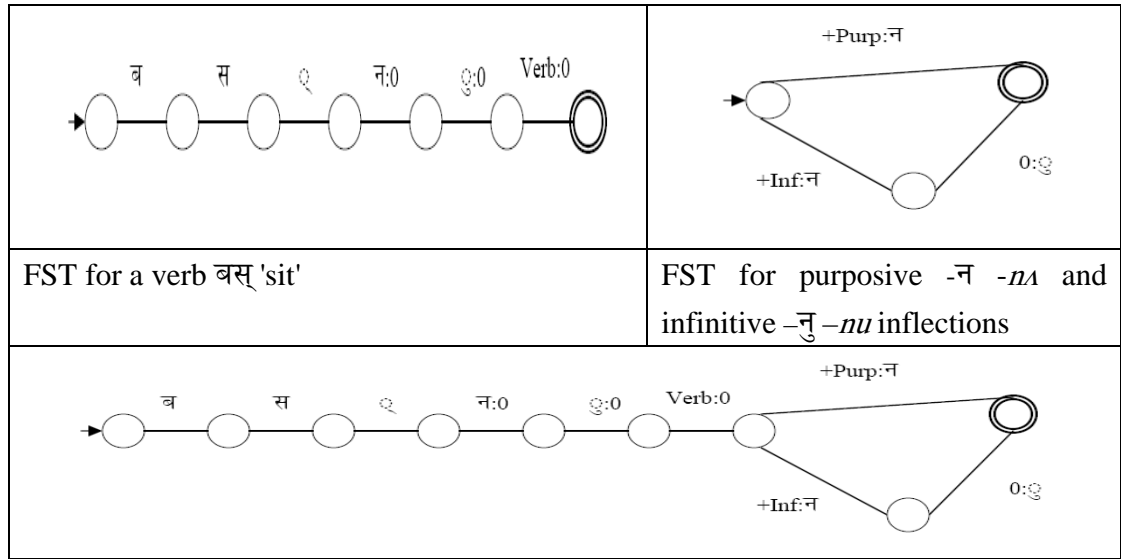


Figure 2.4: A finite state transducer concatenated from two FSTs above

In Figure 2.4, there are two FSTs in the upper part of the Figure 2.4, one for a Nepali verb बस् *bAs* 'sit' and another for purposive -न -*nA* 'PURP' and infinitive -नु -*nu* 'INF' suffixes. And in the lower part of the Figure 2.4, there is an FST resulted from concatenating two FSTs, which can analyze and generate purposive and infinitive forms of the verb बस् *bAs* 'sit'. This concatenation operation is useful in handling the verb stems and inflectional and derivational suffixes.

c. Composition

Composition is an operation on two or more languages or relations. It is usually denoted as $[A \circ B]$. In fact, this operation removes the common elements from the networks used for composing (Beesley and Karttunen 2003). For exemplification of this operation, there is a rule FST at the left top in Figure 2.5 which changes *ो* into *ा* for plural feature. An arbitrary symbol +MP is used for creating the environment so that the rule can be applied to specific group of nouns. Figure 2.5 illustrates the process of composition.

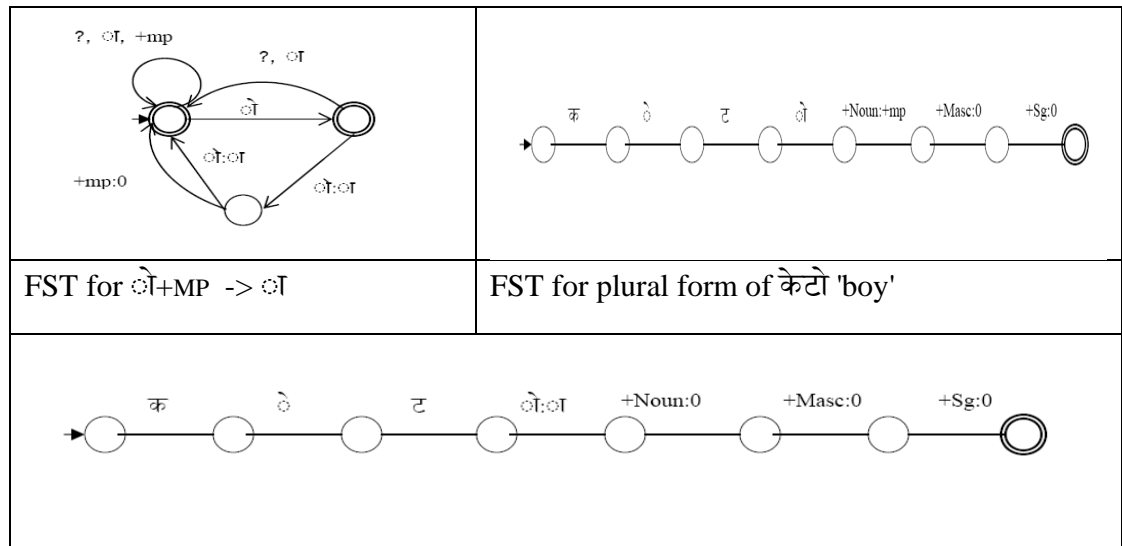


Figure 2.5: A finite state transducer from composing two FSTs above

At the right top of the Figure 2.5, there is an FST for $केटो$ 'boy' with +mp symbol. When these FSTs are composed, it results into a single FST in lower part of the Figure 2.5 which is capable of changing $ो$ *o* into $ा$ *a* for plural feature and also removes the arbitray symbol +mp without any intermediate FSTs.

In fact, composition operation forms a sequence of transducers. It builds a cascade of FSTs into a single one by eliminating the common intermediate outputs, so, it allows working for a modular structure. Because of this feature of composition, it has been very much useful for composing rules with lexicon to obtain the correct surface forms.

d. Intersection

The intersection of two networks contains the set containing all the members that are common to both. It is usually denoted as $[A \& B]$. This operation might not be used as a major operation in this work. But it can be used to find the common words between two sets of words.

e. Subtraction

The subtraction of two networks contains the set containing elements that are in A but not in B. It is denoted as $[A-B]$. This operation is normally performed to find the words in a network which are not in another network.

f. Complementation or negation

The complement language of a network A is the set of all strings that are not in the language A. It is usually denoted as $\sim A$. This operation is very useful for filtering the words from a network.

Among the operations discussed above, operations union, concatenation and composition are used while implementing the analyzed morphological categories and rules to create a single lexical transducer while others are used elsewhere.

2.6 FST in computational morphology

Johnson (1972) was the first to prove that the finite state technology was appropriate and applicable to certain areas of computational linguistics. The most important thing that he discovered was that the language and relations used in traditional rewrite rules of generative phonology were essentially as powerful as the mathematical devices used in the finite state calculus. Kaplan and Kay (1994) showed with a detailed mathematical proof that every rewrite rule corresponds to a regular relation and thus can be modeled by means of an FST. The two-level morphology (Lexical Level and Surface Level) developed by Koskeniemmi (1983) also used the similar concept of an FST.

One of the fundamental results of formal theory (Kleene 1956) is the demonstration that finite-state languages are precisely the set of languages that can be described by a regular expression. Figure 2.6 demonstrates relation among FST, regular language and regular expression.

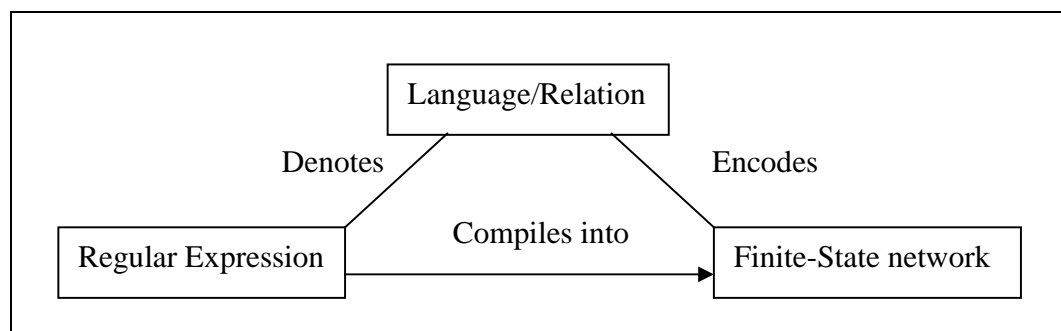


Figure 2.6: The interrelation among language, regular expression and finite state network

Figure 2.6 indicates, a regular expression denotes a set of strings (i.e., a language) or a set of ordered pairs (i.e., a relation). It can be compiled into a finite-state network that compactly encodes the corresponding language or relation that may well be infinite (Beesley and Karttunen 2003:44).

The language of a regular expression includes the common set of operators of Boolean logic and operators such as concatenation that are specific to strings. Each of the regular expression operators for finite-state languages there is a corresponding operation that applies to finite-state network and produces a network for the resulting language.

A finite-state network for a complex language can be built by first constructing a regular expression that describes the language in terms of set operations and then compiling that regular expression into a network. This is, in general, easier than constructing a complex network directly and in fact it is the only practical way for all but the most trivial infinite languages and relation.

2.7 Xerox finite state tool syntax (XFST)

XFST used here for the computational analysis of Nepali morphology, developed at the Xerox Research Center Europe is based on Beesley and Karttunen (2003). It implements the standard finite state operations such as composition, concatenation, complement and union as well as several innovative operations like replacement rules and local sequentialization. XFST includes: *lexc* – a compiler for lexicons in the *lexc* language, which is specifically designed for handling morphotactics (the syntax of the morphemes) in natural languages (see 2.7.1), and *xfst* – the core tool providing an interface to the finite state calculus for building, accessing, manipulating finite state networks and a compiler for regular expressions and replacement rules which will be essential for this work (see 2.7.1). There are other run time tools within it but they are not relevant for this discussion.

XFST defines transducers as relations between two languages. What would be referred to as an upper language could be thought of as the input and the lower language then would be the output when an input is applied to a transducer downwards. If we apply input to the transducer upwards then the roles switch – the input is applied on the lower side and the output comes from the upper side. Although

it seems a bit confusing the terms upper and lower remain constant (Beesley and Kartumnen 2003:85-202). In the definition of a lexical transducer, the upper side language describes the lexical (underlying) forms of the language to be analyzed and the lower side language contains the actual surface forms in the written forms.

XFST has many operators used while analyzing the natural languages, but some important and essential ones are discussed as follows:

(i) “ : ” The crossproduct operator relates every symbol in the language in its left side to every symbol on its right side. For example [$\overline{\text{घर+NOUN+PL}}$: $\overline{\text{घरहरू}}$], the square brackets indicate grouping. The language $\overline{\text{घर+NOUN+PL}}$ is in the upper side of the transducer and $\overline{\text{घरहरू}}$ is in the lower side of the transducer.

(ii) “ -> ” The left-to-right replacement operator is an extended regular expression operator that provides for convenient formulation of rewrite rules in XFST. For example, a rule that replaces every x by a y might be written as $x \rightarrow y$. Every symbol that is not an x will be left unchanged. In generative phonology, rewrite rules have a context part which will cause the rule to apply only if the context is satisfied. XFST provides for that also. For example, if all x s are to be replaced by y s but only in the environment where x s are followed by z . One could formulate the rule like $x \rightarrow y \parallel _ z$. The double bars indicate the beginning of the context.

(iii) “ \$ ” The ‘contains’ operator denotes the language of all the strings containing in a variable. For example, $\$a$ means the language of all strings containing a . This operator can be used to operate on a specific set of strings in the network.

(iv) “ | ” The union operator creates the union of two languages or relations. For example, A and B are two languages and $A|B$ means the language formed from the union of A and B (see 2.5.3).

(v) “ .o. ” The composition operator is very powerful for the combination of two or more transducers into one. Thus, a very complex network is possible by use of this operator. The main use of this operator is to compose the rules FSTs with lexical FST (see 2.5.3).

(vi) “ # ” This symbol is used for two purposes. One in lexc files to indicate the final state (see 2.7.1) and another in the replacement rules to indicate the word boundary where it is surrounded by dots (see 2.7.2).

2.7.1 LEXC grammar

A lexc grammar consists of at least one lexicon (called Root). A lexicon contains a list of entries where each entry has a continuation class. It corresponds to a state. Each entry corresponds to a labeled arc that can be traversed only if the entry is successfully matched against the input string. The entry can be a regular expression. If an entry is matched, the arc is traversed and the continuation class, which is another state, is reached. The procedure is repeated until a final state is reached, denoted by the special continuation class # (Beesley and Karttunen 2003:203-278). The structure and its components of the lexc grammar are presented in Figure 2.7 and discussed.

a. Multichar_Symbols

First of all, there is a set of multicharacter symbols definition such as +NOUN +MASC +HUM +SG +PL +DIM +ERG +DAT +LOC +ABL +NOM +GEN +POP +OBL +FEM where sequences of symbols in the set are treated as atomic symbols. These symbols are primarily used as tags to indicate various grammatical categories and features. They are attached on the upper side that is visible only if morphological analysis is performed and on the lower side, each multicharacter symbol corresponds to respective suffix or epsilon. But, sometimes, additional tags are also used to create an environment for the replace rules and they are removed at the end. Figure 2.7 shows the structure of lexc grammar.

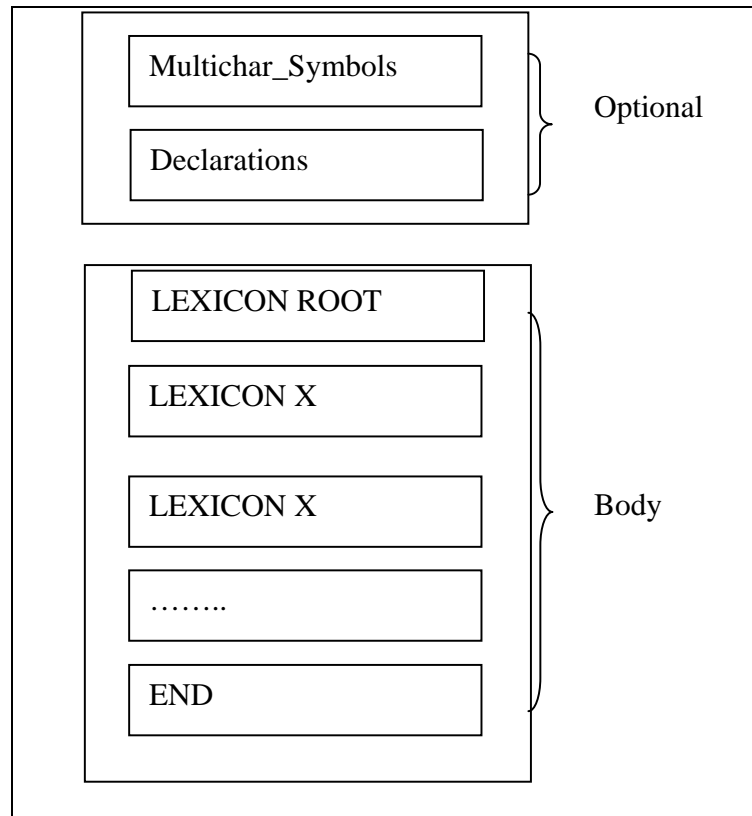


Figure 2.7: The structure of lexc grammar (Beesley and Kartumnen 2003:205)

In Figure 2.7, the lexc grammar begins with optional components Multicharacter symbols and declarations, and body consisting of list lexicons

b. Definitions

After the Multichar_Symbols section, an optional definition section can also appear in the lexc file. It consists of the keyword Definitions followed by one or variable assignment.

c. Lexicon

As discussed in (2.7.1), the lexicon root corresponds to start state of the network to be compiled. There can be any number of other lexicons, but they must follow the lexicon root. Each entry consists of two parts: a form and a continuation class. The form can have formatives (i.e., a stem) or a regular expression and the continuation

class refers to the next sub-lexicon to be followed. The end of the word or final state is indicated by reserved symbol #.

The lexicon is designed according to the format shown in Figure 2.7 above. A sample lexicon of nouns in Nepali is given below which account for both *o*-ending nouns and non-*o*-ending nouns including number, gender, honorificity and augmentative/diminutive features. For the illustration, only one stem of a particular noun type and markers for stated features are included for this purpose.²

Multichar_Symbols +NOUN +MASC +FEM +OBL +PL +SG +DIM +VOC +HON
+MP +FE +PLACE +PROPER ^b³

LEXICON ROOT

Nouns;

LEXICON Nouns

!! Type 1a Nouns:

केटो inflection_1a; ! *keto* 'boy'

!!Type 1b Nouns:

मुसो inflection_1b; ! *muso* 'mouse'

!!Type 1c Nouns:

डालो inflection_1c; ! *dalo* 'basket'

!!Type 1d Nouns:

फोटो inflection_1d; ! *p^hoto* 'photo'

!!Type 21a Nouns:

काका inflection_21a; ! *kata* 'uncle'

!!Type 21b Nouns:

² Classification of nouns is based on their characteristic features. However, the names of the noun class and the sub-lexicon used in the lexc file are purely arbitrary as they are removed during the compilation processes.

³ Multichar_Symbols +MP, +FE and ^b are used in lower language to create environment and they are also removed later.

नाति inflection_21b; ! *nati* 'grandson'
!!Type 21c Nouns:
बाघ inflection_21c; ! *bag^h* 'tiger'
!!Type 21d Nouns:
बिस्ट inflection_21d; ! *bist* 'a surname'
!!Type 22a Nouns:
दाइ inflection_22a; ! *dai* 'elder brother'
!!Type 22b Nouns:
दिदि inflection_22b; ! *didi* 'elder sister'
!!Type 22c Nouns:
राम inflection_22c; ! *ram* 'Ram'
!!Type 22d Nouns:
सीता inflection_22d; ! *sita* 'Sita'
!!Type 22e Nouns:
खेत inflection_22e; ! *k^het* 'farm land'
!!Type 22f Nouns:
पोखरा inflection_22f; ! *pok^hara* 'Pokhara'

LEXICON inflection_1a
+NOUN+MASC+SG:0 #;
+NOUN+MASC+PL:+MP #;
+NOUN+MASC+OBL:+MP #;
+NOUN+MASC+HON:+MP #;
+NOUN+MASC+VOC:+MP #;
+NOUN+FEM:+FE #;
LEXICON inflection_1b
+NOUN+MASC+SG:0 #;
+NOUN+MASC+PL:+MP #;
+NOUN+MASC+OBL:+MP #;

+NOUN+FEM:+FE #;
 LEXICON inflection_1c
 +NOUN+SG:0 #;
 +NOUN+PL:+MP #;
 +NOUN+OBL:+MP #;
 +NOUN+DIM:+FE #;
 LEXICON inflection_1d
 +NOUN+SG:0 #;
 +NOUN+PL:+MP #;
 +NOUN+OBL:+MP #;
 LEXICON inflection_21a
 +NOUN+MASC:0 #;
 +NOUN+FEM:ी #;
 LEXICON inflection_21b
 +NOUN+MASC:0 #;
 +NOUN+FEM:नी #;
 LEXICON inflection_21c
 +NOUN+MASC:0 #;
 +NOUN+FEM:बिनी #;
 LEXICON inflection_21d
 +NOUN+MASC:0 #;
 +NOUN+FEM:ेनी #;
 +NOUN+FEM:िनी #;
 LEXICON inflection_22a
 +NOUN+MASC:0 #;
 LEXICON inflection_22b
 +NOUN+FEM:0 #;
 LEXICON inflection_22c
 +NOUN+PROPER+MASC:0#;
 LEXICON inflection_22d
 +NOUN+PROPER+FEM:0 #;


```
LEXICON  inflection_22e
+NOUN:0  #;
LEXICON  inflection_22f
+NOUN+PLACE:0  #;
END
```

2.7.2 XFST interface

The xfst part of this system is mainly concerned with the realization, i.e., surface forms, and phonological alternation rules. This component takes the output of lexc transducer (lexc grammar) as input, which has stems with grammatical features labeled with tags and it is passed through additional rules to obtain the acceptable surface forms. The xfst component helps to compile the lexc grammar into an FST as well as other rule FSTs using lexc files and rule files respectively. At the same time, other various operations are also performed through the xfst. As demonstrated in Figure 2.8, first the different separate lexicons and rules are compiled, and then they are composed into a single FST.

The lexical level (i.e. upper language) consists of citation form of a word and a sequence of tags indicating various features. The surface level (i.e. lower language) consists of actual spelling of the word. But, the process is not so straightforward. During the process of forming the word by placing the formative through the sub-lexicon in the lexc file and the spelling that is concatenated may differ. Therefore, some replace rules are applied to the lower language so that the final output would be grammatical. The orthographic rules for each FST are formulated and applied using xfst script. Sometimes, to change the sequence of tags, similar rules are applied to upper language also. The entire architecture for creating a finite state transducer that can be used as a morphological analyzer for Nepali is illustrated in Figure 2.8.

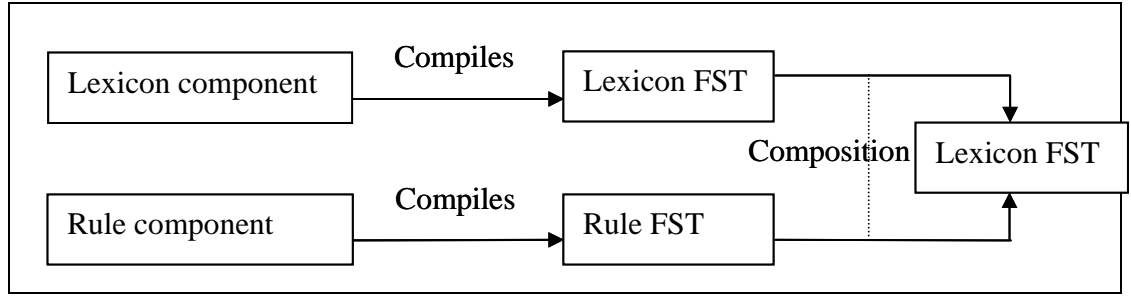


Figure 2.8: xfst interface can compile lexicon and rule and compose them into single FST (Karttunen 2000)

In figure 2.8, the lexicon is compiled to lexicon FST and rules are compiled to rule FST. These two FSTs have been composed to a single FST with the help of xfst interface.

All these functions and operations are systematically carried out through a single script file which defines various kinds of variables for rules and compiles them into an FST. This also compiles the lexicon into an FST and ultimately composes both of them into single FST. A sample of entire process with nouns in Nepali is illustrated below.

```
!! xfst script file
```

```
clear
```

```
define cons कगघडडचछजझञटठडढणतथदाधनापफबभमयारलावासाषशह;
```

```
define liquids रल;
```

```
define change [[ो %+MP -> ा || _ .#.]
```

```
.o.
```

```
[ो %+FE -> ी || _ .#.]
```

```
.o.
```

```
[ा -> [ ] || _ ?* %^b ि न ी .#.]
```

```
.o.
```

```
[ा -> [ ] || _ े न ी .#.]
```

```
.o.
```

```
[ा -> [ ] || _ ि न ी .#.]
```

```

.o.
[ा -> [] || _ ी .#.]

.o.
[ी -> ् || liquids _ न ी .#.]

.o.
[ी -> ि || _ न ी .#.]

.o.
[[. ] -> ् || cons _ न ी .#.]

.o.
[य ा -> [] || _ न ी .#.]

.o.
[ी -> [] || _ ि न ी .#.]

.o.
[् -> ु || _ ?* %^b ि न ी .#.]

.o.
[ी -> [] || _ %^b ि न ी .#.]

.o.
[%^b -> []]

.o.
[ि -> [] || ु _ न ी .#.]

];

```

```

read lexc <nouns.txt
define nouns;
read regex nouns .o. change;
save stack lexicon.fst

```

The finite state technology to process the natural languages has been employed in many languages such as Megerdoomian (2003) for Persian inflectional morphology,

Hussain (2004) for Urdu, Makedonski (2005) for Turkish nouns, Ziai (2006) for Persian simple verbs Islam (2007) for Bangla verbs and nouns, Dasgupta et al. (2007) for compound words in Bangla, Khan and Fatima (2007) for Pashto, Bharati and Kulkarni (2007) for Sanskrit, Bögel et al. (2007) for Urdu and Hindi and many more. Therefore, the finite state approach taken in this study is suitable for morphological analysis of morphology in Nepali.

2.8 Summary

In this chapter, we presented the finite state technology that is successfully applied to computational morphology. The regular expression that can be compiled into finite state network which signifies regular language and the same language can be encoded by the finite state network. The complex finite state network can be built from the smaller networks using various mathematical operations such as union, concatenation, composition, complementation, subtraction and intersection. To achieve the goals taken in this research, how the implementation of the analyzed word level categories is performed in Xerox Finite State Tool (Beesley and Karttunen 2003) is also demonstrated.

CHAPTER 3

NOMINAL MORPHOLOGY

3.0 Outline

This chapter analyzes the nominals in Nepali. It consists of seven sections. Section 3.1 deals with nouns in Nepali and various characteristic features, namely, significant stem finals, number, gender, form, augmentative/diminutive, cases and case markers. In section 3.2, we present the classification of the nouns in Nepali based on features discussed. Each group of nouns are presented with their morphological tags and corresponding finite state transducer. The phonological rules involved are also presented simultaneously. In section 3.3, we present the general characteristic features of the pronouns in Nepali such as person, number, form and honorificity. The pronouns are grouped into various groups and they are presented with their morphological tags and corresponding finite state transducers. Section 3.4 discusses the characteristics of the adjectives in Nepali, namely, number, gender, form, honorificity and degree. The adjectives are grouped into various groups and they are presented with their morphological tags and corresponding finite state transducers. Section 3.5 deals with the classification of numerals in Nepali and presents them with morphological tags and corresponding finite state transducer. Section 3.6 briefly discusses the classifiers in Nepali and also presents them with morphological tags and corresponding finite state transducers. And finally, section 3.7 summarizes the finding of the chapter.

3.1 Nouns in Nepali

Nouns in Nepali, like any other languages, are the names of persons, places, things and abstract entities from notional point of view. The nouns in the sentence function as subject, object and complement (see Pokharel 2053VS, Dahal 1974, Wallace 1985). The nouns in Nepali show various kinds of morphological features which are analyzed in the subsequent sections. The primary approach to classify the nouns is

formal one. However, sometimes the semantics of the certain groups is also considered.¹

3.1.1 Characteristics of nouns in Nepali

a. Significant stem finals

The nouns in Nepali end with various sounds. But from the significance point of view they are in a binary division: *o*-final nouns and non-*o*-final nouns. The *o*-final noun stems show different changes to indicate various kinds of morphological features as discussed in the subsections entitled number, gender, form, augmentative/diminutive, honorificity and vocative case. These features are significant grammatically as they are visible in subject-NP verb agreement; and in head noun and adjective agreement. Some non-*o*-final nouns take gender inflection and the rest do not change their form for anything. Table 3.1 shows some *o*-final and non-*o*-final noun stems.

Table 3.1: The *o*-ending and non-*o*-ending nouns

<i>o</i> -final nouns		Non- <i>o</i> -final nouns	
Stems	Gloss	Stems	Gloss
केटो keʈo	boy	उपहार upʌɦar	gift
बच्चो bʌtstso	child	उपियाँ upijã	flea
बोक्सो bokso	witch (male)	ओखती okʰʌti	medicine
भेडो bʰeɖo	sheep	कथा kʌtʰa	story
गाग्रो gagro	vase	खोल kʰol	cover
चुल्टो tsultʰo	plait	गमला gʌmʌla	flower pot
पिटो pitʰo	flour	चौकी tsʌuki	police post

In Table 3.1, some *o*-final nouns have been listed in the left side and some non-*o*-final nouns in the right side.

b. Number

The nouns in Nepali show two dimensions of number: singular and plural. Unmarked or citation form is the singular whereas the feature plural is indicated either by the change in the citation form (from *o*-ending to *a*-ending) or by a plural/collective

¹ Inflectional morphology of Nepali divided into two parts: nominal morphology and verbal morphology. Chapter 3 deals with nominals and chapter 4 deals with verbs.

marker *-हरू -fiARu:* which is in fact a postposition and is dealt with in 5.3.1. The *o*-ending nouns as *छोरो ts^horo* 'son' in (1a) changes into *a*-ending as *छोरा ts^hora* 'son' in (1b) to mark the plurality. Non-*o*-ending nouns as *घर g^hAr* 'house' in (1d) takes a postposition *-हरू -fiARu:* to indicate the plurality as *घरहरू g^hAr-fiARu:* 'houses' in (1e). The plural marker *-हरू -fiARu:* can also optionally occur with *o*-ending nouns as *छोरा ts^hora* 'son' in (1c). But the *o*-ending nouns change to the *a*-ending also for oblique form and honorific form (see subsections d and e).

(1)a. छोरो काम गर्छ ।

ts^horo kam gAr-ts^hΛ
 son.SG work do-NPST.3SG
 'The son works.'

b. छोरा स्कूल जान्छन् ।

ts^hora iskul dza-n-ts^hΛn
 son.PL school go-φ-NPST.3PL
 'The sons go to school.'

c. छोराहरू स्कूल जान्छन् ।

ts^hora-fiARu: iskul dza-n-ts^hΛn
 son.OBL-PL school go-φ-NPST.3PL
 'The sons go to school.'

d. घर राम्रो छ ।

g^hAr ramro ts^hΛ
 house.SG good.SG.MASC be.NPST.3SG
 'The house is good.'

d. घरहरू राम्रा छन्

g^hAr-fiARu: ramra ts^hΛn
 house-PL good.PL be.NPST.3PL
 'The houses are good.'

The pattern of number system in Nepali nouns with some examples is illustrated in Table 3.2.

Table 3.2: Number: singular and plural

	son	bridegroom	child	house
Singular	छोरो ts ^h oro	बेहुलो beɦulo	बच्चो bʌtstso	घर g ^h ʌr
Plural	छोरा ts ^h ora	बेहुला beɦula	बच्चा bʌtstsa	---
Plural	छोराहरू ts ^h ora-ɦʌru:	बेहुलाहरू beɦula-ɦʌru:	बच्चाहरू bʌtstsa-ɦʌru:	घरहरू g ^h ʌrɦʌru:

In Table 3.2, nouns in Nepali that inflect for number either by changing stem final ओ *o* into आ *a* or by a postposition -हरू *-ɦʌru:* are demonstrated.

In Nepali, it is not always necessary for a plural marker to be present in the non-*o*-ending nouns to indicate the plurality feature. Some other indicators such as numeral modifiers and the plural agreement inflections in the verb are sufficient to indicate the plural meaning. In such cases, the plural/collective marker -हरू *-ɦʌru:* seems to be optional as in example (2). These sorts of syntactic means of indicating plural/collective features have in fact not been considered in this study since they are not morphological phenomena.

(2) मेरा धेरै घर(हरू) छन्।

mera d^herʌi g^hʌr(-ɦʌru:) ts^hʌn
1SG.OBL.GEN.PL many house(-PL) be.NPST.3PL
'I have many houses.'

c. Gender

The nouns in Nepali inflect for two genders: masculine and feminine.² The form and semantics of the gender system is restricted to the animate nouns only. The genders in human nouns have grammatical implications and in non-human nouns it is restricted to morphological form only. From the structural point of view, there are lexical and morphological genders in Nepali.³

² Pokharel (2010a) labels the gender system as Non-Feminine (Masculine) and Feminine.

³ In this study, lexical refers to without any gender marker and morphological means with some sorts of markers.

i. Lexical gender

Some nouns in Nepali do not have any kind of gender marker: their gender can be deduced from their semantics. These nouns inherently carry the gender features. As in (3a) the noun भाइ *b^hai* 'younger brother' does not have any kind of marker in it but it has a masculine feature and in (3b) the noun बहिनी *b^hiini*: 'younger sister' also does not have any marker in it but it is a feminine. Some of the nouns indicating lexical gender are illustrated in Table 3.3. The human nouns have a grammatical implication as the verb agrees with subject NP in terms of number and gender as भाइ *b^hai* 'younger brother' in (3a) and बहिनी *b^hiini*: 'younger sister' in (3b) whereas the non-human feminine nouns do not show this kind of gender agreement with the verb, even if they are naturally male like गोरु *goru* 'ox' in (3c) and female as गाई *gai*: 'cow' in (3d).

(3)a. भाइ बजारतिर जाँदै छ ।

b^hai *b^hadz^har-tir^h* *dz^hād^hi* *ts^hΛ*
younger brother market-ALL go-IMPERF be.NPST.3SG.MASC
'The younger brother is going towards the market.'

b. बहिनी घरमा बस्छे ।

b^hiini: *g^hΛr-ma* *b^has-ts^he*
younger sister.FEM house-LOC sit-NPST.3SG.FEM
'The younger sister stays at home.'

c. गोरु घाँस खान्छ ।

goru *g^hās k^ha-n-ts^hΛ*
ox.MASC grass eat-φ-NPST.3SG.MASC
'The ox eats grass.'

d. गाई घास खान्छ ।

gai: *g^hās k^ha-n-ts^hΛ*
cow.FEM grass eat-Ø-NPST.3SG.MASC
'The cow eats grass.'

Table 3.3 lists some examples of lexical gender in Nepali.

Table 3.3: Lexical gender

	Husband	brother	ox	male
Masculine	लोग्ने logne	भाइ b ^h ai	गोरु goru	भाले b ^h ale
Feminine	स्वास्नी swasni:	बहिनी bahini:	गाई gai:	पोथी pot ^h i:

Thus, the nouns in first two columns of Table 3.3 are significant from an analytic point of view whereas the nouns in last two columns are not.

ii. Morphological gender

Nouns in Nepali in which the gender is indicated either by certain change or by some sorts of marker are said to have morphological gender. The citation form is always masculine gender as *बेहुलो* *behuḷo* 'bridegroom' in (4a) and as *नाति* *nati* 'grandson' in (4e). The *o*-ending nouns change into *i*-ending as *बेहुली* *behuḷi*: 'bride' in (4b). A small set of nouns such as *काका* *kaka* 'paternal uncle' in (4c) takes *-ी-i:* and changes to feminine as *काकी* *kaki*: 'aunt' in (4d). Another set of nouns such as *नाति* *nati* 'grandson' in (4e) changes to its feminine form as *नातिनी* *natinī*: in (4f) by taking suffix *-नी-ni:*. Yet, there is another small set of nouns that take the suffixes *-एनी -eni:* or *-इनी- ini:*. For example *सुब्बा* *subba* 'Subba' --> *सुब्बेनी/सुब्बिनी* *subb-eni:/subb-ini:* 'Female Subba' The majority of non-*o*-ending human nouns which have their corresponding feminine forms are kinship terms, family names (surnames), caste, social status and professions.

(4)a. बेहुलो सुन्दर छे ।

behuḷo sundar ts^hΛ
bridegroom handsome be.NPST.3SG.MASC
'The bridegroom is handsome.'

b. बेहुली कुरूप छे ।

behuḷi: kurup ts^he
bridegroom.FEM ugly be.NPST.3SG.FEM
'The bride is ugly.'

c. काकाले घर बनाए ।

kaka-le g^har bʌn-a-e
uncle-ERG house make-CAUS-PST.3SG.HON
'The uncle made a house.'

d. काकी स्कूलमा काम गर्नु हुन्छ ।

kaki iskul-ma kam gar-nu hun-ts^hʌ
uncle.FEM school-LOC work do-INF be-NPST.3SG.MASC
'The aunt works in the school.'

e. नाति भात खान्छ ।

nati b^hat k^ha-nts^hʌ
grandson.MASC rice eat-NPST.3SG.MASC
'The grandson eats rice.'

f. नातिनी स्कूल जान्छे ।

nati-ni: iskul dza-nts^he
grandson-FEM school go-NPST.3SG.FEM
'The grand daughter goes to school.'

Table 3.4 illustrates the morphological gender system in Nepali.

Table 3.4: Morphological gender

Masculine	Gloss	Feminine	Gloss
छोरो ts ^h oro	son	छोरी ts ^h ori:	daughter
बेहुलो befulo	bridegroom	बेहुली befuli:	bride
काका kaka	uncle	काकी kaki:	aunt
कुमार kumar	lad	कुमारी kumari	lass
नाति nati	grandson	नातिनी natini:	grand daughter
बाघ bag ^h	Tiger	बघिनी bag ^h ini:	tigress
सुब्बा subba	Subba (male)	सुब्बेनी/सुब्बिनी subb-eni:/subb-ini:	Subba (female)

In Table 3.4, *o*-ending and non-*o*-ending nouns in Nepali that inflect for feminine gender by various ways are demonstrated.

d. Form

The nouns in Nepali show two forms morphologically: direct and oblique. In traditional grammars, the nouns which appear as the citation forms are direct and

those appear with postpositions are oblique. The *o*-ending nouns as *बिरालो biralo* 'cat' in (5a) change to *a*-ending as *बिराला birala* in (5b) to take the oblique form. This happens only in *o*-ending nouns when they are followed by postpositions. The non-*o*-ending nouns do not show such changes whether they are followed by postpositions or not such as *रूख ru:k^h* 'tree' in (5c)⁴.

- (5) a. बिरालो दुध खान्छ ।
 biralo dud^h k^ha-n-ts^hΛ
 cat.SG milk eat-φ-NPST.3SG.MASC
 'The cat drinks milk.'
- b. बिरालाले मुसा मार्छ ।
 birala-le musa mar-ts^hΛ
 cat.OBL-ERG mouse.PL kill-NPST.3SG.MASC
 'The cat kills the rats.'
- c. रूखमा एउटा चरो बसेको छ ।
 ru:k^h-ma euṭa tsaro bas-eko ts^hΛ
 tree-LOC one.CLF bird sit-PERF be.NPST.3SG.MASC
 'A bird is sitting in the tree.'

Table 3.5 shows the alternation between direct and oblique forms of some nouns in Nepali.

Table 3.5: Direct and oblique forms

	Horse	cat	mouse	tree
Direct	घोडो g ^h oḍo	बिरालो biralo	मुसो muso	रूख ru:k ^h
Oblique	घोडा g ^h oḍa	बिराला birala	मुसा musa	रूख ru:k ^h

In Table 3.5, Nepali nouns that show the direct and oblique forms in different conditions and those which do not alter are listed.

⁴ For the present purpose, non-*o*-ending nouns when followed by postpositions are not considered as oblique forms.

e. Honorificity

Nouns in Nepali show two levels of honorificity morphologically: non-honorific and honorific. The honorificity distinction can be found only in *o*-ending human nouns. Those nouns with *o*-ending as बेहुलो *beɦulo* 'bridegroom' in (6a) change into *a*-ending as बेहुला *beɦula* 'bridegroom' in (6b) indicating non-honorificity and honorificity respectively. But non-*o*-ending nouns do not show this distinction, therefore, they are not listed here.

(6) a. बेहुलो हात्तीमा थियो ।

beɦulo hatti-ma t^hi-jo
bridegroom.NHON elephant-LOC be-PST.3SG.NHON
'The bridegroom was on the elephant.'

b. बेहुला हात्तीमा थिए ।

beɦula hattima t^hi-e
bridegroom.HON elephant-LOC be-PST.3SG.HON
'The bridegroom was on the elephant.'

Some examples of honorificity and non-honorificity in Nepali *o*-ending human nouns are illustrated in Table 3.6.

Table 3.6: Honorificity: non-honorific and honorific

	Boy	son	bridegroom	child
Non-honorific	केटो keɽo	छोरो ts ^h oro	बेहुलो beɦulo	बच्चो baɽtstso
Honorific	केटा keɽa	छोरा ts ^h ora	बेहुला beɦula	बच्चा baɽtstsa

The examples presented in Table 3.6 show the alternation between honorificity and non-honorificity in *o*-ending human nouns.

f. Augmentative and diminutive

From an evaluative point of view, nouns in Nepali show two distinctions: augmentative and diminutive. This distinction is found only in a small set of *o*-ending inanimate nouns which indicates the size of the object whether it is bigger or smaller.

The *o*-ending as डालो *dalo* 'basket' in (7a) changes into *i*-ending as डाली *dali:* in (7b) indicating augmentative and diminutive forms, respectively. This distinction of bigger and smaller in size is limited to morphology only because there is no augmentative or diminutive agreement in the verbs. The non-*o*-ending inanimate nouns do not show this distinction, therefore, they are not considered here.

- (7) a. राम डालो बनाउन जान्दछ ।
 ram ɖalo bʌnau-nʌ dzan-dʌtsʰʌ
 Ram basket.AUG make-INF know-NPST.3SG.MASC
 'Ram knows to make the basket.'
- b. यो चिज डालीमा राख् ।
 jo tsidz ɖali:-ma rakʰ
 this thing basket.DIM-LOC keep.IMP
 'Keep this thing in the small basket.'

Table 3.7 illustrates some examples of augmentative and diminutive form of Nepali *o*-ending nouns.

Table 3.7: Augmentative and diminutive

	Basket	small hill	bag	bowl
Augmentative	डालो ɖalo	थुम्को thumko	झोलो dzʰolo	बटुको bʌɖuko
Diminutive	डाली ɖali:	थुम्की tʰumki:	झोली dzʰoli:	बटुकी bʌɖuki:

In Table 3.7, Nepali *o*-ending inanimate nouns change to *i*-ending for augmentative and diminutives forms, respectively.

g. Cases and case markers

In Nepali, the cases are marked by postpositions. Even though the case markers are affixed to stems of nouns or pronouns, they are treated as a separate group of linguistic units. Thus, they are tokenized into separate tokens (Hardie et al. 2005, 2009). However, a short traditional description of cases and case markers has been

given here with examples. We have dealt with case markers as postpositions in (see 5.3) for computational purpose.

I. Ergative

The ergative case in Nepali is marked by a postposition *-ले-le* as *रामले ram-le* in (8).

Mostly the ergative case marker occurs with agent subject in perfective transitive constructions.

(8). रामले भात खायो ।

ram-le b^hat k^ha-jo
Ram-ERG rice eat-PST.3SG.MASC
'Ram ate rice.'

II. Dative

The dative case in Nepali is marked by a postposition *-लाई-lai:* as *मलाई m_li-lai:* in

(9). The dative marker appears normally with indirect/direct object noun phrase. Normally the accusative case is not marked, but in some condition it is marked with same dative marker *-लाई-lai:*.

(9) मैले रामलाई पिटेँ ।

m_li-le ram-lai: piṭ-ẽ
1SG.OBL-ERG Ram-DAT beat-PST.1SG
'I beat Ram.'

III. Instrumental

The instrumental case in Nepali is marked by a postposition *-ले-le* as *चम्चाले ts_lamtsa-le* in (10). The ergative and instrumental case markers are same in their forms but ergative case marker appears with agent as in (8), whereas instrumental case marker appears with instruments or objects with which the action is performed.

(10) उसले चम्चाले खाना खायो ।

us-le ts_lamtsa-le k^hana k^ha-jo
3SG.OBL-ERG spoon-INST meal eat-PST.3SG.MASC
'He ate the food with spoon.'

IV. Ablative

The ablative case in Nepali is marked by a postposition -बाट -baṭ as बजारबाट *baḏzar-baṭ* in (11). There is an alternative postposition to the former one, i.e., -देखि -*dek^{hi}* which has almost the same meaning.

(11) ऊ बजारबाट/देखि आयो ।

u: baḏzar-baṭ/dek^{hi} a-jo
3SG market-ABL come-PST.3SG.MASC
'He came from the market.'

V. Locative

The locative case in Nepali is marked by a postposition -मा -*ma* as घरमा *g^har-ma* in (12a). There is another marker as -कहाँ -*kahā* in (12b) which normally occurs with pronouns and occasionally occurs with other nominals indicating the location but it is not frequent.

(12)a. म घरमा बस्छु ।

mʌ g^har-ma bʌs-ts^hu
1SG house-LOC sit-NPST.1SG
'I stay at home.'

b. हरि मकहाँ आयो ।

hʌri məkəhā a-jo
Hari 1SG-LOC come-PST.3SG.MASC
'Hari came to me.'

VI. Allative

The allative case in Nepali is marked by a postposition -तिर -*tira* as बजारतिर *baḏzar-tira* in (13).⁵

(13) तिमी बजारतिर जाऊ!

timi baḏzar-tira dza-u:
2SG.HON market-ALL go-IMP.2SG.HON
'You go towards the market.'

⁵ Most of the traditional Nepali grammars do not assume -तिर -*tira* as a case marker.

VII. Comitative/Associative

The comitative/associative case in Nepali is marked by a postposition -सँग-*sāṅ* and as -सित-*sita* मसँग/सित *mā-sāṅ/sita* in (14).

- (14) तिमि मसँग/सित बस !
timi mā-sāṅ/sita baśa!
2SG.HON 1SG-COM stay-IMP.2SG.HON
'You stay with me.'

VIII. Genitive

The genitive case in Nepali is marked by a postposition -को-*ko* as रामको *ram-ko* in (15). This postposition has three alternate forms -को-*ko*, -का-*ka* and -की-*ki*: for singular masculine, plural and feminine respectively that occurs in most cases. But, the forms -रो-*ro*, -रा-*ra* and -री-*ri*: occur with first person pronouns म *mā* and हामी *hāmī* and second person pronouns तँ *tā* and तिमि *timi*; and forms -नो-*no*, -ना-*na* and -नी-*ni*: occur with the reflexive pronoun आफू *ap^hu* 'self'.

- (15) रामको कलम राम्रो छ ।
ram-ko kalam ramro tsh̄
Ram-GEN.MASC.SG pen good.MASC.SG be.NPST.3SG.MASC
'Ram's pen is good.'

IX. Vocative

The vocative case in Nepali is marked by changing ओ *o* into आ *a* in *o*-ending human nouns as केटा *keta* in (16). Non-*o*-ending nouns do not inflect for this case.

- (16) ए केटा! यो काम गर्!
e keta! jo kam ga!
yeh boy.VOC! this work do.IMP.NHON
'Hey boy! Do this work.'

X. Nominative

The nominative case in Nepali is unmarked as राम *ram-ø* in (17). The subject of an intransitive verb and subject of a transitive verb in a non-perfective construction are in nominative case.

(17) राम सधैँ लेख्छ ।

ram-ø sʌdʰɑĩ lekʰ-tʰʌ
 Ram-NOM always write-NPST.3SG.MASC
 'Ram always writes.'

3.2 Classification of nouns in Nepali

On the basis of the characteristic features discussed in (3.3), nouns in Nepali have been grouped into fourteen classes and the finite state machines or networks have been constructed. The features like stem final segment, number, gender, form, honorificity, augmentative/diminutive and vocative case are considered while grouping the nouns. The features discussed in (3.1) are not consistently present in all the nouns. The basic criteria for grouping the nouns include presence or absence of these features and the semantics of the nouns in some cases (Prasain 2010). The sequence of tags is arbitrary. The tags for default features are not included and the names of the noun classes are arbitrary.

3.2.1 O-ending nouns

a. NounType 1a

In this class, the *o*-ending human nouns which inflect for number, gender, form, honorificity and vocative case are grouped. Some examples with their corresponding morphological tags are given in Table 3.8

Table 3.8: NounType 1a

Morphological Tags	boy	son	bridegroom	child
NOUN+MASC+SG	केटो keʈo	छोरो tsʰoro	बेहुलो befulo	बच्चो bʌtʌtso
NOU N+MASC+PL	केटा keʈʌ	छोरा tsʰora	बेहुला befula	बच्चा bʌtʌtʌ
NOUN+MASC+OBL	केटा keʈʌ	छोरा tsʰora	बेहुला befula	बच्चा bʌtʌtʌ
NOUN+MASC+HON	केटा keʈʌ	छोरा tsʰora	बेहुला befula	बच्चा bʌtʌtʌ
NOUN+MASC+VOC	केटा keʈʌ	छोरा tsʰora	बेहुला befula	बच्चा bʌtʌtʌ
NOUN+FEM	केटी keʈi:	छोरी tsʰori:	बेहुली befuli:	बच्ची bʌtʌtsi:

The finite state transducer illustrated in Figure 3.1 is capable of analyzing and generating the word-forms illustrated in Table 3.8.

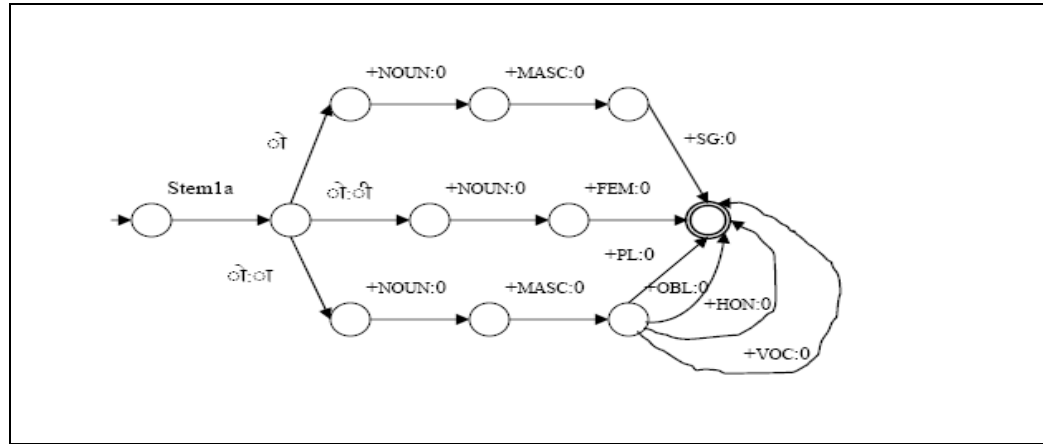


Figure 3.1: A finite state transducer for NounsType 1a

The phonological rules involved in this group of nouns are given in PR 3.1.

Phonological rules

PR 3.1

- i. Stem final vowel ो of the o -ending human nouns of the lower language (i.e., surface level) is changed to vowel ा for plural form, oblique, honorificity and vocative case.

Regular expression: $\text{ो} \rightarrow \text{ा} \parallel _ \#$.

- ii. Stem final vowel ो of the o -ending human nouns of the lower language (i.e., surface level) is changed to vowel ी for feminine gender.

Regular expression: $\text{ो} \rightarrow \text{ी} \parallel _ \#$.

b. NounType 1b

In this class, the o -ending animate nouns which inflect for number, gender and form are grouped. Some examples are listed in Table 3.9 with their corresponding morphological tags.

Table 3.9: NounType 1b

Morphological Tags	horse	goat	cat	rat
NOUN+MASC+SG	घोडो g ^h oɖo	बाख्रो bak ^h ro	बिरालो biralo	मुसो muso
NOUN+MASC+PL	घोडा g ^h oɖa	बाखा bak ^h ra	बिराला biralā	मुसा musa
NOUN+MASC+OBL	घोडा g ^h oɖa	बाखा bak ^h ra	बिराला biralā	मुसा musa
NOUN+FEM	घोडी g ^h oɖi:	बाख्री bak ^h ri:	बिराली biralī:	मुसी musi:

The finite state transducer illustrated in Figure 3.2 is capable of analyzing and generating the word-forms illustrated in Table 3.9.

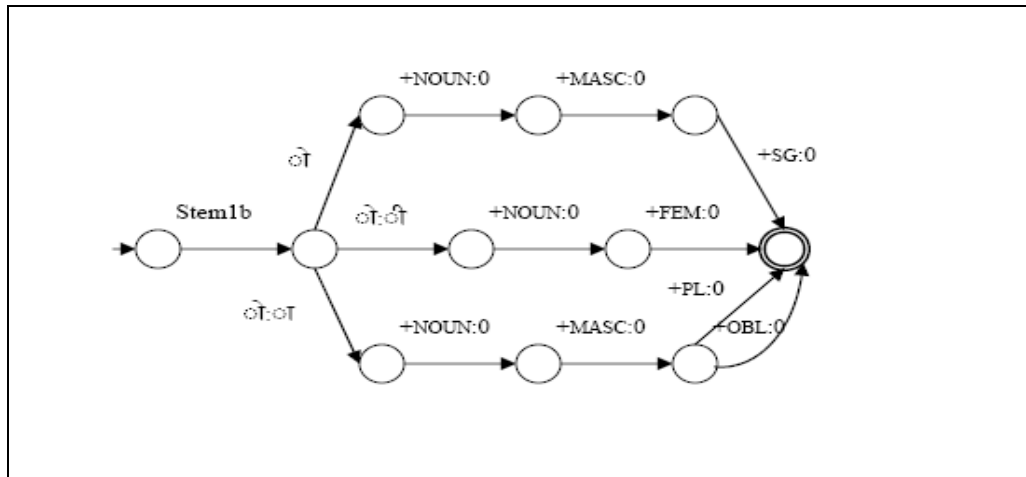


Figure 3.2: A finite state transducer for NounType 1b

The phonological rules that are applied to the finite state transducer illustrated in Figure 3.2 are given in PR 3.2.

Phonological rules

PR 3.2

- i. Stem final vowel ो o of the o-ending animate nouns of the lower language (i.e, surface level) is changed to vowel ा a for plural form and oblique form.

Regular expression: ो -> ा || _ .#.

- ii. Stem final vowel *ो* *o* of the *o*-ending animate nouns of the lower language (i.e, surface level) is changed to vowel *ी* *i* for feminine gender.

Regular expression: *ो* -> *ी* || _ .#.

c. NounType 1c

In this class, the *o*-ending inanimate nouns which inflect for number, form and augmentative/diminutive features are grouped. Some examples are listed in Table 3.10 with their corresponding morphological tags.

Table 3.10: NounType 1c

Morphological Tags	basket	small hill	bag	bowl
NOUN+SG	डालो ɖalo	थुम्को tʰumko	झोलो dzʰolo	बटुको bʌɽuko
NOUN+PL	डाला ɖala	थुम्का tʰumka	झोला dzʰola	बटुका bʌɽuka
NOUN+OBL	डाला ɖala	थुम्का tʰumka	झोला dzʰola	बटुका bʌɽuka
NOUN+DIM	डाली ɖali:	थुम्की tʰumki:	झोली dzʰoli:	बटुकी bʌɽuki:

The finite state transducer illustrated in Figure 3.3 is capable of analyzing and generating the word-forms illustrated in Table 3.10.

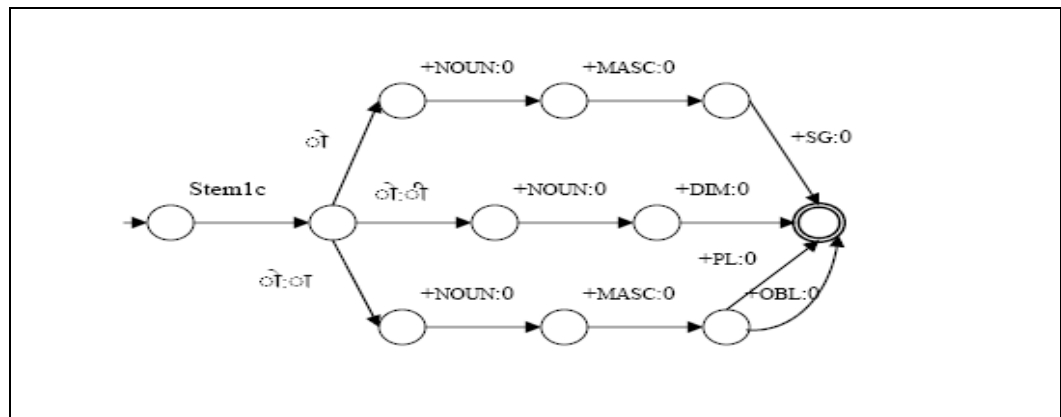


Figure 3.3: A finite state transducer for NounType 1c

The phonological rules listed in PR 3.3 are applied to the finite state transducer illustrated in Figure 3.3.

Phonological rule

PR 3.3

- i. Stem final vowel *o* of the *o*-ending inanimate nouns of the lower language (i.e, surface level) is changed to vowel *a* for plural and oblique form.

Regular expression: $o \rightarrow a \parallel _ \#$.

- ii. Stem final vowel *o* of the *o*-ending inanimate nouns of the lower language (i.e, surface level) is changed to vowel *i* for diminutive feature.

Regular expression: $o \rightarrow i \parallel _ \#$.

d. NounType 1d

In this class, the *o*-ending inanimate nouns which inflect only for number and oblique form are grouped. Some examples are listed in Table 3.11 with their corresponding morphological tags.

Table 3.11: NounType 1d

Morphological Tags	pine	photo	ladder	flesh(dead)
NOUN+SG	सल्लो sallo	फोटो p ^h oʈo	लिस्नो lisno	सिनो sino
NOUN+PL	सल्ला salla	फोटा p ^h oʈa	लिस्ना lisna	सिना sina
NOUN+OBL	सल्ला salla	फोटा p ^h oʈa	लिस्ना lisna	सिना sina

The finite state transducer illustrated in Figure 3.4 is capable of analyzing and generating the word-forms illustrated in Table 3.11.

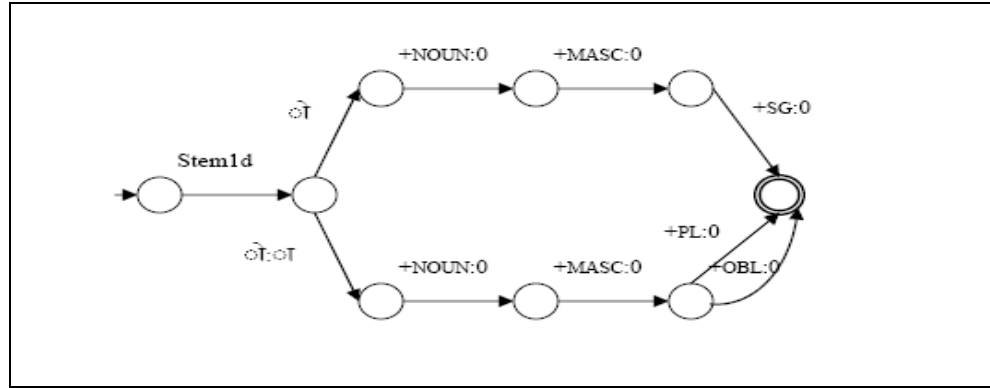


Figure 3.4: A finite state transducer for NounsType 1d

The finite state transducer illustrated in Figure 3.4 is composed with the finite state transducer of rules listed in PR 3.4.

Phonological rule

PR 3.4

- i. Stem final vowel *ो* *o* of the *o*-ending inanimate nouns of the lower language (i.e, surface level) is changed to vowel *ो:ा* for plural and oblique.

Regular expression: *ो* -> *ो:ा* || _ .#.

3.2.2 Non-*o*-ending nouns

I. Marked

a. NounType 21a

In this class, the non-*o*-ending human and animate nouns which inflect only for gender feature with marker *-ी* *-i:* are grouped. Some examples are listed in Table 3.12 with their corresponding morphological tags.

Table 3.12: NounType 21a

Morphological Tags	uncle	lad	pigeon	parrot
NOUN+MASC	काका kaka	कुमार kumar	परेवा pɹɛwɹ:	सुगा sugɹ
NOUN+FEM	काकी kaki:	कुमारी kumari:	परेवी pɹɛwi:	सुगी sugi:

The finite state transducer illustrated in Figure 3.5 is capable of analyzing and generating the word-forms illustrated in Table 3.12.

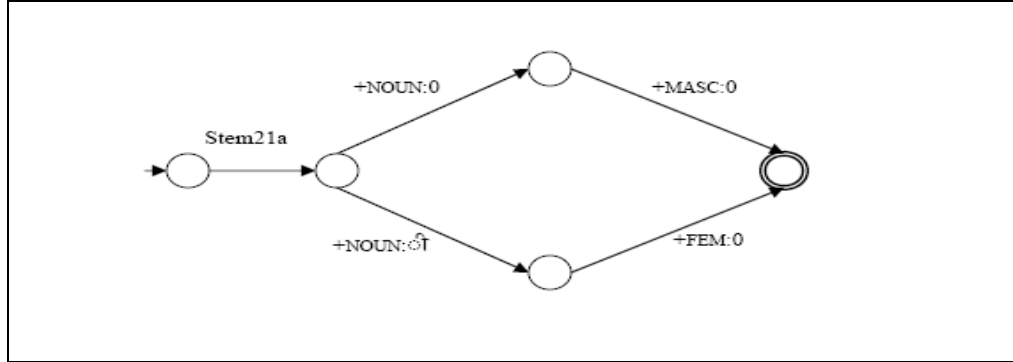


Figure 3.5: A finite state transducer for NounType21a

The phonological rules listed in PR 3.5 are combined at the lower side of the network illustrated in Figure 3.5.

Phonological rule

PR 3.5

- i. Stem final vowel *ā* of the non-*o*-ending animate nouns of the lower language (i.e, surface level) is deleted when followed by gender marker *ी* *i:*.

Regular expression: $\bar{a} \rightarrow [] \parallel _ \bar{i} \#$.

b. NounType 21b

In this class, the non-*o*-ending human nouns which inflect for masculine and feminine features with marker *-नी-ni* are collected. Some examples are listed in Table 3.13 with their corresponding morphological tags.

Table 3.13: NounType 21b

Morphological Tags	grandson	beggar	priest	chief
NOUN+MASC	नाति nati	जोगी dzogi:	पण्डित pāṇḍit	मुखिया muk ^h iya
NOUN+FEM	नातिनी nati-ni:	जोगिनी dzogi-ni:	पण्डित्री pāṇḍit-ni:	मुखिनी muk ^h i-ni:

The finite state transducer illustrated in Figure 3.6 is capable of analyzing and generating the word-forms illustrated in Table 3.13.

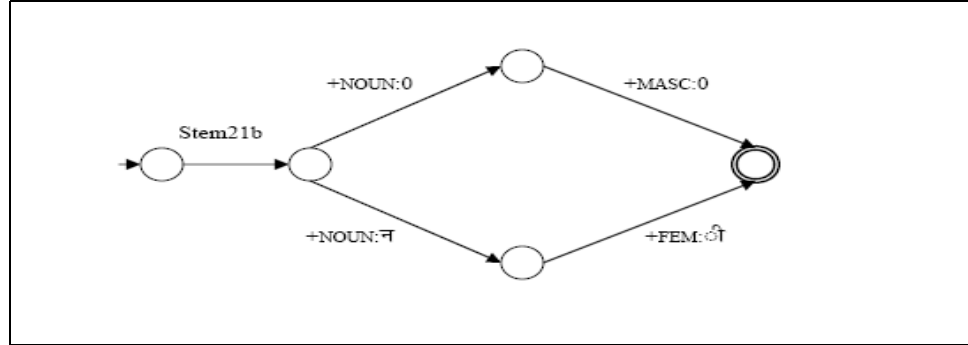


Figure 3.6: A finite state transducer for NounType 21b

The finite state transducer in Figure 3.6 is composed with the phonological rules listed in PR 3.6.

Phonological rule

PR 3.6

- i. Stem final vowel *ी* *i:* of the non-*o*-ending human nouns of the lower language (i.e. surface level) is changed to vowel *ि* *i* before the feminine gender marker *नी* *ni:*.

Regular expression: *ी* -> *ि* || *_न* *ी* .#.

- ii. Halanta *्* is inserted between consonant symbol and feminine gender marker - *नी* *-ni:* at the surface level.⁶

Regular expression: [*.*] -> *्* || *cons* *_न* *ी* .#.

- iii. *य* *ja* is deleted before the feminine gender marker *नी* *ni:* at the surface level.

Regular expression: *य* *ा* -> [*]* || *_न* *ी* .#.

⁶ Halanta *्* is generic term for the diacritic in Devanagari that is used to suppress the inherent vowel that otherwise occurs with every consonant letter.

iv. Stem final vowel *ी* *i:* of the non-*o*-ending nouns of the lower language (i.e. surface level) is replaced by a halanta ् after liquid sounds and before the feminine gender marker *नी* *ni:*.

Regular expression: *ी* -> ् || liquids _ न *ी* .#.

c. NounType 21c

In this class, the non-*o*-ending human and animate nouns which inflect only for the gender feature with the marker *-इनी* *-ini:* are grouped. This group differs from the other NounType 21b because the *अ* *a* sound within the stem changes to the *अ* *ʌ* sound while inflecting for feminine gender. Some examples are listed in Table 3.14 with their corresponding morphological tags.

Table 3.14: NounType 21c

Morphological Tags	tiger	Surname1	Surname2	Surname3
NOUN+MASC	बाघ bag ^h	कर्की karki:	थापा t ^h apa	थारू t ^h aru:
NOUN+FEM	बघिनी bʌg ^h i-ni:	कर्किनी kʌrkini:	थपिनी t ^h ʌp-ini:	थरुनी t ^h ʌru-ni:

The finite state transducer illustrated in Figure 3.6 is capable of analyzing and generating the word-forms illustrated in Table 3.14 when the rules listed in PR 3.7 are applied.

Phonological rule

PR 3.7

i. Vowel *अ* *a* in the non-*o*-ending human and animate noun stems of the lower language (i.e. surface level) is changed to vowel *अ* *ʌ* when the feminine gender marker *-इनी* *-ini:* appears at the end of the word

Regular expression: *अ* -> [] || _ ि न *ी* .#.

ii. Vowel *ी* *i:* is deleted before the feminine gender marker *-इनी* *-ini:*.

Regular expression: ी -> [] || _ ि न ी .#.

iii. Vowel ू *u:* is changed to ु *u* before the feminine gender marker -इनी *-ini:*.

Regular expression: ू -> ु || _ ि न ी .#.

d. NounType 21d

In this class, the non-*o*-ending human nouns which inflect only for gender feature with marker -इनी *-ini:* alternatively -एनी *-eni:* are grouped. Some examples are listed in Table 3.15 with their corresponding morphological tags.

Table 3.15: NounType 21d

Morphological Tags	Ethnic name1	Surname	Ethnic2
NOUN+MASC	खस k ^h ΛS	बिस्ट biʂΛ	सुब्बा subba
NOUN+FEM	खसेनी/खसिनी k ^h ΛS- <i>eni</i> /k ^h ΛS- <i>ini</i>	बिस्टेनी/बिस्टिनी biʂ- <i>eni</i> /biʂ- <i>ini</i>	सुब्बेनी/सुब्बिनी subb- <i>eni</i> /subb- <i>ini</i>

The finite state transducer illustrated in Figure 3.7 is capable of analyzing and generating the word-forms illustrated in Table 3.15.

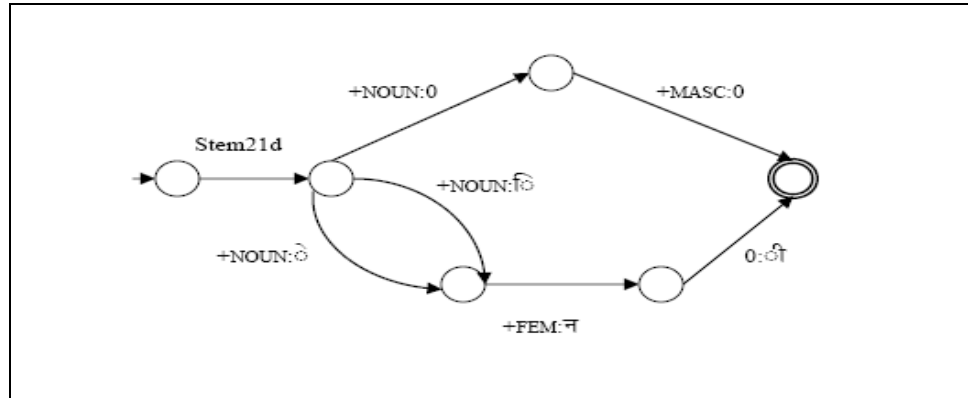


Figure 3.7: A finite state transducer for NounType 21d

The phonological rules involved in this process are listed in PR 3.8 which are compiled and composed with the finite state transducer illustrated in Figure 3.7.

Phonological rule

PR 3.8

- i. Vowel *ā* at the end of non-*o*-ending noun stem is deleted before the feminine gender marker -इनी -*ini*: or -एनी *eni*:

Regular expression: $\text{ā} \rightarrow [\] _ \text{ि न िी | ए न िी} \#$.

II. Unmarked

a. NounType 22a

In this class, the non-*o*-ending human nouns that do not inflect for any features but are inherently masculine are grouped. Some examples are listed in Table 3.16 with their corresponding morphological tags.

Table 3.16: NounType 22a

Morphological Tags	elder brother	younger brother	father	husband
NOUN+MASC	दाइ dai	भाइ b ^h ai	बाबु babu	लोगने logne

The finite state transducer illustrated in Figure 3.8 is capable of analyzing and generating the word-forms illustrated in Table 3.16.

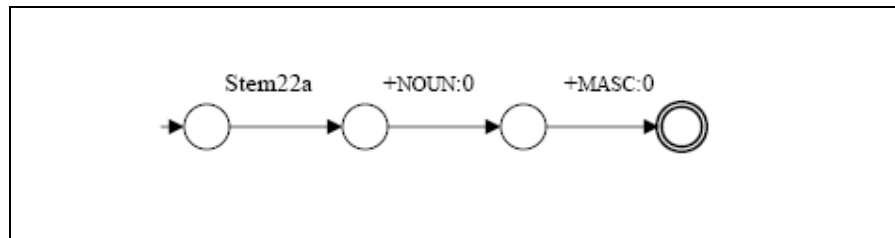


Figure 3.8: A finite state transducer for NounType 22a

b. NounType 22b

In this group, the non-*o*-ending human nouns do not inflect for any features, but are inherently feminine gender are collected. Some examples are listed in Table 3.17 with their corresponding morphological tags.

Table 3.17: NounType 22b

Morphological Tags	elder sister	younger sister	mother	wife
NOUN+FEM	दिदी didi:	बहिनी bʌhini:	आमा ama	स्वास्नी swasni:

The finite state transducer illustrated in Figure 3.9 is capable of analyzing and generating the word-forms illustrated in Table 3.17.

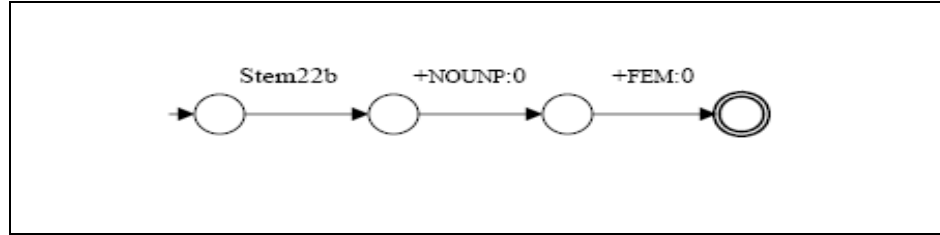


Figure 3.9: A finite state transducer for NounsType 22b

c. NounType 22c

In this group, the proper names of males which never inflect for anything irrespective of their final sound segments, but grammatically agree with verb for masculine gender if they are in subject-NP position. Some examples are listed in Table 3.18 with their corresponding morphological tags.

Table 3.18: NounType 22c

Morphological Tags	Pname1	Pname2	Pname3	Pname4
NOUN+PROPER+MASC	कर्णखर kʌrɳak ^h ʌr	हरि ɦʌri	श्याम sjam	बलराम bʌlʌram

The finite state transducer illustrated in Figure 3.10 is capable of analyzing and generating the word-forms illustrated in Table 3.18.

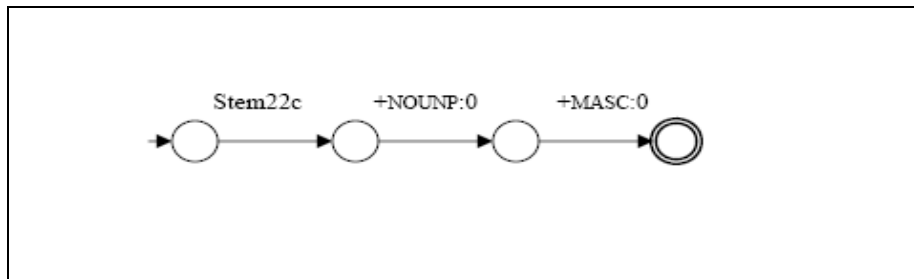


Figure 3.10: A finite state transducer for NounType 22c

d. NounType 22d

In this group, the proper names of females, which never inflect for anything irrespective of their final sound segments, but grammatically agree with verb for feminine gender if they are in subject NP position, are collected. Some examples are listed in Table 3.19 with their corresponding morphological tags.

Table 3.19: NounType 22d

Morphological Tags	Pname1	Pname2	Pname3	Pname4
NOUN+PROPER+FEM	सीता sita	गीता gita	जानकी dzanaki	निर्मला nirmala

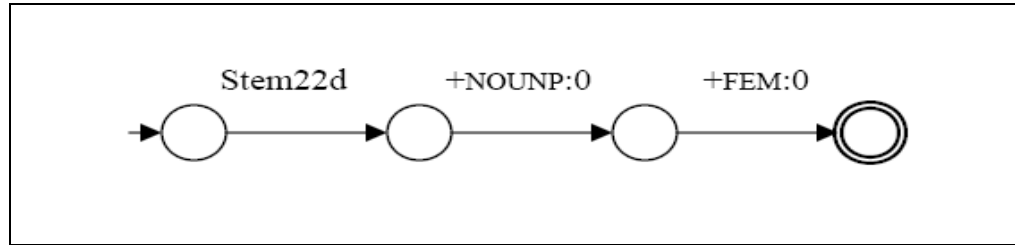


Figure 3.11: A finite state transducer for NounType 22d

The finite state transducer illustrated in Figure 3.11 is capable of analyzing and generating the word-forms illustrated in Table 3.19.

e. NounType 22e

In this group, all the common nouns which are non-*o*-ending are collected. These nouns never inflect for anything irrespective of their final sound segments, but grammatically agree with verb for default feature, i.e., third person masculine singular if they are in subject NP position. Some examples with their corresponding morphological tags are given in Table 3.20

Table 3.20: NounType 22e

Morphological Tags	Promise	shoulder	farm-land	book
+NOUN	कसम kASAM	काँध kādh	खेत k ^h et	किताब kitab

The finite state transducer illustrated in Figure 3.12 is capable of analyzing and generating the word-forms illustrated in Table 3.20.

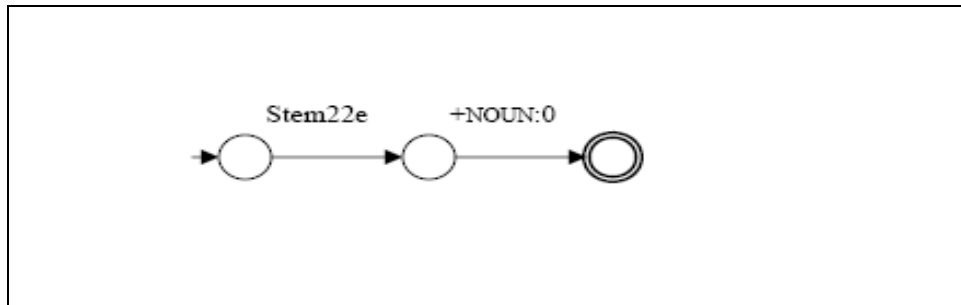


Figure 3.12: A finite state transducer for NounType 22e

f. NounType 22f

In this class, all the place names are grouped. Some examples are listed in Table 3.21 with their corresponding morphological tags.

Table 3.21: NounType 22f

Morphological Tags	PlaceName1	PlaceName2	PlaceName3	PlaceName4
NOUN+PLACE	झापा dz ^h apa	भोजपुर b ^h odzipur	नेपाल nepal	जापान dzapan

The finite state transducer illustrated in Figure 3.13 is capable of analyzing and generating the word-forms illustrated in Table 3.21.

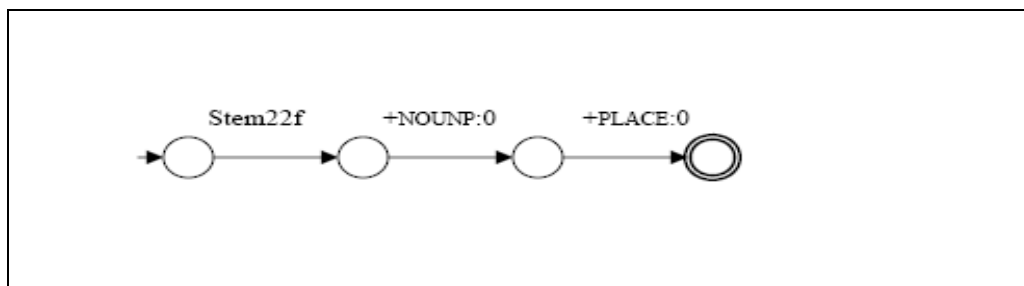


Figure 3.13: A finite state transducer for NounType 22f

3.3 Pronouns

3.3.1 Characteristics of pronouns in Nepali

a. Person

Pronouns in Nepali have three persons: first, second and third. They are listed in Table 3.22.

Table 3.22: Pronouns with respect to persons

Person	Pronouns
First	म <i>mʌ</i> 'I', हामी <i>fami</i> 'we'
Second	तैं <i>tɛ̃</i> 'you', तिमी <i>timi</i> : 'you', तपाईं <i>tʌpaĩ</i> : 'you', यहाँ <i>jʌɦã</i> 'you', हजुर <i>ɦɛ̃dzur</i> 'you', मौसुफ <i>mʌusup^h</i> 'royal you'
Third	यो <i>jo</i> 's/he', यिनी <i>jini</i> : 'she', यी <i>ji</i> : 'they', त्यो <i>tjo</i> 'that', तिनी <i>tini</i> : 's/he', ती <i>ti</i> : 'they', ऊ <i>u</i> : 'he', उनी <i>uni</i> : 'she', उहाँ <i>uhã</i> 'he'

b. Number

Personal pronouns in Nepali show two dimensions of number: singular and plural. The number feature in pronouns is also indicated by a plural/collective postposition - हारू *-ɦaru*: but some of them such as म *mʌ* 'I', तैं *tɛ̃* 'you' do not take any number maker. They have corresponding suppletive forms for the plural feature e.g., हामी *fami* 'we', तिमी *timi*: 'you'. Table 3.23 lists the personal pronouns in Nepali with number distinctions.

Table 3.23: Personal pronouns in number distinctions

	Singular	Plural
First	म mɐ 'I'	हामी fiami: 'we', हामीहरू fiami:fiɐru: 'we-Pl'
Second	तैं tã, 'you' तिमी timi: 'you', तपाईं tɐpã: 'you', यहाँ jɐhã 'you', हजुर fiɐdzur 'you', मौसुफ mɐusup ^h 'you'	तिमीहरू timi:-fiɐru:, तपाईंहरू tɐpã:-fiɐru:, यहाँहरू jɐhã-fiɐru:, हजुरहरू fiɐdzur-fiɐru:, मौसुफहरू mɐusup ^h -fiɐru:
Third	यो jo 'this', यी ji: 'this' यिनी jini: 's/he', त्यो tjo 'that', ती ti: 'those' तिनी tini: 'those', ऊ u: 's/he' उनी uni: 's/he' उहाँ uhã 's/he'	यिनीहरू jini:fiɐru:, तिनीहरू tini:fiɐru:, उनीहरू uni:fiɐru:, उहाँहरू uhãfiɐru:

c. Form

Pronouns in Nepali show two morphological forms: direct and oblique. When a pronoun is followed by postpositions, it changes into oblique forms. The oblique forms are found in personal, demonstrative, relative, reflexive pronouns; and sporadically in interrogative, definite and indefinite pronouns. Table 3.24 lists the direct and oblique form of some pronouns.

Table 3.24: Forms of pronouns: direct and oblique

Direct form	Oblique form
म mɐ 'I'	मै mɐi 'I.OBL'
हामी fiami 'we'	हाम् fiam 'we.OBL'
तैं tã 'you'	तैं tã 'you.OBL'
तिमी timi: 'you'	तिम् tim 'you.OBL'
यो jo 'this'	यस् jɐs 'this.OBL'
ऊ u: 's/he'	उन् un 's/he.OBL'
जो dzo 'who.REL'	जस् dzɐs 'who.REL.OBL'
त्यो tjo 'that'	त्यस् tjɐs 'that.OBL'
को ko 'who.INTERO'	कस् kɐs 'who.INTERO.OBL'

d. Honorificity

The second and third person pronouns in Nepali show five levels of honorificity. There is no particular honorific marker but the hierarchy is maintained at the lexical level. The honorificity in the third person pronouns is marginally marked whereas in second person pronouns it is not morphologically significant. Table 3.25 lists the pronouns in terms of honorific levels. The honorific agreement with the verb at the morphological level occurs only for non-honorific (level 0) and mid honorific (level 1) pronouns and other higher honorific levels (levels 2, 3 and 4) have the syntactic means for encoding the honorificity.¹

Table 3.25: Honorific levels in Nepali pronouns

Honorificity	level	Second Person	Third Person
Non-honorific	0	तैं <i>tã</i>	यो <i>jo</i> , त्यो <i>tjo</i> , ऊ <i>u</i>
Mid-honorific	1	तिमी <i>timi</i> :	यी <i>ji</i> ; ती <i>ti</i> ; यिनी <i>jini</i> ; तिनी <i>tini</i> ; उनी <i>uni</i> ;
High-honorific	2	तपाईं <i>tapaĩ</i> :	उहाँ <i>uɦã</i>
HHigh-honorific	3	यहाँ <i>jɦã</i> , आफू <i>ap^hu</i> ., हजुर <i>ɦɔdzur</i>	उहाँ <i>uɦã</i> , आफू <i>ap^hu</i> ., हजुर <i>ɦɔdzur</i>
Royal-honorific	4	मौसुफ <i>mɔsup^h</i>	मौसुफ <i>mɔsup^h</i>

3.3.2 Grouping of pronouns

The pronouns cannot be grouped like nouns. Each pronoun in Nepali is unique in form and meaning. Therefore, they are treated and illustrated individually. However, for convenience, we have grouped them in terms of their forms to demonstrate the finite-state network.

a. Personal pronouns

First person: First person pronouns have two forms: singular *म* *mɔ* and plural *हामी* *ɦami*. Both first person singular and plural have oblique forms. First person singular pronoun has direct, oblique, emphatic forms, and genitive: masculine, feminine, plural

¹ Though the pronouns in Nepali in terms of honorificity are not morphologically significant, they have been tagged into five levels for computational purpose in this study.

and emphatic forms. But, first person plural pronoun has direct, oblique forms and genitive: masculine, feminine, plural and emphatic forms. Table 3.26 lists first person singular forms with their corresponding morphological tags.

Table 3.26: First person singular pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+1SG	म	mʌ	I
PRON+1SG+OBL	मै	mʌi	I
PRON+1SG+EMPH	मै	mʌi	I
PRON+1SG+OBL+GEN+MASC	मेरो	mero	my
PRON+1SG+OBL+GEN+FEM	मेरी	meri:	my
PRON+1SG+OBL+GEN+PL	मेरा	mera	my
PRON+1SG+OBL+GEN+HON	मेरा	mera	my
PRON+1SG+OBL+GEN+OBL	मेरा	mera	my
PRON+1SG+OBL+GEN+EMPH	मेरै	merʌi	my

The finite state transducer in Figure 3.14 encodes the first person singular pronouns in Nepali presented in Table 3.26. The finite state transducer in Figure 3.14 is capable of analyzing and generating the pronouns of Table 3.26.

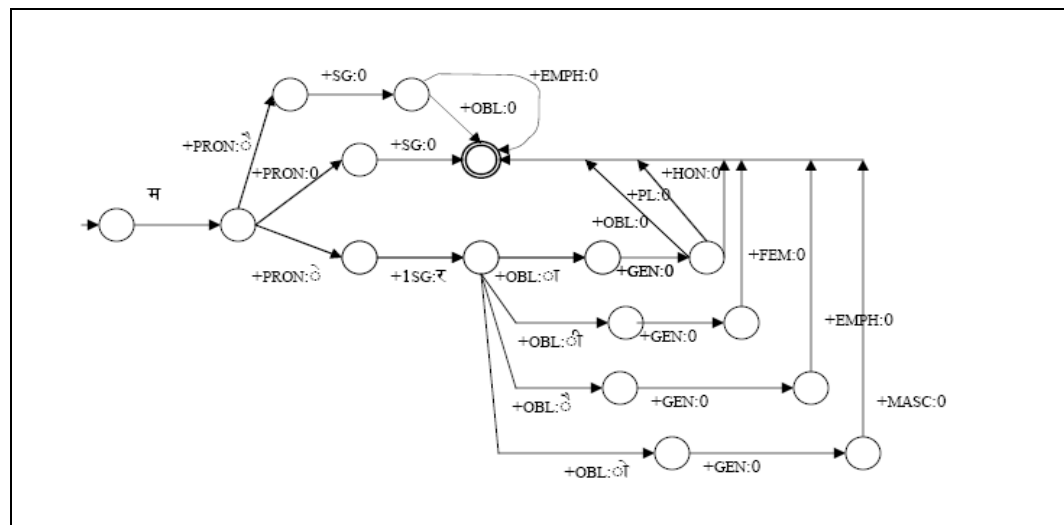


Figure 3.14: A finite state transducer for first person singular pronouns

The first person plural pronouns in Nepali are presented in Table 3.27 with their corresponding morphological tags.

Table 3.27: First person plural pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+1PL	हामी	fiami:	we
PRON+1PL+OBL+GEN+MASC	हाम्रो	fiamro	our
PRON+1PL+OBL+GEN+FEM	हाम्री	fiamri:	our
PRON+1PL+OBL+GEN+PL	हाम्रा	fiamra	our
PRON+1PL+OBL+GEN+HON	हाम्रा	fiamra	our
PRON+1PL+OBL+GEN+OBL	हाम्रा	fiamra	our
PRON+1PL+OBL+GEN+EMPH	हाम्रै	fiamrʌi	our

The finite state transducer illustrated in Figure 3.15 is capable of analyzing and generating the plural pronouns illustrated in Table 3.27.

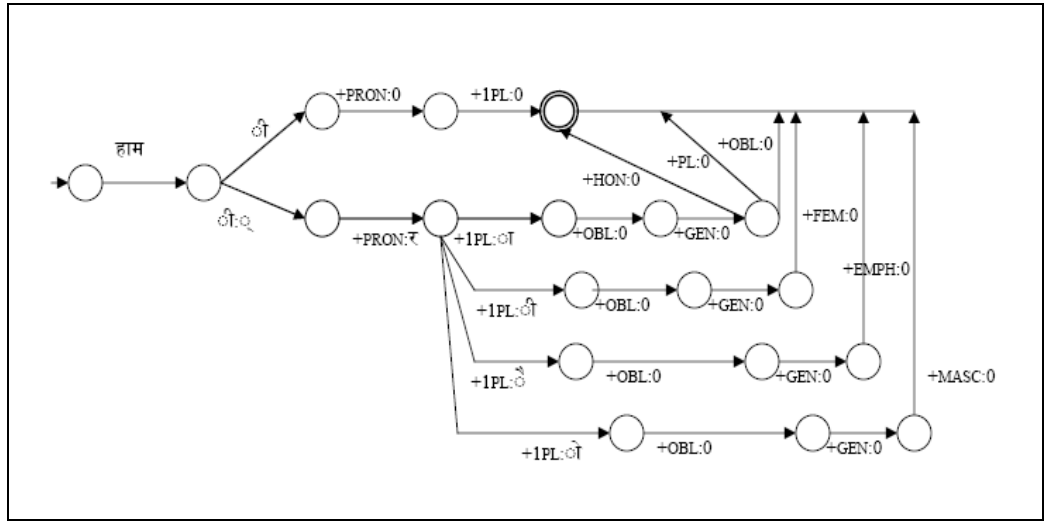


Figure 3.15: A finite state transducer for first person plural pronouns

Second person: The second person pronouns can be grouped into two classes. One consists of *तँ* *tã* 'you' and *तिमी* *timi* 'you' which have various forms for direct, oblique, emphatic and genitive: masculine, feminine, plural and emphatic. And another group consists of *तपाईँ* *tapaĩ*, *उहाँ* *uhã*, *यहाँ* *jaɦã*, *आफू* *ap^hu*, *हजुर* *ɦadzur* and *मौसुफ* *mausup^h* which do not have any other forms. Table 3.28 lists second person non-honorific singular forms with their corresponding morphological tags.

Table 3.28: Second person singular non-honorific pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+2SG	तँ	tã	you
PRON+2SG+OBL	तँ	tãĩ	you
PRON+2SG+EMPH	तँ	tãĩ	you
PRON+2SG+OBL+GEN+MASC	तेरो	tero	your
PRON+2SG+OBL+GEN+FEM	तेरी	teri:	your
PRON+2SG+OBL+GEN+PL	तेरा	tera	your
PRON+2SG+OBL+GEN+HON	तेरा	tera	your
PRON+2SG+OBL+GEN+OBL	तेरा	tera	your
PRON+2SG+OBL+GEN+EMPH	तेरै	terai	your

The second person singular non-honorific pronouns in Nepali are encoded into a finite state transducer as demonstrated in Figure 3.16 which is capable of analyzing and generating the pronouns listed in Table 3.28.

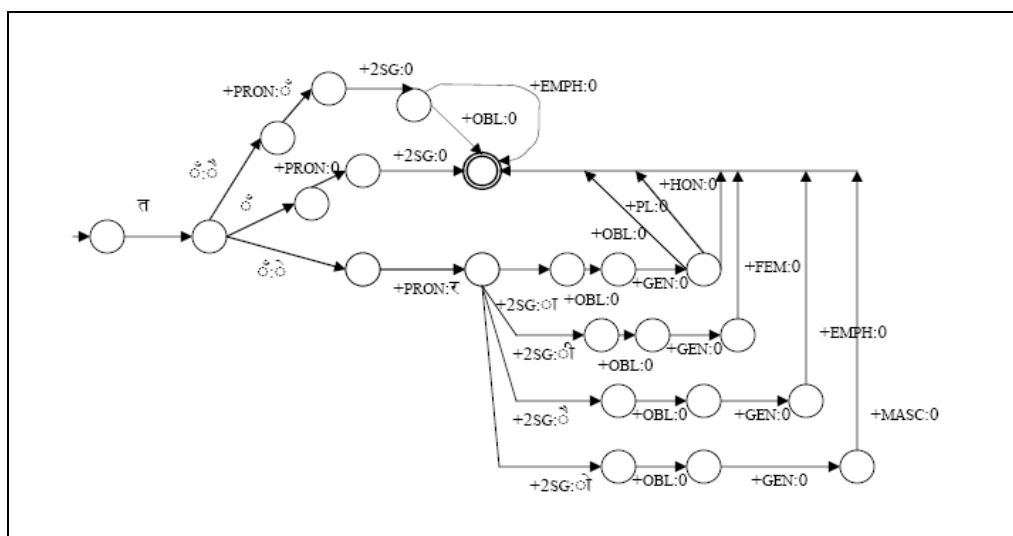


Figure 3.16: A finite state transducer for second person singular non-honorific pronouns

Table 3.29 lists second person singular honorific forms with their corresponding morphological tags.

Table 3.29: Second person honorific pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+2SG+HON	तिमी	timi:	you
PRON+2SG+OBL+HON+GEN+MASC	तिम्रो	timro	your
PRON+2SG+OBL+HON+GEN+FEM	तिमी	timri:	your
PRON+2SG+OBL+HON+GEN+PL	तिम्रा	timra	your
PRON+2SG+OBL+HON+GEN+HON	तिम्रा	timra	your
PRON+2SG+OBL+HON+GEN+OBL	तिम्रा	timra	your
PRON+2SG+OBL+HON+GEN+EMPH	तिम्रै	timrai	your

The finite state transducer illustrated in Figure 3.17 encodes the second person honorific pronouns in Nepali and it is capable of analyzing and generating the pronouns listed in Table 3.29.

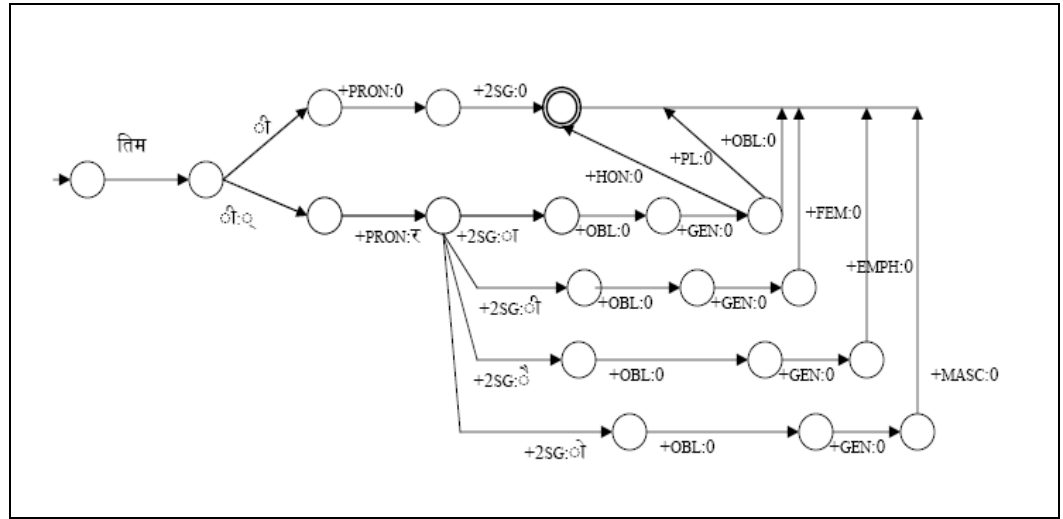


Figure 3.17: A finite state transducer for second person honorific pronouns

Table 3.30 lists second person high honorific singular forms with their corresponding morphological tags.

Table 3.30: Second person high honorific pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+2SG+HHON	तपाईं	tʌpaĩ:	you
PRON+2SG+HHON	यहाँ	jʌɦã:	you
PRON+2SG+HHON	उहाँ	uɦã	you
PRON+2SG+HHON	वहाँ	wʌɦã	you
PRON+2SG+HHON	हजुर	ɦʌdzur	you

The finite state transducer demonstrated in Figure 3.18 encodes the second person high honorific pronouns in Nepal and it is capable of analyzing and generating the pronouns listed in Table 3.30.

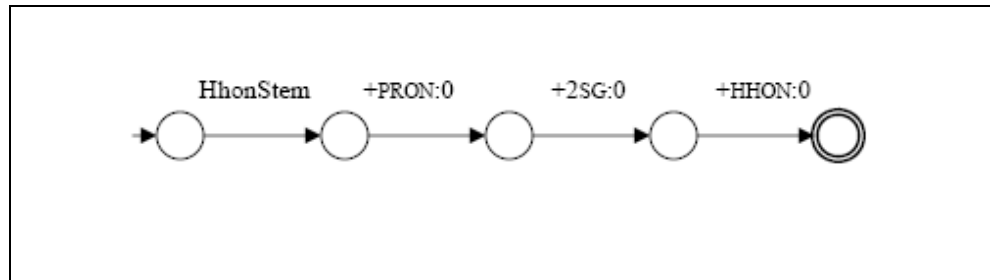


Figure 3.18: A finite state transducer for second person higher honorific pronouns

A second person royal honorific pronoun in Nepali is given in Table 3.31 with its corresponding morphological tags.

Table 3.31: Second person royal honorific pronoun

Morphological Tags	Devanagari	IPA	Gloss
PRON+2SG+RHON	मौसुफ	mΛusup ^h Λ	you.royal

The finite state transducer in Figure 3.19 encodes the royal honorific pronoun and it is capable of analyzing and generating it.

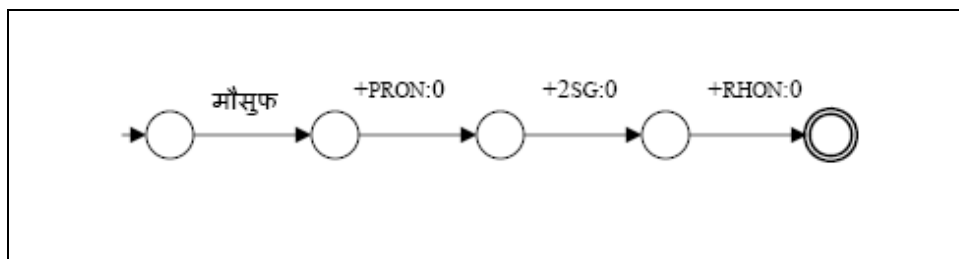


Figure 3.19: A finite state transducer for second person highest honorific pronoun

Third person: The third person pronouns can be grouped into three distinct sets. The first one is ऊ *u:* and its various forms. ऊ *u:* inflects for form: direct and oblique, honorificity: non-honorific and honorific; and emphatic. Table 3.32 lists the pronoun ऊ *u:* and its various forms with their corresponding morphological tags.

Table 3.34: Third person pronouns यो *jo* and यी *ji*:

Morphological Tags	Devanagari	IPA	Gloss
PRON+3SG+PROX	यो	jo	s/he
PRON+3SG+PROX+EMPH	यही	jʌfi	s/he
PRON+3SG+OBL+PROX	यस	jʌsʌ	s/he
PRON+3SG+OBL+PROX+EMPH	यसै	jʌsʌi	s/he
PRON+3SG+PROX+HON	यी	ji:	s/he
PRON+3PL+PROX	यी	ji:	s/he
PRON+3SG+PROX+HON	यिनी	jini:	s/he
PRON+3SG+PROX+OBL+HON	यिन	jɪnʌ	s/he
PRON+3SG+PROX+OBL+HON+EMPH	यिनै	jɪnʌi	s/he

The finite state transducer illustrated in Fig 3.22 encodes the pronouns listed in Table 3.34 and it is capable of analyzing and generating the third person pronouns यो *jo*, यी *ji* and their various forms illustrated in Table 3.34.

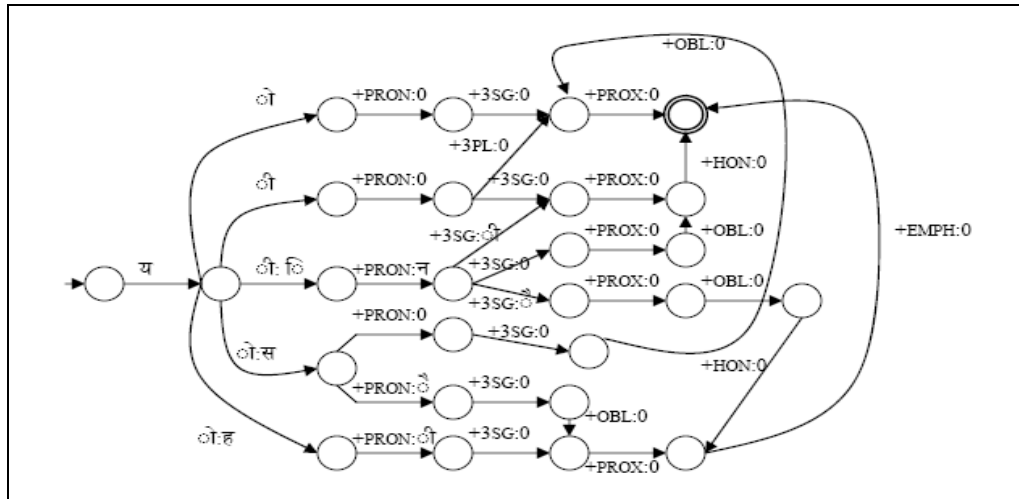


Figure 3.22: A finite state transducer for third person pronouns यो *jo* and यी *ji*:

b. Reflexive pronoun

There is a single reflexive pronoun आफू *ap^hu*: 'self' in Nepali. But it has various forms. It inflects for form: direct and oblique, genitive case: singular, plural, honorific, oblique and feminine, and emphatic. The Table 3.35 lists आफू *ap^hu*: 'self' and its various forms with their corresponding morphological tags.

Table 3.35: The reflexive pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+REFL	आफू	ap ^h u:	self
PRON+REFL+OBL+EMPH	आफै	ap ^h ʌi	self
PRON+REFL+OBL+EMPH	आफै	ap ^h ʌĩ	self
PRON+REFL+OBL+GEN+SG	आफ्नो	ap ^h no	own
PRON+REFL+OBL+GEN+PL	आफ्ना	ap ^h na	own
PRON+REFL+OBL+GEN+HON	आफ्ना	ap ^h na	own
PRON+REFL+OBL+GEN+OBL	आफ्ना	ap ^h na	own
PRON+REFL+OBL+GEN+FEM	आफ्नी	ap ^h ni:	own
PRON+REFL+OBL+GEN+EMPH	आफ्नाै	ap ^h nʌi	own

The finite state transducer illustrated in Figure 3.23 is capable of analyzing and generating the reflexive pronoun *आफू* ap^hu: and its various forms illustrated in Table 3.35.

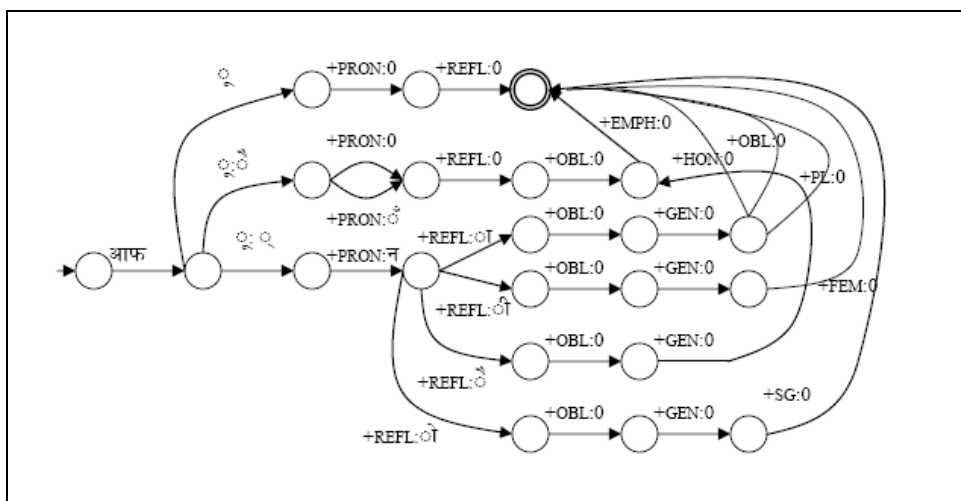


Figure 3.23: A finite state transducer for reflexive pronouns

c. Demonstrative pronouns

The demonstrative pronouns can be grouped into four distinct sets. The first one is *यो* *jo* and *यी* *ji*: and their various forms. *यो* *jo* and *यी* *ji*: inflect for form: direct and oblique and emphatic. Table 3.36 lists the demonstrative pronouns *यो* *jo* and *यी* *ji*: and their various forms with their corresponding morphological tags.

Table 3.36: The demonstrative pronouns यो *jo* and यी *ji*:

Morphological Tags	Devanagari	IPA	Gloss
PRON+DEM+PROX	यो	jo	this
PRON+DEM+PROX+EMPH	यही	jʌfi:	this one
PRON+DEM+PROX	यी	ji:	these
PRON+DEM+PROX+HON	यिनी	jini:	these
PRON+DEM+PROX+OBL	यिन	jɪnʌ	these
PRON+DEM+PROX+OBL+EMPH	यिनै	jɪnʌi	these ones
PRON+DEM+PROX+HON	यहाँ	jʌfã	you

The finite state transducer illustrated in Figure 3.24 is capable of analyzing and generating the demonstrative pronouns यो *jo*, यी *ji*: and their various forms illustrated in Table 3.36.

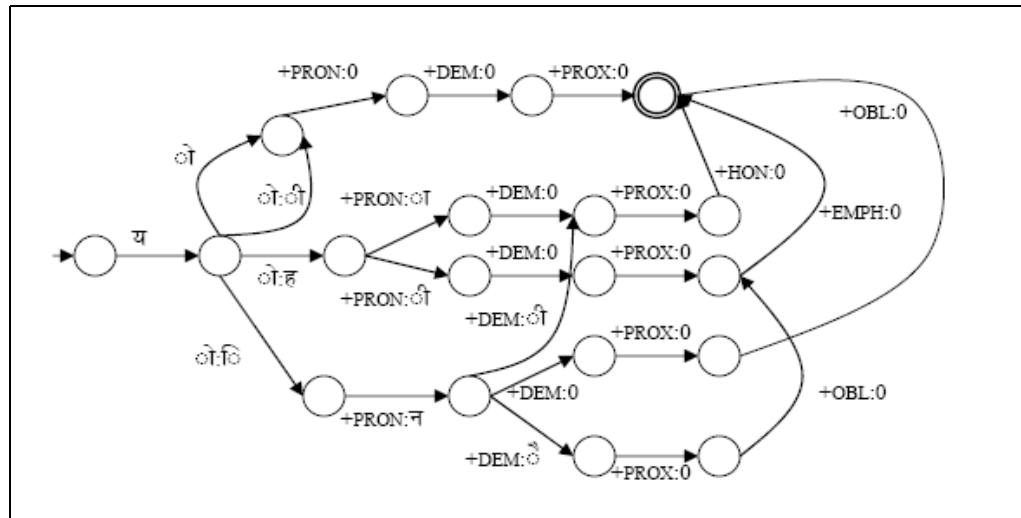


Figure 3.24: A finite state transducer for demonstrative pronouns यो *jo* and यी *ji*:

The second one is त्यो *tjo* and ती *ti*: and their various forms. त्यो *tjo* and ती *ti*: inflect for form: direct and oblique; and emphatic. Table 3.37 lists the demonstrative pronoun त्यो *tjo* and ती *ti*: and their various forms with their corresponding morphological tags.

Table 3.37: The demonstrative pronouns *त्यो tjo* and *ती ti:*

Morphological Tags	Devanagari	IPA	Gloss
PRON+DEM+DIST	त्यो	tjo	that
PRON+DEM+DIST+EMPH	त्यही	tjʌfi:	that one
PRON+DEM+DIST	ती	ti:	those
PRON+DEM+DIST+OBL+HON	तिनी	tini:	those
PRON+DEM+DIST+OBL	तिन	tinʌ	those
PRON+DEM+DIST+OBL+EMPH	तिनै	tinʌi	those

The finite state transducer illustrated in Figure 3.25 is capable of analyzing and generating the demonstrative pronouns *त्यो tjo* and *ती ti:* and their various forms illustrated in Table 3.37.

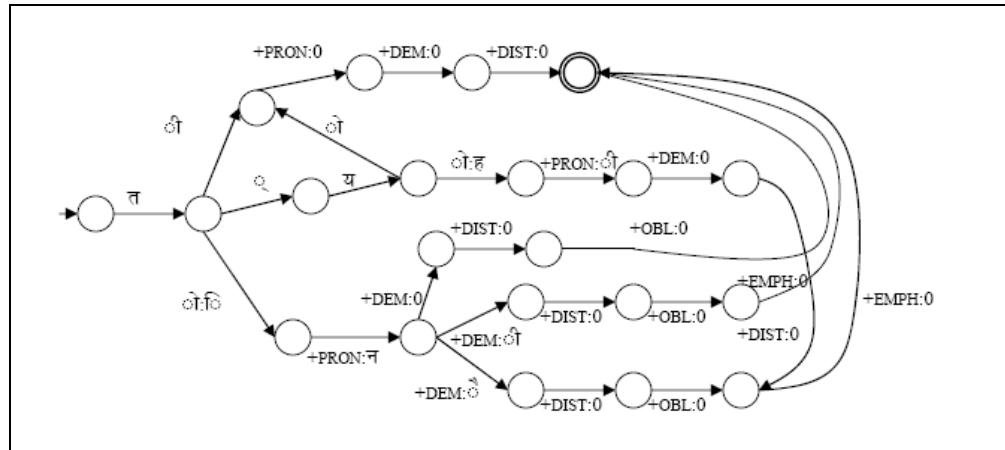


Figure 3.25: A finite state transducer for demonstrative pronouns *त्यो tjo* and *ती ti:*

The third one is *ऊ u:* and its various forms. *ऊ u:* inflects for form: direct and oblique, and emphatic. Table 3.38 lists the pronoun *ऊ u:* and its various forms with their corresponding morphological tags.

Table 3.38: The demonstrative pronouns ऊ u:

Morphological Tags	Devanagari	IPA	Gloss
PRON+DEM+DIST	ऊ	u	that
PRON+DEM+DIST+EMPH	उही	ufi:	that same
PRON+DEM+DIST+HON	उनी	uni:	that
PRON+DEM+DIST+OBL	उन	unʌ	that
PRON+DEM+DIST+OBL+EMPH	उनै	unʌi	that
PRON+DEM+DIST+HON	उहाँ	ufhã	there
PRON+DEM+DIST+HON	वहाँ	wʌfhã	there

The finite state transducer illustrated in Fig 3.26 is capable of analyzing and generating the demonstrative pronouns ऊ u: and its various forms illustrated in Table 3.38.

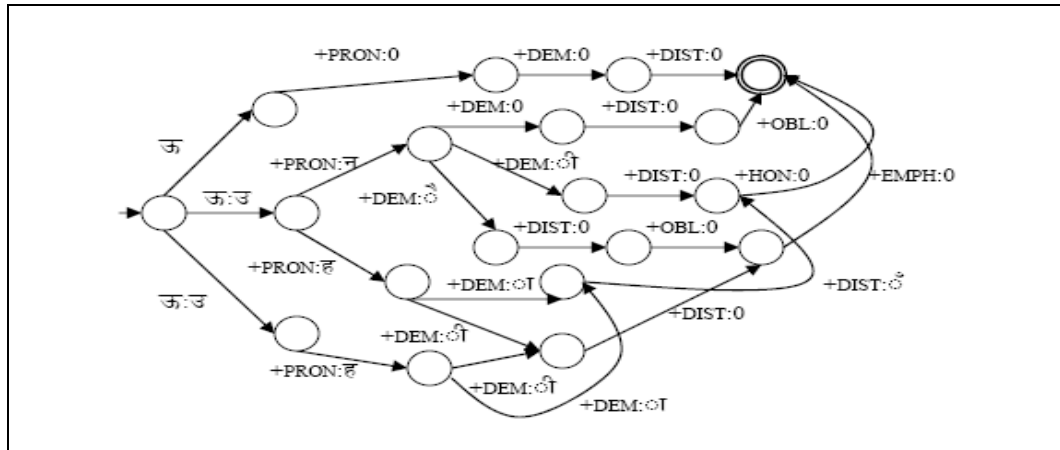


Figure 3.26: A finite state transducer for demonstrative pronouns ऊ u:

The fourth one is remaining demonstratives and their various forms that inflect only for emphatic. Table 3.39 lists the remaining demonstrative pronouns and their various forms with their corresponding morphological tags.

Table 3.39: The remaining demonstrative pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+DEM+DIST	सो	so	that
PRON+DEM+DIST+EMPH	सोही	sofi	that
PRON+DEM+PROX	निज	nidzΛ	him/her
PRON+DEM+PROX+EMPH	निजै	nidzΛi	him/her
PRON+DEM+PROX	उक्त	uktΛ	that

The finite state transducer illustrated in Figure 3.27 is capable of analyzing and generating the remaining demonstrative pronouns and their various forms illustrated in Table 3.39.

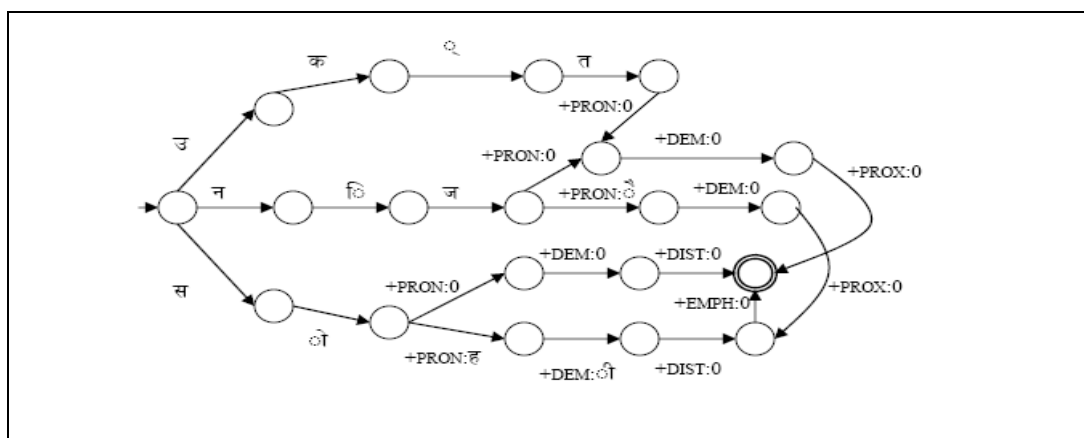


Figure 3.27: A finite state transducer for remaining demonstrative pronouns

d. Relative pronouns

There are three relative pronouns जो *dzo*, जे *dze* and जुन *dzunΛ* in Nepali. These relative pronouns inflect only for oblique and emphatic forms. Table 3.40 lists relative pronouns and their various forms with their corresponding morphological tags.

Table 3.40: The Relative Pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+REL+HUM	जो	dzo	who
PRON+REL+OBL+HUM	जस	dzΛsΛ	who
PRON+REL+OBL+HUM+EMPH	जसै	dzΛsΛi	who
PRON+REL+NHUM	जे	dze	which
PRON+REL	जुन	dzunΛ	which
PRON+REL+EMPH	जुनै	dzunΛi	which

The finite state transducer illustrated in Figure 3.28 is capable of analyzing and generating the relative pronouns and their various forms illustrated in Table 3.40.

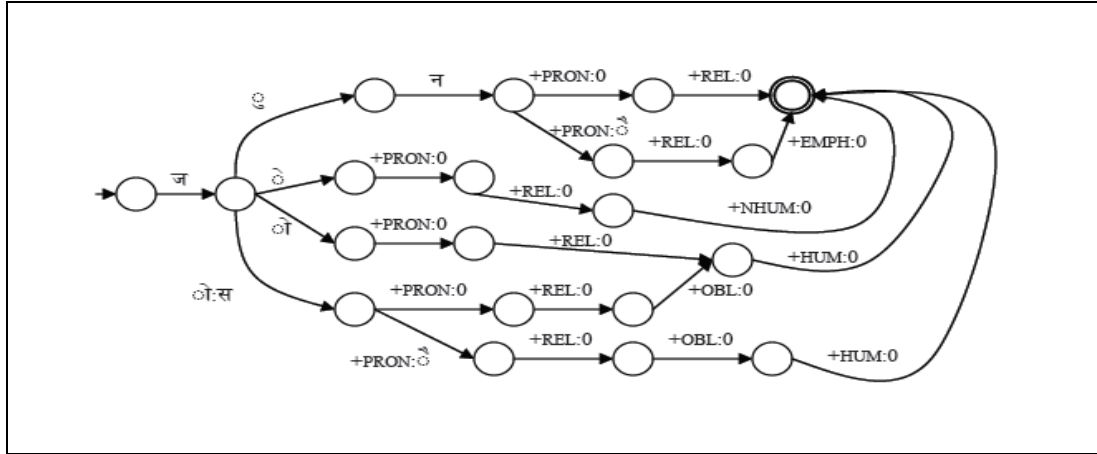


Figure 3.28: A finite state transducer for relative pronouns

e. Interrogative pronouns

There are three interrogative pronouns *को ko*, *के ke* and *कुन kun* in Nepali. But two adverbs which act as interrogative form *किन kin* and *कसरी kasari* are also included here. These interrogative pronouns inflect only for oblique and emphatic forms. Table 3.40 lists relative pronouns and their various forms with their corresponding morphological tags.

Table 3.41a: The interrogative pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+INTERRO+HUM	को	ko	who
PRON+INTERRO+HUM+OBL	कस्	kʌs	who
PRON+INTERRO+HUM+OBL	कसै	kʌsʌi	who
PRON+INTERRO+NHUM	के	ke	what
PRON+INTERRO	कुन	kun	which
PRON+INTERRO	किन	kinʌ	why
PRON+INTERRO	कसरी	kʌsʌri	how

The finite state transducer illustrated in Figure 3.29 is capable to analyze and generate the relative pronouns and their various forms illustrated in Table 3.41a.

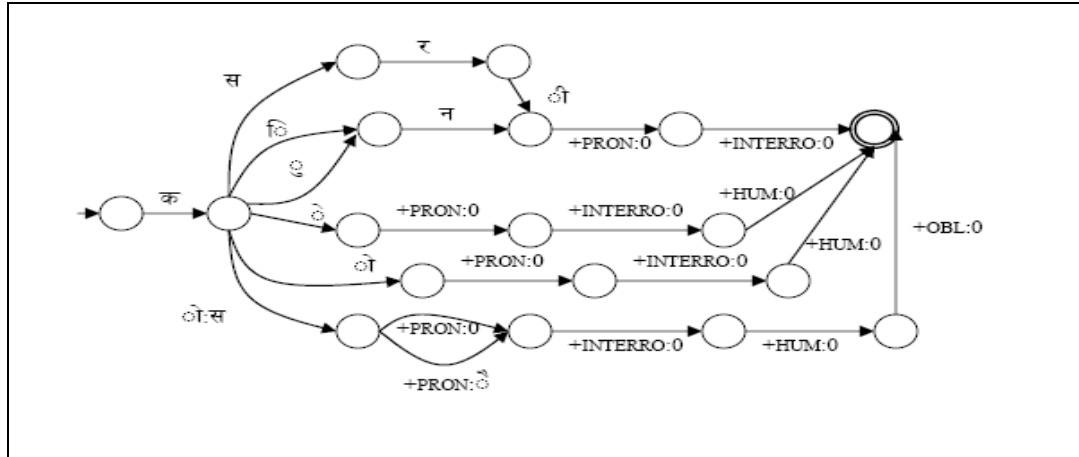


Figure 3.29: A finite state transducer for interrogative pronouns

f. Indefinite pronouns

The indefinite pronouns are derived from interrogative and relative pronouns. The indefinite pronouns derived from interrogative pronouns take *ही* *hi:* and *ै* *ai* as an emphatic marker. And those derived from relative pronouns take *सुकै* *sukai* as an emphatic marker. Table 3.41b lists indefinite pronouns derived from interrogative pronouns with their corresponding morphological tags.

Table 3.41b: The indefinite pronouns derived from interrogative pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+INDEF+HUM	कोही	kofi	someone
PRON+INDEF+NHUM	केही	kefi	something
PRON+INDEF+NEU	कुनै	kunai	anything

The finite state transducer illustrated in Figure 3.30 is capable of analyzing and generating the indefinite pronouns listed in Table 3.41b.

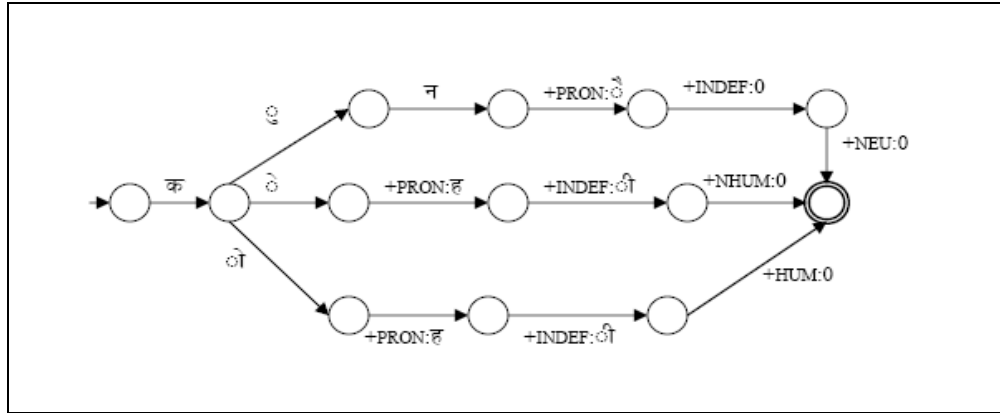


Figure 3.30: A finite state transducer for indefinite pronouns derived from interrogative pronouns

Table 3.42 lists indefinite pronouns derived from relative pronouns with their corresponding morphological tags.

Table 3.42: The indefinite pronouns derived from relative pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+INDEF+HUM	जोसुकै	dzosukʌi	whoever
PRON+INDEF+NHUM	जेसुकै	dzesukʌi	whatever
PRON+INDEF+NEU	जुनसुकै	dzunsukʌi	whichever

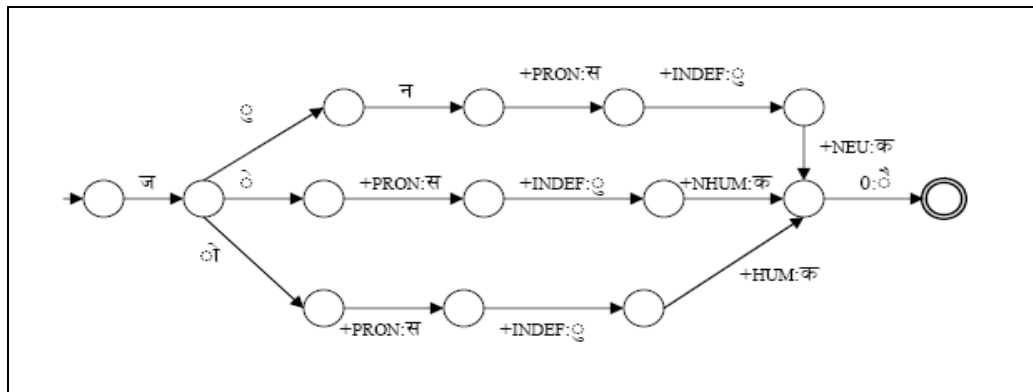


Figure 3.31: A finite state transducer for indefinite pronouns derived from relative pronouns

The finite state transducer illustrated in Figure 3.31 is capable of analyzing and generating the indefinite pronouns and their various forms illustrated in Table 3.42.

g. Definite pronouns

There is a small set of definite pronouns, which does not show any kind of inflections except अर्को *Arko*. अर्को *Arko* inflects for number, honorificity and form: oblique. Table 3.43 lists the definite pronouns with their corresponding morphological tags.

Table 3.43a: The definite pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+DEF	प्रत्येक	prʌtekʌ	everyone
PRON+DEF	हरेक	hʌrekʌ	each one
PRON+DEF	सबै	sʌbʌi	all
PRON+DEF	अरू	ʌru:	other

The finite state transducer in Figure 3.32a encodes the definite pronouns listed in Table 3.43a and it is capable of analyzing and generating those pronouns.

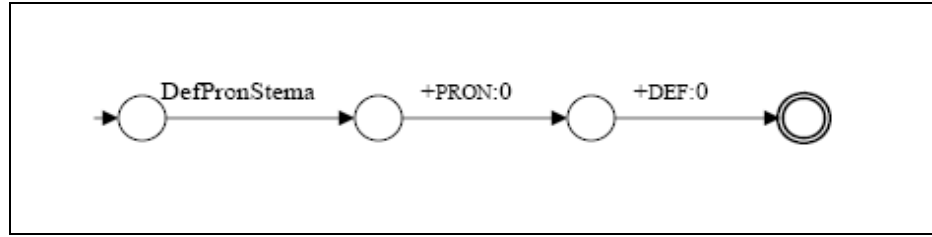


Figure 3.32a: A finite state transducer for definite pronouns

The definite pronoun अर्को along with its various forms and their corresponding morphological tags are listed in Table 3.43b.

Table 3.43b: The definite pronoun अर्को

Morphological Tags	Devanagari	IPA	Gloss
PRON+DEF+SG	अर्को	ʌrko	another
PRON+DEF+PL	अर्का	ʌrka	another
PRON+DEF+HON	अर्का	ʌrka	another
PRON+DEF+OBL	अर्का	ʌrka	another
PRON+DEF+FEM	अर्की	ʌrki:	another
PRON+DEF+EMPH	अर्कै	ʌrkʌi	another

The definite pronoun *अर्को* *ariko* and its various forms listed in Table 3.43b have been compiled into a finite state transducer as demonstrated in Figure 3.32b and it is capable of analyzing and generating them.

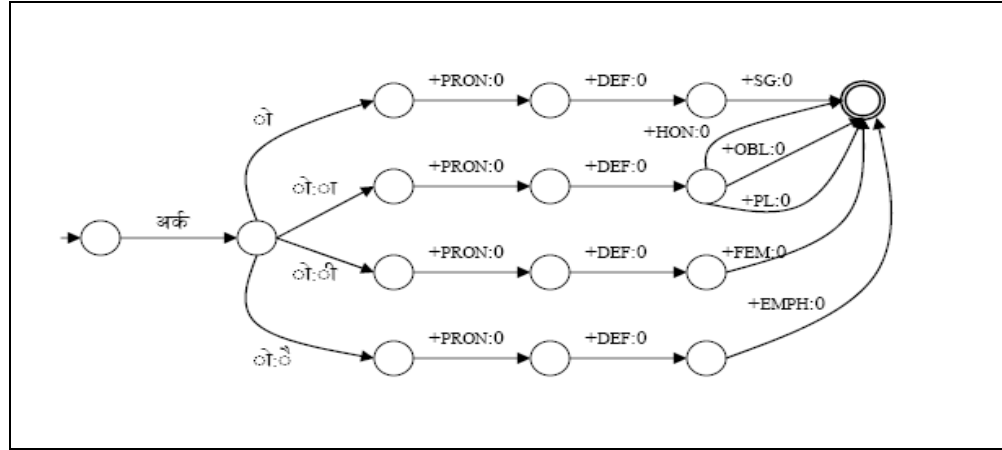


Figure 3.32b: A finite state transducer for definite pronouns

h. Reciprocal pronouns

The reciprocal pronouns in Nepali are compound forms except one, i.e., *आपस* *apasa*. A reciprocal pronoun *एकअर्को* *ekariko* 'each other' inflect for form: oblique, honorificity, number: plural and gender: feminine. Table 3.44 lists the reciprocal pronoun *एकअर्को* *ekariko* and its various forms with their corresponding morphological tags.

Table 3.44a: The reciprocal pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+RECIP	एकअर्को	ekariko	each other
PRON+RECIP+OBL	एकअर्का	ekarika	each other
PRON+RECIP+HON	एकअर्का	ekarika	each other
PRON+RECIP+PL	एकअर्का	ekarika	each other
PRON+RECIP+FEM	एकअर्की	ekarki:	each other
PRON+RECIP+EMPH	एकअर्कै	ekarikai	each other

The finite state transducer demonstrated in Figure 3.33a encodes the reciprocal pronouns listed in Table 3.44a and is capable of analyzing and generating them.

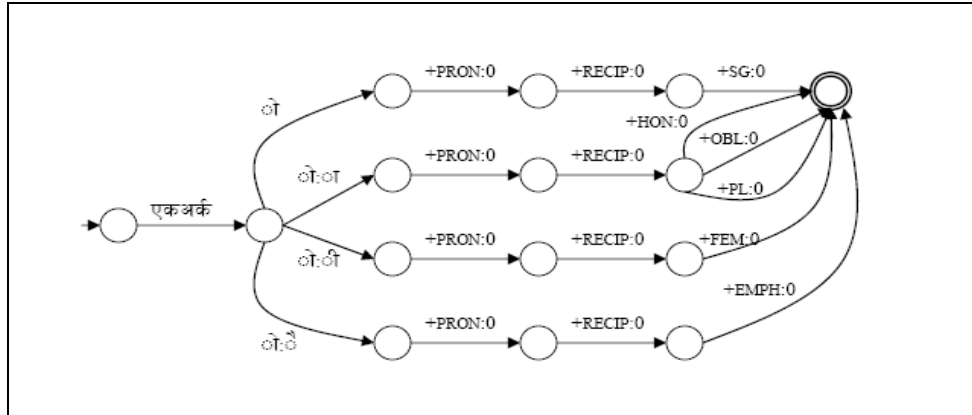


Figure 3.33a: A finite state transducer for reciprocal pronouns

Some other reciprocal pronouns are listed in Table 3.44b with their corresponding morphological tags.

Table 3.44b: The reciprocal pronouns

Morphological Tags	Devanagari	IPA	Gloss
PRON+RECIP	एक आपस	ekapʌs	each other
PRON+RECIP	आपस	apʌs	each other
PRON+RECIP	आआफू	aphu:	each other

The finite state transducer illustrated in Figure 3.33b is capable of analyzing and generating the reciprocal pronouns and their various forms illustrated in Table 3.44b.

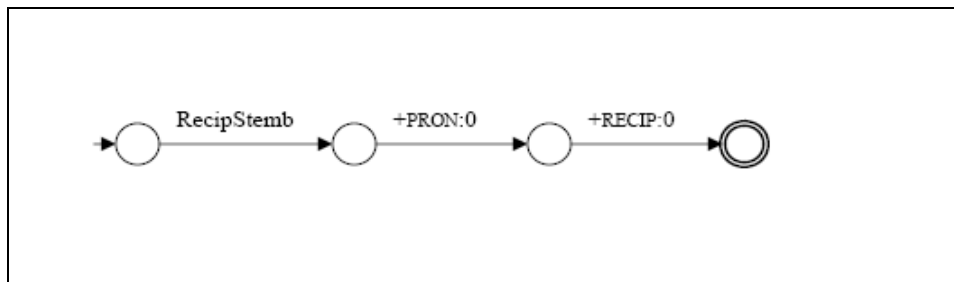


Figure 3.33b: A finite state transducer for reciprocal pronouns

3.4 Adjectives

Adjectives in Nepali are the words indicating quality, quantity and frequency generally modifying the nouns. The adjectives show various kinds of morphological features which are discussed in the following sections.

3.4.1 Characteristics of adjectives in Nepali

a. Significant stem finals

The adjectives in Nepali, like that of nouns, show the binary division between *o*-ending adjectives and non-*o*-ending adjectives. The *o*-ending adjectives inflect for number, gender, form and honorificity. These adjectives agree with the features carried over by the head nouns that they modify. The non-*o*-ending adjectives are not consistent in their formal behavior. Rather a sub-group of non-*o*-ending adjectives take feminine gender marker and another sub-group, especially Sanskrit loan adjectives, inflects for comparative and superlative forms. Table 3.45 lists some *o*-ending and some non-*o*-ending adjectives.

Table 3.45: *O*-ending and non-*o*-ending adjectives

<i>O</i> -ending Adjectives		Non- <i>o</i> -ending Adjectives	
Stems	Gloss	Stem	Gloss
राम्रो ramro	good	असल Δ sAl	good
कालो kalo	black	चतुर ts Δ tur	clever
खस्रो k ^h Δ sro	coarse	लघु lAg ^h u	small
मिठो m ^h ito	sweet	पुर्वीया purwija	related to east

b. Number

Adjectives in Nepali show two dimensions of number: singular and plural. The number distinction is found only in *o*-ending adjectives. The citation form of *o*-ending adjective as राम्रो ramro in (18a) changes to the *a*-ending as राम्रा ramra in (18b) for plural.

(18) a. एउटा राम्रो केटो आयो ।

euṭa ramro keṭo a-jo
one.CL good.SG boy.SG come-P.3SG.MASC
'A handsome boy came.'

- b. दुइटा राम्रा केटा आए ।
 duiṭa ramra keṭa a-je
 two.CL good.PL boy.PL come-P.3PL
 'Two handsome boys came.'

Table 3.46 lists some adjectives that show the singular and plural form and this number feature in the adjectives agree with the number feature of the head noun in the noun phrase.

Table 3.46: Number: singular and plural

	good	black	coarse	old
Singular	राम्रो ramro	कालो kalo	खस्रो k ^h ʌsro	बुढो buḍ ^h o
Plural	राम्रा ramra	काला kala	खस्रा k ^h ʌsra	बुढा buḍ ^h a

c. Gender

Adjectives in Nepali that are *o*-ending show masculine and feminine gender. The *o*-ending adjective such as राम्रो *ramro* in (19a) changes to the *i*-ending as राम्री *ramri*: in (19b) showing masculine and feminine alternation. Some of the non-*o*-ending adjectives change into feminine adjective with the suffix -नी-*ni*: (alternately -इनी-*ini*: and -एनी-*eni*).

- (19) a. एउटा राम्रो केटो आयो ।
 euṭa ramro keṭo a-jo
 one.CL good.MASC.SG boy.MASC.SG come-P.3SG.MASC
 'A handsome boy came.'
- b. एउटी राम्री केटी आई ।
 euṭi ramri keṭi a-i:
 one.CL.FEM good.FEM.SG boy.FEM.SG come-P.3SG.FEM
 'A beautiful girl came.'

Table 3.47 lists some examples of adjectives showing the gender change. The gender distinction depends on the head noun. If head noun refers to human, then only the gender is functional.

Table 3.47: Gender: masculine and feminine

	Good	black	clever	rural
Masculine	राम्रो ramro	कालो kalo	चतुर tsaturΛ	पाखे pak ^h e
Feminine	राम्री ramri:	काली kali:	चतुर्नी tsaturni:	पखिनी pak ^h ini:

d. Form

Adjectives in Nepali show two forms: direct and oblique. The *o*-ending adjective as राम्रो *ramro* in (20a) shows oblique form and it changes to *a*-ending as राम्रा *ramra* in (20b) showing oblique form.

(20) a. एउटा राम्रो केटो आउँदै छ ।

eʊṭa ramro keṭo a-ũdʌi ts^hΛ
 one.CL good.SG boy come-IMPERF be.NP.3SG.MASC
 'A handsome boy is coming.'

b. एउटा राम्रा केटाले प्रस्ताव राखेको छ ।

eʊṭa ramra keṭa-le prastaw rak^h-eko ts^hΛ
 one.CL good.OBL boy.OBL-ERG proposal keep-PERF be.NP.3SG.MASC
 'A handsome boy has proposed.'

Table 3.48 lists some examples of adjectives showing the direct and oblique forms

Table 3.48: Form: direct and oblique

	good	black	coarse	old
Direct	राम्रो ramro	कालो kalo	खस्रो k ^h Λsro	बुढो buḍ ^h o
Oblique	राम्रा ramra	काला kala	खस्रा k ^h Λsra	बुढा buḍ ^h a

e. Honorificity

Adjectives in Nepali show two levels of honorificity: non-honorific and honorific. The *o*-ending adjectives as राम्रो *ramro* in (21a) changes into *a*-ending as राम्रा *ramra* in (21b) showing non-honorific and honorific, respectively.

(21) a. तँ राम्रो छस्।

tã ramro ts^hʌs
 you.NHON good.NHON be.NP.2SG.NHON
 'You are good.'

b. तिमी राम्रा छौ।

timi ramra ts^hʌu
 you.HON good.HON be.NP.2SG.HON
 'You are good.'

Table 3.49 lists some examples of adjectives showing the honorificity.

Table 3.49: Honorificity: non-honorific and honorific

	good	black	coarse	old
Non-honorific	राम्रो ramro	कालो kalo	खस्रो k ^h ʌsro	बुढो buḍ ^h o
Honorific	राम्रा ramra	काला kala	खस्रा k ^h ʌsra	बुढा buḍ ^h a

f. Degree

Native adjectives in Nepali do not inflect for degree. The degrees in adjectives are handled syntactically. But the Sankrit loan adjectives show three levels of degree morphologically: positive, comparative and superlative. The positive adjective is unmarked as न्यून *nju:nʌ* in (22a). The comparative degree is indicated by a suffix -तर-*tar* as न्यूनतर *nju:nʌ-tar* in (22b) and superlative by a suffix -तम *-tam* as न्यूनतम *nju:nʌ-tam* in (22c).

(22) a. हाम्रो आम्दानी न्यून छ।

fiamro amdani nju:nʌ ts^hʌ
 our income less be.NP.3SG.MASC
 'Our income is less.'

b. हाम्रो आम्दानी न्यूनतर छ।

fiamro amdani nju:nʌ-tar ts^hʌ
 our income less-COMP be.NP.3SG.MASC
 'Our income is lesser.'

- c. हाम्रो आम्दानी न्यूनतम छ ।
 fiamro amdani nju:nA-tam ts^hA
 our income less-SUPER be.NP.3SG.MASC
 'Our income is the least.'

Table 3.50 lists some examples of Sanskrit loan adjectives that show three degrees.

Table 3.50: Degree: positive, comparative and superlative

	less	low	rigorous	small
Positive	न्यून nju:nA	निम्न nimnA	गहन gAhanA	लघु lAg ^h u
Comparative	न्यूनतर nju:nA-tarA	निम्नतर nimnA-tarA	गहनतर gAhanA-tarA	लघुतर lAg ^h u-tarA
Superlative	न्यूनतम nju:nA-tamA	निम्नतम nimnA-tamA	गहनतम gAhanA-tamA	लघुतम lAg ^h u-tamA

3.4.2 Classification of adjectives

On the basis of characteristic features of adjectives in Nepali as discussed in (3.4.1), the adjectives are classified into two major groups. The first one is *o*-ending adjectives whereas the second one is non-*o*-ending adjectives.

a. *O*-ending adjectives

All the *o*-ending adjectives are grouped in a class. The adjectives in this group inflect for number, gender, form and honorificity. The inflection in the adjectives has direct relation with the head noun which it modifies because there is feature agreement between head noun and modifier adjective. Table 3.51 lists some examples of *o*-ending adjectives.

Table 3.51: *O*-ending adjectives

Morphological Tags	good	Black	coarse	old
+ADJ+SG	राम्रो ramro	कालो kalo	खस्रो k ^h ʌsro	बुढो buḍ ^h o
+ADJ+PL	राम्रा ramra	काला kala	खस्रा k ^h ʌsra	बुढा buḍ ^h a
+ADJ+OBL	राम्रा ramra	काला kala	खस्रा k ^h ʌsra	बुढा buḍ ^h a
+ADJ+HON	राम्रा ramra	काला kala	खस्रा k ^h ʌsra	बुढा buḍ ^h a
+ADJ+FEM	राम्री ramri:	काली kali:	खस्री k ^h ʌsri:	बुढी buḍ ^h i:

The finite state transducer illustrated in Figure 3.33 is capable of analyzing and generating the *o*-ending adjectives and their forms illustrated in Table 3.44.

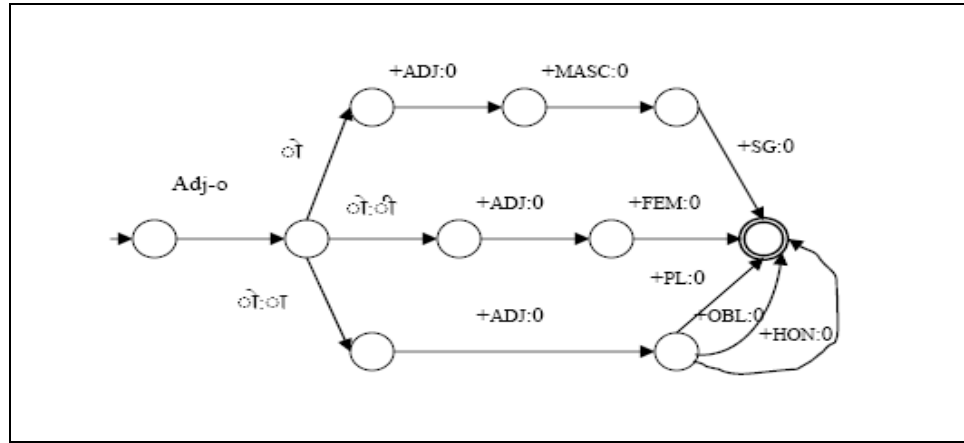


Figure 3.34: A finite state transducer for *o*-ending adjectives

The phonological rules given in PR 3.9 are compiled into a finite state transducer and composed with finite state transducer illustrated in Figure 3.34.

Phonological rule

PR 3.9

- i. Stem final vowel *ो* *o* of the *o*-ending adjectives of the lower language (i.e, surface level) is changed to vowel *ा* *a* for plural, oblique and honorificity.

Regular expression: ो -> ी || _ .#.

- ii. Stem final vowel ो *o* of the *o*-ending adjectives of the lower language (.i.e, surface level) is changed to vowel ी *i* for feminine gender.

Regular expression: ो -> ी || _ .#.

b. Non-*o*-ending adjectives

Non-*o*-ending adjectives in Nepali form a group which includes both marked and unmarked adjectives. Marked adjectives mean those which take some sort of marking such as feminine marker, comparative marker and superlative maker.

i. Marked adjectives

Type 1: Those non-*o*-ending adjectives in Nepali that inflect for gender: masculine and feminine have been grouped in this class. The citation form is masculine in gender and marker -नी /-इनी -*ni:-ini*: when suffixed to changes to feminine gender.

Table 3.52 lists some adjectives of this group.

Table 3.52: Type 1 marked adjectives

Morphological Tags	clever	cunning	of east	rural
ADJ+MASC	चतुर tsaturΛ	धुर्त d ^h urta	पुर्विया purwija	पाखे pak ^h e
ADJ+FEM	चतुर्नी tsaturni:	धुर्तिनी d ^h urtini:	पुर्विनी purwini:	पखिनी pAk ^h eni:

The finite state transducer illustrated in Figure 3.34 is capable of analyzing and generating the non-*o*-ending type 1 adjectives and their forms illustrated in Table 3.52.

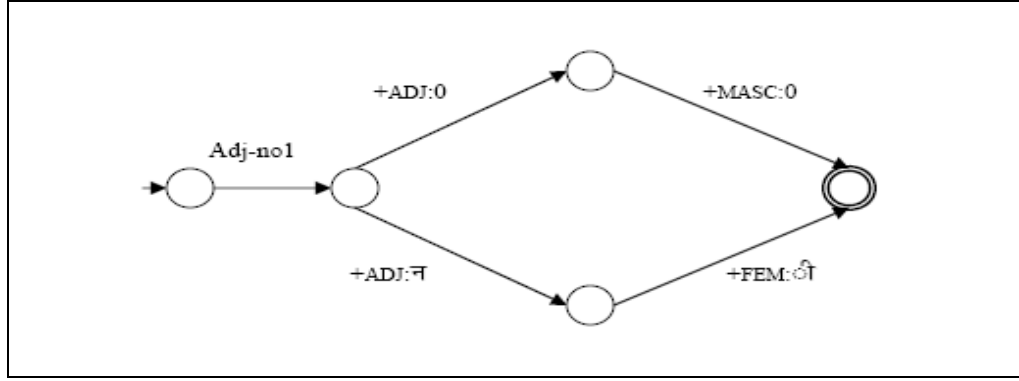


Figure 3.35: A finite state transducer for Type 1 marked adjectives

The phonological rules involved in this process are given in PR 3.10 which are compiled and composed with finite state transducer illustrated in Figure 3.35.

Phonological rule

PR 3.10

- i. Halant ् is inserted between consonant symbol and feminine gender marker *नी*

ni: at the surface level.

Regular expression: [.] -> ् || liquids _ न ी .#.

- ii. *या* *ja* is deleted before the feminine gender marker *नी* *ni:* at the surface level.

Regular expression: य ा -> [] || _ न ी .#.

Type 2 Those non-*o*-ending adjectives in Nepali that inflect for comparative and superlative forms are grouped in this class. The adjectives in this group, in fact, are Sanskrit loan adjectives. The adjectives in this group take the comparative marker -*तर* -*tara* and superlative maker -*तम* -*tama* forming the comparative and superlative forms respectively. Table 3.53 lists some examples of Sanskrit loan adjectives.

Table 3.53: Type 2 marked adjectives

Morphological Tags	less	low	regorous	small
+ADJ+POSIT	न्यून nju:nA	निम्न nimnA	गहन gAɦANA	लघु lAg ^h u
+ADJ+COMP	न्यूनतर nju:nA-tARA	निम्नतर nimnA-tARA	गहनतर gAɦANA-tARA	लघुतर lAg ^h u-tARA
+ADJ+SUPER	न्यूनतम nju:nA-tAMA	निम्नतम nimnA-tAMA	गहनतम gAɦANA-tAMA	लघुतम lAg ^h u-tAMA

The finite state transducer illustrated in Figure 3.35 is capable of analyzing and generating the non-*o*-ending type 2 adjectives and their forms illustrated in Table 3.53. In this class of adjectives, no rules are involved.

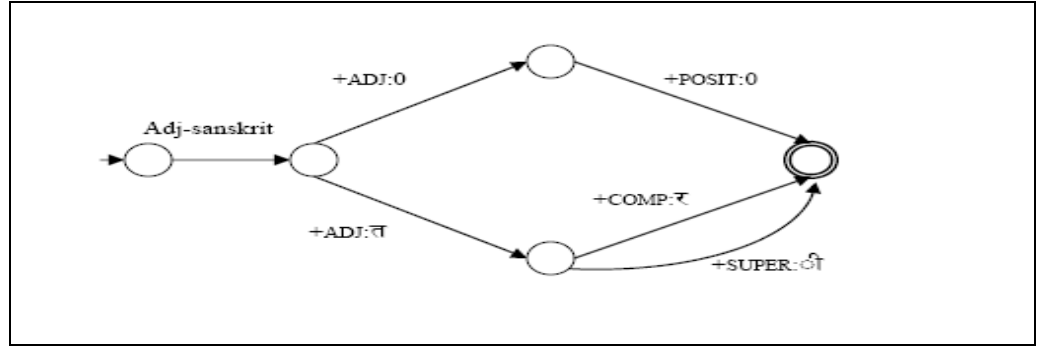


Figure 3.36: A finite state transducer for Sanskrit loan adjectives

ii. Unmarked adjectives

All those non-*o*-ending adjectives in Nepali which never take any marker are grouped in this class. The adjective in this class remains unaltered. Table 3.54 lists some examples of unmarked adjectives.

Table 3.54: Unmarked mdjectives

Morphological Tags	gentle	bad	new	rich
+ADJ	असल ASAl	खराब k ^h Arab	नयाँ nAjã	धनी d ^h Ani

The finite state transducer illustrated in Figure 3.35 is capable of analyzing and generating the non-*o*-ending unmarked adjective forms illustrated in Table 3.54.

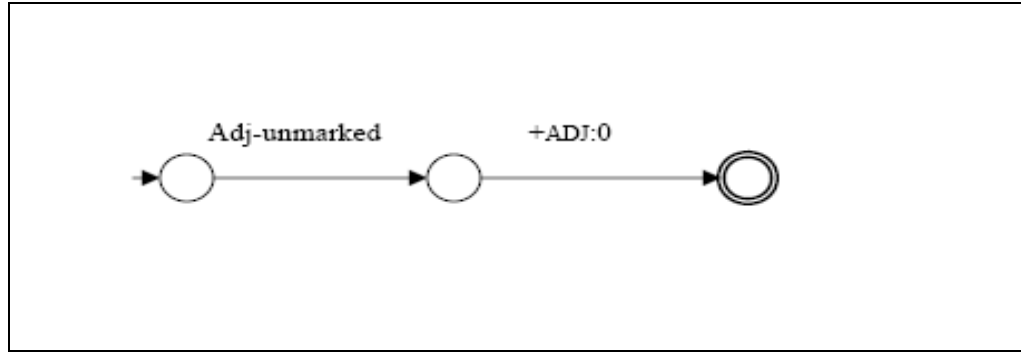


Figure 3.37: A finite state transducer for unmarked adjectives

3.5 Numerals

The numerals in Nepali are of two types: cardinal numbers and ordinal numbers.

3.5.1 Cardinal numbers

Cardinal number in Nepali from one to hundred and some other such as *हजार fiAdzar* 'thousand', *लाख lak^h* 'hundred thousand', *करोड kArOd* 'ten million', *अरब ArAb* 'ten billion' and *खरब k^hArAb* 'ten trillion' are written as a single word. The cardinal numbers appear with numeral classifiers and modify the head nouns. Table 3.55 lists some examples of cardinal numbers.

Table 3.55: Some cardinal numbers

Morphological Tags	Devanagari	IPA	Gloss
+NUM	शून्य	ʃu:njʌ	zero
+NUM+CARD	एक	ek	one
+NUM+CARD	दुई	dui:	two
+NUM+CARD	तीन	ti:n	three
+NUM+CARD	चार	tsar	four
+NUM+CARD	पाँच	pāts	five
+NUM+CARD	छ	ts ^h ʌ	six
+NUM+CARD	सात	sat	seven
+NUM+CARD	आठ	at ^h	eight
+NUM+CARD	नौ	nʌu	nine
+NUM+CARD	दस	dʌs	ten
+NUM+CARD	एघार	eg ^h arʌ	eleven

+NUM+CARD	बाह	bahrΛ	twelve
+NUM+CARD	तेह	tehrΛ	thirteen
+NUM+CARD	चौध	tsΛud ^h Λ	fourteen
+NUM+CARD	पन्ध्र	pΛnd ^h rΛ	fifteen
+NUM+CARD	सोह	sohrΛ	sixteen
+NUM+CARD	सत्र	sΛtrΛ	seventeen
+NUM+CARD	अठार	Λt ^h arΛ	eighteen
+NUM+CARD	उन्नाइस	unnais	nineteen
+NUM+CARD	बीस	bi:s	twenty
+NUM+CARD	एक्काइस	ekkais	twenty one
+NUM+CARD	पच्चीस	pΛtstsi:s	twenty five
+NUM+CARD	तीस	ti:s	thirty
+NUM+CARD	चालीस	tsali:s	fourty
+NUM+CARD	पचास	pΛtsas	fifty
+NUM+CARD	साठी	sa ^h ṭi:	sixty
+NUM+CARD	सत्तरी	sΛttari:	seventy
+NUM+CARD	असी	Λsi:	eighty
+NUM+CARD	नब्बे	nΛbbe	ninenty
+NUM+CARD	सय	sΛjΛ	hundred
+NUM+CARD	हजार	ḥΛdzar	thousand
+NUM+CARD	लाख	lak ^h	hundred thousand
+NUM+CARD	करोड	kΛroḍ	ten million
+NUM+CARD	अरब	ΛrΛb	ten billion
+NUM+CARD	खरब	k ^h ΛrΛb	ten trillion

3.5.2 Ordinal number

The ordinal numbers in Nepali are of two types: regular and irregular.

a. Regular ordinal number: Numbers one, two, three, four and six constitute an exceptional set in the formation of the ordinal numbers from the cardinal numerals.

Except the exceptional set, all the numerals take -औँ -Λū as a suffix and form the ordinal numbers. Some examples are illustrated in Table 3.56.

Table 3.56: Some regular ordinal numbers

Morphological Tags	Devanagari	IPA	Gloss
+NUM+ORD	पाँचौँ	pãtsΛũ	fifth
+NUM+ORD	सातौँ	satΛũ	seventh
+NUM+ORD	आठौँ	at ^h Λũ	eighth
+NUM+ORD	दसौँ	dΛsΛũ	tenth
+NUM+ORD	बीसौँ	bi:sΛũ	twentieth
+NUM+ORD	सयौँ	sΛjΛũ	hundredth
+NUM+ORD	हजारौँ	hΛdzarΛũ	thousandth
+NUM+ORD	लाखौँ	lak ^h Λũ	hundred thousandth
+ORD+NUM	करोड़ौँ	kΛrodΛũ	ten millionth

The finite state transducer for cardinal numbers listed in Table 3.55 and ordinal number listed in Table 3.56 except the exceptional set is illustrated in the Figure 3.37 which is capable of analyzing and generating these numeral forms.

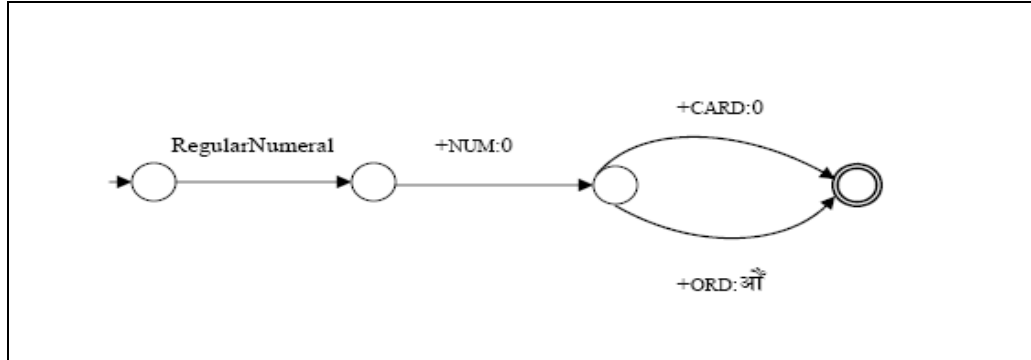


Figure 3.38 A finite state transducer for cardinal numbers and regular ordinal numbers

The phonological rules involved in the regular numerals are given in PR 3.11, which are compiled and composed with finite state transducer illustrated in Figure 3.38.

Phonological rule

PR 3.11

- i. Vowel sequence औँΛũ is changed to its corresponding dependent vowel symbol

ौँΛũ if the numeral ends with consonant.

Regular expression: औँ -> ौँ || cons _ .#.

b. Irregular ordinal numbers: The corresponding ordinal numerals from number one, two, three and four are different from the regular ordinal numerals. They inflect for number, gender, form and honorificity. Table 3.57, Table 3.58, Table 3.59, Table 3.60 list the ordinal numerals and their corresponding morphological tags of number one, two, three and four respectively.

Table 3.57: Irregular ordinal numbers of one

Morphological Tags	Devanagari	IPA	Gloss
+NUM+ORD+MASC	पहिलो	pʌhilo	first
+NUM+ORD+PL	पहिला	pʌhila	first
+NUM+ORD+OBL	पहिला	pʌhila	first
+NUM+ORD+HON	पहिला	pʌhila	first
+NUM+ORD+FEM	पहिली	pʌhili:	first

Table 3.58: Irregular ordinal numbers of two

Morphological Tags	Devanagari	IPA	Gloss
+NUM+ORD+MASC	दोस्रो	dosro	second
+NUM+ORD+PL	दोस्रा	dosra	second
+NUM+ORD+OBL	दोस्रा	dosra	second
+NUM+ORD+HON	दोस्रा	dosra	second
+NUM+ORD+FEM	दोस्री	dosri:	second

Table 3.59: Irregular ordinal numbers of three

Morphological Tags	Devanagari	IPA	Gloss
+NUM+ORD+MASC	तेस्रो	tesro	third
+NUM+ORD+PL	तेस्रा	tesra	third
+NUM+ORD+OBL	तेस्रा	tesra	third
+NUM+ORD+HON	तेस्रा	tesra	third
+NUM+ORD+FEM	तेस्री	tesri:	third

Table 3.60: Irregular ordinal numbers of four

Morphological Tags	Devanagari	IPA	Gloss
+NUM+ORD+MASC	चौथो	tsʌut ^h o	fourth
+NUM+ORD+PL	चौथा	tsʌut ^h a	fourth
+NUM+ORD+OBL	चौथा	tsʌut ^h a	fourth
+NUM+ORD+HON	चौथा	tsʌut ^h a	fourth
+NUM+ORD+FEM	चौथी	tsʌut ^h i:	fourth

The finite state transducer illustrated in Figure 3.39 is capable of analyzing and generating the ordinal numerals from numbers one, two, three and four and their corresponding forms illustrated in Table 3.57, Table 3.58, Table 3.59, Table 3.60.

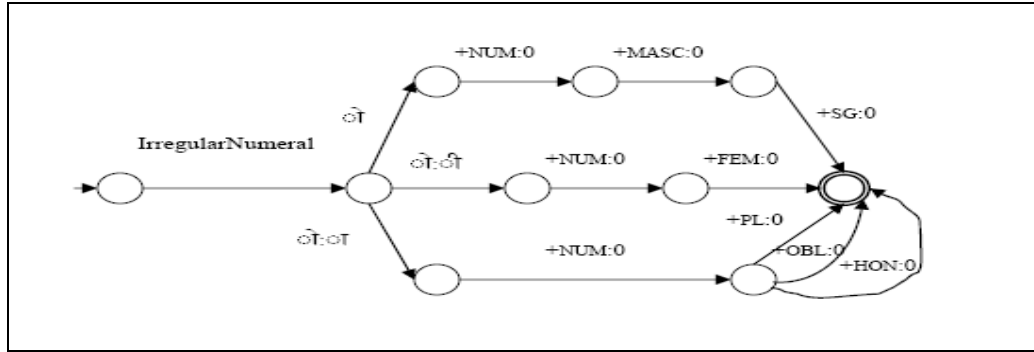


Figure 3.39: A finite state transducer for irregular ordinal numerals

The phonological rules involved in irregular ordinal numerals are given in PR 3.12, which are compiled and composed with finite state transducer illustrated in Figure 3.39.

Phonological rule

PR 3.12

- i. Stem final vowel $\overset{\circ}{o}$ of the o -ending irregular numeral of the lower language (i.e, surface level) is changed to vowel $\overset{\circ}{a}$ for plural, oblique and honorificity

Regular expression: $\overset{\circ}{o} \rightarrow \overset{\circ}{a} \parallel _ \#$.

- ii. Stem final vowel $\overset{\circ}{o}$ of the o -ending irregular numeral of the lower language (i.e, surface level) is changed to vowel $\overset{\circ}{i}$ for feminine gender.

Regular expression: $\overset{\circ}{o} \rightarrow \overset{\circ}{i} \parallel _ \#$.

c. Ordinal numbers loaned from Sanskrit

Some ordinal numbers in Nepali are loan words from Sanskrit. They are listed in Table 3.61.

Table 3.61: Some ordinal numbers from Sanskrit loan

Morphological Tags	Devanagari	IPA	Gloss
+NUM+ORD	प्रथम	prʌtʰʌm	first
+NUM+ORD	द्वितीय	dwitiːjʌ	second
+NUM+ORD	तृतीय	tritiːjʌ	third
+NUM+ORD	चतुर्थ	tsʌturtʰʌ	fourth
+NUM+ORD	पञ्चम	pʌntsʌm	fifth

The ordinal numbers borrowed from Sanskrit are encoded in the finite state transducer as demonstrated in Figure 3.40 and it is capable of analyzing and generating them.

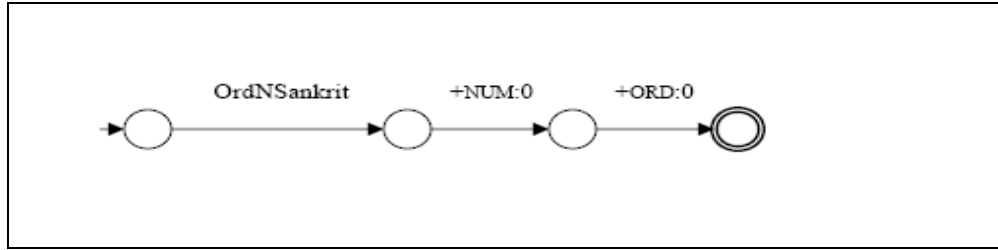


Figure 3.40: A finite state transducer for ordinal numerals from Sanskrit loan

3.5.2 Other numerals

Some numerals in Nepali indicate the frequency and also modify the head nouns. Such numerals grouped into four classes and they are listed in Table 3.62, Table 3.63, Table 3.64 and Table 3.65.

Table 3.62: Frequency numerals (I)

Morphological Tags	Devanagari	IPA	Gloss
+NUM+FREQ	एकोहोरो	ekohoro	one
+NUM+ FREQ	दोहोरो	dohoro	two
+NUM+FREQ	तेहोरो	tefioro	three

Table 3.63: Frequency numerals (II)

Morphological Tags	Devanagari	IPA	Gloss
+NUM+FREQ	एकसरो	eksʌro	one layer
+NUM+ FREQ	दुईसरो	dui:sʌro	two layer
+NUM+FREQ	तीनसरो	ti:nsʌro	three layer

Table 3.64: Frequency numerals (III)

Morphological Tags	Devanagari	IPA	Gloss
+NUM+FREQ	दोबर	dobʌr	twice/double
+NUM+ FREQ	तेबर	tebʌr	thrice
+NUM+FREQ	चौबर	tsʌubʌr	four times

Table 3.65: Frequency numerals (IV)

Morphological Tags	Devanagari	IPA	Gloss
+NUM+FREQ	दुईगुना	dui:guna	two times
+NUM+ FREQ	तीनगुना	ti:nguna	three times
+NUM+FREQ	चौगुना	tsʌuguna	four times

The finite state transducer illustrated in Figure 3.41 is capable of analyzing and generating the frequency numerals illustrated in Table 3.62, Table 3.63, Table 3.64 and Table 3.65.

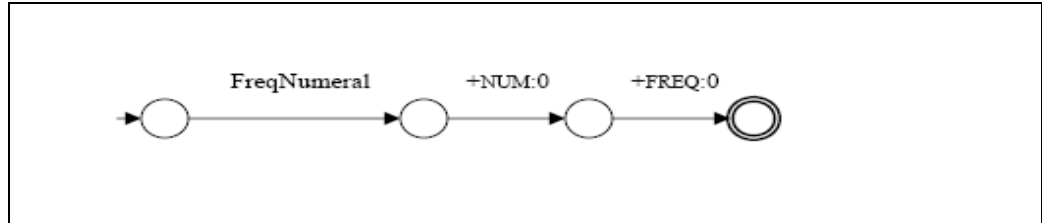


Figure 3.41: A finite state transducer for frequency numerals

There are few numerals which indicate part of the measurement of things, time and space. Some of the portion numerals are listed in Table 3.66.

Table 3.66: Some portion numerals

Morphological Tags	Devanagari	IPA	Gloss
+NUM+PORT	आधा	adha	half
+NUM+PORT	पौने	pʌune	(a number less) a quarter
+NUM+PORT	सवा	sʌwʌ	one and quarter
+NUM+PORT	डेढ	deḍʰʌ	one and half
+NUM+PORT	साढे	sʌḍʰe	(a number and) half
+NUM+PORT	अढाइ	ʌḍʰai	two and half
+NUM+PORT	चौथाइ	tsʌutʰai	one fourth

The finite state transducer illustrated in Figure 3.42 is capable of analyzing and generating the portion numerals illustrated in Table 3.66.

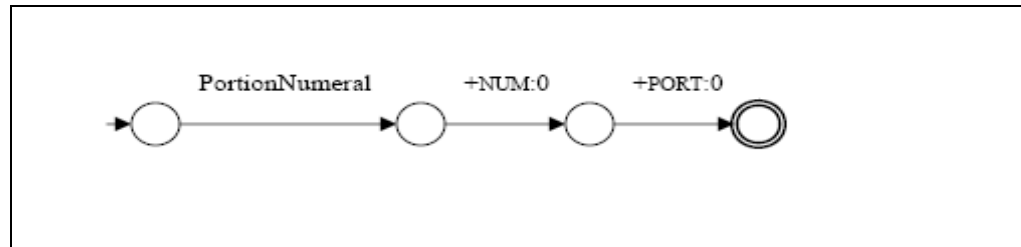


Figure 3.42: A finite state transducer for portion numerals

3.6 Classifiers in Nepali

3.6.1 Numeral classifiers

There are two numeral classifiers in Nepali. *-जना -dzana* is human masculine classifiers and it does not inflect for anything. *-वटा -wata* is a non-human classifier but it inflects for human feminine. The numeral classifiers appear only with countable nouns. Table 3.67 lists these two numeral classifiers and their various forms.

Table 3.67: Numeral classifiers

Morphological Tags	Devanagari	IPA	Gloss
+CLF+HUM	जना	dzana	
+CLF+NHUM	वटा/ओटा	wata/ota	
+CLF+FEM	वटी/ओटी	wati/oti	

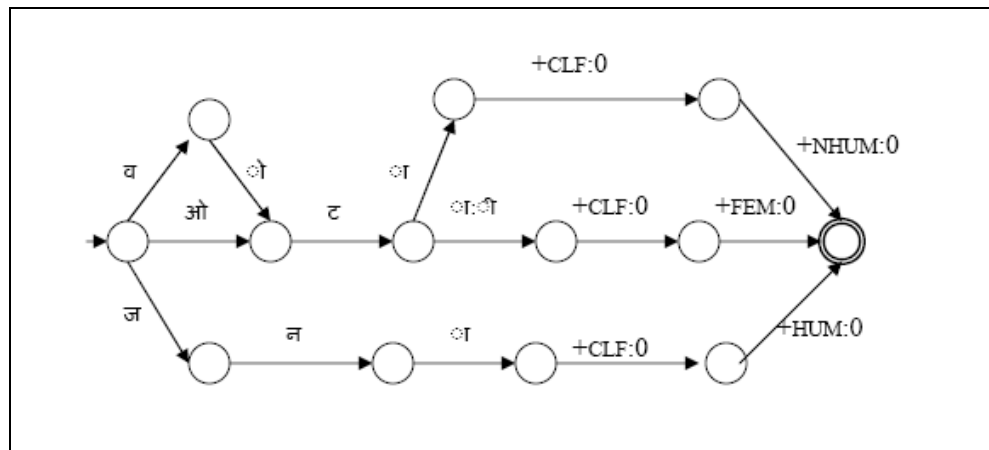


Figure 3.43: A finite state transducer for numeral classifiers

The finite state transducer illustrated in Figure 3.43 is capable of analyzing and generating the numeral classifiers illustrated in Table 3.67.

3.6.2 Quasi-classifiers

Quasi-classifiers in Nepali have their lexical content as well as the properties of being the classifier. Each item in the list classifies a small set of nouns and also follows the numerals.

Quasi-classifiers are related to mensurality or sortality. Such classifiers end in either *o* or non-*o* like nouns and adjectives in Nepali. *o*-ending quasi-classifiers inflect for number and oblique features. Some examples of *o*-ending classifiers are given in Table 3.68.

Table 3.68: *o*-ending quasi-classifiers

Morphological Tags	Classifier1	Classifier2	Classifier3
+CL+SG	कोसो koso	दानो dano	थोपो t ^h opo
+CL+PL	कोसा kosa	दाना dana	थोपा t ^h opa
+CL+OBL	कोसा kosa	दाना dana	थोपा t ^h opa

The *o*-ending quasi-classifiers in Nepali are compiled into a finite state transducer as demonstrated in Figure 3.44 and it is capable of analyzing and generating the quasi-classifiers illustrated in Table 3.68.

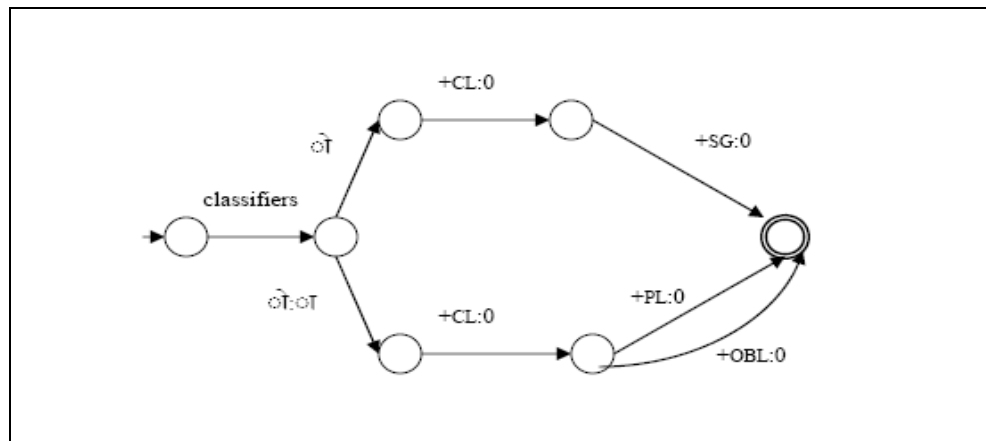


Figure 3.44: A finite state transducer for general classifier type 1

The phonological rules involved in this set of quasi-classifiers are given in PR 3.13, which are compiled and composed with finite state transducer illustrated in Figure 3.44.

Phonological rule

PR 3.13

- i. Stem final vowel ॆ o of the o -ending quasi-classifiers of the lower language (i.e, surface level) is changed to vowel ॆ a for plural and oblique.

Regular expression: $\text{ॆ} \rightarrow \text{ॆ} \mid _ \#$.

The finite state transducer in Figure 3.44 is capable of analyzing and generating the quasi-classifiers illustrated in Table 3.68.

Non- o -ending quasi-classifiers do not inflect for anything. Table 3.69 presents some examples of non- o -ending quasi-classifiers in Nepali.

Table 3.69: General non- o -ending classifiers

Morphological Tags	Devanagari	IPA	
+CL	पोटी	poṭi	
+CL	थुन	t ^h un	
+CL	जुवा	dzuwa	
+CL	गाँस	gās	
+CL	चोइली	tsoili	
+CL	खिल्ली	k ^h illi	
+CL	घरी	g ^h ari	

The finite state transducer in Figure 3.45 is capable of analyzing and generating the quasi-classifiers illustrated in Table 3.69.

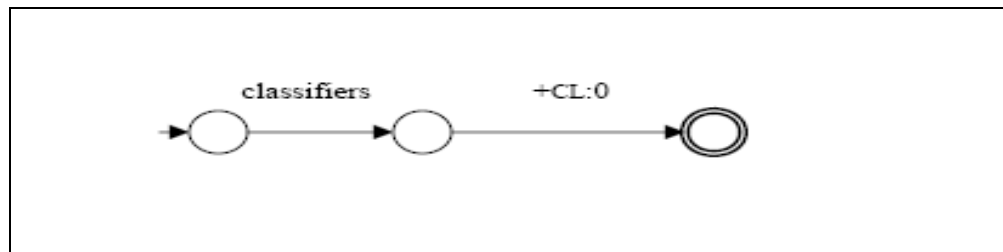


Figure 3.45: A finite state transducer for general classifier type 2

3.7 Summary

This chapter analyzed that nouns in Nepali. They can be grouped into two classes: *o*-ending and non-*o*-ending nouns. The *o*-ending are further sub-grouped into four classes and non-*o*-ending nouns are further sub-grouped into two classes, viz. marked and unmarked classes. Marked non-*o*-ending nouns are of four types and unmarked nouns are of six types. The basis on which the classification is done to match and implement the word categories into finite state technology is made up of the formal characteristic features possessed by the nouns in Nepali. Some of phonological rules for one group of nouns are repeated for another group; they are minimized, delimiters are used if required and implemented as regular expression and finally composed with the main noun lexicon.

Personal pronouns in Nepali possess person, number, form and honorific features. Demonstratives, reflexives, reciprocal, definite and indefinite pronouns inconsistently possess number, form and honorific features. The formal grouping of the pronouns is significant for the illustration and demonstration of their finite state transducers. Since the number of pronouns is limited and their behavior is more or less idiosyncratic, they are directly encoded for creating the finite state network.

Adjectives in Nepali are mainly of two major types: *o*-ending and non-*o*-ending. Non-*o*-ending adjectives are of two types: marked and unmarked. One group of marked adjectives shows the distinction in masculine and feminine gender whereas another group containing Sanskrit loans shows three levels of degree: positive, comparative and superlative. And unmarked adjectives remain unaltered.

The numerals in Nepali are mainly grouped into three classes; they are cardinal, ordinal and other numerals. Except some, all ordinal numerals are derived from the cardinal numerals. Some irregular ordinal numbers show the distinctions for the features like number, gender, honorificity and form.

The classifiers in Nepali are grouped into two classes; true classifiers and quasi-classifiers. The true classifiers inflect for gender whereas some of the quasi-classifiers inflect for number and form.

CHAPTER 4

VERBAL MORPHOLOGY

4.0 Outline

This chapter presents the analysis of verb stems in Nepali. It consists of six sections. Section 4.1 discusses the characteristic features of verbs, namely, significant verb stem finals, transitivity, syllabicity and sound *a*. In section 4.2, we discuss the morphological processes like causativization, passivization and negativization. The stem formation concept is presented in section 4.3. Section 4.4 groups the verbs into various groups based on the features discussed above and presents them with their morphological tags. The finite state transducer of each group is illustrated. Section 4.5 deals with verbal inflections which include tense, aspect and mood. For every group of inflections the morphological tags and finite state transducers are illustrated. Section 4.5 deals with verbal inflections which include tense, aspect, mood and participial forms. For every group of inflections the morphological tags and finite state transducers are illustrated. Section 4.6 summarizes the findings of the chapter.

4.1 Characteristics of verb in Nepali

4.1.1 Significant verb stem finals

The basic verb stems end with different sound segments. Some of the final segments are noteworthy from the morphological point of view. The morphological processes that are under consideration such as passivization, causativization, negativization and other affixation processes need the information of the final segment of the verb to produce the acceptable surface forms. The stem of the basic verb is identified by removing the past tense third person singular marker -यो *-jo* from the verb forms and then the remaining segment is analyzed with reference to various phenomena. Those final segments which are significant from our point of view are discussed as follows:¹

¹ Pokharel (2010a) has mentioned the various strategies to derive the verb stems. Among them imperative singular form as the basic stem has been adopted here for the simplicity, although it leaves some exceptions.

a. Vowel final stems

i. *i*-ending verb stems: A set of verb stems which end in vowel इ *i* are listed in Table 4.1. The majority of the verb stems in this class are intransitive verbs but some of them are transitive also. Some examples are listed in Table 4.2. The verbs उफ्रि *up^hri* 'jump' and पक्रि *pakri* 'arrest' in (1a) and (1b), respectively end with vowel इ *i*.

(1) a. केटो उफ्रियो ।

keṭo up^hri-jo
boy jump-PST.3SG.MASC
'The boy jumped.'

b. प्रहरीले चोरलाई पक्रियो ।

prahāri-le tsor-lai pakri-jo
police-ERG thief-DAT arrest-PST.3SG.MASC
'The police arrested the thief.'

Table 4.1: *i*-ending intransitive verb stems

Verb stem	IPA	Gloss
उफ्रि	up ^h ri-	'jump'
खुम्चि	k ^h umtsi-	'shrink'
चोइँटि	tsoiṭi-	'be pieces'
भत्कि	b ^h atki-	'be broken'

The *i*-ending intransitive verb stems listed in Table 4.1 and *i*-ending transitive stems listed in Table 4.2 look similar in their form. But they differ in their further morphology.

Table 4.2: *i*-ending transitive verb stems

Base form	IPA	Gloss of stem
पक्रि	pakri-	'arrest'
पखिँ	paṛk ^h i-	'wait'
बिसिँ	birsi-	'forget'
मन्सि	maṅsi-	'throw away'
सम्झि	saṃdʒ ^h i-	'remember'
कुल्चि	kultsi-	'tread'
उइँटि	uṭṭi-	'spindle'
दि	di-	'give'
लि	li-	'take'

i-ending verb stems listed in Table 4.2a behave differently. The vowel उ *u* is obligatorily inserted between the stems and suffix if the suffix that follows the stems begins with न् *n* and उँ *ũ* if the suffix begins with छ् *ts^h* and थ् *t^h*.

Table 4.2a: *i*-ending transitive verb stems

Verb stem	IPA	Gloss
पि	pi	'drink'
सि	si	'sew'
जि	dzi	'live'

The vowel इ *i* at the end of the verb stem optionally drops without change in meaning. The verb stem पग्लि *pAgli* 'melt' in (2a) has retained vowel इ *i* and verb stem पगल् *pAgI* 'melt' in (2b) vowel इ *i* is dropped.

- (2) a. हिउँ पग्लियो ।
 hiũ pAgli-jo
 ice melt-PST.3SG.MASC
 'The ice melted.'
- b. हिउँ पग्ल्यो ।
 hiũ pAgI-jo
 ice melt-PST.3SG.MASC
 'The ice melted.'

This vowel इ *i* at the end of the verb stems also is optionally changed to अ *A* especially when the suffix beginning with न् *n*, द् *d*, and ए *e*. For example, when -नु -*nu* '-INF' gets attached to verb stem, इ *i* optionally changes to अ *A*. Table 4.3 lists these alternative forms due to change of इ *i* to अ *A* in *i*-ending verb stems.

Table 4.3: Alternative forms of *i*-ending verb stems

-i forms	IPA	-a forms	IPA
उफ्रि	up ^h ri-	उफ्र	up ^h rΛ-
खुम्चि	k ^h umtsi-	खुम्च	k ^h umtsΛ-
चोइटि	tsoiṭi-	चोइट	tsoiṭΛ-
भत्कि	b ^h Λtki-	भत्क	b ^h ΛtkΛ-
सिउरि	siuri-	सिउर	siurΛ-
बिग्रि	bigri-	बिग्र	bigrΛ-
सप्रि	sΛpri-	सप्र	sΛprΛ-
उघ्रि	ug ^h ri-	उघ्र	ug ^h rΛ-
पग्लि	pΛgli-	पगल	pΛglΛ-
उग्लि	ugli-	उगल	uglΛ-
उक्लि	ukli-	उकल	uklΛ-
पक्रि	pΛkri-	पक्र	pΛkrΛ-
पर्खि	pΛrk ^h i-	पर्ख	pΛrk ^h Λ-
बिर्सि	birsi-	बिर्स	birsΛ-

ii. ***a*-ending verb stems:** Some of the verb stems ending with the vowel *अ* *a* are listed in Table 4.4 and Table 4.5. Verb stems in this group are of both intransitive and transitive types. The verb stem *कमा* *kΛma*- 'earn' in (3a) and *आ* *a*- 'come' in (3b) end with vowel *अ* *a*.

(3) a. उसले धेरै पैसा कमाएको छ ।

us-le d^herΛi pΛisa kΛma-eko ts^hΛ
 3SG.OBL-ERG more money earn-PERF be-PST.3SG.MASC
 'He has earned a lot of money.'

b. राम स्कूलबाट घर आयो ।

ram skul-baṭΛ g^hΛr a-jo
 Ram school-ABL house come-PST.3SG.MASC
 'Ram came home from school.'

Table 4.4: *a*-ending verb stems (group 1)

Verb stem	IPA	Gloss
अघा	ʌg ^h a-	'satisfy'
कमा	kʌma-	'earn'
टकरा	tʌkʌra-	'be broken'
मुस्कुरा	muskura-	'insert'
पा	pa-	'get'
आ	a-	'come'
छा	ts ^h a-	'cover the roof'
बा	ba-	'open (mouth)'
पा	pa-	'get'
ला	la-	'put on'
भ्या	b ^h ja-	'manage'
ब्या	bja-	'give birth'

Table 4.5: *a*-ending verb stems (group 2)

Verb stem	IPA	Gloss
खा	k ^h a-	'eat'
जा	dza-	'go'

The *a*-ending verb stems are also of two kinds, a set of verbs in which vowel उ *u* is inserted between stem and suffix if the following suffix begins with न् *n*, and ऊँ *ũ* with छ् *ts^h*, and थ् *t^h* as in Table 4.4. Those verb stems as listed in Table 4.5 do not take उ *u* in the condition as stated above. In this group न् *n* is inserted in the non-past tense and past habitual aspect.

ii. *o*-ending verbs stems: There are a few verb stems which end with ओ *o*. The stem final ओ *o* obligatorily changes to उ *u* if the following suffix begins with छ् *ts^h*, द् *d*, थ् *t^h* then न् *n* sound segments and न् *n* is obligatorily inserted in non-past tense. Table 4.6 lists some of the *o*-ending verb stems and Table 4.7a shows the change of ओ *o* to उ *u* in the condition mentioned above.

Table 4.6: o-ending verb stems

Verb stem	IPA	Gloss
रो	ro-	'weep'
धो	d ^h o-	'wash'
छो	ts ^h o-	'touch'

Table 4.7a: Change of o to u in o-ending verb stems

Verb stem	IPA	Gloss
रुनु	ru-nu	'to weep'
धुनु	d ^h u-nu	'to wash'
छुनु	ts ^h u-nu	'to touch'

iii. **Λ-ending verbs stems:** There is a small set of verbs which end with the vowel अΛ. The vowel अΛ in the end of the vowel stem drops if the following suffixes beginning with ए e, इ i, उ u and ओ o are attached. Table 4.7a lists some Λ-ending verb stems and Table 4.7b shows some dropping of vowel Λ.

Table 4.7b: Λ-ending verb stems

Verb stem	IPA	Gloss
सह	sΛfiΛ-	'tolerate'
रह	rΛfiΛ-	'remain'

Table 4.7c: Λ-ending verb stems (Λ-dropped)

Verb stem	IPA	Gloss
सहेर	sΛfi-erΛ	'tolerate-CONJUNCT'
रहेर	rΛfi-erΛ	'remain-CONJUNCT'

In the vowel ending verb stems, except verbs in Table 4.2a and Table 4.4, semantically null element ँ n is inserted between stem and suffix if the suffix begins with छ् ts^h or थ् t^h sounds. But, in the case of the verb stems in Table 4.2a and Table 4.4, only ँँ dΛ is inserted after उ u is inserted for some other purpose.

b. Consonant final stems

i. Voiceless consonant ending stems: The verb stems that end with voiceless consonants are both intransitive and transitive types. Some examples of the verb stems ending with voiceless consonants are listed in Table 4.7d.

Table 4.7d: Verb stems ending with a voiceless consonant

Verb stem	IPA	Gloss
कस्	k _{AS} -	'tighten'
काँप्	kāp-	'tremble'
घसेट्	g ^h _{ASeṭ} -	'drag'
जाक्	dzak-	'insert'
फ्याँक्	p ^h jāk-	'throw'
नाच्	nats-	'dance'

In this group of verb stems, semantically null elements *त* *t*_A or *द* *d*_A are inserted optionally between the stem and suffix if the suffix begins with *छ* *ts^h* and *थ* *t^h*. These forms are used only in non-past tense and past habitual aspect. These alternative forms of the stems are listed in Table 4.8.

Table 4.8: Alternative forms from stems ending with voiceless consonant

Base stem	form1	form2
कस् k _{AS}	कस्त k _{AS} t _A -	कसद k _{AS} d _A -
काँप् kāp	काँप्त kāp _t _A -	काँपद kāp _d _A -
घसेट् g ^h _{ASeṭ}	घसेट्त g ^h _{ASeṭ} t _A -	घसेट्द g ^h _{ASeṭ} d _A -
जाक् dzak	जाक्त dzak _t _A -	जाक्द dzak _d _A -
फ्याँक् p ^h jāk	फ्याँक्त p ^h jāk _t _A -	फ्याँक्द p ^h jāk _d _A -
नाच् nats	नाक्त nat _t _A -	नाक्द nats _d _A -

ii. Voiced consonant ending stems: The verb stems that end with voiced consonants are of both types intransitive and transitive. Some examples of the verb stems ending with voiced consonants are listed in Table 4.9.

Table 4.9: Verb stems ending with voiced consonant

Verb stem	IPA	Gloss
बोल्	bol-	'speak'
पिँद्	pĩd-	'grind'
थुन्	t ^h un-	'close'
पछार्	pʌts ^h ar-	'throw down'
डुब्	ḍub-	'sink'
छाम्	ts ^h am-	'feel'
खोज्	k ^h odz-	'search'

In this group of stems also, a semantically null element δ is inserted optionally between the stem and suffix if the suffix begins with $\text{छ} ts^h$ or $\text{थ} t^h$. These forms are used only in non-past tense and past habitual aspect. These alternative forms of the stems are listed in Table 4.10.

Table 4.10: Alternative forms from stems ending with voiced consonant

Base stem	Alternative form
बोल् bol-	बोल्द bold Δ -
पिँद् pĩd-	पिँद्द पिँद्द Δ -
थुन् t ^h un-	थुन्द t ^h und Δ -
पछार् pʌts ^h ar-	पछार्द pʌts ^h ard Δ -
डुब् ḍub-	डुब्द ḍubd Δ -
छाम् ts ^h am-	छाम्द ts ^h amd Δ -
खोज् k ^h odz-	खोज्द k ^h odz Δ -

4.1.2 Transitivity

Transitivity is the number of argument that a verb takes (Katamba 1993:256-62; Pyane 1997:171). The transitivity is significant in verbs. Morphology of the verbs can be further analyzed in term of this feature.

a. Intransitive verbs

Those verbs which take only one argument as subject noun phrase are intransitive verbs. In example (4) the verb उठ् $u\text{t}^h$ 'get up' has taken only one argument उ u 'he' as

a subject and in example (5) the verb बस् *bAS* 'sit' has taken only one argument उ *u* 'he' as subject, therefore, they are intransitive verbs.

(4) उ बिहानै उठ्यो ।

u bihan-ai ut^h-jo
 he morning-EMP rise-PST.3SG.MASC
 'He got up early in the morning.'

(5) उ सँधै घरमा बस्छ ।

u sãdhai g^har-ma bAS-ts^hA
 3SG always home-LOC sit-NPST.3SG.MASC
 'He always stays at home.'

The other verbs listed in Table 4.11 such as कुद् *kud* 'run', बस् *bAS* 'sit', सुत् *sut* 'sleep', etc. also take only one argument as the subject.

Table 4.11: Intransitive verbs

Intransitive verb	IPA	Gloss
उठ्	ut ^h -	'wake up'
कुद्	kud-	'run'
बस्	bAS-	'sit'
लड्	lAd-	'fall down'
सुत्	sut-	'sleep'
अघा	Agh ^a -	'satisfied'

b. Transitive/ditransitive verbs

Those verbs which take two arguments are said to be transitive and those verbs which take three arguments are said to be ditransitive verbs. Both types of verbs are kept here under the same group as they behave in the same way at the morphological level. The Tables 4.12 and 4.13 list the transitive verbs and ditransitive verbs, respectively. The verb काट् *kat* 'cut' in (6) has taken two arguments श्याम *sjam* 'Shyam' and रुख *ruk^h* 'tree' as subject and object of the sentence, respectively. And the verb दि *di* 'give' in (7) has taken three arguments म *mai* '1SG', उस् *us* 'he.OBL' and किताब *kitab* 'book' as subject, indirect and direct object of the sentence, respectively.

(6) श्यामले रुख काट्यो ।
 sjam-le ruk^h kaṭ-jo
 Shyam-ERG tree cut-PST.3SG.MASC
 'Shyam cut the tree.'

(7) मैले उसलाई किताब दिएँ ।
 maɪ-le us-lai kitab di-ẽ
 1SG.OBL-ERG 3SG.OBL-DAT book give-PST.1SG.MASC
 'I gave him a book.'

Some transitive verbs are listed in Table 4.12 which take only two arguments as subject and object and some ditransitive verbs as listed in Table 4.13 take three arguments as subject, indirect and direct objects.

Table 4.12 Some transitive verbs

Transitive verb	IPA	Gloss
काट्	kaṭ-	'cut'
खा	k ^h a-	'eat'
चुस्	tsu-	'suck'
पढ्	pʌḍ ^h -	'read'
टोक्	ṭok-	'bite'

Table 4.13: Some ditransitive verbs

Ditransitive verb	IPA	Gloss
तिर्	tir-	'pay'
बेच्	bets-	'sell'
दि	di-	'give'
लेख्	lek ^h -	'write'
सोध्	sod ^h -	'ask'

4.1.3 Syllabicity

Nepali verb stems can be grouped into two classes based on the number of syllables in a stem. This feature is significant especially in the causative stem formation.

a. Monosyllabic verb stems

Those verb stems which have only one syllable are said to be monosyllabic verb stems. Some examples are listed in Table 4.14.

Table 4.14: Monosyllabic verb stems

Verb stem	IPA	Gloss
बोल्	bol-	'speak'
खा	k ^h a-	'eat'
पिँद्	pĩd-	'grind'
थुन्	t ^h un-	'close'
कस्	kʌs-	'tighten'
डुब्	ɖub-	'sink'
छाम्	ts ^h am-	'feel'
खोज्	k ^h odz-	'search'
खोल्	k ^h ol-	'open'
सुक्	suk-	'be dried'
खा	k ^h a-	'eat'
जा	dza-	'go'
दि	di-	'give'
धो	d ^h o-	'wash'
रो	ro-	'weep'
सि	si-	'sew'
पि	pi-	'drink'

b. Polysyllabic verb stems

Those verb stems which are formed from two or more syllables are said to be polysyllabic verb stems. Some examples are illustrated in Table 4.15.

Table 4.15: Polysyllabic verb stems

Verb stem	IPA	Gloss
उफ्रि	up ^h ri-	'jump'
खुम्चि	k ^h umtsi-	'shrink'
भत्कि	b ^h ʌtki-	'be broken'
पछार	pʌts ^h ar-	'throw down'
घसेट्	g ^h ʌset-	'drag'
मुस्कुरा	muskura-	'insert'
निचोर्	nitsor-	'squeeze'
निमोठ्	nimot ^h -	'twist'
चिथोर्	tsit ^h or-	'scratch'
छिमल्	ts ^h imʌl-	'prune'

4.1.4 Sound अ a

The sound अ a appears in Nepali verb stems in two manifestations, one as a normal vowel phoneme अ a ; and another as a causative marker -अ -a while forming the causative verb stems. The presence and absence of अ a sound in the base verb stem is very significant for forming the causative stems. Therefore, the basic verb stems can be grouped into two classes, i.e., stems with अ a sound and stems without अ a sound. Some examples of former group are listed in Table 4.16 and of latter group are listed in Table 4.17.

Table 4.16 Verb stems with a sound

Verb stem	IPA	Gloss
खाँद्	k ^h ãd-	'press down'
गाल्	gal-	'melt'
छान्	ts ^h an-	'choose'
पछार्	pʌts ^h ar-	'throw down'
कोच्ार्	kotsar-	'insert into'
डकार्	ɖʌkar-	'bulch'

Table 4.17: Verb stems without a sound

Verb stem	IPA	Gloss
तिर्	tir-	'pay'
बल्	bʌl-	'burn'
खोप्	k ^h op-	'cut deep'
घट्	g ^h ʌt-	'be less'
चिमट्	tsimʌt-	'pinch'
छिमल्	ts ^h imʌl-	'prune'

4.2 Morphological processes

4.2.1 Causativization/transitivization

In transitivization, an argument is added irrespective of the role of the argument but in causativization, the added argument is definitely the causer. The morphological change in the verb stem and syntactic make up are the same in both the processes,

however, the interpretation may differ semantically (Katamba 1993:274-5; Pokharel 2054VS:6-16). But, in this study, both are treated as a single process. In sentence (8a), the verb सुत्यो 'slept-PST.3SG' is non-causative which has taken बच्चो *batstso* 'child' as subject of the sentence. When it is causativized as सुताइन् *sut-a-in* 'sleep-CAUS-PST.3SG.FEM.HON' in (8b), it has taken a new subject आमा *ama* 'mother' as a causer and the subject of the non-causative construction is demoted to the object of causativized verb. So, in the process of causativization, a morphological causative marker is suffixed to the verb stem and is followed by the agreement markers. Table 4.18 lists some examples of such causative verb stems.

(8) a. बच्चो सुत्यो ।

batstso sut-jo
 child.SG.MASC sleep-PST.3SG.MASC
 'The child slept.'

b. आमाले बच्चालाई सुताइन् ।

ama-le batstsa-lai sut-a-in
 mother-ERG child-DAT sleep-CAUS-PST.3SG.FEM.HON
 'The mother made the child sleep.'

Table 4.18 Causative verb stems

Casulative verb	IPA	Gloss
उठा	u ^h -a	'cause to wake up'
सुता	sut-a	'cause to sleep'
तिरा	tir-a	'cause to pay'
लेखा	le ^h -a	'cause to write'
भना	b ^h ʌn-a	'cause to say'

Some ways of causative formation

a. by -आ -a suffix

The causativization by a causative marker -आ -a is the most regular and the bulk of the non-causative stems become causative stem by this process. The verb stems listed in Table 4.18 are formed by this method.²

b. by both -आ -a and आल् -al suffixes

A small set of verb stems which, instead of taking marker -आ -a, also take marker आल् -al to form the causative stems. For example, verb stem खस् $k^h\Lambda S$ 'drop' in (9a), gets causativized by marker -आ -a in (9b) and by आल् -al in (9c). Table 4.19 lists some examples of this type of causative stem formation.

- (9) a. दुङ्गा खस्यो ।
 $\text{d}^h\text{u}\text{ŋ}\text{ga}$ $k^h\Lambda S\text{-jo}$
 stone drop-PST.3SG.MASC
 'The stone dropped.'
- b. केटाले दुङ्गा खसायो ।
 keta-le $\text{d}^h\text{u}\text{ŋ}\text{ga}$ $k^h\Lambda S\text{-a-j}\text{o}$
 boy-ERG stone drop-CAUSE-PST.3SG.MASC
 'The boy dropped a stone.'
- c. केटाले दुङ्गा खसाल्यो ।
 keta-le $\text{d}^h\text{u}\text{ŋ}\text{ga}$ $k^h\Lambda S\text{-al-j}\text{o}$
 boy-ERG stone drop-CAUSE-PST.3SG.MASC
 'The boy dropped a stone.'

Table 4.19: Verb stems forming causatives with -आ -a and आल् -al

Base	Gloss	Causative	Gloss
बस् $b\Lambda S\text{-}$	sit	बसा/बसाल् $b\Lambda Sa\text{-}/b\Lambda Sa\text{-}$	cause to sit
खस् $k^h\Lambda S\text{-}$	drop	खसा/खसाल् $k^h\Lambda Sa\text{-}/k^h\Lambda Sa\text{-}$	cause to drop
चुँड् $ts\ddot{u}\text{d}\text{-}$	snatch	चुँडा/चडाल् $ts\ddot{u}\text{d}\text{-}/ts\ddot{u}\text{d}\text{-}$	cause to snatch
छिन् $ts^h\text{in}\text{-}$	chop off	छिना/छिनाल् $ts^h\text{ina}\text{-}/ts^h\text{ina}\text{-}$	cause to chop off

² Most of the Nepali grammarians believe that the basic causative marker is -आउ -au. But in this study, -आ -a is assumed to be the basic causative marker simply for computing purpose.

c. by अ₁ → आ_a

A small set of monosyllabic verb stems having the vowel अ₁ in between consonants (i.e. C₁C structure) form the causative stem by changing the vowel अ₁ to आ_a. The verb stem मर_{mar} 'die' in (10a) is causativized as मार_{mar} 'kill' in (10b). Some of the verb stems in which causative stems are formed by this way are listed in Table 4.20.

(10) a. मृग मर्यो ।

mriga mar-jo
deer die-PST.3SG.MASC
'The deer died.'

b. बाघले मृग मार्यो ।

bag^h-le mriga mar-jo
tiger-ERG deer die-CAUSE-PST.3SG.MASC
'The tiger killed the deer.'

Table 4.20: Verb stems forming causatives by changing अ₁ to आ_a

Base verb	Gloss	Causative	Gloss
मर _{mar} -	die	मार _{mar} -	kill
सर _{sar} -	shift	सार _{sar} -	cause to shift
चल् _{tsal} -	move	चाल् _{tsal} -	cause to move
टर् _{tar} -	pass over	टार _{tar} -	cause to pass over
पर _{par} -	fall	पार _{par} -	cause to fall
गल् _{gal} -	melt	गाल् _{gal} -	cause to melt
बल् _{bal} -	burn	बाल् _{bal} -	cause to burn

d. by उ_u → ओ_o

Another set of monosyllabic verb stems having vowel उ_u in between the consonants (i.e. C_uC structure) forms the causative stem by changing the vowel *u* to *o*. The verb stem खुल्_{k^hul} 'open' in (11a) is causativized as खोल्_{k^hol} 'open.CAUSE' in (11b). Some of the verb stems in which causative stems are formed by this way are listed in Table 4.21a.

(11) a. ढोका खुल्यो ।

ḍʰoka kʰul-jo
 door open-PST.3SG.MASC
 'The door opened.'

b. पालेले ढोका खोल्यो ।

pale-le ḍʰoka kʰol-jo
 gate-keeper-ERG door open.CAUSE-PST.3SG.MASC
 'The gate keeper opened the door.'

Table 4.21a: Verb stems forming causatives by changing उ to ओ

Base verb	Gloss	Causative	Gloss
छुट् tsʰuṭ-	be left behind	छोड् tsʰoḍ-	cause to be left behind
खुल् kʰul-	open	खोल् kʰol-	cause to open
फुट् pʰuṭ-	break	फोड् pʰoḍ-	cause to break
घुल् gʰul-	dissolve	घोल् gʰol-	cause to dissolve

Interestingly, both the verb stems listed in Table 4.21a can also be causativized with causative marker -आ -a like the verb stems as listed in Table 4.18. The causative verb stems of this set are listed in Table 4.21b.³

Table 4.21b: Verb stems forming causatives by suffixing -आ -a

Base verb	Gloss	Causative	Gloss
छुट् tsʰuṭ-	be left behind	छुटा tsʰuṭ-a	cause to be left behind
खुल् kʰul-	open	खुला kʰul-a	cause to open
फुट् pʰuṭ-	break	फुटा pʰuṭ-a	cause to break
घुल् gʰul-	dissolve	घुला gʰul-a	cause to dissolve
छोड् tsʰoḍ-	be left behind	छोडा tsʰoḍ-a	cause to be left behind
खोल् kʰol-	open	खोला kʰol-a	cause to open
फोड् pʰoḍ-	break	फोडा pʰoḍ-a	cause to break
घोल् gʰol-	dissolve	घोला gʰol-a	cause to dissolve

³ In Table 4.21a, the change of ट् to ड् has not been discussed here (see Pokharel 2054VS). The causativizations shown in Table 4.21a and Table 4.21b have slightly different semantics.

e. by आ a insertion

A subset of polysyllabic *i*-ending verb stems containing consonant cluster form the causative stem by inserting the vowel आ *a* in between the consonants in the cluster.

The verb stem पग्लि *pʌgli* 'melt' in (12a) is causativized as पगाल् *pʌgal* 'melt.CAUSE' in (12b). Some examples of verb stems in this process are listed in Table 4.22.

(12) a. हिउँ पग्लियो

fiũ pʌgli-jo
ice melt-PST.3SG.MASC
'The ice melted.'

b. घामले हिउँ पगाल्यो ।

g^ham-le fiũ pʌgal-jo
sun-ERG ice melt.CAUSE-PST.3SG.MASC
'The sun melted the ice.'

Table 4.22: Verb stems form causatives by inserting आ a

Base verb	Gloss	Causative	Gloss
उफ्रि up ^h ri-	jump	उफार् up ^h ar-	cause to jump
बिग्रि bigri-	spoil	बिगार् bigar-	cause to spoil
सप्रि sʌpri-	flourish	सपार् sʌpar-	cause to flourish
उघ्रि ug ^h ri-	open	उघार् ug ^h ar-	cause to open
पग्लि pʌgli-	melt	पगाल् pʌgal-	cause to melt
उक्लि ukli-	climb up	उकाल् ukal-	cause to climb up

Now, it has been clear that the causative stem formation from base verb stems depends on various features available in the verb stems such as syllabicity, presence and absence of आ *a* sound in the stem, transitivity and stem final segments.⁴

4.2.2 Passivization

Passivization is an opposite phenomenon to causativization in terms of syntax. When passivization takes place, the subject noun phrase is either demoted to postpositional phrase or dropped (Katamba 1993:268-9; Pokharel 2054VS:1-5) In Nepali,

⁴ See Adhikari (2062VS) and Pokharel (2054VS) for detail information.

passivization from intransitive verbs is also possible, but it is restricted only to default agreement (i.e, third person singular), and to some other morphology and interpretation as well (Pokharel 2054VS; Adhikari 2062VS). But the passivization from transitive/causative verbs undergoes for full morphological paradigm and in its interpretations. However, in both cases, the passive marker is the same, i.e., ई *-i* that follows the non-passive stem. The verb as सुत् *sut* 'sleep' in (13a) is intransitive and सुति *sut-i* 'sleep-PASS' is the passive form in (13b). The verb लेख् *lek^h* 'write' in (13c) is a transitive verb and लेखि *lek^h-i* 'write-PASS' in (13d) is the passive form, लेखा *lek^h-a* 'write-CAUSE' in (13e) is causative stem and लेखाइ *lek^h-a-i* 'write-CAUSE-PASS' in (13f) is the causative-passive stem. Therefore, the passive stem of a verb is at least theoretically possible to be derived from intransitive, transitive and causative verb stems. Table 4.23 lists some passive forms of the verbs.

(13) a. म आज राम्ररी सुतैं ।

mʌ adʌ ramrʌri sut-ẽ
 1SG today nice sleep-PST.1SG
 'I slept nicely today.'

b. आज राम्ररी सुतियो ।

adʌ ramri sut-i-jo
 today nice sleep-PASS-PST.3SG.MASC
 '(Myself) slept nicely today.'

c. उसले एउटा चिठी लेख्यो ।

us-le eʊṭa tsit^ht^hi: lek^h-jo
 3SG-ERG one.CLF letter write-PST.3SG.MASC
 'He wrote a letter.'

d. उसबाट एउटा चिठी लेखियो ।

us-baṭʌ eʊṭa tsit^ht^hi: lek^h-i-jo
 3SG-ABL one.CLF letter write-PASS-PST.3SG.MASC
 'A letter was written by him.'

e. उसले एउटा चिठी लेखायो ।

us-le euṭa tsit^hhi: lek^h-a-jo

3SG-ERG one.CLF letter write-CAUS-PST.3SG.MASC

'He caused to write a letter.'

f. उसबाट एउटा चिठी लेखाइयो ।

us-baṭa euṭa tsit^hhi: lek^h-a-i-jo

3SG-ABL one.CLF letter write-CAUS-PASS-PST.3SG.MASC

'He was made to write a letter.'

Table 4.23: Some passive verb stems

Passive verb	IPA	Gloss
उठि	uṭ ^h -i-	'be waken up'
सुति	sut-i-	'be slept'
तिराइ	tir-a-i-	'cause to be paid'
लेखाइ	lek ^h -a-i-	'cause to be written'
अघाइ	ʌg ^h a-i-	'be satisfied'

4.2.3 Negativization

Negativization in Nepali is primarily an affixation process which includes both prefixation and suffixation. Basically the negative marker is ऋ *na* 'NEG' is used in both cases; it is consistent in its form in prefixation process whereas it gets slightly modified in suffixation due to morphophonemic changes (Pokharel 2054 VS:40-6).

a. Prefixation

The negativization by prefixation takes place in moods: potential, optative and imperative, aspects: perfect and imperfect and participial forms: absolute, conjunctive, infinitive, purposive, perfective, prospective and conditional as shown in Table 4.24, the negative by prefixation in a verb खा *k^ha-* 'eat'.

Table 4.24: Negation by the prefixation of negative marker न न-

Grammatical categories	Positive	Negative
Potential	खाला k ^h ala	नखाला nΛ-k ^h ala
Optative	खाएस् k ^h aes	नखाएस् nΛ-k ^h aes
Imperative	खा k ^h a	नखा nΛ-k ^h a
Perfect Aspect	खाएको k ^h a-eko	नखाएको nΛ-k ^h a-eko
Imperfect Aspect	खाँदै k ^h ã-dΛi	नखाँदै nΛ-k ^h ã-dΛi
Absolutive	खाई k ^h a-i:	नखाई nΛ-k ^h a-i:
Conjunctive Participle	खाएर k ^h a-erΛ	नखाएर nΛ-k ^h a-erΛ
Infinitive	खानु k ^h a-nu	नखानु nΛ-k ^h a-nu
Purposive	खान k ^h a-nΛ	नखान nΛ-k ^h a-nΛ
Conditional	खाए k ^h a-e	नखाए nΛ-k ^h a-e
Perfective	खाए k ^h a-e	नखाए nΛ-k ^h a-e
Prospective	खाने k ^h a-ne	नखाने nΛ-k ^h a-ne

b. Suffixation

The negativization by suffixation takes place in tense: past and non-past and aspects: past habitual and inferential as shown in Table 4.25 in a verb खा k^ha- 'eat'. The negative marker न न- 'NEG' always follow the tense marker and precedes the agreement markers.⁵

Table 4.25: Negation by the suffixation of negative marker -न न-

Grammatical categories	Positive	Negative
Non-Past Tense	खान्छ k ^h a-nts ^h Λ	खाँदै न k ^h ã-dΛinΛ
Past Tense	खायो k ^h a-jo	खाए न k ^h a-enΛ
Past Habitual Aspect	खान्थ्यो k ^h a-nt ^h jo	खाँदै न थ्यो k ^h ã-dΛinΛ-t ^h jo
Inferential Aspect	खाएछ k ^h a-ets ^h Λ	खाए न छ k ^h a-e-nΛ-ts ^h Λ

4.3 Stem formation

As discussed in (4.2.1) the causativization is very productive in Nepali verbs at morphological level. The causative stems are formed from both intransitive and transitive verb stems. Thus, from a causativization process, the stems can be divided

⁵ In non-past tense and past habitual aspect, negative marker is preceded by दै dΛi, it's status is yet to be discovered.

into two types of stems: base verb stems and causative stems. However, there are some verb stems from which the causative verb stems can not be formed due to either phonological or semantic constraints.

The passivization as discussed in (4.2.2) is a very productive phenomenon in Nepali morphology. That means, almost all the verb stems either intransitive or transitive verb stems can be passivized. Above all, the causative stems formed from the non-causative stems can still be passivized. This means, causative-passive stems have also been possible. Therefore, it can be generalized that a verb can have at least four different forms as shown in Table 4.26.

Table 4.26: Pattern of the stem formation

Category	Form	Example 'write'
Basic verb stem	V	लेख् lek ^h
Passive verb stem	V- <i>i</i>	लेखि lek ^h -i
Causative verb stem	V- <i>a</i>	लेखा lek ^h -a
Causative Passive verb stem	V- <i>a-i</i>	लेखाइ lek ^h -a-i

4.4 Grouping of verb stems in Nepali

Characteristic features of Nepali verbs discussed in (4.1) and (4.2) are taken as the bases for grouping of Nepali verbs into various classes, so that the syntax of morphemes can be described and implemented to create the finite state network. At the same time, classification of verb stems also helps in branching of sub-lexicons to their respective inflectional paradigms. The phonological rules that are identified can also be systematically implemented.

4.4.1 Intransitive verb stems

a. Verb stem Type1a

a-ending polysyllabic verbs in Nepali which have only two forms: base stem and passive stem. Some such verb stems with both the forms are listed in Table 4.27 with their corresponding morphological tags and gloss of base stems.

Table 4.27: Type1a verb stems

Base form	Tags	Passive form	Tags	Gloss of base
अघा ᱠg ^h a	+VERB	अघाइ ᱠg ^h a-i	+VERB+PASS	'to be satisfied'
करा kᱠra	+VERB	कराइ kᱠra-i	+VERB+PASS	'to shout'
निदा nida	+VERB	निदाइ nida-i	+VERB+PASS	'to sleep'
बहुला bᱠhula	+VERB	बहुलाइ bᱠhula-i	+VERB+PASS	'to be mad'
मुस्कुरा muskura	+VERB	मुस्कुराइ muskura-i	+VERB+PASS	'to smile'
लजा lᱠdza	+VERB	लजाइ lᱠdza-i	+VERB+PASS	'to shy'
टुसा ᱦusa	+VERB	टुसाइ ᱦusa-i	+VERB+PASS	'to sprout'

The finite state transducer illustrated in Figure 4.1 encodes the verb stems listed in Table 4.27 and it is capable of analyzing and generating verb stems listed in Table 4.27.

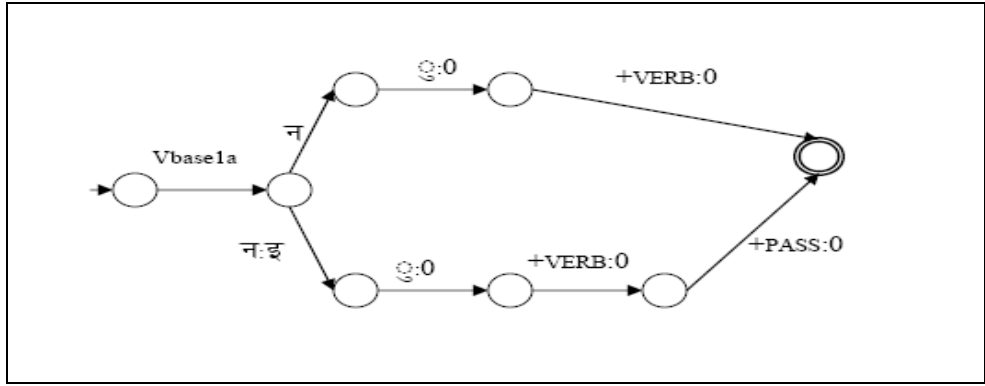


Figure 4.1: A finite state transducer for Type1a verb stems

b. Verb stem Type1b

i-ending polysyllabic verbs in Nepali which have four forms: base stems, passive stems, causative stems and causative-passive forms. Some such verbs with all the forms are listed in Table 4.28 with their corresponding morphological tags and gloss of base stems.

Table 4.28: Type1b verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
चोखि tsok ^h i	चोखिइ tsok ^h i-i	चोख्या tsok ^h j-a	चोख्याइ tsok ^h j-a-i	'to be pure'
गुम्सि gumsi	गुम्सिइ gumsi-i	गुम्स्या gumsj-a	गुम्स्याइ gumsj-a-i	'to be suffocated'
घोष्टि g ^h opti	घोष्टिइ g ^h opti-i	घोष्ट्या g ^h optj-a	घोष्ट्याइ g ^h optj-a-i	'to be overturned'
टुक्रि tukri	टुक्रिइ tukri-i	टुक्र्या tukrj-a	टुक्र्याइ tukrj-a-i	'to be broken into pieces'

The finite state transducer illustrated in Figure 4.2 encodes the verb stems listed in Table 4.28 and it is capable of analyzing and generating them.

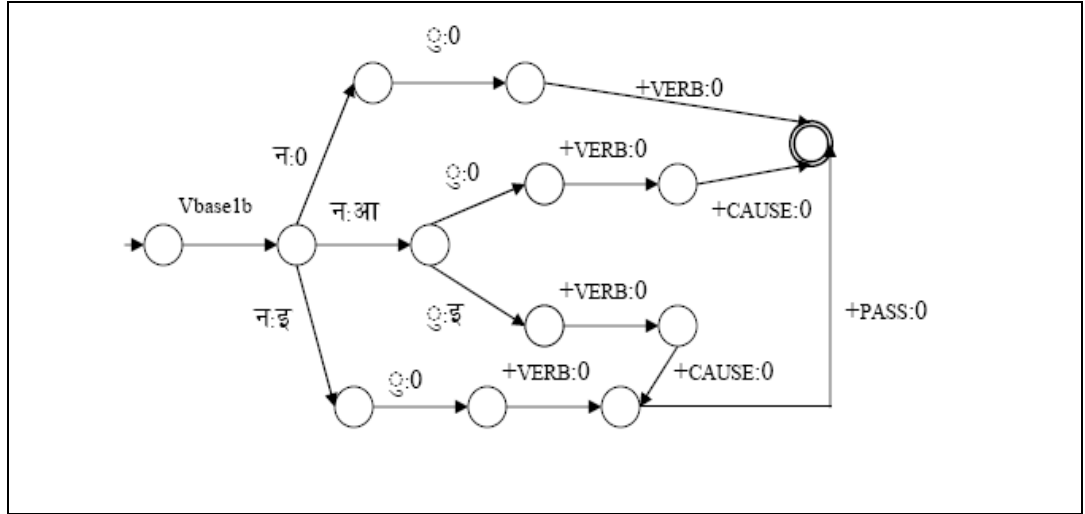


Figure 4.2: A finite state transducer for Type1b verb stems

The rules listed in PR 4.1 are compiled into a finite state transducer and composed with the finite state transducer illustrated in Figure 4.2.

Phonological rule

PR 4.1

- i. Stem final vowel ि *i* of the *i*-ending intransitive verbs at the surface level is changed to vowel य *j* before the causative marker आ *a*.

Regular expressions: ि ->्य || __ आ

- ii. Independent vowel आ *a* changes to its corresponding dependent vowel ा *a* after य *j*.

Regular expression: आ -> ा || य __ ;

c. Verb stem Type1c

i-ending polysyllabic verbs in Nepali which have four forms: base stems, passive stems, causative stem and causative-passive stems. In this group of verbs, causative marker *-a* is inserted between the consonants in consonant cluster while forming the causative stems and final vowel इ *i* is dropped. Some examples are listed in Table 4.29 with their corresponding morphological tags.

Table 4.29: Type1c verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
उक्लि ukli	उक्लिइ ukli-i	उकाल् ukal	उकालि ukal-i	'step up'
उघ्नि ug ^h ri	उघ्निइ ug ^h ri-i	उघार् ug ^h ar	उघारि ug ^h ar-i	'be opened'
उफ्रि up ^h ri	उफ्रिइ up ^h ri-i	उफार् up ^h ar	उफारि up ^h ar-i	'jump'
घस्रि g ^h ʌsri	घस्रिइ g ^h ʌsri-i	घसार् g ^h ʌsar	घसारि g ^h ʌsar-i	'scrawl'
थुप्रि t ^h upri	थुप्रिइ t ^h upri-i	थुपार् t ^h upar	थुपारि t ^h upar-i	'be piled up'
निखि nik ^h ri	निखिइ nik ^h ri-i	निखार् nik ^h ar	निखारि nik ^h ar-i	'be empty'
पग्लि pʌgli	पग्लिइ pʌgli-i	पगाल् pʌgal	पगालि pʌgal-i	'melt'
सप्रि sʌpri	सप्रिइ sʌpri-i	सपार् sʌpar	सपारि sʌpar-i	'grow well'
सुध्रि sud ^h ri	सुध्रिइ sud ^h ri-i	सुधार् sud ^h ar	सुधारि sud ^h ar-i	'improve'

The verb stems listed in Table 4.29 are compiled into a finite state transducer as illustrated in Figure 4.3 and it is capable of analyzing and generating them.

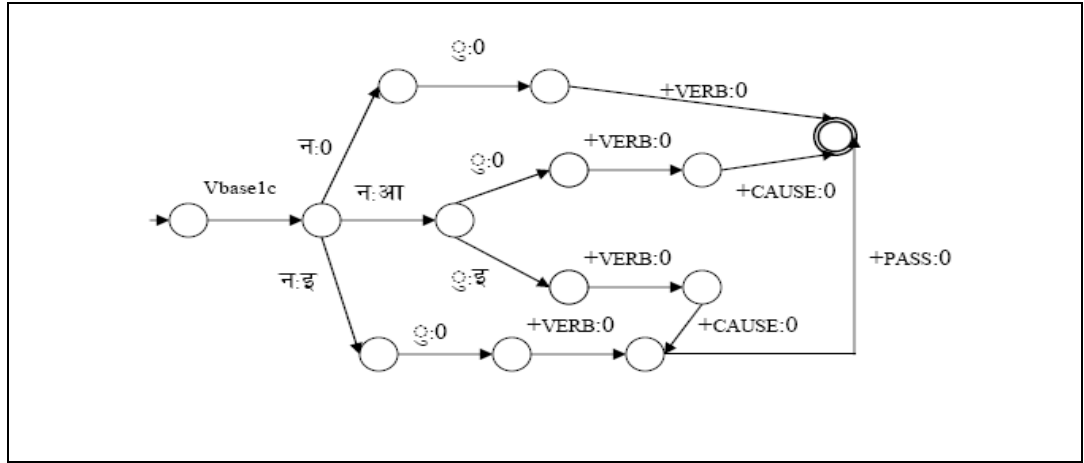


Figure 4.3: A finite state transducer for Type1c verb stems

The phonological rules in PR 4.2 are compiled into a finite state transducer and composed with the finite state transducer demonstrated in Figure 4.3.

Phonological rule

PR 4.2

- i. Causative marker आ a inserted between the consonants cluster of the *i*-ending intransitive verbs at the surface level.

Regular expression: [.] -> ा || cons __ cons;

- ii. Stem final vowel ि i of the *i*-ending intransitive verbs is deleted for causative stem.

Regular expression: ि -> [] || __;

- iii. Independent vowel आ a is changed to its corresponding dependent vowel ा a.

Regular expression: आ -> ा || __ .#;

d. Verb stem Type1d

Monosyllabic verb stems in Nepali having CaC structure which have four forms: base stem, passive stems, causative stems and causative-passive stems. Vowel ा a of the verb stems changes to vowel आ A when causative maker आ a follows the stem. Some examples are illustrated in Table 4.30 with their corresponding morphological tags.

Table 4.30: Type1d verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
काँप् kãp	काँपि kãp-i	कॅपा kãp-a	कॅपाइ kãp-a-i	'shiver'
हाँस् hãs	हाँसि hãs-i	हँसा hãs-a	हँसाइ hãs-a-i	'laugh'

The verb stems listed in Table 4.30 are compiled into a finite state transducer as illustrated in Figure 4.4 and it is capable of analyzing and generating them.

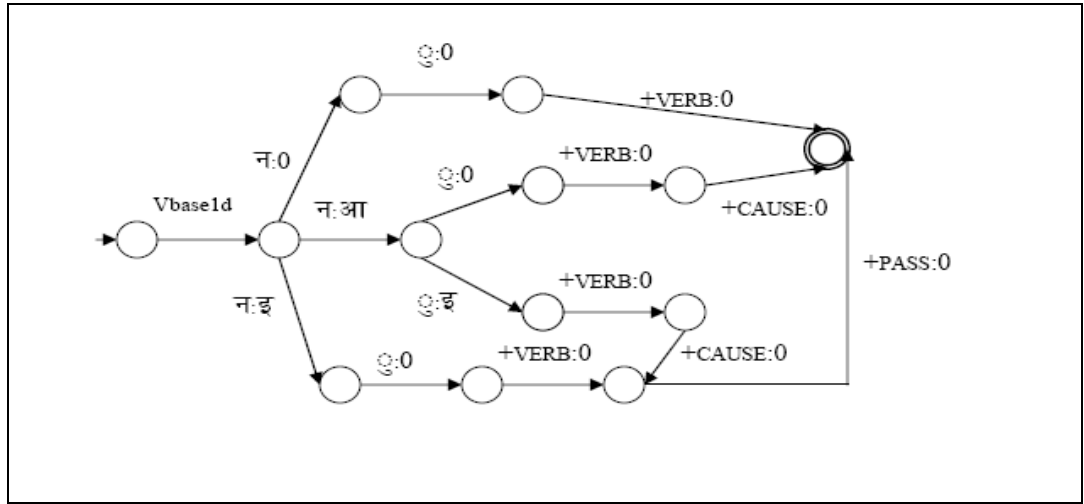


Figure 4.4 A finite state transducer for Type1d verb stems

The phonological rules involved in this process are listed in PR 4.3 which are compiled and composed with the transducer illustrated Figure 4.4.

Phonological rule

PR 4.3

- i. Vowel ा of the verb stems having CaC structure is changed to vowel अ if the stem is followed by causative maker आ at the surface level.

Regular expression: आ -> [] || cons __ cons;

- ii. Independent vowel आ changes to its corresponding dependent vowel ा.

Regular expression: आ -> ा || __ .#;

e. Verb stem Type1e

Monosyllabic consonant ending intransitive verb stems in Nepali that have four forms: base stem, passive stems, causative stems and causative-passive stems have been grouped. Some examples of verb stems of are illustrated in Table 4.31 and Table 4.32 with their corresponding morphological tags.

Table 4.31: Type1e verb stems (i)

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
बस् b Δ S	बसि b Δ S-i	बसा b Δ S-a	बसाइ b Δ S-a-i	sit
खस् k ^h Δ S	खसि k ^h Δ S-i	खसा k ^h Δ S-a	खसाइ k ^h Δ S-a-i	drop
छिन् ts ^h in	छिनि ts ^h in-i	छिना ts ^h in-a	छिनाइ ts ^h in-a-i	chop off

Table 4.32: Type1e verb stems (ii)

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
मर् m Δ r	मरि m Δ r-i	मरा m Δ r-a	मराइ m Δ r-a-i	'to kill'
गल् g Δ l	गलि g Δ l-i	गला g Δ l-a	गलाइ g Δ l-a-i	'to melt'
चल् ts Δ l	चलि ts Δ l-i	चला ts Δ l-a	चलाइ ts Δ l-a-i	'to move?'
जँच् dz $\tilde{\Delta}$ ts	जँचि dz $\tilde{\Delta}$ ts-i	जँचा dz $\tilde{\Delta}$ ts-a	जँचाइ dz $\tilde{\Delta}$ ts-a-i	'to examine'
झर् dz ^h Δ r	झरि dz ^h Δ r-i	झरा dz ^h Δ r-a	झराइ dz ^h Δ r-a-i	'to drop'
टर् t Δ r	टरि t Δ r-i	टरा t Δ r-a	टराइ t Δ r-a-i	'to escape artfully'
सर s Δ r	सरि s Δ r-i	सरा s Δ r-a	सराइ s Δ r-a-i	'to move aside'

The verb stems listed in Table 4.31 and Table 4.32 are compiled into a finite state transducer as demonstrated in Figure 4.5 and it is capable of analyzing and generating them.

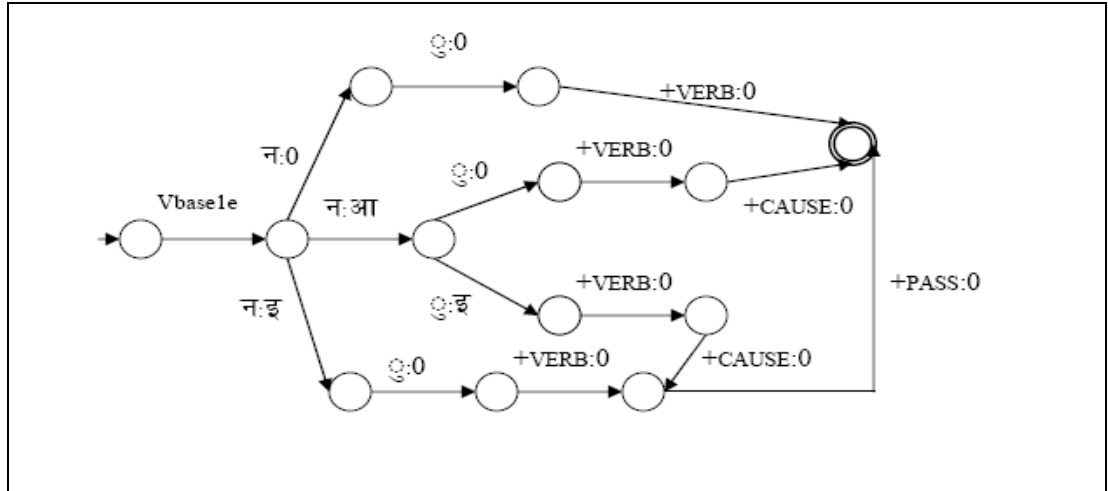


Figure 4.5: A finite state transducer for Type 1e verb stems

The phonological rules in PR 4.4 are compiled and composed with the transducer illustrated in Figure 4.5.

Phonological rule

PR 4.4

- i. Halanta ् at the end of the consonant ending verb stem is deleted before the causative marker आ a and passive marker ि i at the surface level.

Regular expression: ् -> [] || __ ई|आ;

- ii. Independent vowel आ a and इ i change to their corresponding dependent vowels ा a and ि i, respectively.

Regular expressions: आ -> ा || __ .#;

इ -> ि || __. #.;

4.4.2 Transitive verb stem

a. Verb stem Type2a

Polysyllabic verbs which contain vowel आ a within the stems and have only two forms: base stem and passive stem are grouped in this class. Some examples of such verbs are listed in Table 4.33 and Table 4.34 with their corresponding morphological tags and the gloss of base stems.

Table 4.33: Type2a verb stems (i)

	Tags		Tags	Gloss of stem
उचाल् utsal	+VERB	उचालि utsal-i	+VERB+PASS	'to lift'
अर्जाप् ᳵrdzap	+VERB	अर्जापि ᳵrdzap-i	+VERB+PASS	'to sharpen'
झपार dz ^h apar	+VERB	झपारि dz ^h apar-i	+VERB+PASS	'to scold'
पछार pats ^h ar	+VERB	पछारि pats ^h ar-i	+VERB+PASS	'to make upside down'
सराप् sarap	+VERB	सरापि sarap-i	+VERB+PASS	'to curse'

Table 4.34: Type2a verb stems (ii)

VI-form	Tags	Passive form	Tags	Gloss of stem
मार mar	+VERB	मारि mar-i	+VERB+PASS	'to lift'
गाल् gal	+VERB	गालि gal-i	+VERB+PASS	'to melt'
चाल् tsal	+VERB	चालि tsal-i	+VERB+PASS	'to make move'
जाँच् dzats	+VERB	जाँचि dzats-i	+VERB+PASS	'to examine'
झार dz ^h ar	+VERB	झारि dz ^h ar-i	+VERB+PASS	'to drop'
टार tar	+VERB	टारि tar-i	+VERB+PASS	'to '
सार sar	+VERB	सारि sari	+VERB+PASS	'to shift'

The verb stems listed in Table 4.33 and Table 4.34 are compiled into a finite state transducer as demonstrated in Figure 4.6 and it is capable of analyzing and generating them.

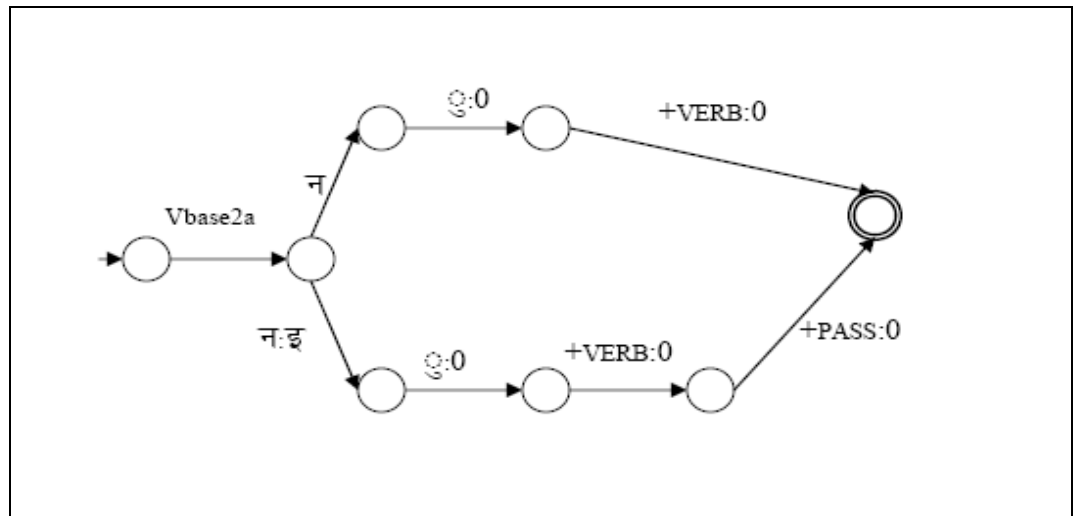


Figure 4.6: A finite state transducer for Type2a verb stems

The phonological rules involved in this process are listed in PR 4.5 and they are compiled and composed with the finite state transducer as demonstrated in Figure 4.6.

Phonological rule

PR 4.5

- i. Halanta ् at the end of the consonant ending verb stems is deleted before the causative marker आ *a* and passive marker ई *i* at the surface level.

Regular expression: ् -> [] || __ ई|आ;

- ii. Independent vowel आ *a* and ई *i* are changed to their corresponding dependent vowels ा and ि, respectively.

Regular expressions: आ -> ा || __ .#.;

ई -> ि || __ .#.;

b. Verb stem Type2b

i-ending polysyllabic basic verb stems which have four forms: base stems, passive stems, causative stems and causative-stems have been grouped in this class. Some examples are listed in Table 4.35 with their corresponding morphological tags.

Table 4.35: Type2b verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
पक्रि pʌkri	पक्रिइ pʌkri-i	पक्रा pʌkr-a	पक्राइ pʌkr-a-i	'arrest'
पर्खि pʌrk ^h i	पर्खिइ pʌrk ^h i-i	पर्खा pʌrk ^h -a	पर्खाइ pʌrk ^h -a-i	'wait'
बिर्सि birsi	बिर्सिइ birsi-i	बिर्सा birs-a	बिर्साइ birs-a-i	'forget'
मन्सि mans-i	मन्सिइ mans-i-i	मन्सा mans-a	मन्साइ mans-a-i	'throw away'
सम्झि sʌmdz ^h i	सम्झिइ sʌmdz ^h i-i	सम्झा sʌmdz ^h -a	सम्झाइ sʌmdz ^h -a-i	'remember'
कुल्चि kultsi	कुल्चिइ kultsi-i	कुल्चा kults-a	कुल्चाइ kults-a-i	'tread'
उईंति uĩti	उईंतिइ uĩti-i	उईंटा uĩt-a	उईंटाइ uĩt-a-i	'spindle'

The verb stems listed in Table 4.35 are compiled into a finite state transducer as demonstrated in Figure 4.7 and it is capable of analyzing and generating them.

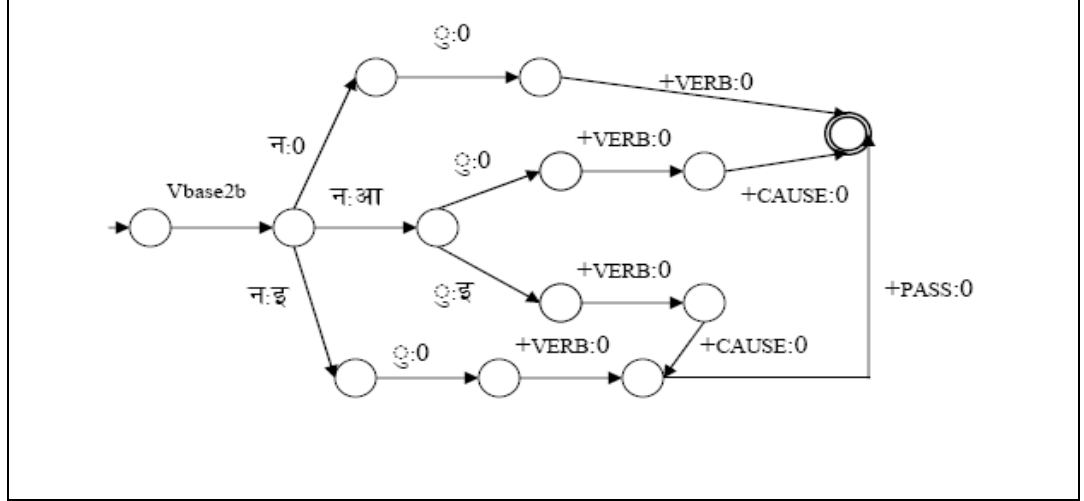


Figure 4.7: A finite state transducer for Type2b verb stems

The finite state transducer in Figure 4.7 is composed with the network of phonological rules listed in PR 4.6.

Phonological rules

PR 4.6

- i. Vowel ि *i* of the *i*-ending transitive verbs is deleted before the causative marker आ *a* at the surface level.

Regular expression: ि -> [] || __ आ;

- ii. Independent vowel आ *a* is changed to its corresponding dependent vowel ा *aa*.

Regular expression: आ -> ा || __ .#;

c. Verb stem Type2c

Monosyllabic verb stems having C a C structure which have four forms: base stem, passive stems, causative stems and causative-passive stems. Vowel आ *a* of the basic

verb stems changes \mathcal{A} when the causative maker \mathcal{A} a follows the base stem. Some examples are illustrated in Table 4.36 with their corresponding morphological tags.

Table 4.36: Type2c verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
खाप् k ^h ap	खापि k ^h ap-i	खपा k ^h Λp-a	खपाइ k ^h Λp-a-i	'pile up'
गाड् gaḍ	गाडि gaḍ-i	गडा gΛḍ-a	गडाइ gΛḍ-a-i	'bury'
छाप् ts ^h ap	छापि ts ^h ap-i	छपा ts ^h Λp-a	छपाइ ts ^h Λp-a-i	'print'
टाँस् tã̃s	टाँसि tã̃s-i	टँसा tã̃s-a	टँसाइ tã̃s-a-i	'stick up'
तान् tan	तानि tan-i	तना tΛn-a	तनाइ tΛn-a-i	'pull'
नाच् nats	नाचि nats-i	नचा nΛts-a	नचाइ nΛts-a-i	'dance'
बाँच् bã̃ts	बाँचि bã̃ts-i	बँचा bã̃ts-a	बँचाइ bã̃ts-a-i	'survive'
हाँक् fiãk	हाँकि fiãk-i	हँका fiãk-a	हँकाइ fiãk-a-i	'drive'

The verb stems listed in Table 4.36 are compiled into a finite state transducer as illustrated in Figure 4.8 and it is capable of analyzing and generating them.

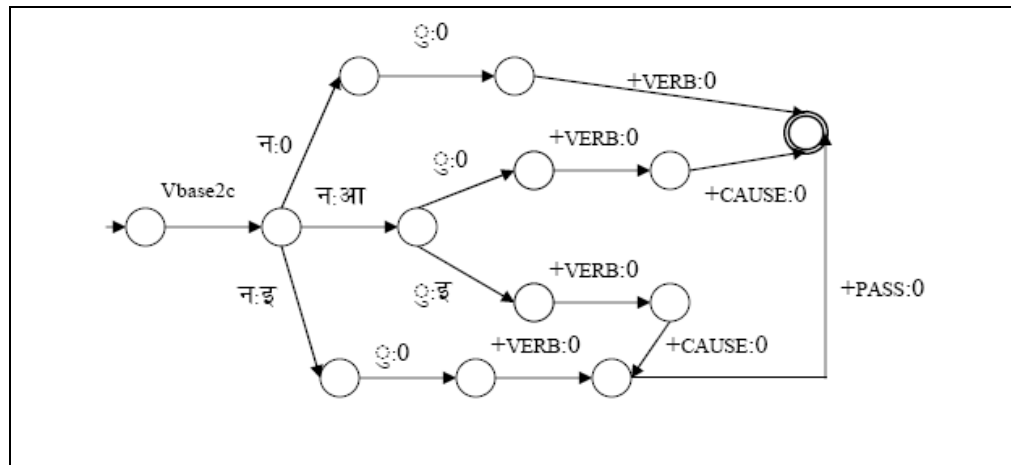


Figure 4.8: A finite state transducer for Type2c verb stems

The phonological rules involved in this process are listed in PR 4.7 which are compiled and composed with the transducer illustrated in Figure 4.8.

Phonological rule

PR 4.7

- i. Vowel $\text{ɔ}a$ of the verb stems having C a C structure is changed to vowel a if the stem is followed by causative maker $\text{आ}a$ at the surface level.

Regular expression: $\text{आ} \rightarrow [] \parallel \text{cons} _ \text{cons}$;

- ii. Independent vowel $\text{आ}a$ is changed to its corresponding dependent vowel $\text{ɔ}a$.

Regular expression: $\text{आ} \rightarrow \text{ा} \parallel _ \#.$;

d. Verb stem Type2d

Monosyllabic consonant ending transitive basic verb stems which have four forms: base stem, passive stems, causative stems and causative-passive stems have been grouped in this class. Some examples are illustrated in Table 4.37 with their corresponding morphological tags.

Table 4.37: Type2d verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
पढ् $p\Lambda d^h$	पढि $p\Lambda d^h-i$	पढा $p\Lambda d^h-a$	पढाइ $p\Lambda d^h-a-i$	read
किन् kin	किनि $kin-i$	किना $kin-a$	किनाइ $kin-a-i$	buy
जोत् $dzot$	जोति $dzot-i$	जोता $dzot-a$	जोताइ $dzot-a-i$	plough
घस् $g^h\Lambda s$	घसि $g^h\Lambda s-i$	घसा $g^h\Lambda s-a$	घसाइ $g^h\Lambda s-a-i$	massage

The finite state transducer illustrated in Figure 4.9 encodes the verb stems listed in Table 4.37 and it is capable of analyzing and generating them.

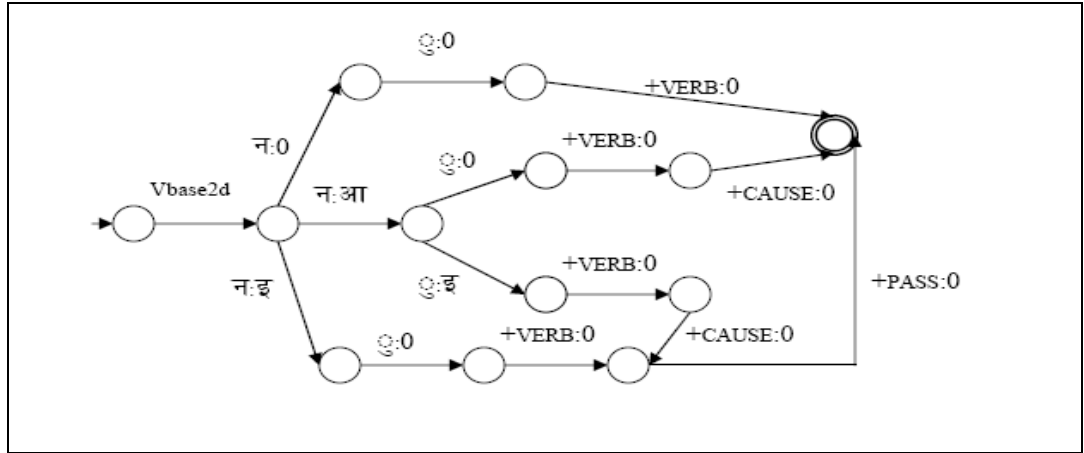


Figure 4.9: A finite state transducer for Type2d verb stems

The phonological rules listed in PR 4.8 are compiled and composed with the transducer demonstrated in Figure 4.9.

Phonological rule

PR 4.8

- i. Halanta ् at the end of the consonant ending verb stem is deleted before the causative marker आ and passive marker इ at the surface level.

Regular expression: ् -> [] || __ इ|आ;

- ii. Independent vowel आ and इ are changed to their corresponding dependent vowels ा and ि, respectively.

Regular expression: आ -> ा || __ .#;

इ -> ि || __ .#;

4.4.3 Irregular verb (intransitive and transitive) stems

Some of the intransitive and transitive verb stems which are not regular in their stem formation process. Some of the verbs of this type are listed in Table 4.38.

Table 4.38: Irregular verb stems

+VERB	+VERB+PASS	+VERB+CAUSE	+VERB+CAUSE+PASS	Gloss of base
आ a	आइ a-i			'come'
जा dza	जाइ dza-i			'go'
रो ro	रोइ ro-i	रुवा ru-wa	रुवाइ ru-wa-i	'cry'
हो ho	होइ ho-i			'be'
खा k ^h a	खाइ k ^h a-i	ख्वा k ^h w-a	ख्वाइ k ^h w-a-i	'eat'
पा pa	पाइ pa-i			'get'
दि di	दिइ di-i	दिला di-la	दिलाइ di-la-i	'give'
लि li	लिइ li-i	ल्या l-ja	ल्याइ l-ja-i	'take'
धो d ^h o	धोइ d ^h o-i	धुला d ^h u-la	धुलाइ d ^h u-la-i	'wash'
बस् b _{AS} -		बसाल् b _{AS} -al-	बसालि b _{AS} -al-i	'sit'
खस् k ^h _{AS} -		खसाल् k ^h _{AS} -al-	खसालि k ^h _{AS} -al-i	'drop'
चुँडि tsũḍi-		चुडाल् tsũḍal-	चुडालि tsũḍ-al-i	'snatch'
छिन् ts ^h in-		छिनाल् ts ^h inal-	छिनालि ts ^h in-al-i	'chop off'

The finite state transducer for irregular can not be generalized as done in earlier cases. Therefore, their network is not demonstrated rather they will directly be encoded and implemented.

4.4.4 Suppletive verb stems

There are two pairs of suppletive verb stems; they are हुँ- *hiu*- 'become' vs. भ- *b^hA*- 'became' and जा- *dza*- 'go' vs. ग- *gA*- 'went'. First and second members of the suppletive pairs follow the different tracks in the inflectional paradigm and this is illustrated in Table 4.39.

Table 4.39: Suppletive verb stems

हुँ <i>fiu-</i> 'become'!	भ <i>b^hΛ-</i> 'became'	जा <i>dza-</i> 'go'	ग <i>gΛ-</i> 'went'
Non-past tense		Non-past tense	
	Past tense		Past tense
	Perfect aspect		Perfect aspect
Imperfect aspect		Imperfect aspect	
Habitual aspect		Habitual aspect	
	Inferential aspect		Inferential aspect
Imperative		Imperative	
Optative	Optative	Optative	Optative
Potential		Potential	
	Absolutive		Absolutive
Infinitive		Infinitive	
Purposive		Purposive	
Prospective		Prospective	
Durative		Durative	
	Conditional		Conditional
	Perfective		Perfective
	Conjunctive		Conjunctive

The phonological rules involved in altering the suppletive forms are listed in PR 4.9.

Phonological rule:

PR 4.9

- i. The verb stem *हुँ fiu* changes to *भ b^hΛ* if the following suffix begins with *ए e* or *इ i* or *ई i:* or *य ja*.

Regular expression: हुँ -> भ || __ एइईय;

- ii. The verb stem *dza* is changed to *gΛ* if the following suffix begins with *ए e* or *इ i* or *ई i:* or *य ja*.

Regular expression: जा -> ग || __ एइईय;

4.5 Verbal inflections

In Nepali, verbal inflections are suffixes attached to verb stems. A verb stem can be a base form, a causative form or a passive form. The inflectional suffixes in general encode the inherent verbal features such as tense, aspect and mood. Besides these inherent features, these inflectional suffixes also encode the agreement features such as person, number, gender and honorificity with reference to the subject of the sentence.

Inherent features and agreement features are not clearly distinguishable in terms of symbols rather represented by a set of symbols. Therefore, these inflectional suffixes form a paradigm with respect to above mentioned features. The suffixal negative marker gets intermixed with these inflectional suffixes in some forms. This leads the formation of both positive and negative paradigms of the verbal inflection. The second type of negative marker is a prefix which appears in front of the verb stems (see 4.2.3). The auxiliary verbs in Nepali are more or less equivalent to the inflections in encoding the verbal features, therefore, they are discussed in this section.

4.5.1 Auxiliary verbs in Nepali

Nepali has two kinds of auxiliaries, namely, non-past existential auxiliary $\text{छ } ts^h\lambda$ and non-past identificational auxiliary $\text{हो } ho$. But the past auxiliary $\text{थि } t^hi$ is a different stem which inflects like main verbs (Dahal 1974 and Adhikari 2062VS). The existential and indentificational auxiliary verbs for both non-past and past tenses (i and ii) have been discussed with their inflections.¹

	<u>Non-past form</u>	<u>Past form</u>
i. Existential	छ	थि-
ii. Indentificational	हो	--

a. Non-past existential auxiliary verb $\text{छ } ts^h\lambda$ 'be' inflects for person, number, gender and honorific agreement as in $\text{छिन् } ts^hin$ 'be.NPST.3SG.FEM.HON' in (14). In the default case, the auxiliary verb form $\text{छ } ts^h\lambda$ 'be' itself represents the third person singular masculine and carries non-past tense and for other cases agreement inflections follow

¹ See Sharma (1980) for detailed description of auxiliary verbs in Nepali.

it. All together, this existential verb छ *ts^hʌ* 'be' has twelve forms and the inflections are listed in Table 4.39 with their corresponding morphological tags.

(14) दिदि घरमा छिन्।

didi g^hʌr-ma ts^hʌin
 elder sister home-LOC be.NPST.3SG.FEM.HON
 'The elder sister is at home.'

Table 4.39: Inflections for non-past existential verb छ *chʌ* 'be' (affirmative)

Grammatical category	Inflections	IPA	Tags
First person singular	उ	u	NPST.1SG
First person plural	औँ	ʌũ	NPST.1PL
Second person masculine singular	स्	s	NPST.2SG.MASC
Second person feminine singular	एस्	es	NPST.2SG.FEM
Second person masculine singular hon	औ	ʌu	NPST.2SG.MASC.HON
Second person feminine singular hon	यौ	jʌu	NPST.2SG.FEM.HON
Second person plural	औ	ʌu	NPST.2PL
Third person singular masculine	ϕ	ϕ	NPST.3SG.MASC
Third person feminine singular	ए	e	NPST.3SG.FEM
Third person masculine singular hon	न्	ʌn	NPST.3SG.MASC.HON
Third person feminine singular hon	इन्	in	NPST.3SG.FEM.HON
Third person plural	न्	ʌn	NPST.3PL

The finite state transducer in Figure 4.10 encodes the auxiliary verb छ *chʌ* 'be' and its various forms and it is can analyze and generate them. The rules involved in this case are directly encoded into the finite state transducer.

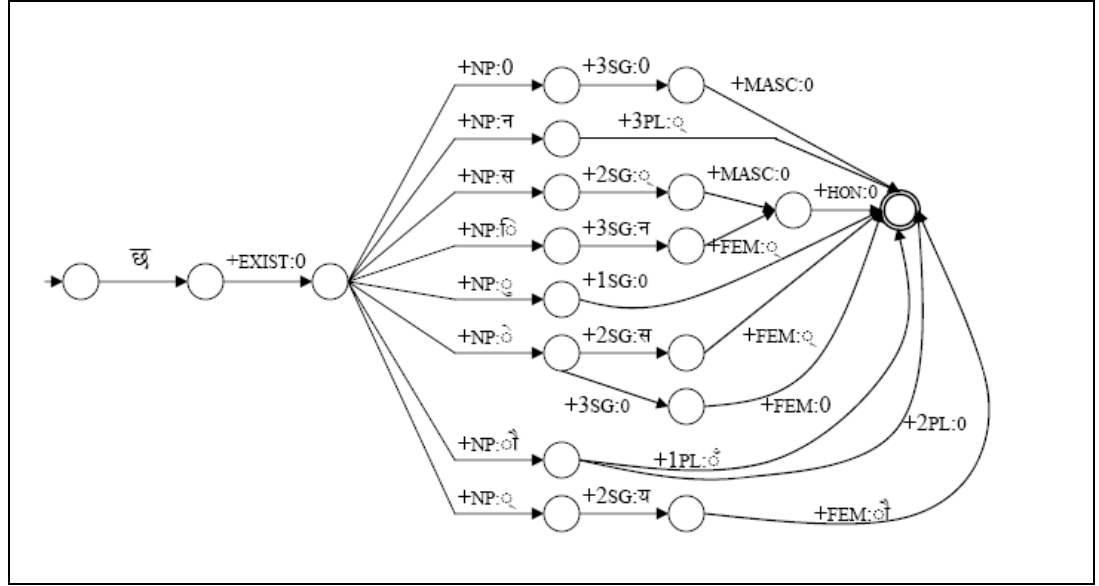


Figure 4.10: A finite state transducer for inflections of non-past existential verb छ *cha* 'be' (affirmative)

b. In the negative formation of the existential verb छ *ts^hʌ*, the negative suffix -न -*na* 'NEG' is inserted between the auxiliary stem छ *ts^hʌ* 'be' and agreement inflections. During the process of negativization some morphophonemic changes occur as छैनन् *ts^hʌinʌn* 'be-NEG' in (15). There are eight negative forms where there is no distinction in gender. Table 4.40 lists the inflections with their corresponding morphological tags.

- (15) दिदि घरमा छैनन्।
 didi g^hʌr-ma ts^hʌinʌn
 elder sister home-LOC be.NPST.NEG.3SG.FEM.HON
 'The elder sister is not at home.'

Table 4.40: Inflection for non-past existential verb छ *cha* 'be' (Negative)

Grammatical category	Inflections		Tags
First person singular	इनँ	inã	NPST.NEG.1SG
First person plural	इनौं	inãũ	NPST.NEG.1PL
Second person singular	इनस्	inʌs	NPST.NEG.2SG
Second person singular hon	इनौ	inʌu	NPST.NEG.2SG.HON
Second person plural	इनौ	inʌu	NPST.NEG.2PL
Third person singular	इन	inʌ	NPST.NEG.3SG
Third person singular hon	इनन्	inʌn	NPST.NEG.3SG.HON
Third person plural	इनन्	inʌn	NPST.NEG.3PL

The finite state transducer in Figure 4.11 encodes both inflections and rules involved. It can analyze and generate the negative forms of the auxiliary verb छे *tsʰɛ* 'be'.

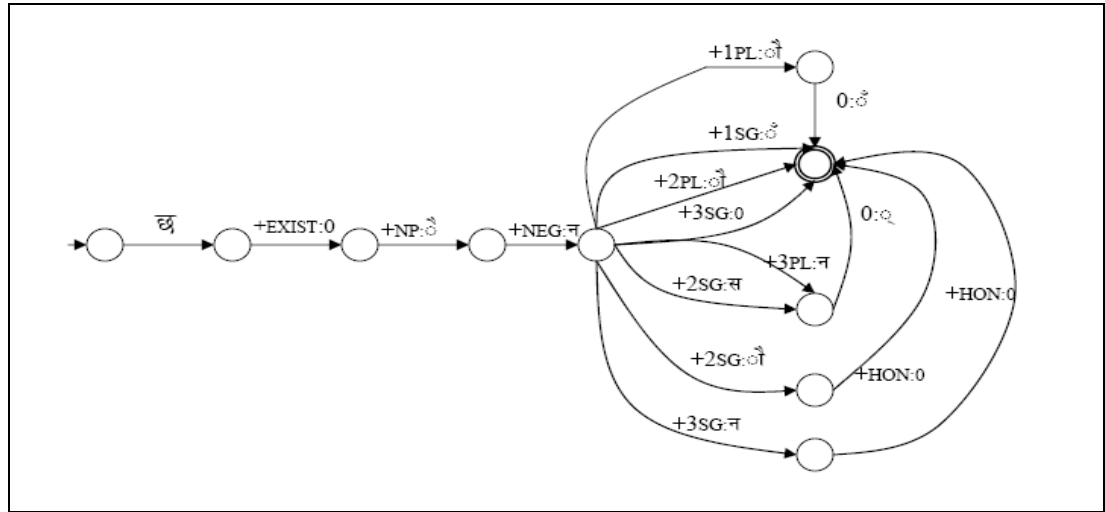


Figure 4.11: A finite state transducer for inflections of non-past existential verb छे *chɛ* 'be' (negative)

c. Non-past identificational auxiliary verb हो *ho* 'be' takes similar agreement inflections but this verb does not have gender distinction. That means, there are no feminine forms. The verb itself represents the third person singular form. Some morphophonemic changes occur when inflections combine with हो *ho* 'be' as हो *ɦɔu* 'be.NPST.2SG.HON' in (16). There are altogether eight forms of this verb and the inflections are listed in Table 4.41 with their corresponding morphological tags.

- (16) तिमी शिक्षक हो।
 timi siktsʰɛk ɦɔu
 2SG.HON teacher be.NPST.2SG.HON
 'You are a teacher.'

Table 4.41: Inflections for non-past identificational verb हो *ho* ‘be’ (affirmative)

Grammatical category	Inflections	IPA	Tags
First person singular	उँ	ũ	NPST.1SG
First person plural	औँ	ʌũ	NPST.1PL
Second person singular	स्	s	NPST.2SG
Second person singular hon	औ	ʌu	NPST.2SG.HON
Second person plural	औ	ʌu	NPST.2PL
Third person singular	ϕ	ϕ	NPST.3SG
Third person singular hon	न्	n	NPST.3SG.HON
Third person plural	न्	n	NPST.3PL

The auxiliary verb हो *ho* ‘be’ and its various forms are compiled into a finite state transducer as demonstrated in Figure 4.12 and it is capable of analyzing and generating them.

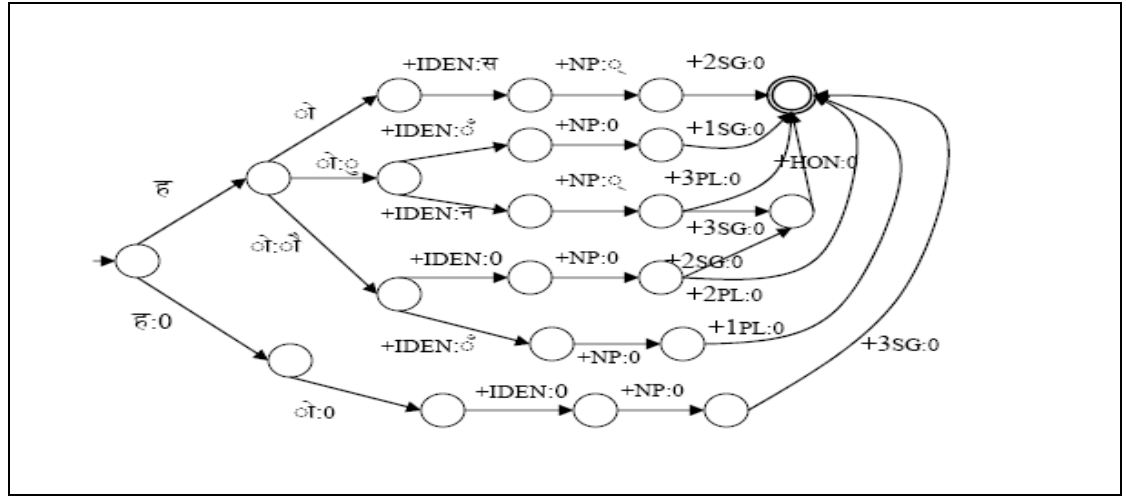


Figure 4.12: A finite state transducer for inflections of non-past identificational verb हो *ho* ‘be’ (affirmative)

The phonological rules involved in this process are listed in PR 4.10 and they have been directly encoded into the finite state transducer illustrated in Figure 4.12.

Phonological rules

PR 4.10

- i. Independent vowels उu and औau are changed to their corresponding dependent vowels ुu and ौau, respectively after हो.

Regular expressions: उ -> उ॒ || हो __;
 औ -> औ॒ || हो __;
 ो -> ो॒ || ह __ न् .#.
 ो -> [] || ह __ ो॒;

d. In the negative formation of identificational auxiliary verb *हो हो*, the negative suffix *न-न* (*इन in*) is inserted between the auxiliary stem *हो हो* and agreement inflections. In this process of negativization, no changes occur in the stem itself as *होइनो होइनो* *होइनो* *होइनो* in (17). There are eight negative forms parallel to the positive ones. They are listed in Table 4.42 with their corresponding morphological tags.

(17) तिमी शिक्षक होइनो।
timi sikts^hak ho in
 2SG.HON teacher be.NPST.NEG.2SG.HON
 'You are not a teacher.'

Table 4.42: Inflection for non-past identificational verb हो हो 'be' (Negative)

Grammatical category	Inflections		Tags
First person singular	इनँ	inã	NPST.NEG.1SG
First person plural	इनौं	inãũ	NPST.NEG.1PL
Second person singular	इनस्	inãs	NPST.NEG.2SG
Second person singular hon	इनो	inãu	NPST.NEG.2SG.HON
Second person plural	इनो	inãu	NPST.NEG.2PL
Third person singular	इन	inã	NPST.NEG.3SG
Third person singular hon	इनन्	inãn	NPST.NEG.3SG.HON
Third person plural	इनन्	inãn	NPST.NEG.3PL

The finite state transducer in Figure 4.13 is capable of analyzing and generating the negative forms of auxiliary verb *हो हो* 'be'.

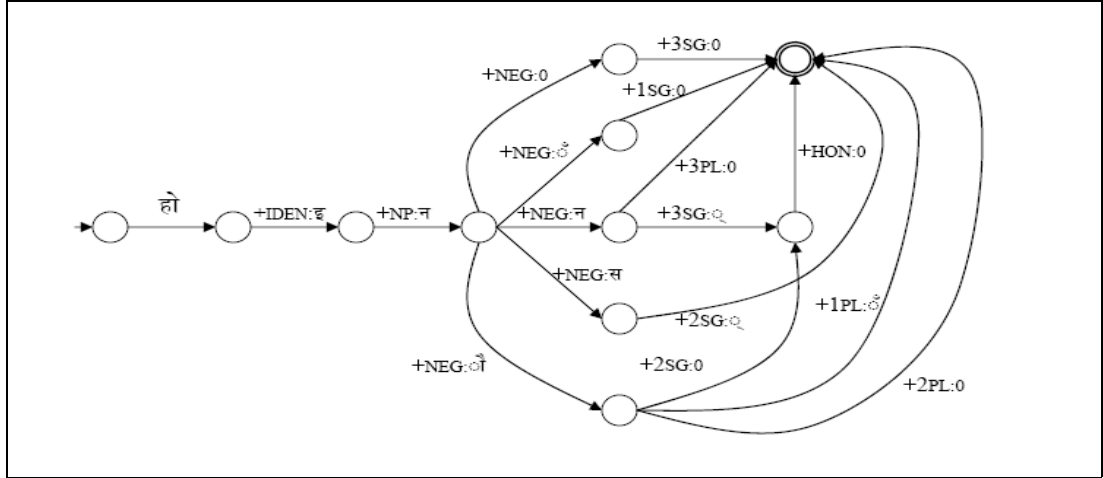


Figure 4.13: A finite state transducer for inflection of non-past identificational verb हो *ho* 'be' (Negative)

a. Past existential auxiliary verb थि t^hi 'be.PST' inflects for person, number, gender and honorific agreement as in थियोँ $t^hi:j\Lambda\tilde{u}$ in (18). The auxiliary verb form थि t^hi itself carries past tense and the agreement inflections follow it. In this case also the auxiliary verb stem does not change when suffixes are attached. All together, this existential verb has ten forms and the inflections are listed in Table 4.43 with their corresponding morphological tags.

- (18) हामी तनावमा थियोँ।
 fiami taṇab-ma $t^hi-j\Lambda\tilde{u}$
 2PL tension-LOC be.PST.2PL
 'We were at tension.'

Table 4.43: Inflections for past existential verb थि t^hi 'be' (affirmative)

Grammatical category	Inflections	IPA	Tags
First person singular	एँ	\tilde{e}	PST.1SG
First person plural	योँ	$j\Lambda\tilde{u}$	PST.1PL
Second person singular	इस्	is	PST.2SG
Second person singular hon	यो	$j\Lambda u$	PST.2SG.HON
Second person plural	यो	$j\Lambda u$	PST.2PL
Third person masculine singular	यो	jo	PST.3SG.MASC
Third person feminine singular	ई	i:	PST.3SG.FEM
Third person masculine singular hon	ए	e	PST.3SG.MASC.HON
Third person feminine singular	इन्	in	PST.3SG.FEM.HON
Third person plural	ए	e	PST.3PL

The finite state transducer illustrated in Figure 4.14 can analyze and generate the positive forms of auxiliary verb थि t^hi 'be'.

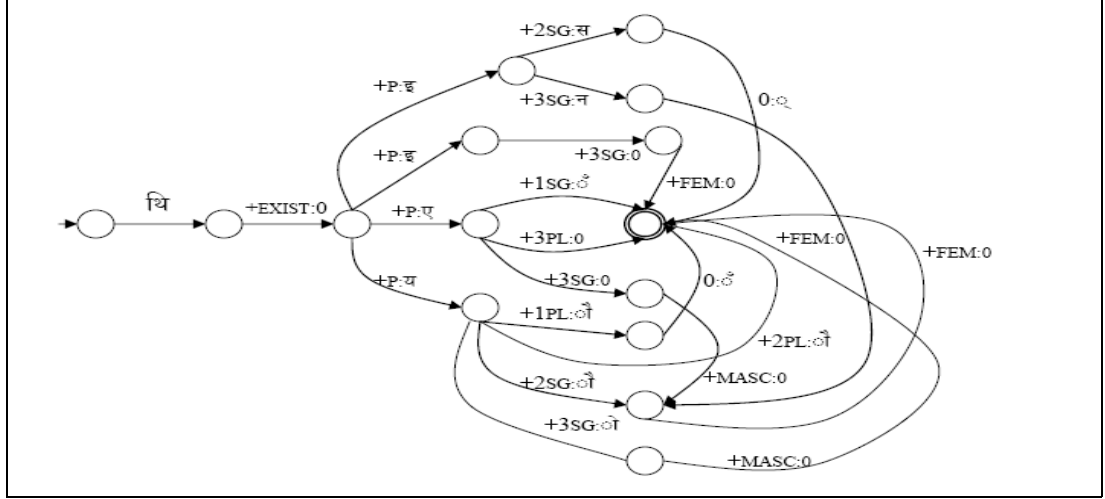


Figure 4.14: A finite state transducer for inflections of past existential verb थि t^hi 'be' (affirmative)

b. In the negative formation of the existential verb थि t^hi-, the negative suffix -न -na is inserted between the auxiliary stem थि t^hi- and agreement inflections. During the process of negativization no morphophonemic changes occur auxiliary verb stem as थिएनौ t^hi-enajau in (19). There are eleven negative forms one more than positive ones and Table 4.44 lists the inflections with their corresponding morphological tags.

- (19) हामी तनावमा थिएनौ।
 fiami t^hanab-ma t^hi-enajau
 2PL tension-loc be.P-NEG.2PL
 'We were not at tension.'

Table 4.44: Inflections for past existential verb थि *tʰi* 'be' (negative)

Grammatical category	Inflections	IPA	Tags
First person singular	इनँ	inã	PST.NEG.1SG
First person plural	एनौ	enAu	PST.NEG.1PL
Second person singular	इनस्	inAs	PST.NEG.2SG
Second person singular masculine hon	एनौ	enAu	PST.NEG.2SG.MASC.HON
Second person singular hon	इनौ	enAu	PST.NEG.2SG.FEM.HON
Second person plural	एनौ	enAu	PST.NEG.2PL
Third person masculine singular	एन	enA	PST.NEG.3SG.MASC
Third person feminine singular	इन	inA	PST.NEG.3SG.FEM
Third person masculine singular hon	एनन्	enAn	PST.H.NEG.3SG.MASC
Third person feminine singular hon	इनन्	inAn	PST.NEG.3SG.FEM.HON
Third person plural	एनन्	enAn	PST.NEG.3PL

The finite state transducer in Figure 4.15 can analyze and generate the negative forms of auxiliary verb थि *tʰi* 'be'.

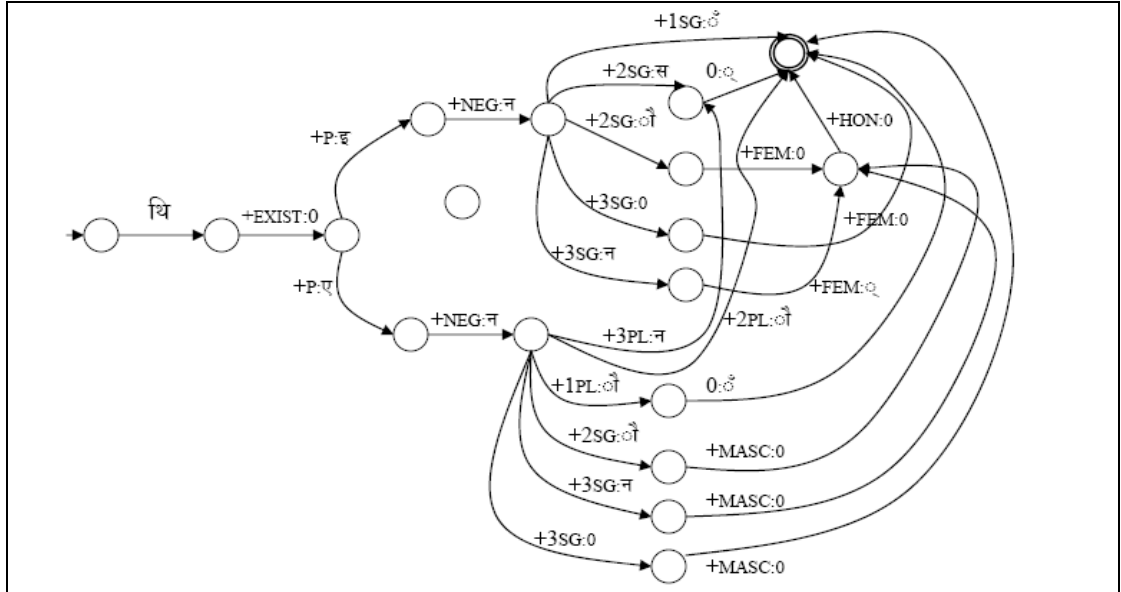


Figure 4.15: A finite state transducer for inflections of past existential verb थि *tʰi* 'be' (negative)

4.5.2 Tense

The Nepali morphologically exhibits two tenses: past and non-past. Past tense refers to the action that is completed prior to the speech event and non-past tense refers to

the action that happens at the time of speech event or later (Schmidt 1993; Pokharel 2054VS and Adhikari 2055VS).

a. Non-past Tense

Non-past tense in Nepali covers both present and future. There is no such future tense marker; however, the same present (non-past) makers and sometimes combines with prospective marker referring to the future tense. In fact, there are no such definite non-past tense markers in Nepali. The non-past existential auxiliary verb $\overline{\text{छ}} ts^h\Lambda$ as discussed in (4.5.1) behaves as non-past tense marker. The auxiliary meaning 'be' seems to be completely absorbed and agreement features are retained. This feature can be seen in (22), the verb $\overline{\text{ख}} k^h a$ 'eat' combines with auxiliary verb $\overline{\text{छ}} ts^h\Lambda$ and with its agreement inflection but the form $\overline{\text{छन्}} ts^h\Lambda n$ only indicates non-past tense and agreement features. The non-past tense and inflections for person, number, gender and honorificity altogether are in ten forms and they are listed in Table 4.45 with their corresponding morphological tags.²

- (22) भाइहरू भात खान्छन् ।
b^hai-fi_{PL}u: b^hat k^ha-n-ts^h_{NPST.3PL}
brother-PL rice eat- ϕ -NPST.3PL
'The brothers eat rice.'

² Most of the traditional grammarians treat $\overline{\text{छ}} ts^h\Lambda$ as an auxiliary verb separately. So, the most of the verb stems, except past form, are considered as the compound stems.

Table 4.45: Inflections for non-past tense (affirmative)

Grammatical category	Inflections	IPA	Tags
First person singular	छु	ts ^h u	NPST.1SG
First person plural	छौ	ts ^h Λũ	NPST.1PL
Second person masculine singular	छस्	ts ^h Λs	NPST.2SG.MASC
Second person feminine singular	छेस्	ts ^h es	NPST.2SG.FEM
Second person masculine singular hon	छौ	ts ^h Λu	NPST.2SG.MASC.HON
Second person feminine singular hon	छ्यौ	ts ^h jΛu	NPST.2SG.FEM.HON
Second person plural	छौ	ts ^h Λu	NPST.2PL
Third person masculine singular	छ	ts ^h Λ	NPST.3SG.MASC
Third person feminine singular	छे	ts ^h e	NPST.3SG.FEM
Third person masculine singular hon	छन्	ts ^h Λn	NPST.3SG.MASC.HON
Third person feminine singular hon	छिन्	ts ^h in	NPST.3SG.FEM.HON
Third person plural	छन्	ts ^h Λn	NPST.3PL

The finite state transducer illustrated in Figure 4.16 can analyze and generate the non-past positive forms of the inflections. This transducer is concatenated with finite state transducer of verb stems.

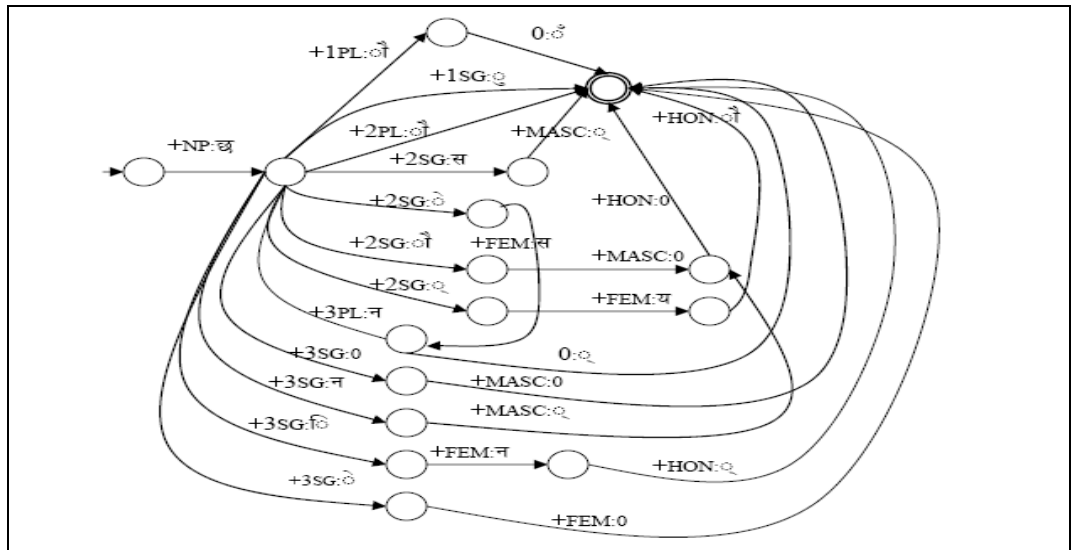


Figure 4.16: A finite state transducer for inflections of non-past tense

Non-past negative forms are formed from the suffixation of negative marker -न नऱ with verb stems. The non-tense marker छ ts^hΛ 'be' is completely absorbed and semantically null element दि di or दै dΛi are inserted before the negative marker न नऱ;

and the agreement markers follow it. Even though the non-past tense marker is not overtly present, the tense is indicated by the sequence. There are altogether twelve non-past negative forms which are listed in Table 4.46 with their corresponding morphological tags.

Table 4.46: Inflections for non-past tense negative 1

Grammatical category	Inflections	IPA	Tags
First person singular	दिनँ	dinã	NPST.NEG.1SG
First person plural	दिनौ	dΛinΛũ	NPST.NEG.1PL
Second person masculine singular	दिनस्	dΛinΛs	NPST.NEG.2SG.MASC
Second person feminine singular	दिनस्	dinΛs	NPST.NEG.2SG.FEM
Second person masculine singular hon	दिनौ	dΛinΛu	NPST.NEG.2SG.MASC.HON
Second person feminine singular hon	दिनौ	dinΛu	NPST.NEG.2SG.FEM.HON
Second person plural	दिनौ	dΛinΛu	NPST.NEG.2PL
Third person masculine singular	दिन	dΛinΛ	NPST.NEG.3SG.MASC
Third person feminine singular	दिन	dinΛ	NPST.NEG.3SG.FEM
Third person masculine singular hon	दिनन्	dΛinΛn	NPST.NEG.3SG.MASC.HON
Third person feminine singular hon	दिनन्	dinΛnΛn	NPST.NEG.3SG.FEM.HON
Third person plural	दिनन्	dΛinΛn	NPST.NEG.3PL

The finite state transducer in Figure 4.17 can analyze and generate the non-past negative forms of the inflections set 1. This transducer is concatenated with transducer of verb stems.

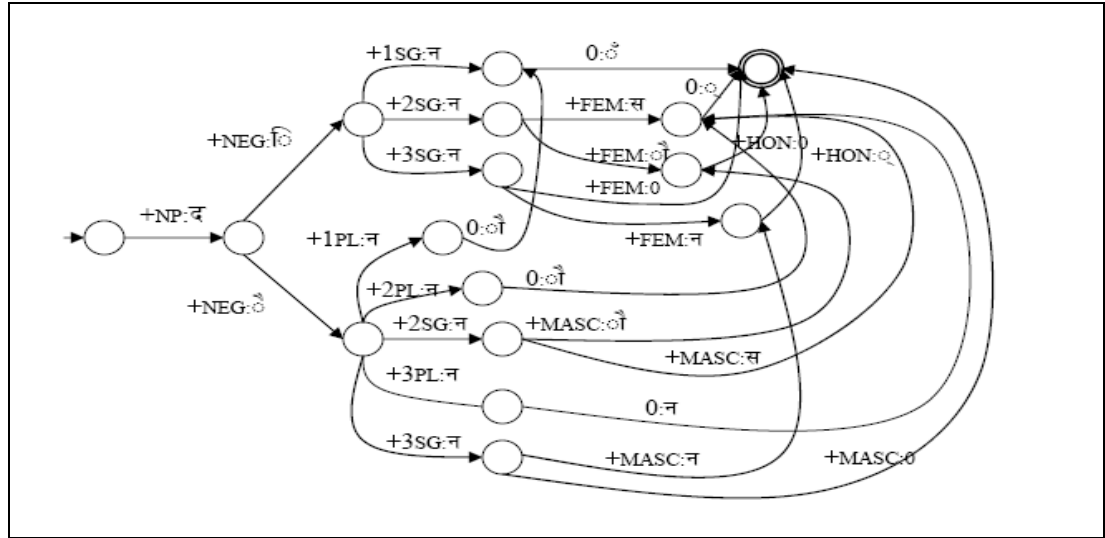


Figure 4.17: A finite state transducer for inflections of non-past tense negative 1

There is another set of negative marker $-न$ $-n\Lambda$ and agreement inflections which appear exclusively only with vowel ending verb stems. As discussed above the non-past tense maker $\bar{ञ}$ $ts^h\Lambda$ 'be' is completely absorbed; however, the tense feature is indicated by the sequence. In this case, inflections indicate person, number and honoricity but not the gender. Therefore, there are only eight non-past negative forms in this set as listed in Table 4.47 with their corresponding morphological tags.

Table 4.47: Inflections for non-past tense negative 2

Grammatical category	Inflections		Tags
First person singular	नँ	$n\tilde{\Lambda}$	NPST.NEG.1SG
First person plural	नौँ	$n\Lambda\tilde{u}$	NPST.NEG.1PL
Second person singular	नस्	$n\Lambda s$	NPST.NEG.2SG
Second person singular hon	नौ	$n\Lambda u$	NPST.NEG.2SG.HON
Second person plural	नौ	$n\Lambda u$	NPST.NEG.2PL
Third person singular	न	$n\Lambda$	NPST.NEG.3SG
Third person singular hon	नन्	$n\Lambda n$	NPST.NEG.3SG.HON
Third person plural	नन्	$n\Lambda n$	NPST.NEG.3PL

The finite state transducer in Figure 4.17a can analyze and generate the non-past negative forms of the inflections set 2. This transducer is concatenated with transducer of verb stems.

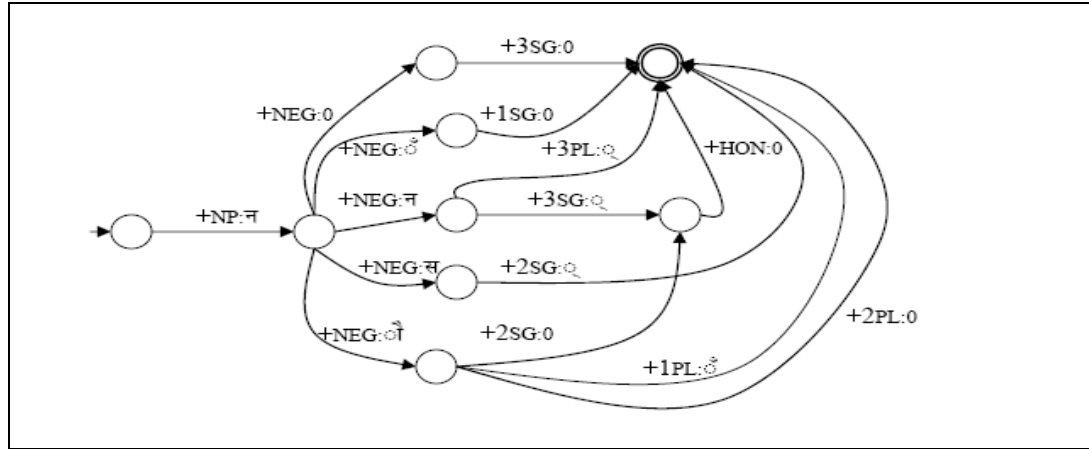


Figure 4.17a: A finite state transducer for inflections of non-past tense negative 2

b. Past Tense

The past tense in Nepali is indicated by the same inflectional paradigm as that of the past tense auxiliary verb stem *रि* *tʰi*- 'be'. In example (23), the verb stem *रो* *ro* 'cry' is followed by the past tense and agreement inflections -*यो* *-jo* 'PST.3SG.MASC'. Although, the past tense marker is not clear, *ए* *e*, *य* *j* and *इ* *i* in the paradigm indicate the past tense and the inflections indicating the agreement features follow them. For the convenience, the paradigm listed in Table 4.43 is reproduced in Table 4.48 with their corresponding morphological tags.

(23) बच्चो सारै रोयो ।

bʌtstso	sarʌi	ro-jo
child	more	cry-PST.3SG.MASC
'The child cried a lot.'		

Table 4.48: Inflections for past tense (affirmative)

Grammatical category	Inflections		Tags
First person singular	एँ	ẽ	PST.1SG
First person plural	यौँ	jΛũ	PST.1PL
Second person singular	इस्	is	PST.2SG
Second person singular hon	यौ	jΛu	PST.2SG.HON
Second person plural	यौ	jΛu	PST.2PL
Third person masculine singular	यो	jɔ	PST.3SG.MASC
Third person feminine singular	ई	i:	PST.3SG.FEM
Third person masculine singular hon	ए	e	PST.3SG.MASC.HON
Third person feminine singular hon	इन्	in	PST.3SG.FEM.HON
Third person plural	ए	e	PST.3PL

The finite state transducer in Figure 4.18 can analyze and generate the positive inflections of past tense. This transducer is also concatenated with transducer of verb stems.

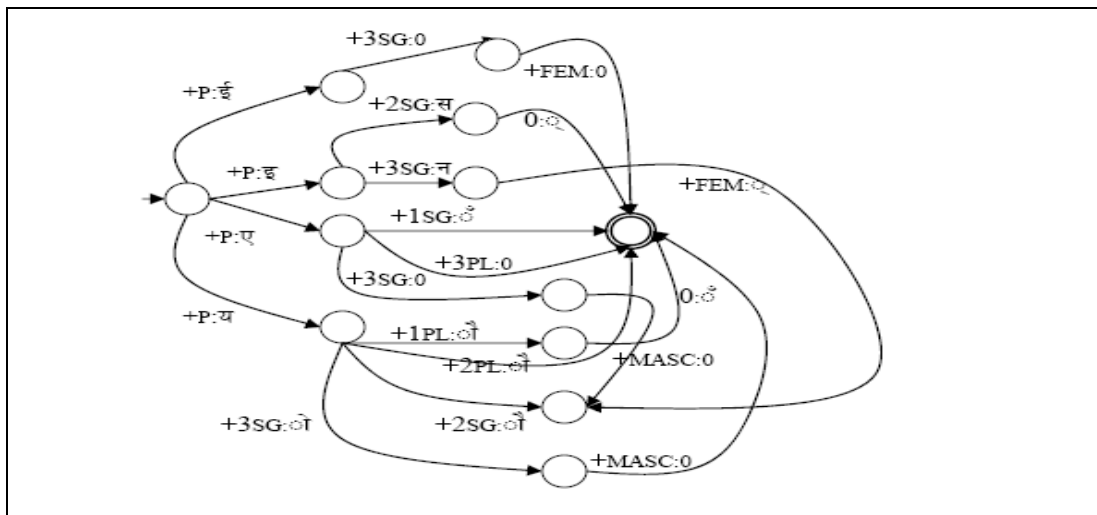


Figure 4.18: An Finite State Transducer for inflections of past tense (affirmative)

The phonological rules listed in PR 4.11 are directly encoded into the finite state transducer demonstrated in Figure 4.18.

Phonological rule

PR 4.11

- i. Independent vowels ए *e*, इ *i* and ई *i:* change to their corresponding dependent vowels ऐ *e*, ि *i* and ी *i:*, respectively if the verb stems end with consonants.

Regular expressions: ए -> ऐ || cons __;

इ -> ि || cons __;

ई -> ी || cons __;

In the past tense negative forms, the negative marker न् *n* '-NEG' is inserted between the past markers ए *e* or य् *j* or इ *i* and agreement inflections. In example (24), the negative maker न् *n* '-NEG' is in between past tense marker ए *e* and agreement inflection न् *n*. Altogether, there are eleven past tense negative forms as listed in Table 4.49 with their corresponding morphological tags.

(24) बच्चाहरू रोएनन् ।

batstsa-flaru: ro-enan
 child-PL cry-PST.NEG.3PL
 'The children did not cry.'

Table 4.49: Inflections for past tense (negative)

Grammatical category	Inflections	IPA	Tags
First person singular	इनँ	inã	PST.1SG
First person plural	एनौ	enãũ	PST.1PL
Second person singular	इनस्	inãs	PST.2SG
Second person singular hon	एनौ	enãu	PST.2SG.HON
Second person singular Female hon	इनौ	enãu	PST.NEG.2SG.FEM.HON
Second person plural	एनौ	enãu	PST.2PL
Third person masculine singular	एन	enã	PST.3SG.MASC
Third person feminine singular	इन	inã	PST.3SG.FEM
Third person masculine singular hon	एनन्	enãn	PST.3SG.MASC.HON
Third person feminine singular hon	इनन्	inãn	PST.3SG.FEM.HON
Third person plural	एनन्	enãn	PST.3PL

The finite state transducer in Figure 4.19 can analyze and generate the negative inflections of past tense. This transducer is also concatenated with a transducer of verb stems.

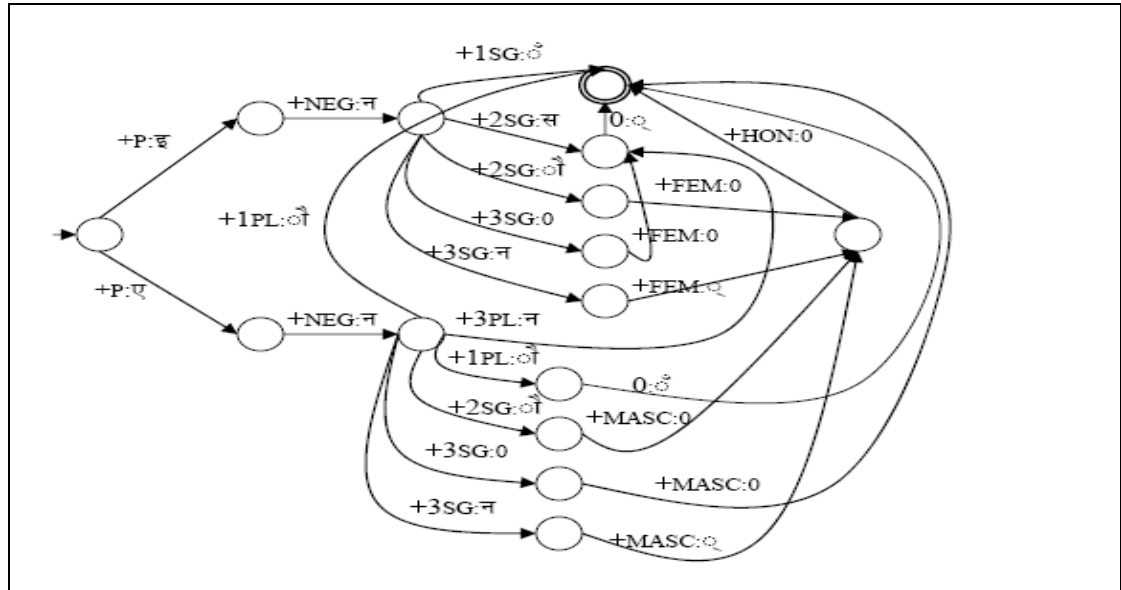


Figure 4.19: A finite state transducer for inflections of past tense (negative)

The phonological rules involved are listed in PR 4.12 have been directly encoded in the finite state transducer illustrated in Figure 4.19.

Phonological rules

PR 4.12

- i. Independent vowels ए, इ, and ई: change to their corresponding dependent vowels ऐ, ि, and ी: if the verb stems end with consonants.

Regular expressions: ए -> ऐ || cons __;

इ -> ि || cons __;

ई -> ी || cons __;

4.5.3 Aspects

The internal temporal orientation in a language is said to be aspect and this phenomena is expressed in Nepali morphologically through inflections. Traditionally four aspects, namely perfect, imperfect, habitual and inferential (unknown) aspect

(Pokharel 2054VS and Adhikari 2055VS) have been illustrated and discussed in the subsequent sections.

a. Perfect Aspect

In Nepali, the perfect aspect is indicated by a suffix -एको -eko 'PERF.SG.MASC'. The aspect marker inflects for number and gender. The verb form गरेको *gar-eko* 'do-PERF.SG.MASC' in (25a) is singular masculine, the verb form गरेका *gar-eka* 'do-PERF.PL' in (25b) is plural and the verb form गरेकी *gar-eki* 'do-PERF.SG.FEM' in (25c) is singular feminine. The inflections of perfect aspects are listed in Table 4.50.

(25) a. मैले यो काम गरेको छु ।

mɛi-le jo kam gar-eko ts^hu
 1SG.OBL-ERG this work do-PERF.SG.MASC be.NPST.1SG
 'I have done this work.'

b. हामीले यो काम गरेका छौं ।

ɦami-le jo kam gar-eka ts^hɛũ
 1PL-ERG this work do-PERF.PL be.NPST.1PL
 'We have done this work.'

c. सीताले यो काम गरेकी छे ।

si:ta-le jo kam gar-eki ts^he
 Sita.FEM-ERG this work do-PERF.SG.FEM be.NPST.3SG.FEM
 'Sita has done this work.'

Table 4.50: Inflections for perfect aspect

Grammatical category	Inflections	IPA	Tags
Perfect singular masculine	एको	eko	PERF.SG.MASC
Perfect plural	एका	eka	PERF.PL
Perfect singular feminine	एकी	eki:	PERF.SG.FEM
Perfect singular feminine Emphatic	एकै	ekɛi	PERF.SG.FEM.EMPH

The finite state transducer in Figure 4.20 can analyze and generate the inflections of perfect aspect. This transducer is also concatenated with the finite state transducer of verb stems.

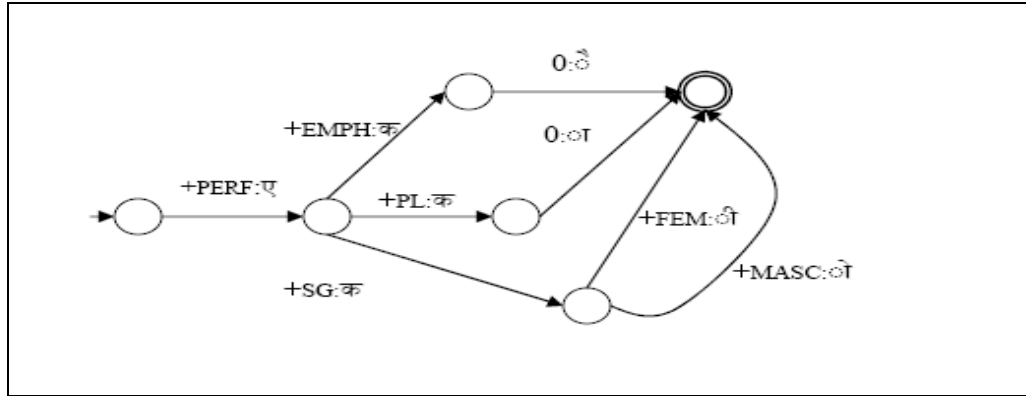


Figure 4.20: A finite state transducer for inflections of perfect aspect

The phonological rule listed in PR 4.13 has been compiled and composed with the transducer of Figure 4.20.

Phonological rules

PR 4.13

- i. Independent vowel ए *e* changes to its corresponding dependent vowel ऐ *e* if the verb stem ends with consonant.

Regular expression: ए → ऐ || cons __;

The perfect aspect negative forms are formed by prefixing the negative marker न *na* with the perfect aspect form of the verb stem as नगरेका *na-gar-eka* 'NEG-do-PERF.HON' in (26).

- (26) तिमीले यो काम नगरेका भए पैसा पाउँदैनौ ।
 timi:-le jo kam na-gar-eka b^hΛ-e pa:sa
 2SG.HON-ERG this work NEG-do-PERF.HON be-COND money
 pa-ũ-dΛinΛu
 get-NPST.NEG.2.HON
 'If you have not done this work, (you) won't get money.'

b. Imperfect Aspect

The imperfect aspect in Nepali is indicated by a marker -दै *-dai* '-IMPERF' as in (27) the verb form गर्दै *gar-dai* 'do-IMPERF' is imperfect aspect. The maker -दै *-dai* is

neutral with respect to number and gender. However, it has other three forms which distinguish between number and gender. All four imperfect aspect markers are listed in Table 4.51.³

(27) केटो यो काम गर्दै छ ।

keṭo jo kam gaɾ-dʌi tʂʰʌ
 boy.SG.MASC this work do-IMPERF be.NPST.3SG.MASC
 'The boy is doing this work.'

Table 4.51: Inflections for imperfect aspect

Grammatical category	Inflections		Tags
Imperfect singular masculine	दो	do	IMPERF.SG.MASC
Imperfect singular feminine	दी	di:	IMPERF.SG.FEM
Imperfect plural	दा	da	IMPERF.PL
Imperfect	दै	dʌi	IMPERF

The finite state transducer in Figure 4.21 can analyze and generate the inflections of imperfect aspect.

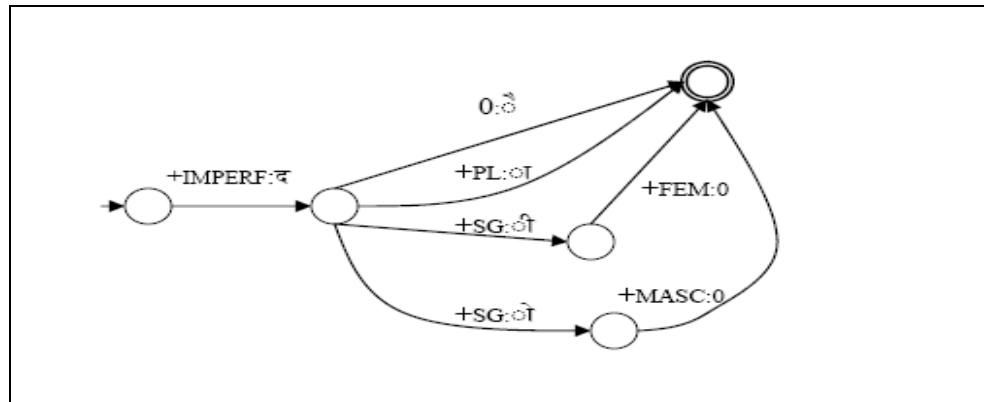


Figure 4.21 A finite state transducer for Inflections of imperfect aspect

The finite state transducer demonstrated in Figure 4.21 is composed with the finite state transducer of phonological rules listed in PR 4.14.

³ The suffixes -दो *do*, -दा *da* and -दी *di*: may have some other syntactic and semantic status, however, the computational purpose, they are treated as imperfect suffixes.

Phonological rules

PR 4.14

- i. Nasal ॐ is inserted if the verb stems end with vowels and the following affix begins with *ɖd*.

Regular expression: [.] -> ॐ || vowel __ ɖ;

The imperfect aspect negative forms are formed by prefixing the negative marker *न* *nA* 'NEG-' with the imperfect aspect form of the verb stem as *नखाँदै* *nA-k^hã-dAi* 'NEG-eat-IMPERF' in (28).

(28) चिया नखाँदै उहाँले फोन उठाउनुभयो ।

tsija nA-k^hã-dAi uɦã-le p^hon ut^h-au-nu b^h-jo
tea NEG-eat-IMPERF 3SG.HON-ERG phone lift-CAUSE-INF be.p-3SG
'He lifted the phone while not drinking tea.'

c. Habitual Aspect

There are two habitual aspects in the Nepali: present habitual and past habitual. The present habitual aspect is encoded by the non-past tense marker *छ* *ts^hA* and its inflections as *पिउँछु* *pi-ũ-t^hu* 'drink-φ-NPST.1SG' in (29a) whereas the past habitual aspect is indicated by a marker *थ* *t^h* plus inflections for agreement such as person, number, gender and honorificity. In example (29c), the verb form *पिउँथे* *pi-ũ-t^hẽ* 'drink-φ-HAB.PST.1SG' is in past habitual form in which *थ* *t^h* is past habitual marker which is followed by the agreement inflection. There are altogether ten past habitual forms of inflections. They are listed in Table 4.52.

(29) a. म धेरै रक्सि पिउँछु ।

mA d^herAi rAksi pi-ũ-t^hu
1SG more alcohol drink-φ-NPST.1SG
'I drink a lot of alcohol.'

b. म धेरै रक्सि पिउँथे ।

mA d^herAi rAksi pi-ũ-t^hẽ
1SG more alcohol drink-φ-HAB.PST.1SG
'I used to drink a lot of alcohol.'

Table 4.52: Inflections for past habitual aspect (Affirmative)

Grammatical category	Inflections	IPA	Tags
First person singular	थें	t ^h ẽ	PST.HAB.1SG
First person plural	थ्यौ	t ^h jʌũ	PST.HAB.1PL
Second person singular	थिस्	t ^h is	PST.HAB.2SG
Second person singular hon	थ्यौ	t ^h jʌu	PST.HAB.2SG.HON
Second person plural	थ्यौ	t ^h jʌu	PST.HAB.2PL
Third person masculine singular	थ्यो	t ^h jo	PST.HAB.3SG.MASC
Third person feminine singular	थि	t ^h i	PST.HAB.3SG.FEM
Third person masculine singular hon	थे	t ^h e	PST.HAB.3SG.MASC.HON
Third person feminine singular hon	थिन्	t ^h in	PST.HAB.3SG.FEM.HON
Third person plural	थे	t ^h e	PST.HAB.3PL

The finite state transducer in Figure 4.22 can analyze and generate the inflections of past habitual aspect. This finite state transducer is concatenated with the finite state transducer of the verb stems.

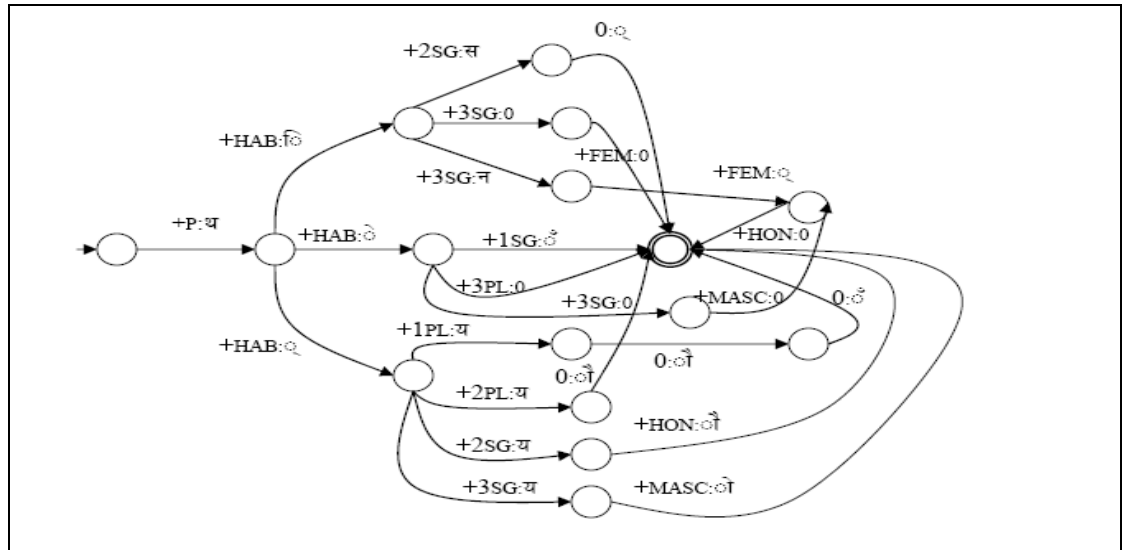


Figure 4.22: A finite state transducer for inflections of habitual aspect (affirmative)

The past habitual negative forms are formed by inserting the negative marker न *nA* between verb stems and past habitual maker थ् *t^h* plus agreement inflections. In this case, semantically null elements दै *dai* and दि *di* are inserted between the negative maker न *nA* '-NEG' and verb stem as पिउँदैनेथें *pi-ũ-dai nA-t^hẽ* 'drink-φ-HAB.NEG.PST.1SG'

in (30). There are altogether eleven past habitual negative forms and they are listed in Table 4.53.

30. म रक्सि पिउँदैनेथे ।

mΛ rΛksi pi-ũ-dΛinΛ-t^hẽ
 1SG alcohol drink-ϕ-HAB.NEG.PST.1SG
 'I did not used to drink alcohol.'

Table 4.53: Inflections for habitual aspect (negative)

Grammatical category	Inflections		Tags
First person singular	दैनेथे	dΛinΛt ^h ẽ	PST.NEG.HAB.1SG
First person plural	दैनेथ्यौ	dΛinΛt ^h jΛũ	PST.NEG.HAB.1PL
Second person singular	दैनेथिस्	dΛinΛt ^h is	PST.NEG.HAB.2SG
Second person singular hon	दैनेथ्यौ	dΛinΛt ^h jΛu	PST.NEG.HAB.2SG.HON
Second person plural	दैनेथ्यौ	dΛinΛt ^h jΛu	PST.NEG.HAB.2PL
Second person feminine singular	दिनेथिस्	dinΛt ^h is	PST.NEG.HAB.2SG.FEM
Second person feminine singular hon	दिनेथ्यौ	dinΛt ^h jΛu	PST.NEG.HAB.2SG.FEN.HON
Third person masculine singular	दैनेथ्यो	dΛit ^h jo	PST.NEG.HAB.3SG.MASC
Third person feminine singular	दिनेथिस्	dinΛt ^h is	PST.NEG.HAB.3SG.FEM
Third person masculine singular hon	दिनेथे	dinΛt ^h e	PST.NEG.HAB.3SG.MASC.HON
Third person plural	दैनेथे	dΛinΛt ^h e	PST.NEG.HAB.3PL

The finite state transducer in Figure 4.23 can analyze and generate the negative inflections of habitual aspect and it is also concatenated with the finite state transducer of the verb stems.

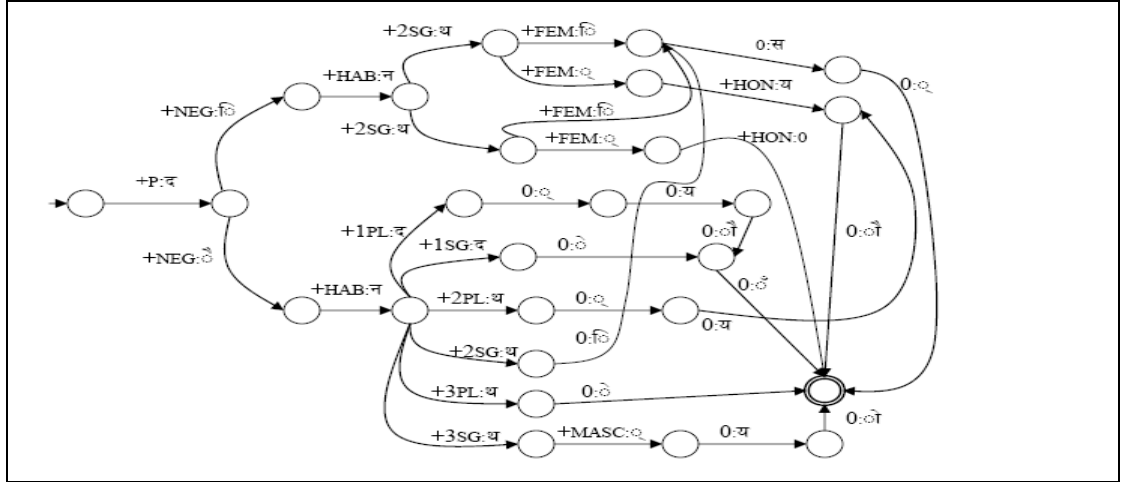


Figure 4.23: A finite state transducer for inflections of habitual aspect (negative)

d. Inferential (unknown) aspect

Inferential (unknown) aspect indicates the event that took place in past but it is known at present based on some evidence or clues. The inferential form of a verb is formed by inserting the inferential aspect marker $\text{ए} e$ and $\text{ई} i$ between the verb stems and non-past tense plus agreement inflections. The verb form $\text{सुतेछौ} sut-ets^{h\Lambda u}$ 'sleep-INFER.2.PL' in (31) is the inferential aspect form and Table 4.54 lists twelve inferential aspect inflections.

(31) तिमीहरू त हिजो बजारमा सुतेछौ ।

timi:-fi Λ ru: t Λ fidzo b Λ dzar-ma sut-ets $^{h\Lambda u}$
 2-PL PART yesterday market-LOC sleep-INFER.2.PL
 '(I came to know that) you slept in the market yesterday.'

Table 4.54: Inflections for inferential aspect (affirmative)

Grammatical category	Inflections	IPA	Tag
First person singular	एछु	ets ^h u	PST.INFER.1SG
First person plural	एछौ	ets ^h Λũ	PST.INFER.1PL
Second person masculine singular	एछस्	ets ^h Λs	PST.INFER.2SG.MASC
Second person feminine singular	इछस्	its ^h Λs	PST.INFER.2SG.FEM
Second person masculine singular hon	एछौ	ets ^h Λu	PST.INFER.2SG.MASC.HON
Second person feminine singular hon	इछौ	its ^h Λu	PST.INFER.2SG.FEM.HON
Second person plural	एछौ	ets ^h Λu	PST.INFER.2PL
Third person masculine singular	एछ	ets ^h Λ	PST.INFER.3SG.MASC
Third person feminine singular	इछ	its ^h Λ	PST.INFER.3SG.FEM
Third person masculine singular hon	एछन्	ets ^h Λn	PST.INFER.3SG.MASC.HON
Third person feminine singular hon	इछन्	its ^h Λn	PST.INFER.3SG.FEM.HON
Third person plural	एछन्	ets ^h Λn	PST.INFER.3PL

The finite state transducer in Figure 4.24 can analyze and generate the positive inflections of inferential aspect and it is concatenated with the finite state transducer of verb stems.

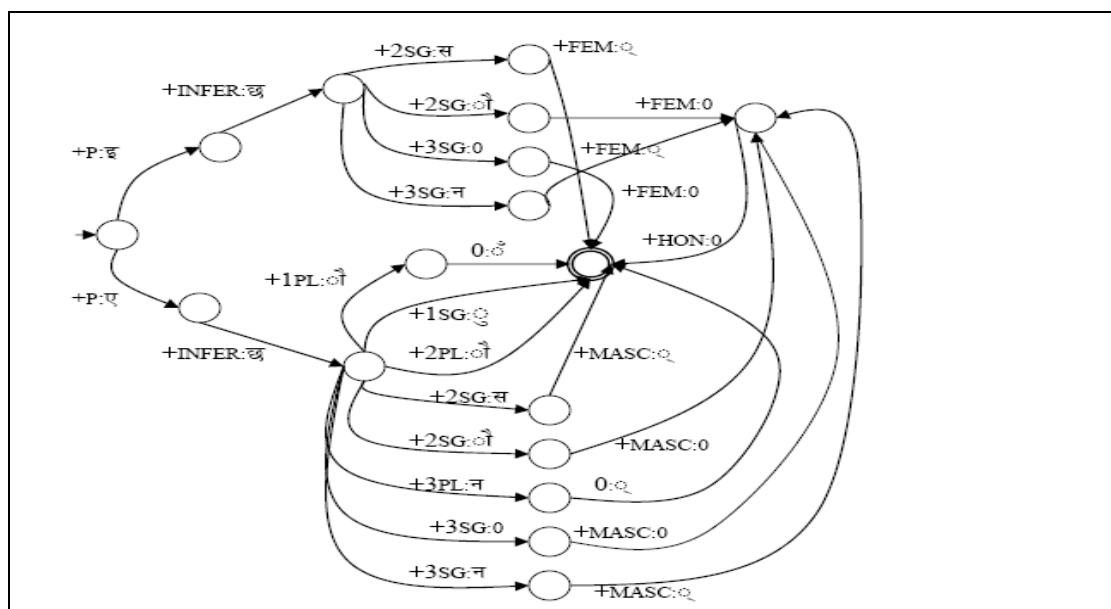


Figure 4.24: A finite state transducer for inflections of inferential aspect (affirmative)

The phonological rules listed PR 4.15 are compiled and composed with the transducer demonstrated in 4.24.

Phonological rules

PR 4.15

- i. Independent vowels ए *e* and इ *i* change to their corresponding dependent vowels े *e* and ि *i*, respectively if the verb stems end with consonants.

Regular expression: ए -> े || cons __;

इ -> ि || cons __;

The inferential aspect negative forms are formed by inserting the negative marker न *n* between inferential aspect maker ए *e* or इ *i* and agreement inflections as सुतेनछौ *sut-enAts^hau* 'sleep-INFER.NEG.2.PL' in (32). There are twelve inferential aspect negative forms which are listed in Table 4.55.

(32) तिमीहरू त हिजो बजारमा सुतेनछौ।

timi:-fiaru: t_Λ fiidzo bAdzar-ma sut-enAts^hau
 2-PL PART yesterday market-LOC sleep-INFER.NEG.2PL
 '(I came to know that) you did not sleep in the market yesterday.'

Table 4.55: Inflections for inferential aspect (negative)

Grammatical category	Inflections	IPA	Tags
First person singular	एनछु	enAts ^h u	PST.INFER.NEG.1SG
First person plural	एनछौ	enAts ^h au	PST.INFER.NEG.1PL
Second person masculine singular	एनछस्	enAts ^h As	PST.INFER.NEG.2SG.MASC
Second person feminine singular	इनछेस्	inAts ^h es	PST.INFER.NEG.2SG.FEM
Second person masculine singular hon	एनछौ	enAts ^h au	PST.INFER.NEG.2SG.MASC.HON
Second person feminine singular hon	इनछौ	inAts ^h au	PST.INFER.NEG.2SG.FEM.HON
Second person plural	एनछौ	enAts ^h au	PST.INFER.NEG.2PL
Third person masculine singular	एनछ	enAts ^h Λ	PST.INFER.NEG.3SG.MASC
Third person feminine singular	इनछ	inAts ^h Λ	PST.INFER.NEG.3SG.FEM
Third person masculine singular hon	एनछन्	enAts ^h Λn	PST.INFER.NEG.3SG.M.HON
Third person feminine singular hon	इनछन्	inAts ^h Λn	PST.INFER.NEG.3SG.FEM.HON
Third person plural	एनछन्	enAts ^h Λn	PST.INFER.NEG.3PL

The finite state transducer in Figure 4.25 can analyze and generate the negative inflections of inferential aspect. This transducer is also concatenated with transducer of verb stems.

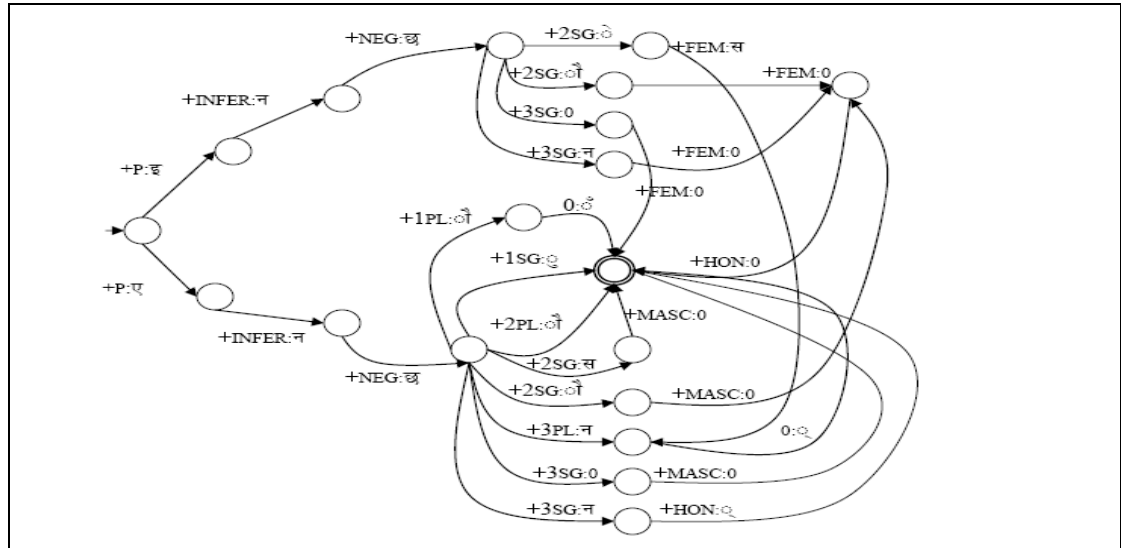


Figure 4.25: A finite state transducer for inflections of inferential aspect (negative)

The rules involved in this process are listed in PR 4.16 which are compiled and composed with finite state transducer illustrated in Figure 4.25.

Phonological rules

PR 4.16

- i. Independent vowels ए*e*, इ*i* and ई*i*: change to their corresponding dependent vowels ऐ*e*, ि*i* and ी*i*., respectively if the verb stems end with consonants.

Regular expressions: ए -> ऐ || cons __;

इ -> ि || cons __;

ई -> ी || cons __;

4.5.4 Moods

Morphologically, Nepali has two types of moods, namely, declarative and non-declarative. The former one does not have a distinct marker to indicate the mood, rather they are indicated by the default system; and the latter is further sub-divided

into imperative, optative and potential moods (Pokharel 2054VS). Each of them is indicated by their respective markers.

a. Imperative Mood

The imperative form of a verb has number and honorific distinctions. The base stem of the verb indicates singular non-honorific imperative form. Other two forms are singular honorific and plural. The imperative markers differ depending upon the end segment of the verb stems. The consonant ending verb stems take $-\phi$ for singular non-honorific and $\text{अ-}\Lambda$ for singular honorific and plural. In the case of vowel ending verb stems, i -ending verb stems take $-\text{ई-}i:$ for singular non-honorific and $-\text{अ-}\Lambda$ for singular honorific and plural; and other vowel ending take $-\phi$ for singular non-honorific, $-\text{ऊ-}u:$ for singular honorific and for plural. The imperative inflections are listed in Table 4.56. In example (33), the imperative verb form $\text{जाओ } dza-o$ 'go-IMP.2PL' indicates plural form for instance.

- (33) तिमीहरू बजारतिर जाओ!
 timi:-fiaru: badzar-tira dza-o!
 2-PL market-DIR go-IMP.2PL
 '(You) go towards the market.'

Table 4.56: Inflections for imperative mood

Grammatical Category	Inflections		Tag
Second person singular	$-\phi/-\text{ई}$	$-\phi/-i:$	IMP.2SG
Second person singular hon	$-\text{अ-}\text{ऊ}$	$-\Lambda/u:$	IMP.2SG.HON
Second person plural	$-\text{अ-}\text{ओ}$	$-\Lambda/-o$	IMP.2PL

The finite state transducer in Figure 4.26 can analyze and generate the inflections of imperative mood. This transducer is also concatenated with transducer of verb stems.

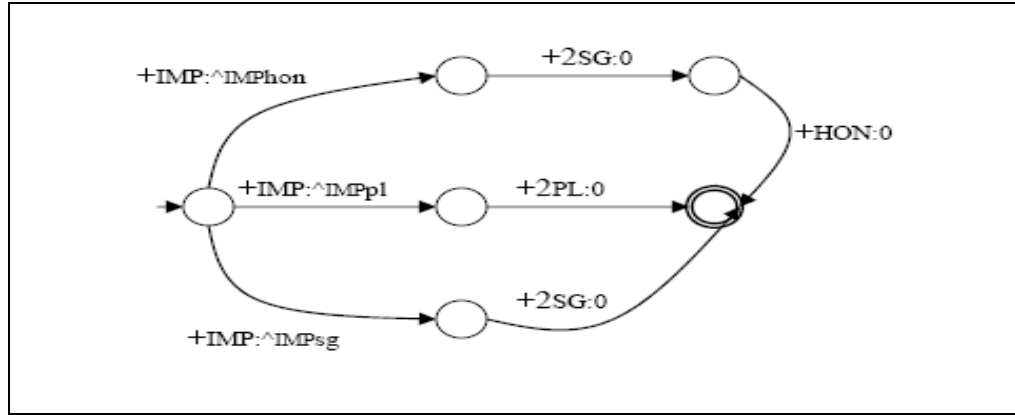


Figure 4.26: A finite state transducer for inflections of imperative mood

The phonological rules listed in PR 4.17 are compiled into a network and composed with the finite state transducer as demonstrated in Figure 4.26.

Phonological rules

PR 4.17

- i. ^IMPsg removed for non-honorific imperative

Regular expression: ^IMPsg -> [],

- ii. ^IMPpl changes to ऊ for honorific imperative after आ a ending verb stems

Regular expression: ^IMPpl -> ऊ || आ _ ,

- iii. ^IMPPhon changes to ओ for plural imperative after आ a ending verb stems

Regular expression: ^IMPpl -> ऊ || ओ _ ,⁴

Negative Imperative forms are obtained from prefixing the negative marker न nA 'NEG-' to the imperative form of the verb stems. The example sentence in (34) is negative sentence of example (33) in which form नजाओ nA-dza-o 'NEG-go-IMP.2PL' is negative imperative one.

(34) तिमीहरू बजारतिर नजाओ !

timi:-fiaru: bAdzar-tirA nA-dza-o!
 2-PL market-DIR NEG-go-IMP.2PL
 '(You) go towards the market.'

⁴ Arbitrary tags ^IMPsg and ^IMPpl are used for creating the environment and are finally eliminated from the network.

b. Optative Mood

The optative forms are obtained from the combination of verb stems and the optative inflections. The verb stems in optative mood inflect for person, number, gender and honorificity. The example in (35), the verb form *गरेस्* *gar-es* 'do-OPT.2SG' indicates second person singular optative form. There are altogether eight optative inflections and they are listed in Table 4.57.

- (35) ल परदेशमा राम्रोसँग काम गरेस् ।
 lA paɾades-ma ramro-sāṅa kam gaɾ-es
 PART foreign-LOC good-COM work do-OPT.2SG
 'I wish, (you) work nicely in foreign country.'

Table 4.57: Inflections for optative mood (affirmative)

Grammatical category	Inflections		Tags
First person singular	ऊँ	ũ	OPT.1SG
First person plural	औँ	Λũ	OPT.1PL
Second person singular	एस्	es	OPT.2SG
Second person singular hon	ए	e	OPT.2SG.HON
Second person plural	ए	e	OPT.2PL
Third person singular	ओस्	os	OPT.3SG
Third person singular hon	ऊन्	u:n	OPT.3SG.HON
Third person plural	ऊन्	u:n	OPT.3PL

The finite state transducer in Figure 4.27 can analyze and generate the inflections of optative mood. This transducer is also concatenated with transducer of verb stems.

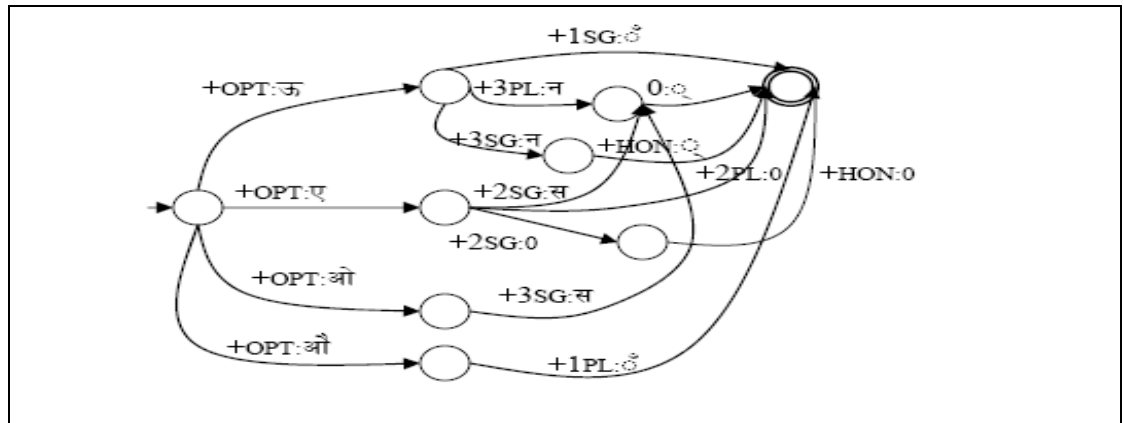


Figure 4.27: A finite state transducer for inflections of optative mood

The phonological rules listed in PR 4.18 are compiled and composed with the transducer illustrated in Figure 4.27.

Phonological rules

PR 4.18

- i. Independent vowels ए *e*, औ *au*, ओ *o* and ऊ *u*: change to their corresponding dependent vowels ऐ *e*, औ *au*, ओ *o* and ू *u*., respectively if the verb stems end with consonants.

Regular expression: ए -> ऐ || cons __;

औ -> औ || cons __;

ओ -> ओ || cons __;

ऊ -> ू || cons __;

The negative optative forms are obtained by prefixing the negative marker न *na* 'NEG-' to the optative forms of the verb. In the sentence (36) the form नगरोस् *na-gar-os* 'NEG-do-OPT.3SG' is the third person singular negative optative form.

36. उसले त्यस्तो काम नगरोस्।
 us-le tjas̥to kam na-gar-os
 3SG.OBL-ERG like that work NEG-do-OPT.3SG
 'I wish him not to do work like that.'

c. Potential Mood

Potential forms of the verbs are obtained from the combination of verb stems and potential inflections. The potential forms make the distinction on person, number, gender and honorificity. In example (37), the verb stem गर *gar* 'do' and potential mood marker -ल- *la* '-POT' form third person singular potential verb form. Altogether, there are twelve potential mood inflections and they are listed in Table 4.58.

(37) यो केटाले परिक्षा पास गर्ला ।

jo keṭa-le pʌriksja pas gʌr-la
 DEM.PROX boy-ERG examination pass do-POT.3SG
 'This boy may pass the examination.'

Table 4.58: Inflections for potential mood (affirmative)

Grammatical category	Inflections	IPA	Tag
First person singular	उँला	ũla	POT.1SG
First person plural	औंला	ʌũla	POT.1PL
Second person masculine singular	लास्	las	POT.2SG.MASC
Second person feminine singular	लिस्	lis	POT.2SG.FEM
Second person masculine singular hon	औला	ʌula	POT.2SG.MASC.HON
Second person feminine singular hon	औली	ʌuli	POT.2SG.FEM.HON
Second person plural	औला	ʌula	POT.2PL
Third person masculine singular	ला	la	POT.3SG.MASC
Third person feminine singular	ली	li	POT.3SG.FEM
Third person masculine singular hon	लान्	lan	POT.3SG.MASC.HON
Third person feminine singular hon	लिन्	lin	POT.3SG.FEM.HON
Third person plural	लान्	lan	POT.3PL

The finite state transducer in Figure 4.28 can analyze and generate the inflections of potential mood.

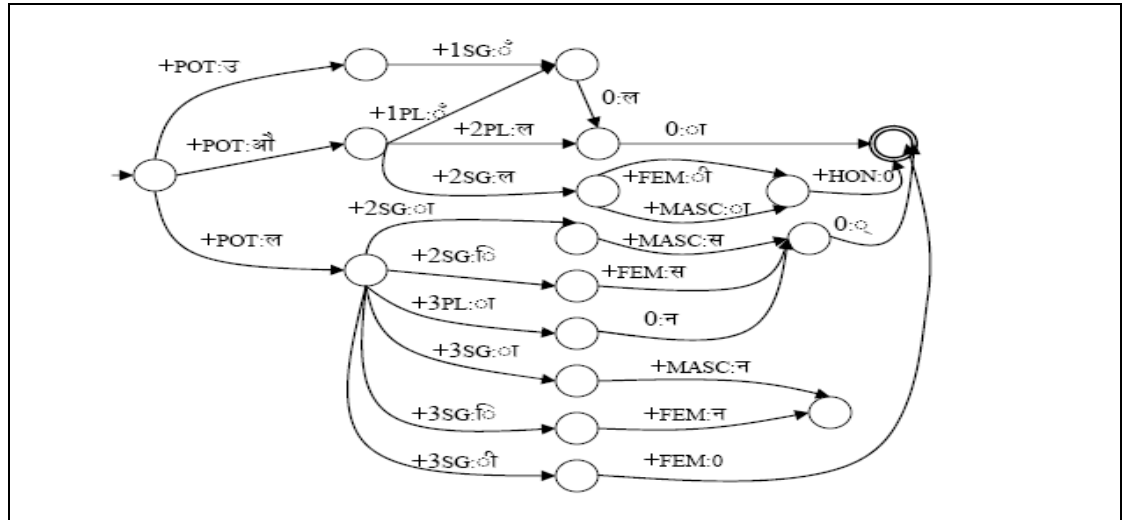


Figure 4.28: A finite state transducer for inflections of potential mood

The transducer presented in Figure 4.28 is composed with the network of the phonological rules in PR 4.19.

Phonological rules

PR 4.19

- i. Independent vowels उ *u* and औ *au* change to their corresponding dependent vowels ु *u* and ौ *au* if the verb stems end with consonants.

Regular Expression: उ -> ु || cons __;

औ -> ौ || cons __;

The negative potential mood forms are obtained from prefixing the negative marker न *na* 'NEG-' to the verb stems. The verb form नगर्ला *na-gar-la* 'NEG-do-POT.3SG' in (38) is third person negative potential form.

(38) यो केटाले परिक्षा पास नगर्ला ।

jo keṭa-le paṛiksja pas na-gar-la
DEM.PROX boy-ERG examination pass NEG-do-POT.3SG
'This boy may not pass the examination.'

4.5.5 Participial forms

a. Absolutive

The absolutive form is formed from the combination of verb stems and the absolutive marker ई *i:* '-ABS'. The absolutive form normally occurs with other forms of the verbs forming the compound verbs. In example (39) the verb form पठाई *paṭa-i:* 'send-ABS' is the absolutive form. The inflection for absolutive participle is listed in Table 4.59 with its morphological tag and its finite state transducer is demonstrated in Figure 4.29.

(39) उसले चिठी पठाई दियो ।

us-le tsitʰiṭʰ: paṭa-i: di-jo
3SG.OBL-ERG letter send-ABS give-PST.3SG
'He sent a letter.'

Table 4.59: Inflection for absolutive participle

Grammatical category	Inflections	IPA	Tags
Absolutive	ई	i:	ABS

The finite state transducer presented in Figure 4.29 can analyze and generate the absolutive form when it is concatenated with the finite state transducer of the verb stems.

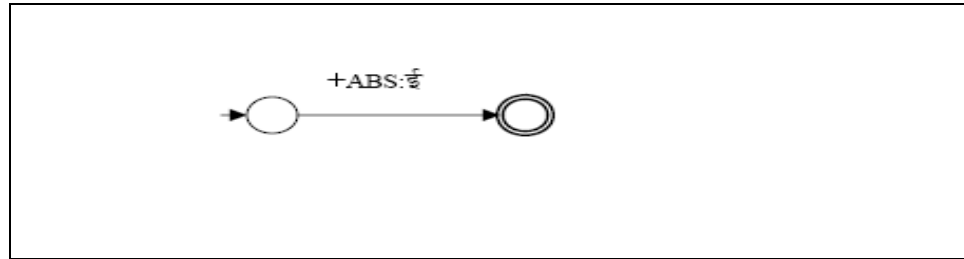


Figure 4.29: A finite state transducer for inflection of absolutive form

The phonological rules involved are compiled and composed with finite state transducer as demonstrated in Figure 4.29.

Phonological rules

PR 4.20

- i. Independent vowel ई *i*: changes to its corresponding dependent vowel ी *i*: if the verb stems end with consonants.

Regular expression: ई -> ी || cons __;

The negative absolutive form is obtained from prefixing the negative marker न न¹ 'NEG-' to the absolutive verb form. The verb form नपठाई *n¹-paṭ^ha-i*: 'NEG-send-ABS' in (40) is negative absolutive form.

(40) उसले चिठी नपठाई राख्यो ।

us-le tsit^ht^h: n¹-paṭ^ha-i: rak^h-jo
 3SG.OBL-ERG letter NEG-send-ABS keep-PST.3SG
 'He kept the letter without reading it.'

b. Infinitive

The infinitive form of a verb, in fact, is the dictionary entry in Nepali. It is obtained from suffixing an infinitive marker नु *-nu* '-INF' to the verb stem. In example (41), the verb form हिँड्नु *hiṅḍ-nu* 'walk-INF' is the infinitive form. The infinitive has three forms:

infinitive, oblique and emphatic. The infinitive is the default one with the marker नु -
nu '-INF', the oblique form with marker -ना -*na* occurs with case markers and the
 emphatic form with marker -नै -*nai* occurs at pragmatic level. The infinitive markers
 are listed in Table 4.60 with their morphological tags.

- (41) बिहान हिँड्नु राम्रो कुरो हो ।
 bifan fiḥd-nu ramro kuro ho
 morning walk-INF good.SG.MASC thing be.ID.NPST.3SG
 'To walk in the morning is a good thing.'

Table 4.60: Inflections for infinitive participle

Grammatical category	Inflections		Tags
Infinitive	नु	nu	INF
Infinitive Oblique	ना	na	INF.OBL
Infinitive Emphatic	नै	nai	INF.EMPH

The finite state transducer illustrated in Figure 4.30 can analyze and generate the
 infinite forms when it is concatenated with the finite state transducer of the verb
 stems.

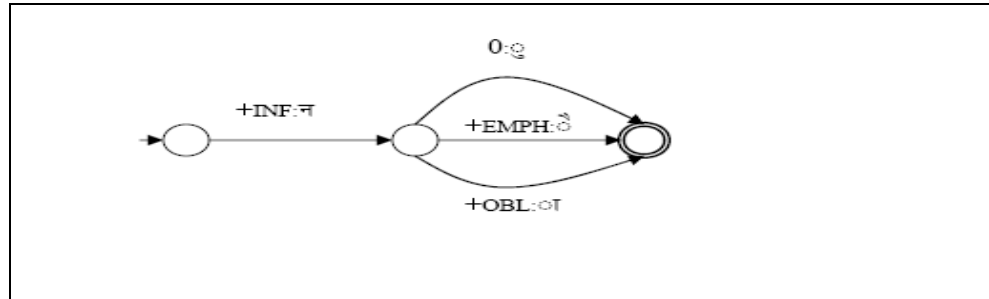


Figure 4.30: A finite state transducer for inflections of infinitive participial
 form

The negative infinitive form is obtained from prefixing the negative marker न- *na*-
 'NEG-' to the verb stems. The verb form नहिँड्नु *na-fiḥd-nu* 'NEG-walk-INF' in (42) is
 negative infinitive form.

- (42) बिहान नहिँड्नु राम्रो कुरो होइन ।
 bifan na-fiḥd-nu ramro kuro hoina
 morning NEG-walk-INF good.SG.MASC thing be.ID.NPST.NEG.3SG
 'Not to walk in the morning is not a good thing.'

c. Purposive

Purposive form of the verb is obtained from the combination of the verb stem and purposive marker -न -nΛ '-PURP'. The verb form कित्त kin-nΛ 'buy-PURP' in (43) is the purposive form. The purposive marker inflects for emphasis also. The inflections for purposive participle are listed in Table 4.61 with their morphological tags.

(43) म समान कित्त बजार जान्छु ।

mΛ saman kin-nΛ bΛdzar dza-n-tshu
 1SG thing buy-PURP market go-φ-NPST.1SG
 'I will go the market (in order) to buy things.'

Table 4.61: Inflections for purposive participle

Grammatical category	Inflections	IPA	Tags
Purposive	न	nΛ	PURP
Purposive emphasis	नै	nΛi	PURP.EMPH

The finite state transducer illustrated in Figure 4.31 encodes purposive participial inflections listed in Table 4.61 and it is capable of analyzing and generating the purposive forms when it is concatenated with the finite state transducer of the verb stems.

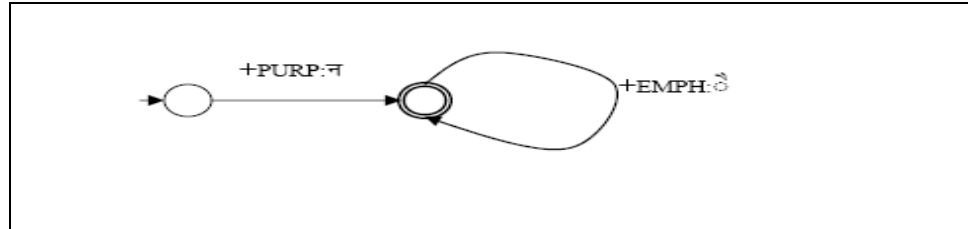


Figure 4.31: A finite state transducer for inflections of purposive participial form

d. Prospective

The prospective form of the verb is formed from the suffixation of the prospective marker -ने -ne with the verb stems. The verb form गुजार्ने gudzar-ne in (44) is the prospective form. The prospective inflection is listed in Table 4.62 with its morphological tag.

(44) उनिहरूले यो वर्ष पनि त्रिपालमुनि नै गुजार्ने छन् ।

uni:-ɦΛru:-le jo warsΛ pani tripalmuni nΛi gudzar-ne
 3.OBL-PL-ERG this year also tent PART spend-PROSP

ts^hʌn
 be.NPST.3PL
 'This year also, they will spend under the tent.'

Table 4.62: Inflection for prospective participle

Grammatical category	Inflections	IPA	Tags
Prospective	ने	ne	PROSP

The finite state transducer demonstrated in Figure 4.32 encodes the prospective participial forms and it can analyze and generate the prospective forms when it is concatenated with the finite state transducer of the verb stems.

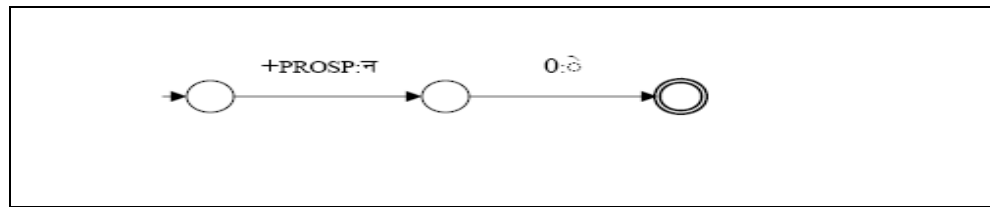


Figure 4.32: A finite state transducer for inflection of prospective participial form

e. Durative

The durative form is formed from the combination of verb stems and the durative markers *-दा -da*. In example (45), the verb *आउँदा a-ũ-da* form is the durative form which indicates the duration of the action. The durative form also inflects for emphasis. The inflections for durative participles are listed in Table 4.63 with their morphological tags.

- (45) उनिहरू आउँदा म सुतिरहेको थिएँ ।
 uni:-fiʌru: a-ũ-da mʌ sut-i-rʌfi-eko t^hiẽ
 3.OBL-PL come-ϕ-DUR 1SG sleep-ABS-remain-PERF.SG.M be.P1.SG
 'I was sleeping while they arrived.'

Table 4.63: Inflections for durative participle

Grammatical category	Inflections	IPA	Tags
Durative	दा	da	DUR
Durative emphatic	दै	dʌi	DUR.EMPH

The durative inflections listed in Table 4.63 are compiled into a network as illustrated in Figure 4.33 and it becomes capable of analyzing and generating the durative forms of the verbs when it is concatenated with the finite state transducer of verb stems.

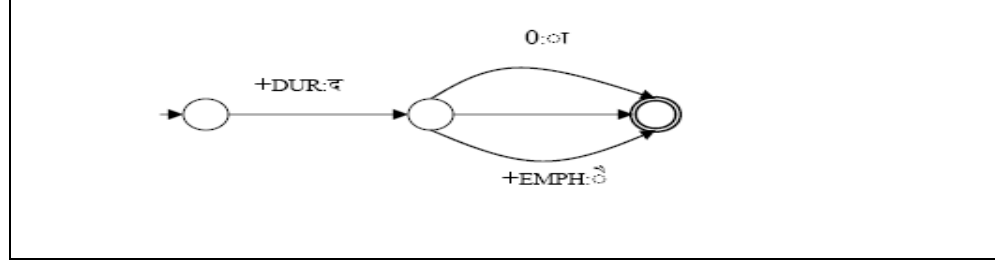


Figure 4.33 A finite state transducer for inflections of durative participial forms

f. Conjunctive

The conjunctive participle form of the verb is obtained from the combination of the verb stems and a conjunctive marker -एर $-er\Lambda$. In example (46), the पढेर $p\Lambda d^h-er\Lambda$ is the conjunctive participle form. This conjunctive form inflects for emphasis at pragmatic level with marker -ऐ Λi . There is another conjunctive marker -इकन $-i\Lambda kn\Lambda$ but it has low frequency use. The inflections for conjunctive participle with their corresponding morphological tags are listed in Table 4.64 with their morphological tags.

- (46) राम स्कूलमा पढेर आयो ।
 ram iskul-ma p $\Lambda d^h-er\Lambda$ a-jo
 Ram school-LOC read-CONJ come-PST.3SG.M
 'Ram studied in the school and came.'

Table 4.64: Inflections for conjunctive participle

Grammatical category	Inflections		Tags
Conjunctive	एर	$er\Lambda$	CONJ
Conjunctive emphasis	एरै	$er\Lambda i$	CONJ.EMPH
Conjunctive	इकन	$i\Lambda kn\Lambda$	CONJ
Conjunctive emphasis	इकनै	$i\Lambda kn\Lambda i$	CONJ.EMPH

The finite state transducer illustrated in Figure 4.34 encodes the conjunctive participial forms and it can analyze and generate them when it is concatenated with the finite state transducer of the verb stems.

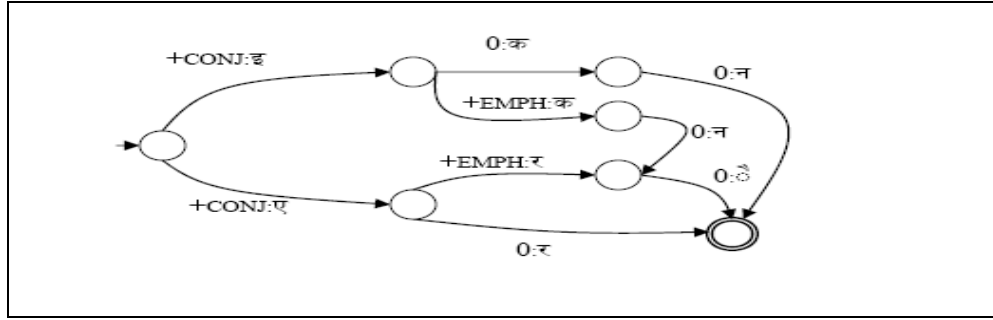


Figure 4.34: A finite state transducer for inflections of conjunctive participial form

The phonological rules involved in this process are listed in PR 4.21. They are compiled and composed with the finite state transducer illustrated in Figure 4.34.

Phonological rules

PR 4.21

- i. Independent vowels ए *e* and इ *i* change to their corresponding dependent vowels ऐ *e* and ि *i* if the verb stems end with consonants.

Regular expression: ए -> ऐ || cons __;

इ -> ि || cons __;

g. Conditional

The conditional form of the verb is obtained from the combination of the verb stems and the conditional marker ए *-e*. In example (47) गरे *gar-e* is the conditional form of the verb. The inflection for conditional participle is listed in Table 4.65 with its morphological tags.

(47) तिमिले सहयोग गरे म पास हुन्छु ।

timi:-le sahiAjog gaR-e ma pas fu-n-tshu
 2SG-ERG help do-COND 1SG pass be-φ-NPST.1SG
 'If you help I will pass.'

Table 4.65: Inflection for conditional participle

Grammatical category	Inflections		Tag
Conditional	ए	e	COND

The finite state transducer illustrated in Figure 4.35 encodes the conditional inflection and in association with transducer of verb stems, it can analyze and generate the conditional forms.

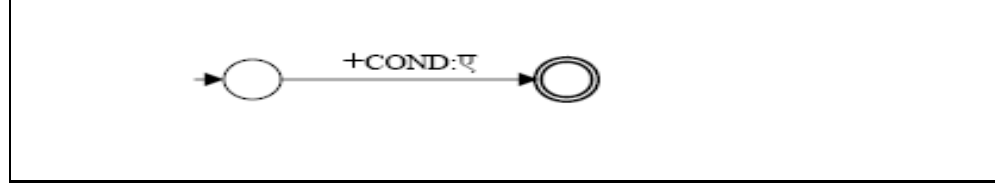


Figure 4.35: A finite state transducer for conditional participial form

The phonological rules in PR 4.22 are compiled and composed with the finite state transducer illustrated in Figure 4.35.

Phonological rules

PR 4.22

- i. Independent vowel ए *e* changes to its corresponding dependent vowel ऐ *e* if the verb stems end with consonants.

Regular expression: ए -> ऐ || cons __;

h. Perfective

The perfective form of the verb is obtained from the combination of the verb stems and the perfective marker ए *-e*. In example (48) गरे *gar-e* is the perfective form of the verb. But this form generally occurs with some postpositions. The inflection for perfective participle is listed in Table 4.65 with its morphological tag.⁵

- (48) तिमीले भनेअनुसार मैले काम गरे ।
 timi:-le b^han-e-anusar mai-le kam gar-ẽ
 2SG-ERG say-PERFT-POSTP 1SG.OBL-ERG work do-PST.1SG
 'I did the work as you said.'

Table 4.65: Inflection for conditional participle

Grammatical category	Inflection		Tag
Perfective	ए	e	PERFT

⁵ Pokharel (2054VS) has treated this form as oblique of past tense forms.

The finite state transducer illustrated in Figure 4.36 encodes the perfective inflection. It can analyze and generate the perfective forms of the verbs when concatenated with the finite state transducer of the verb stems.

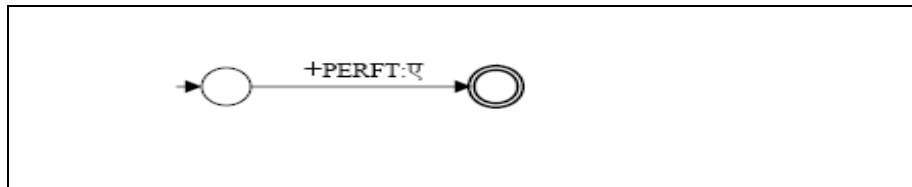


Figure 4.36: A finite state transducer for inflection of conditional participial form

The phonological rules listed in PR 4.23 are compiled and composed with the finite state transducer illustrated in Figure 4.36.

Phonological rules

PR 4.23

- i. Independent vowel ए changes to its corresponding dependent vowel ऐ if the verb stems end with consonants.

Regular expression: ए -> ऐ || cons __;

4.6 Summary

In this chapter, we presented the verb stems that are classified into two major groups: intransitive based stems and transitive based stems. Intransitive based stems are further grouped into five classes and transitive based stems are grouped into four classes. Each class is distinct at least in one feature discussed in the process of stem formation. And a set of few irregular verb stems have been discussed as a separate class.

The existential and identificational auxiliary verbs were discussed and illustrated under inflections as they more or less carry the features similar to the inflections. Inflectional paradigms for both affirmative and negative forms of tense, aspects, moods and participle are analyzed and illustrated. The finite state transducers of each type have been demonstrated. The phonological rules are identified, stated and expressed in regular expression format.

CHAPTER 5

ADVERBS, CONJUNCTIONS, POSTPOSITIONS AND PARTICLES

5.0 Outline

This chapter presents the analysis of closed words. It consists of four sections. Section 5.1 groups adverbs and presents them with morphological tags and corresponding finite state transducers. In section 5.2, we present conjunctions with their morphological tags and finite state transducers. Section 5.3 deals with postpositions, namely, plural marker, case markers and adverbial postpositions with their morphological tags. Section 5.4 discusses the particles and interjections and also presents morphological tags and finite state transducers. And section 5.5 finally summarizes the findings.

5.1 Adverbs in Nepali

Adverbs in Nepali indicate manner, place, time and intensity. They do not inflect for anything but appear with postpositions in writing (Adhikari 2062VS). They are not the obligatory elements in the sentence. However, they are classified into various groups based on their semantics for our purpose.¹

5.1.1 Temporal adverbs

Temporal adverbs are those adverbs which indicate time with respect to the action performed as आज *adza* 'today' in (1). Table 5.1 lists some temporal adverbs.

- (1) राम आज स्कुल गयो।
ram adza skul gA-jo
Ram today school go.PST-3SG.MASC
'Today, Ram went to school.'

¹ Though the classification of adverbs in Nepali into various classes is not computationally significance, it has been done simply for the identification.

Table 5.1: Temporal Adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+TEMP	अहिले	ʌhile	now
+ADV+TEMP	आज	adzʌ	today
+ADV+TEMP	अबेर	ʌberʌ	late
+ADV+TEMP	सोमबार	somʌbarʌ	Monday
+ADV+TEMP	हिउँद	hiũdʌ	winter

Since temporal adverbs do not inflect, the finite-state transducer demonstrated in Figure 5.1 is capable of analyzing and generating them.

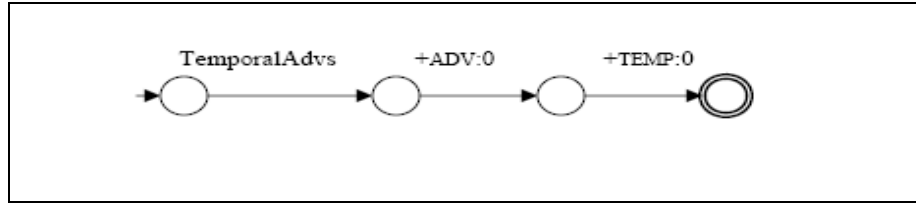


Figure 5.1: A finite state transducer for temporal adverbs

5.1.2 Spatial adverbs

Spatial adverbs are those adverbs which indicate place or location in the space where the action has taken place as निजकै *nʌdzik-ʌi* 'near-EMPH' in (2). Table 5.2 lists some of the spatial adverbs.

(2) मेरो घर निजकै मन्दिर छ ।

mero g^hʌr nʌdzik-ʌi mʌndir ts^hʌ
 1SG.GEN house near-EMPH temple be.NP.3SG.MASC
 'There is a temple near my house.'

Table 5.2: Spatial adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+SPAC	तल	tʌʌ	below
+ADV+SPAC	यहाँ	jʌhã	here
+ADV+SPAC	पछाडि	pʌtsaɖi	behind
+ADV+SPAC	भिन्न	b ^h itrʌ	inside
+ADV+SPAC	नजिक	nʌdzik	near

The finite state transducer illustrated in Figure 5.2 encodes the adverbs listed in Table 5.2 and it can analyze and generate them.

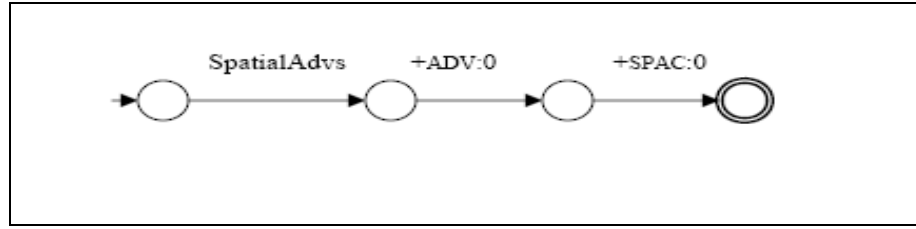


Figure 5.2: A finite state transducer for spatial adverbs

5.1.3 Amount adverbs

Amount adverbs are those words which indicate the amount of the head nouns it modifies as धेरै *derai* 'more' in (3). Table 3 lists some of the amount adverbs.

(3) रामसँग धेरै पैसा छ।

ram-sāṅgΛ derai pāisa ts^hΛ
 Ram-COM more money be.NP.3SG.MASC
 'Ram has a lot of money.'

Table 5.3: Amount adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+AMOUNT	धेरै	d ^h erai	more
+ADV+AMOUNT	थोरै	t ^h orai	less
+ADV+AMOUNT	अलिक	Λlik	little
+ADV+AMOUNT	त्यति	tjΛti	that much
+ADV+AMOUNT	कति	kΛti	how much

Since amount adverbs do not inflect, the finite-state transducer is simple and it is demonstrated in Figure 5.3 and it is capable of analyzing and generating them.

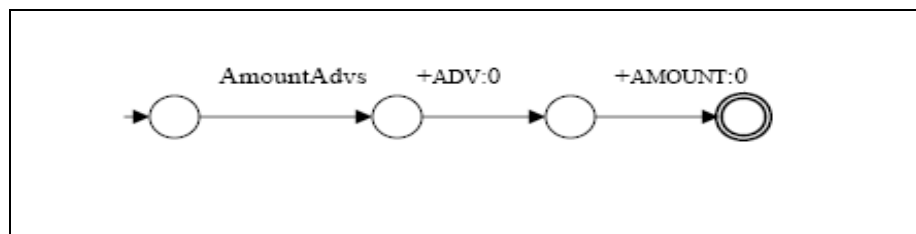


Figure 5.3: A finite state transducer for amount adverbs

5.1.4 Manner adverbs

Manner adverbs are those adverbs which indicate the ways or modes how the action has taken place as बिस्तारो *bistaro* 'slowly' in (4). Table 4 lists the some of the manner adverbs.

- (4) सीता बिस्तारो पढ़े।
 si:ta bistaro pʌdʰ-tsʰe
 Sita.FEM slowly read-NP.3SG.FEM
 'Sita slowly reads.'

Table 5.4: Manner Adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+MANNER	सुस्तरी	sustʌri	slowly
+ADV+MANNER	कसरी	kʌsʌri	how
+ADV+MANNER	सुटुक्क	suʈukkʌ	quietly
+ADV+MANNER	छिटो	tsiʈo	quickly
+ADV+MANNER	यसरी	jʌsʌri	this way

The finite state transducer demonstrated in Figure 5.4 encodes the manner adverbs listed in Table 5.4 and it can analyze and generate them.

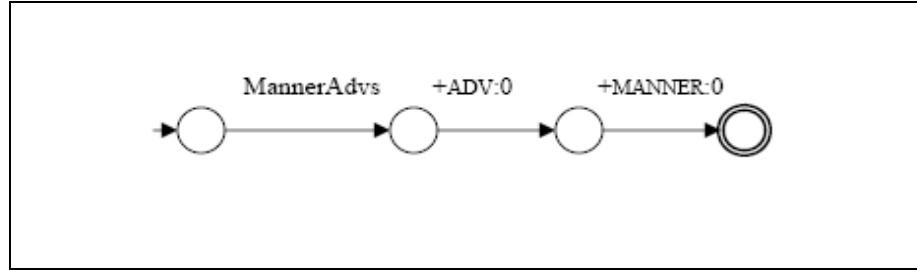


Figure 5.4 A finite state transducer for manner adverbs

5.1.5 Frequency adverbs

Frequency adverbs are those words which indicate the frequency of the action that is performed or takes place as कहिलेकाहीं *kʌhilekahĩ* 'sometimes' in (5). Table 5 lists some of the frequency adverbs.

- (5) हामी कहिलेकाहीं बजार जान्छौं।
 fiami: kʌhilekahĩ bʌdzar dza-n-tsʰʌũ
 1PL sometimes market go-NP.3PL
 'We sometimes go to market.'

Table 5.5: Frequency Adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+FREQ	कहिलेकाहीं	kʌhilekahĩ	sometimes
+ADV+FREQ	सधैँ	sʌdʰɛĩ	always
+ADV+FREQ	बारम्बार	barʌmbar	frequently
+ADV+FREQ	प्रायः	prajʌ	often
+ADV+FREQ	निरन्तर	nirʌntar	continuously

The finite-state transducer is simple since frequency adverbs do not inflect. It is demonstrated in Figure 5.5 and it is capable of analyzing and generating them.

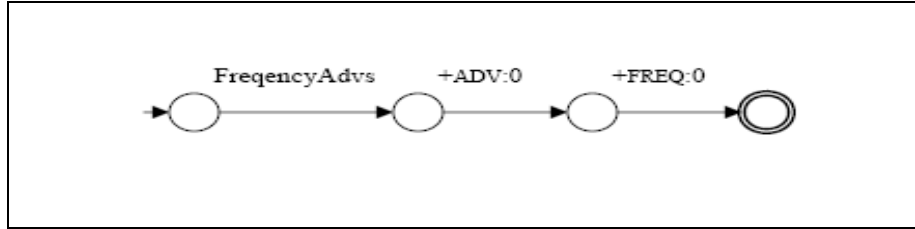


Figure 5.5: A finite state transducer for frequency adverbs

5.1.6 Reason adverbs

Reason adverbs are those which provide the reasons as त्यसकारण *tjʌskarʌŋ* 'therefore' in (6). Table 5.6 lists some of the reason adverbs.

(6) त्यसकारण म माथि दुवै ठाउँको प्रभाव छ ।

tjʌskarʌŋ mʌ matʰi duwʌi tʰaũko prʌbʰawʌ tsʰʌ
 therefore 1SG above both place-GEN influence be-NP.3SG.MASC
 'Therefore, I have the influences of both places.'

Table 5.6: Reason adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+REASON	त्यसकारण	tjʌsʌkarʌŋ	therefore
+ADV+REASON	फलस्वरूप	pʰʌʌswʌrup	as a result
+ADV+REASON	तसर्थ	tʌsʌrtʰʌ	thus

The reason adverbs listed in Table 5.6 are compiled into a finite state transducer as demonstrated in Figure 5.6 and it can analyze and generate them.

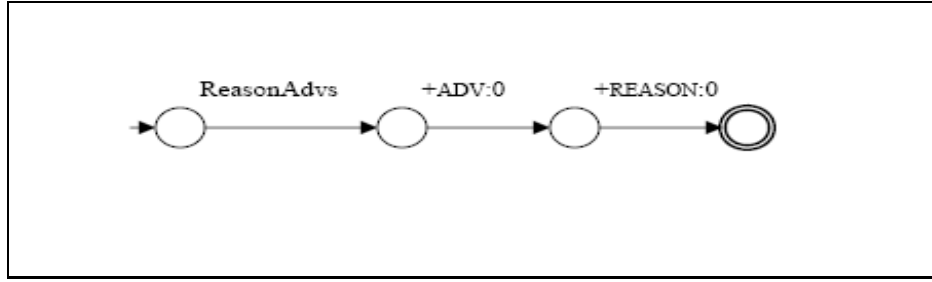


Figure 5.6: A finite state transducer for reason adverbs

5.1.7 Sentential adverbs

Sentential adverbs are those which modify the entire sentence as सायद *sajAdA*

'probably' in (7). Table 5.7 lists some of the sentential adverbs.

(7) सायद यति धेरै खुशीलाई मनमै अटाउन असमर्थ भएँ।

sajAdA jAti d^herAi k^husi-lai: maN-mAi Atauna ASAMart^h
 probably this much more happy-DAT heart-LOC.EMPH keep-INF unable
b^hAē

be-P.1SG

'Probably, (I) could not keep this much happiness within the heart.'

Table 5.7: Sentential adverbs

Morphological tags	Devanagari	IPA	Gloss
+ADV+SENT	सायद	sajAd	probably
+ADV+SENT	अवश्य	AWAʃjA	surely
+ADV+SENT	सामान्यतः	samanjAtA	generally
+ADV+SENT	यद्यपि	jAdjApi	however
+ADV+SENT	साँच्चै	sātstsAi	truly

The finite state transducer illustrated in Figure 5.7 encodes the sentential adverbs listed in Table 5.7 and it is capable of analyzing and generating them.

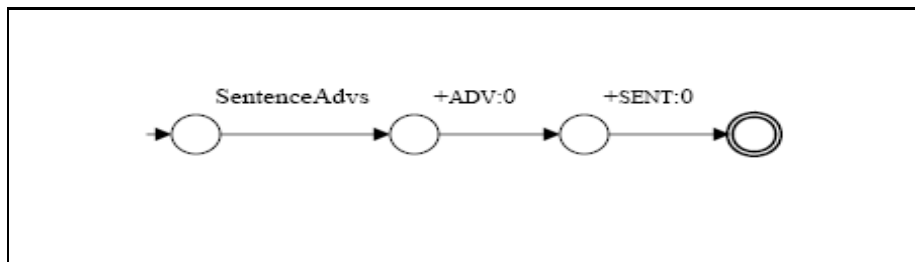


Figure 5.7: A finite state transducer for sentential adverbs

The individual finite-state transducers of the adverbs can be unioned together into a finite state transducer that can handle all the adverbs, that means, it can analyze and generate them.

5.2 Conjunctions in Nepali

Conjunctions in Nepali are of two kinds: coordinate and subordinate. The coordinate conjunctions are simple in their formation whereas the subordinate conjunctions are simple as well as compound types (Pokharel 2054VS; Adhikari 2062VS).

5.2.1 Coordinate conjunctions

Coordinate conjunctions join the constituents of the sentence having equal status as र *ra* 'and' in (8) it has joined कलम *kalam* 'pen' and किताब *kitab* 'book'. Some coordinate conjunctions in Nepali are listed in Table 5.8.

- (8) रामले कलम र किताब किन्यो ।
 ram-le kalam ra kitab kin-jo
 Ram-ERG pen and book buy-P.3SG.MASC
 'Ram bought a pen and a book.'

Table 5.8: Coordinate conjunctions

Morphological tags	Devanagari	IPA	Gloss
+CCONJ	र	ra	and
+CCONJ	वा	wa	or
+CCONJ	तर	taɾa	but
+CCONJ+EMPH	तरै	taɾai	but
+CCONJ	अनि	ani	and
+CCONJ	तथा	ta ^h a	and

The coordinators are compiled into a finite state transducer as demonstrated in Figure 5.8 and it can analyze and generate them.

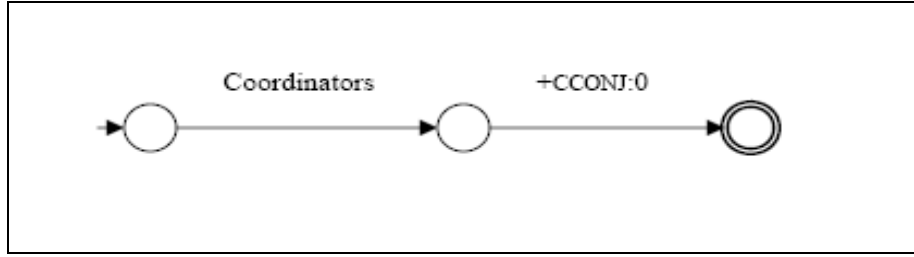


Figure 5.8: A finite state transducer for coordinate conjunctions

5.2.2 Subordinate conjunctions

Coordinate conjunctions join the constituents of the sentence that have unequal status, mainly subordinate clause with matrix clause as कि *ki* 'that' in (9). Some of subordinate conjunctions in Nepali are listed in Table 5.9 and the finite-state transducer to account them is demonstrated in Figure 5.9.

- (9) भण्डारीले भन्नु भयो कि उहाँले संग्राम किन छोड्नु भयो ।
 b^hʌndari-le b^hʌn-nu b^hʌ-jo ki uhā-le saŋgram kinʌ
 Bhandari-ERG say-INF be-P.3SG that 3SG.HON war why
 ts^hod-nu b^hʌ-jo
 leave-INF be-P.3SG
 'Bhandari said that why he gave up the war.'

Table 5.9: Subordinate conjunctions

Morphological Tags	Devanagari	IPA	Gloss
+SCONJ	किनकि	kinʌki	so that
+SCONJ	भने	b ^h ʌne	If
+SCONJ	कि	ki	That
+SCONJ	किनभने	kinʌb ^h ʌne	that's why
+SCONJ	यसकारण	jʌsʌkarʌŋ	Therefore

The subordinate conjunctions are compiled into a finite state transducer as illustrated in Figure 5.9 and it is capable of analyzing and generating them.

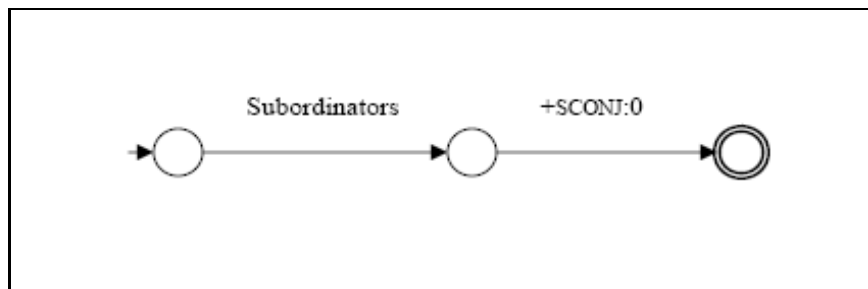


Figure 5.9: A finite state transducer for subordinate conjunctions

5.3 Postpositions in Nepali

Postpositions in Nepali always follow the nominal. According to their functions they perform in the sentence and their semantics, the postpositions can be grouped into three classes, namely, plural/collective marker, case markers and adverbial postpositions (Hardie et al 2005).

5.3.1 Plural/collective marker

-हरू *-haru* is the plural or collective marker in Nepali and it occurs optionally with *o*-ending nouns but systematically occurs with other non-*o*-ending nouns and pronouns. However, this is not an obligatory element to indicate the plural in nominals as there are other mechanisms to express the plural number in Nepali (see 3.1.1). The finite-state transducer is demonstrated in Figure 5.10.

Table 5.10: Collective/plural marker

Morphological Tags	Devanagari	IPA	Gloss
+PL	-हरू	ɦarɯ:	Plural

The plural marker has been compiled into a finite state transducer as illustrated in Figure 5.10 and it can analyze and generate it.

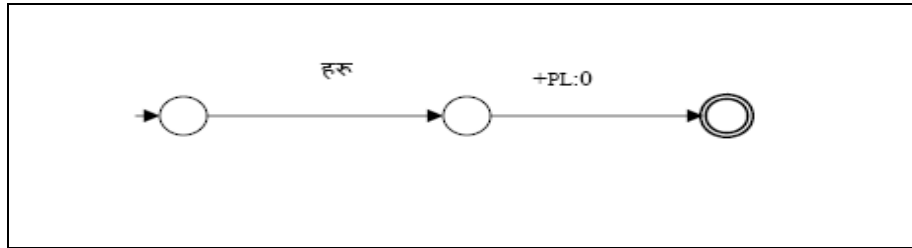


Figure 5.10: A finite state transducer plural/collective marker

5.3.2 Case markers in Nepali

Cases in Nepali morphology are marked with postpositions except nominative case and absolute case. In traditional Nepali grammars, the following cases are identified, namely, ergative, instrumental, dative, ablative, locative, commutative, genitive and allative. The ergative and instrumental cases have same marker -ले *-le*, ablative case is marked by two markers बाट *bat* and देखि *-dek^{hi}*,

commutative/associative case with two markers *सँग* -sãgʌ and *सित* -sitʌ and genitive case is marked with *को* -ko , *का* -ka and *की* -ki for singular, plural and feminine features. The allative case is marked with *तिर* -tirʌ (see 3.1.1). The case markers in Nepali are listed in Table 5.11a and Table 5.11b. The case markers in Table 5.11a do not take an emphatic marker whereas case markers in Table 5.11b take an emphatic marker.

Table 5.11a: Case marker postpositions (i)

Morphological tags	Devanagari	IPA	Gloss
+ERG	ले	-le	Ergative
+INST	ले	-le	Instrumental
+DAT	लाई	-lai:	Dative
+ABL	देखि	-dek ^{hi}	Ablative

They are separately compiled into a finite state transducer which can analyze and generate them.

Table 5.11b: Case marker postpositions (ii)

+ABL	बाट	-baʈ	Ablative
+ABL+EMPH	बाटै	-baʈʌi	Ablative
+LOC	मा	-ma	Locative
+LOC+EMPH	मै	-mʌi	Locative
+COM	सँग	-sãgʌ	Commutative
+COM+EMPH	सँगै	-sãgʌi	Commutative
+COM	सित	-sitʌ	Commutative
+COM+EMPH	सितै	-sitʌi	Commutative
+GEN+SG	को	-ko	Genitive
+GEN+PL	का	-ka	Genitive
+GEN+FEM	की	-ki	Genitive
+GEN+EMPH	कै	-kʌi	Genitive
+DIR	तिर	-tirʌ	Directional
+DIR+EMPH	तिरै	-tirʌi	Directional

The case markers listed in Table 5.11b also take emphatic marker. Therefore, they are compiled into a finite state transducer along with emphatic marker. It can analyze and generate them.

5.3.3 Adverbial postpositions

There are a number of forms which behave like postpositions but they also have the content meaning like that of adverbs. These forms usually occur with nominals providing the information about time, space, amount, frequency and manner. Some adverbial postpositions which do not take emphatic marker are listed in Table 5.12a.

Table 5.12a: Adverbial postpositions (a)

Morphological tags	Devanagari	IPA	Gloss
+POSTP	माथि	-mat ^h i	above
+POSTP	कहाँ	-kΛhã	where
+POSTP	मुनि	-muni	below, under
+POSTP	वारि	-wari	this side
+POSTP	पारि	-pari	other side
+POSTP	वरि	-wΛri	this side
+POSTP	परि	-pΛri	other side
+POSTP	पट्टि	-pΛtti	towards
+POSTP	प्रति	-prΛti	towards
+POSTP	पालि	-pali	time
+POSTP	खेरि	-k ^h eri	while doing
+POSTP	पालि	-pali	time
+POSTP	छेउ	-ts ^h eu	at the side
+POSTP	पछि	-pΛts ^h i	later
+POSTP	पछाडि	-pΛts ^h aɖi	back side
+POSTP	अघि	-Λg ^h i	before
+POSTP	अगाडि	-Λgaɖi	before
+POSTP	भरि	-b ^h Λri	full of
+POSTP	पहिले	-pΛhile	before
+POSTP	निम्ति	-nimti	for
+POSTP	लागि	-lagi	for
+POSTP	बारे	-bare	about
+POSTP	सामु	-samu	before
+POSTP	सरि	-sΛri	like
+POSTP	मध्ये	-mΛd ^h je	among
+POSTP	जति	-dzΛti	whatever
+POSTP	पिच्छे	-pitsts ^h e	each
+POSTP	पालि	-pali	time

The finite-state transducer is demonstrated in Figure 5.12a encodes the adverbial postpositions listed in Table 5.12a which is capable of analyzing and generating them.

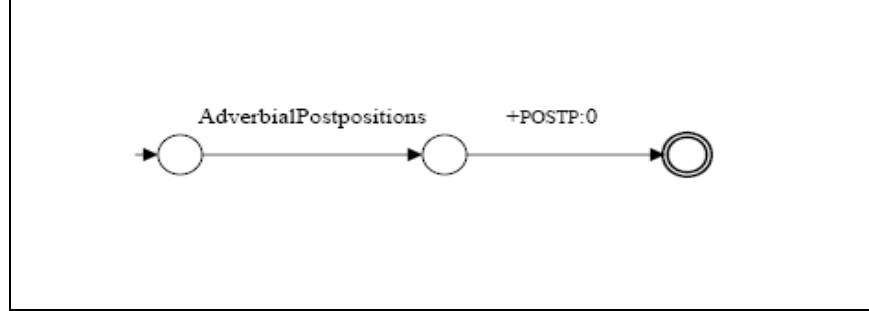


Figure 5.12a: A finite state transducer for adverbial postpositions that do not take emphatic marker

Some postpositions take emphatic marker, they are listed in Table 5.12b.

Table 5.12b: Adverbial postpositions (b)

Morphological tags	Devanagari	IPA	Gloss
+POSTP	सहित	-sΛhitΛ	along with
+POSTP	साथ	-sat ^h Λ	with
+POSTP	सम्म	-sΛmmΛ	until, upto
+POSTP	बाहिर	-bafirΛ	out side
+POSTP	वार	-warΛ	
+POSTP	पार	-parΛ	
+POSTP	वर	-warΛ	little near
+POSTP	पर	-pΛrΛ	little further
+POSTP	उँभो	-ũb ^h o	up
+POSTP	उँधो	-ũd ^h o	down
+POSTP	तर्फ	-tΛrp ^h Λ	towards
+POSTP	नेर	-nerΛ	near
+POSTP	निर	-nirΛ	near
+POSTP	समक्ष	-sΛmkΛk	before
+POSTP	पर्यन्त	-pΛrjΛntΛ	till
+POSTP	खेर	-k ^h erΛ	moment
+POSTP	उप्रान्त	-uprantΛ	then after
+POSTP	साथ	-sat ^h Λ	with
+POSTP	पख	-pΛk ^h Λ	time
+POSTP	ताक	-takΛ	time
+POSTP	ताका	-taka	time

+POSTP	पाला	-pala	time
+POSTP	पटक	-paṭaka	times
+POSTP	पल्ट	-paṭa	times
+POSTP	पश्चात्	-paṣṭat	after
+POSTP	छेक	-ts ^h ek	at the time
+POSTP	भिन्न	-b ^h itr	inside
+POSTP	नजिक	-naḍzika	near
+POSTP	भर	-b ^h ara	full of
+POSTP	तक	-taka	till
+POSTP	यता	-jata	here
+POSTP	उता	-uta	there
+POSTP	बीच	-bi:tsa	between
+POSTP	निमित्त	-nimitta	for the sake of
+POSTP	खातिर	-k ^h atir	for
+POSTP	अन्तर्गत	-antargata	within
+POSTP	बमोजिम	-bomodzima	according to
+POSTP	माफिक	-map ^h ika	according to
+POSTP	मुताबिक	-mutabika	according to
+POSTP	अनुसार	-anusara	according to
+POSTP	उपर	-upara	on
+POSTP	मार्फत	-marph ^h ata	via
+POSTP	अलावा	-alawa	beside
+POSTP	अतिरिक्त	-atirikta	in addition
+POSTP	बाहेक	-baheka	except
+POSTP	जस्तो	-dzasto	like
+POSTP	सरह	-saraha	same as
+POSTP	बाबजुद	-babadzuda	
+POSTP	विरुद्ध	-wirudd ^h	against
+POSTP	बापत	-bapata	for
+POSTP	सट्टा	-saṭṭa	instead of
+POSTP	बदला	-badala	instead of
+POSTP	लेखा	-lek ^h a	
+POSTP	सुद्ध	-sudda	
+POSTP	समेत	-sameta	along with

The adverbial postpositions which take emphatic marker listed in Table 5.12b are compiled into a network as demonstrated in Figure 5.12b which can analyze and generate them.

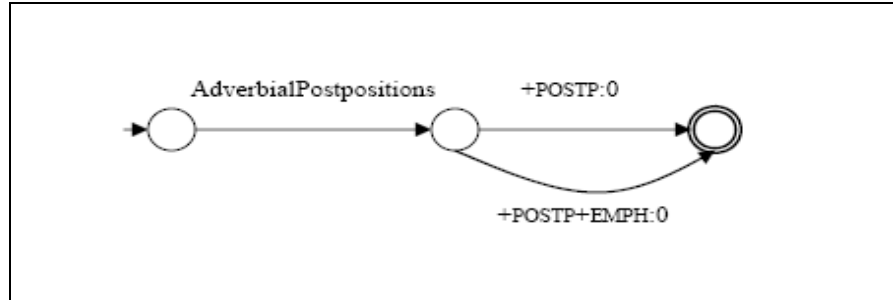


Figure 5.12b: A finite state transducer for adverbial postpositions that take emphatic marker

Phonological Rule:

PR 5.1

1. Vowels $\overset{\circ}{a}$, $\overset{\circ}{o}$ and halant $\overset{\circ}{\text{्}}$ at the end of the adverbial postpositions of this group are removed when the emphatic marker $\overset{\circ}{\text{ः}}$ is attached.

Regular expression $\overset{\circ}{a}|\overset{\circ}{o}|\overset{\circ}{\text{्}} \rightarrow [] \quad || \text{ cons } _ \overset{\circ}{\text{ः}} \#;$

5.4 Particles and interjections in Nepali

5.4.1 Particles

In Nepali, particles are the residual class comprising those stems which do not enter into inflectional constructions and stand as free forms (Dahal 1974). They appear before or after any lexical word and add an abstract meaning to the word that they are associated with. The extra meaning added can only be predicted from the context where they are used. The Nepali particles are monosyllabic or disyllabic words and their behaviors are different from other indeclinable words such as adverbs, postpositions and conjunctions. In written Nepali, the particles are written separately. Some particles are listed in Table 5.13 with morphological tags.

Table 5.13: Particles in Nepali

Morphological tags	Devanagari	IPA
+PARTICLE	नै	nɛi
+PARTICLE	मात्र	matrʌ
+PARTICLE	चाहिँ	tsahĩ
+PARTICLE	पनि	pʌni
+PARTICLE	ल	lʌ
+PARTICLE	है	hɛi
+PARTICLE	न	nʌ
+PARTICLE	नि	ni
+PARTICLE	त	tʌ
+PARTICLE	पो	pɔ
+PARTICLE	र	rʌ

The finite state transducer as illustrated in Figure 5.13 encodes the particles listed in Table 5.13 and it can analyze and generate them.

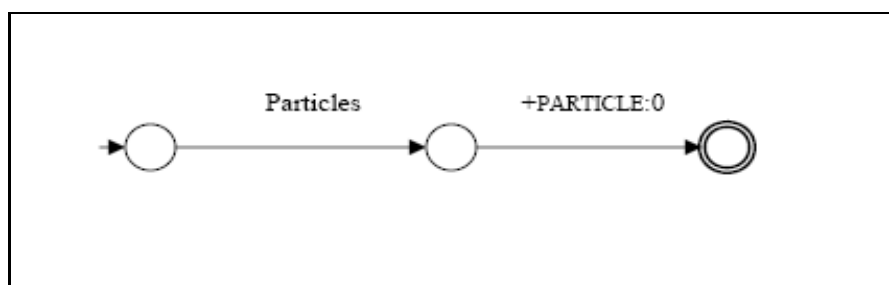


Figure 5.13: A finite state transducer for particles

5.4.2 Emphatic markers

a. ऐ *ai* ! emphasizes or draws an attention to or focuses a sentence or a part of the sentence. In addition, the techniques such as (a) sentence stress, (b) use of particles, (c) dislocation of the sentence constituents and (d) intonation are used for emphasizing the sentence or part of the sentence. The technique (a & d) are phonological and technique (b) is syntactic where the particles follow the word that gets focused. The technique (c) involves topicalization and some sorts of movements. The use of emphatic marker ऐ *ai* is not restricted to a particular class of words. Except some phonological constraints, it gets attached to any words irrespective of its parts of

speech. So, this marker can be said as global marker (Pokharel 2053VS). The following are the conditions where, this emphatic marker can or cannot appear.

- i. It doesn't appear with इ *i*, ए *e* and उ *u* (except आफू *ap^hu*) ending words.
- ii. When it gets attached to the words ending with Λ , *a* and *o*, these vowels get deleted as illustrated below.

म $m\Lambda$ 'I'	$m\Lambda + \Lambda i$	= $m\Lambda i$	मै
केटो <i>keto</i> 'boy'	<i>keto</i> + Λi	= $ket\Lambda i$	केटै
राजा <i>radza</i> 'king'	<i>radza</i> + Λi	= $radz\Lambda i$	राजै

- iii. But, with the words ending with consonants, this marker gets attached without any change as given below.

रूख <i>ru:k^h</i> 'tree'	<i>ru:k^h</i> + Λi	= $ru:kh\Lambda i$	रूखै
किताब <i>kitab</i> 'book'	<i>kitab</i> + Λi	= $kitab\Lambda i$	किताबै

In contrast, when this marker gets attached to the adjective, it deemphasizes the attribute possessed by adjectives.

(10) रामको बानी राम्रै छ ।

ram-ko bani ramr Λ i ts^h Λ
 Ram-GEN habit good.DEMPH be.NP.3SG.MASC
 'Ram's habit is okay (lesser than good).'

b. ही *hi* emphasizes only some of the demonstratives. But, the same marker with interrogative pronouns changes it to the indefinite pronouns. The examples given below demonstrate this phenomenon.

- यो 'this' + ही = यही 'this-EMPH'
 त्यो 'that' + ही = त्यही 'that-EMPH'
 ऊ 'this' + ही = उही 'this-EMPH'
 सो 'the same' + ही = सोही 'the same-EMPH'
 को 'who' + ही = कोही 'who-INDEF'
 के 'what' + ही = केही 'what-INDEF'

5.4.3 Interjections in Nepali

Interjections are a subset of particles which generally appear at the sentence initial position and express speaker's emotions such as surprise, pain, disgust, joy,

excitement and enthusiasm (Pokharel 2054VS:125-31). Interjections are also indeclinable words. Some examples of interjections are listed in Table 5.14 with morphological tags.

Table 5.14: Interjections in Nepali

Morphological Tags	Devanagari	IPA
+INTERJ	ओहो	oho
+INTERJ	आत्था	at ^h a
+INTERJ	छि	ts ^h i
+INTERJ	थुइक्क	t ^h uikkʌ
+INTERJ	हरे	hʌre
+INTERJ	बिचरा	bitsʌra
+INTERJ	ज्या	dzja
+INTERJ	ए	e
+INTERJ	ज्यू	dzju
+INTERJ	नाई	naĩ
+INTERJ	कुन्नि	kunni

The interjections listed in Table 5.14 are compiled into a finite state transducer as demonstrated in Figure 5.14 and it is capable of analyzing and generating them.

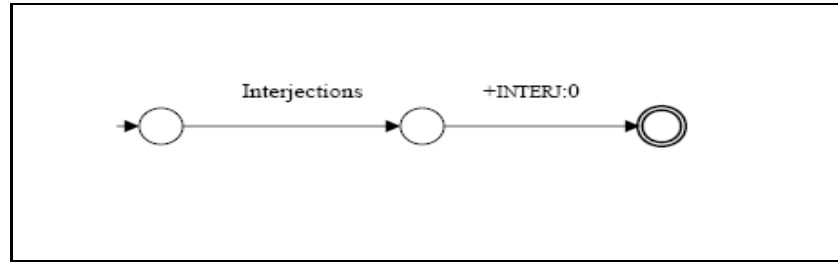


Figure 5.14: A finite state transducer for interjections

5.5 Summary

In this chapter, we discussed and presented the grouping of adverbs, conjunctions, case markers, particles and interjections. Adverbs in Nepali are grouped into seven semantic classes: temporal, spatial, amount, frequency, manner, reason and sentential. Conjunctions are of two types: coordinate and subordinate. Postpositions are of three types: plural marker, case markers and adverbial postpositions. Particles in general are in the single class but two emphatic markers which can be applied globally except in the phonologically constrained case is equivalent to some particles. The interjections form a single class.

CHAPTER 6

DERIVATIONAL MORPHOLOGY

6.0 Outline

This chapter presents analysis of derivational processes in Nepali morphology. It consists of three sections. Section 6.1 presents the prefixation that includes noun to noun derivation, noun to adjective derivation, noun to adverb derivation and adjective to adjective derivation. In each types, a model for the finite state transducer and a tag for prefixaton is provided. In section 6.2, we present the suffixation process that includes noun to noun derivation, noun to noun adjective derivation, noun to noun/adjective derivation, adjective to noun derivation, adjective/noun to noun derivation, verb to noun derivation, verb to adjective derivation, verb to adverb derivation, adverb to adjective derivation, verb to noun conversion, verb to adjective/noun conversion and verb to noun derivation. In each type of derivation, the morphological tags and finite state transducers for each group are illustrated. And finally, section 6.3 summarizes the chapter.

6.1 Prefixation

Derivation is a morphological process of word formation. It involves the additon of bound affix forms to an existing lexeme/stem, whereby the addition of the affix derives a new word or a lexeme (Katamba 1993; Payne 1997). The word class of newly formed word is generally different from the original word from which it is derived. Sometimes, this not always true, i.e. the word class remains the same, however, the semantics of the word definitely changes. The meaning of a derived new word may have clear meaning change, addition of speciality, technicality and stylistics. The derivation of a word from the same word class and from different class is possible. In most of the prefixing derivation, the word class remains same except few cases whereas in suffixing derivation the word class changes except few cases (Adhikari 2062VS).

In prefixation, an affix is prefixed to a base stem and a new word is derived. In Nepali, a number of prefixes as listed in the Tables (6.1-6.4) with base stem derived a

word. The various types of derivation using prefixes are discussed in subsequent sections.

6.1.2 Noun to noun derivation

In this type of derivation, 24 prefixes are involved and they derive a noun from a noun stem. The semantics of the prefixes is not predictable, so they are simply marked as prefix with a tag PFX. Table 6.1 lists those prefixes with base stem and derived word.

Table 6.1: Noun to noun derivation

Prefix	Base noun stem	Gloss	Derived noun	Gloss
प्र pra-	चलन tsalAn	tradition	प्रचलन pratsalAn	Tradition
परा para-	जय dzAjA	victory	पराजय paradzAjA	Defeat
अप aPa-	शब्द sAbdA	word	अपशब्द aPa sAbdA	abusing word
सम् sam-	मान man	honor	सम्मान samman	respect
अनु anu-	शासन sAsAn	governance	अनुशासन anu sAsAn	discipline
अव awa-	गुण guṇ	quality	अवगुण awaguṇ	demerit
दुस् dus-	परिणाम pariṇam	result	दुस्परिणाम duspariṇam	bad result
दुर् dur-	घटना g ^h AtAna	event	दुर्घटना durg ^h AtAna	accident
वि wi-	नाश naś	loss	विनाश winaś	damage
अधि ad ^{hi} -	राज्य radzjA	state	अधिराज्य ad ^{hi} radzjA	kingdom
अति ati-	वृष्टि wristi	rain	अतिवृष्टि atiwristi	over rain
अभि ab ^{hi} -	रुचि rutsi	interest	अभिरुचि ab ^{hi} rutsi	interest
प्रति prati-	ध्वनि d ^h wAni	sound	प्रतिध्वनि prati d ^h wAni	echo
परि pari-	योजना jodzAna	plan	परियोजना pari jodzAna	project
उप upa-	ग्रह graha	planet	उपग्रह upagraha	satalite
सह saha-	कार्य karjA	work	सहकार्य saha karjA	collaboration
स sa-	परिवार pariwar	family	सपरिवार sa pariwar	whole family
कु ku-	पुत्र putra	son	कुपुत्र kuputra	bad son
अ a-	ज्ञान gjan	knowledge	अज्ञान agjan	ignorance
अन् an-	आस्था ast ^h a	belief	अनास्था anast ^h a	disbelief
बे be-	इज्जत idzdzAt	prestige	बेइज्जत beidzdzAt	insult
बद badA-	नाम nam	name	बदनाम badAnam	bad name
ला la-	वारिस waris	heir	लावारिस lawaris	???
सु su-	समाचार samatsar	news	सुसमाचार susamatsar	good news

The finite-state transducer in Figure 6.1 is common to all the prefixes listed in Table 6.1. It can analyze and generate both base and derived word.

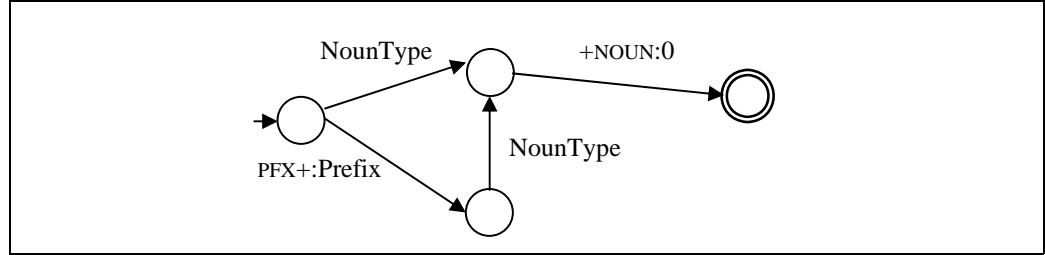


Figure 6.1: A finite state transducer for noun to noun derivation

6.1.3 Noun to adjective derivation

In this type of derivation, 9 prefixes are involved and they derive an adjective from a noun stem. The semantics of the prefixes is not predictable, so they are simply marked as prefix with a tag PFX. Table 6.2 lists those prefixes with base stem and derived word.

Table 6.2: Noun to adjective derivation

Prefix	Base noun stem	Gloss	Derived adjective	Gloss
निर् nir-	दोष dos	fault	निर्दोष nirdos	innocent
निः ni:-	स्वार्थ swarth ^h Λ	self-interest	निःस्वार्थ ni:swarth ^h Λ	selfless
नि ni-	डर dΛr	fear	निडर niḍΛr	bold
वि wi-	मुख muk ^h	mouth	विमुख wimuk ^h	deviated
निस् nis-	फल p ^h Λl	fruit	निस्फल nisp ^h Λl	fruitless
स sΛ-	बल bΛl	force	सबल sΛbΛl	capable
बे be-	घर g ^h Λr	house	बेघर beg ^h Λr	homeless
अ Λ-	मूल्य mu:ljΛ	value	अमूल्य Λmu:ljΛ	valueless
अन ΛnΛ-	मोल mol	price	अनमोल ΛnΛmol	priceless

The finite-state transducer in Figure 6.2 is common to all the prefixes listed in Table 6.2 and it is capable of analyzing and generating them.

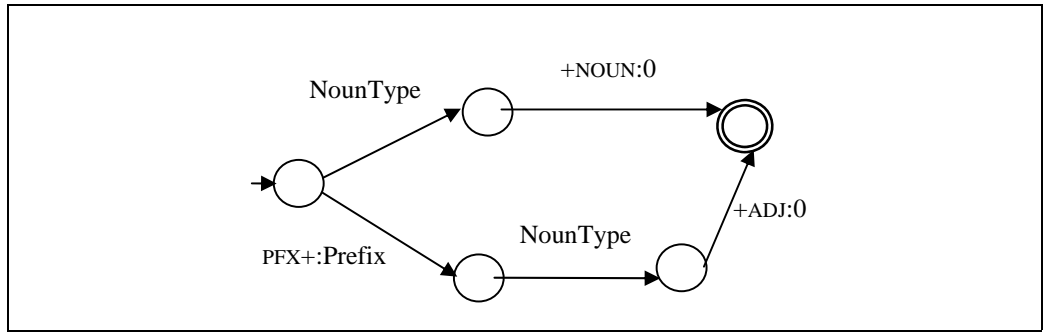


Figure 6.2: A finite state transducer for noun to adjective derivation

6.1.4 Noun to adverb derivation

In this type of derivation, 4 prefixes are involved and they derive an adverb from a noun stem. The semantics of the prefixes is not predictable, so they are simply marked as prefix with a tag PFX. Table 6.3 lists those prefixes with example base stem and derived word.

Table 6.3: Noun to adverb derivation

Prefix	Base Noun stem	Gloss	Derived Adverb	Gloss
आ <i>a-</i>	मरण $m\Lambda r\Lambda n$	death	आमरण $a\Lambda m\Lambda r\Lambda n$	till death
स <i>sa</i>	हर्ष $h\Lambda r\text{ṣ}\Lambda$	happy	सहर्ष $s\Lambda h\Lambda r\text{ṣ}\Lambda$	with happiness
निर् <i>nir-</i>	घात $g^h\text{at}$	stroke	निर्घात $nirg^h\text{at}$	severly
प्रति <i>prati-</i>	हसा $h\Lambda p\text{ta}$	week	प्रतिहसा $pr\Lambda tih\Lambda p\text{ta}$	per week

The finite-state transducer in Figure 6.3 is common to all the prefixes listed in Table 6.3 and it can analyze and generate them.

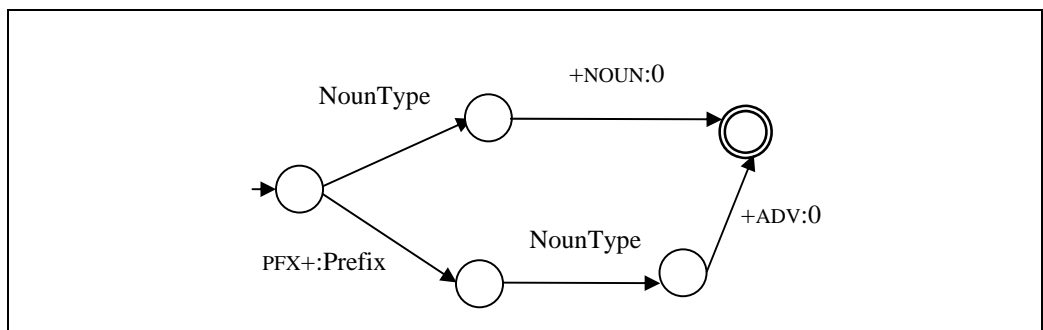


Figure 6.3: A finite state transducer for noun to adverb derivation

6.1.5 Adjective to adjective derivation

In this derivation, 6 prefixes are involved and derive an adjective from an adjective stem. The semantics of the prefixes is not consistent, so they are simply marked as prefix with a tag PFX. Table 6.4 lists those prefixes with base stem and derived word.

Table 6.4: Adjective to adjective derivation

Prefix	Base adjective stem	Gloss	Derived adjective	Gloss
सम् $_{SAM}$ -	पूर्ण $pu:r\eta\Lambda$	complete	सम्पूर्ण $s\Lambda mpu:r\eta\Lambda$	total
वि $_{wi}$ -	शुद्ध $\int udd^h\Lambda$	pure	विशुद्ध $wi\int udd^h\Lambda$	pure
दुर् $_{dur}$ -	भेद्य $b^h edj\Lambda$	vulnerable	दुर्भेद्य $durb^h edj\Lambda$	invulnerable
उन् $_{un}$ -	मुक्त $mukt\Lambda$	free	उन्मुक्त $unmukt\Lambda$	free
सु $_{su}$ -	शिक्षित $\int i kts^h it$	educated	सुशिक्षित $s\int i kts^h it$	educated
परि $_{pari}$ -	पूर्ण $pu:r\eta\Lambda$	complete	परिपूर्ण $p\Lambda r i pu:r\eta\Lambda$	sufficient

The finite-state transducer in Figure 6.4 is common to all the prefixes listed in Table 6.4 and it is capable of analyzing and generating.

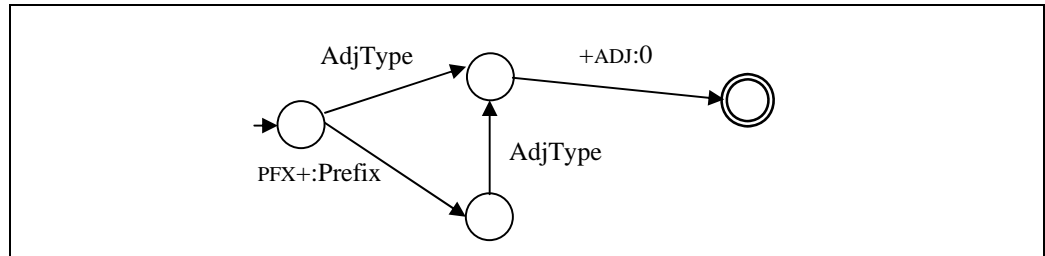


Figure 6.4: A finite state transducer for adjective to adjective derivation

6.2 Suffixation

6.2.1 Noun to noun derivation

In this derivation, 2 suffixes are involved and they derive a noun from a noun stem. The semantics of the suffixes is not considered, so they are simply marked as suffix with a tag SFX. Table 6.5 lists those suffixes with base stem and derived word.

Table 6.5: Noun to noun derivation

Base noun stem	Gloss	Suffix	Derived noun	Gloss
सुन sun	gold	आर- ar	सुनार $sunar$	goldsmith
घाँस $g^h\tilde{a}s$	grass	ई- i :	घाँसी $g^h\tilde{a}si$:	grass cutter

The finite-state transducer in Figure 6.5 is common to all the suffixes listed in Table 6.5 and it is capable of analyzing and generating them.

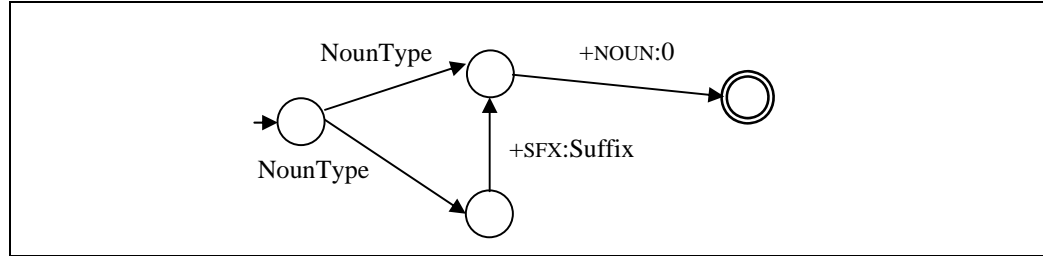


Figure 6.5: A finite state transducer for noun to noun derivation

The phonological rules involved in the derivation are listed in PR 6.1. They are compiled and composed with the finite state transducer illustrated in Figure 6.5.

Phonological rule

PR 6.1

- i. Independent vowels आ and ई change to corresponding dependent vowels ा and ी, respectively after the consonants

Regular expressions: आ -> ा || cons __

ई -> ी || cons __

6.2.2 Noun to adjective derivation

In this derivation, 11 suffixes are involved and they derive an adjective from a noun stem. The semantics of the prefixes is not considered, so they are simply marked as suffix with a tag SFX. Table 6.6 lists those suffixes with example of base stem derived word.

Table 6.6: Noun to adjective derivation

Base noun stem	Gloss	Suffix	Derived adjective	Gloss
दया dʌja	love	अनीय -ʌnijʌ	दयनीय dʌjanijʌ	lovable
लाभ lab ^h	profit	अक -ʌkʌ	लाभक lab ^h ʌkʌ	profitable
सेवा sewa	service	इका-ika	सेविका sewika	service girl
मुगल mugʌl	Mugal	आन -an	मुगलान mugʌlan	Indian
लिम्बु limbu	Limbu	वान -wan	लिम्बुवान limbuwan	Limbu area
दान dan	donation	ई i:	दानी dani:	donor
खर्च k ^h ʌrtsʌ	expense	आलु-alu	खर्चालु k ^h ʌrtsalu	expensive
भिर b ^h ir	cliff	आलो-alo	भिरालो b ^h iralo	steep
रिस ris	anger	आहा-aha	रिसाहा risaha	angry
शहर ʃʌɦʌr	town	इया-ija	शहरिया ʃʌɦʌrija	urban
होस fios	sense	इयार-ijar	होसियार fiosijar	careful

The finite-state transducer illustrated in Figure 6.6 is common to all the suffixes listed in Table 6.6.

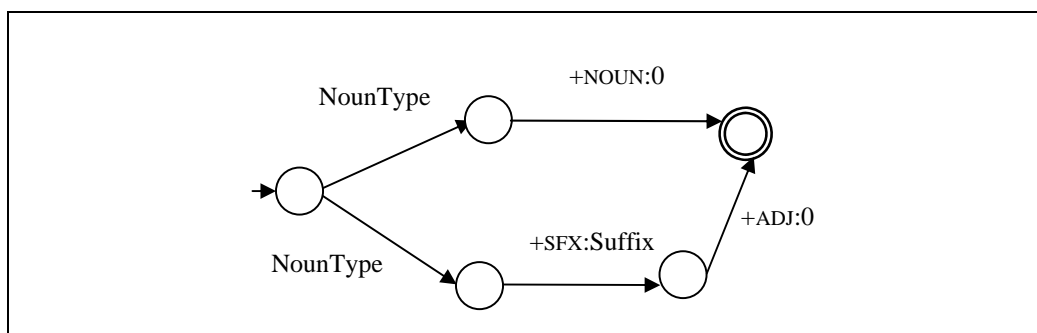


Figure 6.6 A finite state transducer for noun to adjective derivation

The phonological rules involved in this derivation are listed in PR 6.2. They are compiled and composed with the finite state transducer demonstrated in Figure 6.6 which is capable of analyzing and generating them.

Phonological rule

PR 6.2

- i. Independent vowels आ a, ई i: and इ i change to their corresponding dependent vowel ा a, ी i: and ि i after the consonants

Regular expressions: आ -> ा || cons __

ई -> ी || cons __

इ -> ि || cons __

ii. Vowel sequence of dependent ाa and अA changes to अA.

Regular expression: ा अ -> []

iii. Vowel sequence of dependent ाa and independent ईi changes to िi.

Regular expression: ा इ -> ि []

6.2.3 Noun to noun/adjective derivation

In this derivation, 3 suffixes are involved and they derive a noun or adjective from a noun stem. The semantics of the prefixes is not considered, so they are simply marked as suffix with a tag SFX. Table 6.7 lists those suffixes with example of base stem and derived word.

Table 6.7: Noun to noun/adjective derivation

Base noun stem	Gloss	Suffix	Derived noun/adjective	Gloss
झापा dz ^h apa	Jhapa	ली-li	झापली dz ^h apali	of Jhapa
गुल्मी gulmi	Gulmi	एली-eli	गुल्मेली gulumeli	of Gulmi
इलाम ilam	Illam	ए-e	इलामे ilame	of Illam
गाउँ gaū	village	ले-le	गाउँले gaūle	villager
नेपाल nepal	Nepali	ई-i:	नेपाली nepali:	of Nepal

The finite-state transducer in Figure 6.7 is common to all the suffixes listed in Table 6.7.

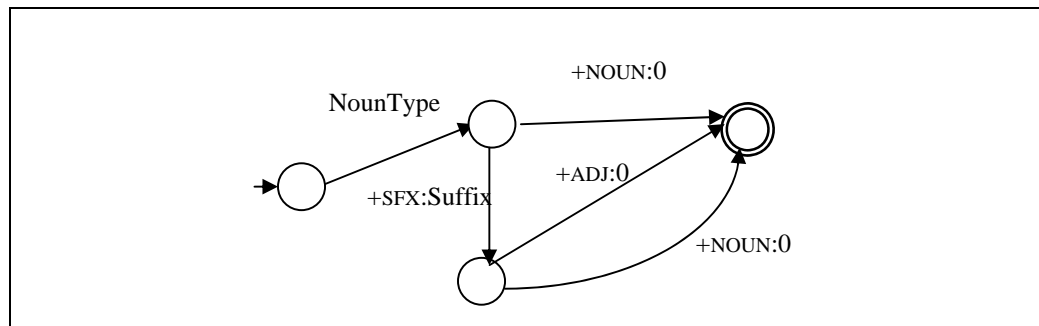


Figure 6.7: A finite state transducer for noun to noun/adjective derivation

The phonological rules involved in this derivation are listed in PR 6.3; they are compiled and composed with the finite state transducer illustrated in Figure 6.7 and it can analyze and generate both stem and derived words.

Phonological rule:

PR 6.3

i. Independent vowels ई*i*: and ए*e* change to their corresponding dependent vowel

ी*i*: and े*e* after the consonants

Regular expressions: ई -> ी || cons __ ;

ए -> े || cons __;

ii. Vowel sequence of dependent vowel ी*i*: and independent vowel ए*e* changes

to े*e*.

Regular expression: ी ए -> े;

iii. Vowel sequence of dependent vowel ा*a* and independent इ*i* changes to ि*i*.

Regular expression: ा इ -> ि []

6.2.4 Adjective to noun derivation

In this derivation, 1 suffix is involved and they derive a noun from a adjective stem. The semantics of the suffixes is not considered but they are marked as suffix with a tag SFX. Table 6.8 lists those suffixes with example of derived and base stem of each.

Table 6.8: Adjective to noun derivation

Base adjective Stem	Gloss	Suffix	Derived noun	Gloss
लामो lamo	long	आइ -ai	लमाइ lamai	length

The finite-state transducer in Figure 6.8 is common to all the suffixes listed in Table 6.8.

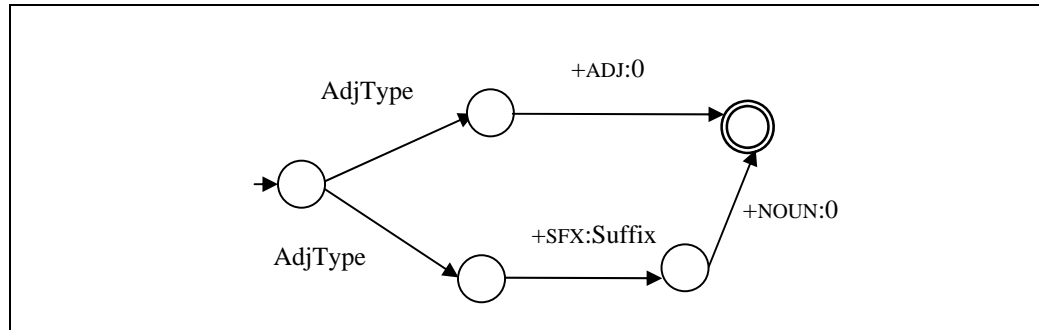


Figure 6.8: A finite state transducer for noun to adjective derivation

The phonological rules involved in this derivation are listed in PR 6.4; they are compiled and composed with the finite state transducer and it can analyze and generate both underived and derived words.

Phonological rule

PR 6.4

- i. Dependent vowel ऌ between consonant changes to ऎ .

Regular expression: $\text{ऌ} \rightarrow [] \parallel \text{cons} _ \text{cons}$;

- ii. Vowel sequence of dependent vowel ऌ and independent vowel ऎ changes to dependent vowel ऌ .

Regular expression: $\text{ऌ ऎ} \rightarrow \text{ऌ}$;

6.2.5 Adjective/noun to noun derivation

In this derivation, 1 suffix is involved and it derives a noun from a noun/adjective stem. The semantics of the suffixes is not considered, so it is simply marked as suffix with a tag SFX. Table 6.9 lists those suffixes with base stem and derived word.

Table 6.9: Adjective/Noun to Noun Derivation

Base noun/adjective stem	Gloss	Suffix	Derived noun	Gloss
गरिब gAr i b	poor	ई-i:	गरिबीgAr i bi	poverty

The finite-state transducer in Figure 6.9 is common to all the suffixes listed in Table 6.9. It is capable of analyzing and generating the derived words.

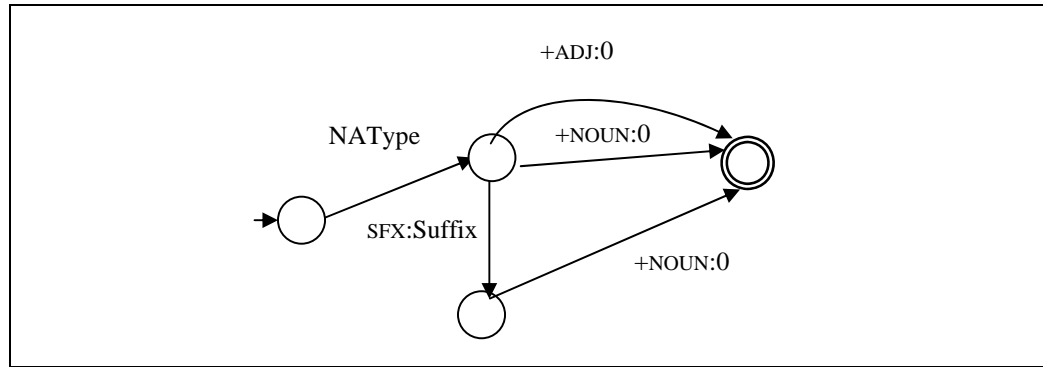


Figure 6.9: A finite state transducer for noun/adjective to noun derivation

The phonological rules involved in this process are listed in PR 6.5; they are compiled into a network and composed with the network illustrated Figure 6.9.

Phonological rule

PR 6.5

- i. Independent vowel $\text{ई}i$: changes to its corresponding dependent vowel $\text{ी}i$: after the consonant

Regular expression: $\text{ई} \rightarrow \text{ी} \parallel \text{cons } _$;

6.2.6 Verb to noun derivation

In this type of derivation, 32 suffixes are involved and they derive a noun from a verb stem. The semantics of the suffixes is not considered but they are marked as suffix with a tag SFX. Table 6.10 lists those suffixes with example of derived and base stem of each.

Table 6.10: Verb to Noun Derivation

Base verb stem	Gloss	Suffix	Derived noun	Gloss
चुन् tsun	elect	आउ- <i>au</i>	चुनाउ tsunau	election
चुन् tsun	elect	आब- <i>ab</i>	चुनाब tsunab	election
किट् kiṭ	decide	आन- <i>an</i>	किटान kiṭan	decision
किट् kiṭ	decide	आनी- <i>ani</i>	किटानी kiṭani	decision
ढाक् ḍʰak	cover	अनी- <i>Ani</i>	ढकनी ḍʰakʌni	lid
जल् dzʌl	burn	अन- <i>An</i>	जलन dzʌlʌn	burning
चोर tsor	steal	ई- <i>i:</i>	चोरी tsori:	theft
हाँस् hā̃s	laugh	ओ- <i>o</i>	हाँसो hā̃so	laughter
पढ् pʌḍʰ	read	आइ- <i>ai</i>	पढाइ pʌḍʰai	reading
थाक् tʰak	tire	आवट- <i>awaṭ</i>	थकावट tʰakawaṭ	tiredness
छाप् tsʰap	print	आ- <i>a</i>	छापा tsʰapa	printing
छान् tsʰan	choose	ओट- <i>oṭ</i>	छनोट tsʰanoṭ	selection
चिच्या tsitsja	shout	हट- <i>ḥaṭ</i>	चिच्याहट tsitsjahaṭ	shouting
झर् dzʌr	drop	अना- <i>Ana</i>	झरना dzʌrʌna	water fall
ढोग् ḍʰog	bow	आउनी- <i>auni</i>	ढोगाउनी ḍʰogauni	bowing
राख् rakʰ	keep	आलो- <i>alo</i>	रखालो rakʰalo	servant
दाब् dab	press	आब- <i>ab</i>	दाबाब dabab	pressure
बच् bʌts	save	अत- <i>At</i>	बचत bʌtsʌt	saving
सड् sʌḍ	decay	अल- <i>Al</i>	सडल sʌḍʌl	decay
रोप् rop	plant	आइँ- <i>aĩ</i>	रोपाइँ ropaiĩ	plantation
छेक् tsʰek	block	आरो- <i>aro</i>	छेकारो tsʰekaro	blockade
चिर् tsir	split	औटो- <i>Auto</i>	चिरौटो tsirʌuto	split
बढ् bʌḍʰ	grow	औती- <i>Auti</i>	बढौती bʌḍʰʌuti	growth
सर् sʌr	move	उवा- <i>uwa</i>	सरुवा sʌruwa	shift
उठ् utʰ	rise	ती- <i>ti</i>	उठ्ती utʰti	credit
चाल् tsal	sieve	नी- <i>ni</i>	चाल्नी tsalni	sieve
बेर ber	roll	नो- <i>no</i>	बेर्नो berno	piles
गा ga	sing	ना- <i>na</i>	गाना gana	song
भिड् biḍ	fight	अन्त- <i>Ant</i>	भिडन्त biḍʌnt	fighting
जित् dzit	win	औरी- <i>Auri</i>	जितौरी dzitʌuri	winning
कोर् kor	scratch	एसो- <i>eso</i>	कोरसो koreso	scratcher
खुल् kʰul	open	अस्त- <i>Asta</i>	खुलस्त kʰulʌsta	open

The finite-state transducer in Figure 6.10 is common to all the prefixes listed in Table 6.10 and it can analyze and generate the derived words.

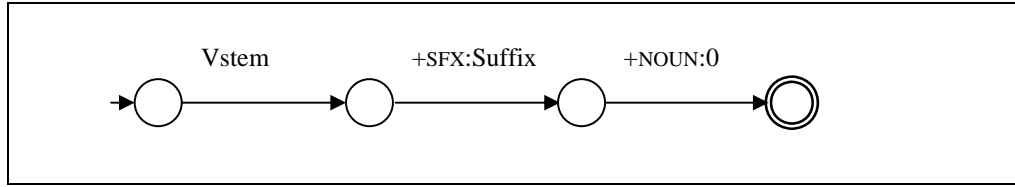


Figure 6.10: A finite state transducer for verb to noun derivation

The phonological rules involved in this derivation process are listed in PR 6.6; they are compiled and composed with the finite state transducer illustrated in Figure 6.10.

Phonological rule

PR 6.6

- i. Independent vowels आ, ई, ओ, इ, उ, ए, ऐ, औ change to their corresponding dependent vowels ा, ी, ो, ि, ु, े, ै, ौ and [] after the consonants, respectively.

Regular expressions: आ -> ा || cons __ ;

ई -> ी || cons __ ;

ओ -> ो || cons __ ;

इ -> ि || cons __ ;

उ -> ु || cons __ ;

ए -> े || cons __ ;

ऐ -> ै || cons __ ;

औ -> ौ || cons __ ;

अ -> [] || cons __ ;

6.2.7 Verb to adjective derivation

In this derivation, 14 suffix are involved and they derive an adjective from a verb stem. The semantics of the suffixes is not considered but they are marked as suffix

with a tag SFX. Table 6.11 lists those suffixes with example of base stem and derived word.

Table 6.11: Verb to adjective derivation

Base verb stem	Gloss	Suffix	Derived adjective	Gloss
मिच् mits	squeeze	आहा-aha	मिचाहा mitsaha	suppressor
भुल् b ^h ul	forget	अकड - <i>AkkAḍ</i>	भुलकड b ^h ulAkkAḍ	forgetful
पोस् pos	feed	इलो-ilo	पोसिलो posilo	nutritious
घुम् g ^h um	roam	अन्ते- <i>Ante</i>	घुमन्ते g ^h umAnte	vagabond
घुम् g ^h um	roam	अन्ता- <i>Anta</i>	घुमन्ता g ^h umAnta	vagabond
खप् k ^h Ap	last	आलु- <i>alu</i>	खपालु k ^h Apalu	long lasting
पढ् pAḍ ^h	read	ऐया- <i>Aija</i>	पढैया pAḍ ^h Aija	studious
छाड् tsAḍ	leave	आ-a	छाडा tsAḍa	wanton
रोप् rop	plant	आर-ar	रोपार ropar	planter
सिक् sik	learn	आरु- <i>aru</i>	सिकारु sikaru	learner
बिक् bik	sell	आउ-au	बिकाउ bikau	salable
भाग् b ^h ag	flee	औटो- <i>Auṭo</i>	भागौटो b ^h agAuṭo	runner
छेर ts ^h er	pass stool	औटी- <i>Auṭi</i>	छेरौटी ts ^h erAuṭi	sufferer
लाग् lag	attach	उ-u	लागु lagu	addicted

The finite-state transducer in Figure 6.11 is common to all the suffixes listed in Table 6.11. It can analyze and generate the derived words of this type.

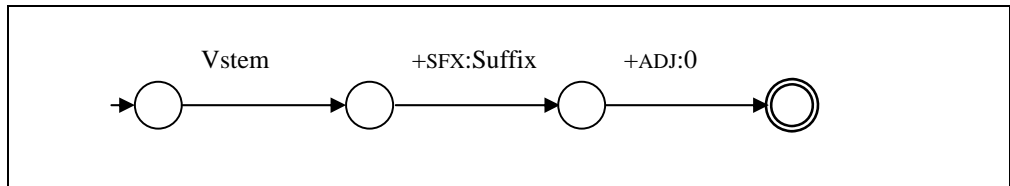


Figure 6.11: A finite state transducer for verb to adjective derivation

The phonological rules involved in this type of derivation are listed in PR 6.7; they are compiled and composed with the finite state transducer illustrated in Figure 6.11.

Phonological rule

PR 6.7

- i. Independent vowels आ_a, ई_i, ओ_o, इ_i, उ_u, ए_e, ऐ_{ai}, औ_{au}, and अ_∅ change to their corresponding dependent vowels ा_a, ी_i, ो_o, ि_i, उ_u, े_e, ै_{ai}, ौ_{au} and [] after the consonants, respectively.

Regular expressions: आ -> ा || cons __ ;

ई -> ी || cons __ ;

ओ -> ो || cons __ ;

इ -> ि || cons __ ;

उ -> उ || cons __ ;

ए -> े || cons __ ;

ऐ -> ै || cons __ ;

औ -> ौ || cons __ ;

अ -> [] || cons __ ;

6.2.8 Verb to adverb derivation

In this derivation, 2 suffixes are involved and they derive an adverb from a verb stem. The semantics of the suffixes is not considered, so they are simply marked as suffix with a tag SFX. Table 6.12 lists those suffixes with example of base stem and derived word.

Table 6.12: Verb to adverb derivation

Base verb stem	Gloss	Suffix	Derived adverb	Gloss
गर _{gAr}	do	उन्जेल -undzel	गरुन्जेल _{gArundzel}	till doing
गर _{gAr}	do	इन्जेल -indzel	गरिन्जेल _{gArindzel}	till doing

The finite-state transducer in Figure 6.12 is common to all the suffixes listed in Table 6.12. It can analyze and generate the derived words.

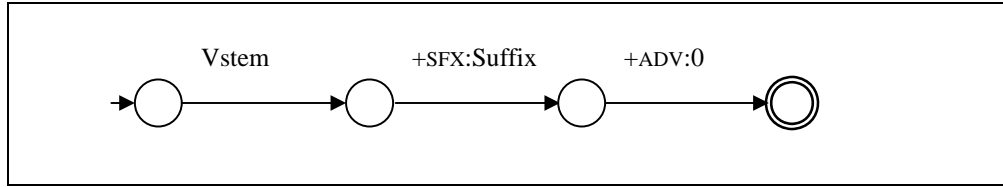


Figure 6.12: A finite state transducer for verb to adverb derivation

The phonological rules involved in this derivation are listed in PR 6.8; they are compiled and composed with the finite state transducer illustrated in Figure 6.12.

Phonological rule

PR 6.8

- i. Independent vowels उu and इi change to their corresponding dependent vowel ुu and िi after the consonants

Regular expressions: उ -> ु || cons __ ;

इ -> ि || cons __ ;

6.2.9 Adverb to adjective derivation

In this derivation, 1 suffix is involved and it derives an adjective from an adverb stem. The semantics of the suffixes is not considered but it is marked as suffix with a tag SFX. Table 6.12 lists those suffixes with example of base stem derived word.

Table 6.13: Adverb to adjective derivation

Base adverb stem	Gloss	Suffix	Derived adjective	Gloss
भिन्न b ^h itrΛ	inside	ई-i:	भित्री b ^h itri:	inner
बाहिर bahirΛ	outside	ई-i:	बाहिरी bahiri	ourter

The finite-state transducer in Figure 6.12 is common to all the suffixes listed in Table 6.13. It is capable of analyzing and generating the base stems and derived words.

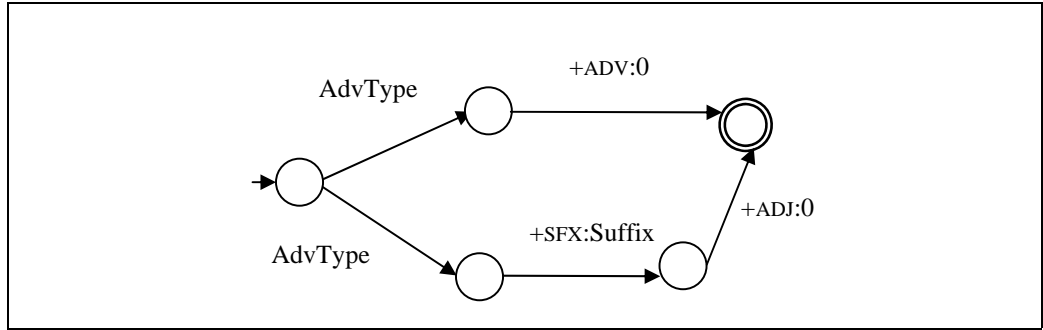


Figure 6.13: A finite state transducer for noun to adjective derivation

The phonological rules involved in this process is listed in PR 6.9; it is compiled and composed with the finite state transducer illustrated in Figure 6.13.

Phonological rule

PR 6.9

- i. Independent vowel ई*i*: changes to corresponding dependent vowel ी*i*: after the consonants

Regular expression: ई -> ी || cons __;

6.2.10 Verb to noun conversion

Some of the verb stems alter between verb and noun. They are same phonologically but differ in written form. In the noun form, a diacritic halanta is dropped. Some examples of such stems are listed in Table 6.14.

Table 6.14: Verb to noun conversion

Base verb stem	Gloss	Derived noun	Gloss
खेल् ^h el	'play'	खेल ^h el	'game'
खोज् ^h odz	'search'	खोज ^h odz	'research'

The finite-state transducer in Figure 6.14 encodes stems listed in Table 6.14 and it is capable of analyzing and generating the base stem and derived words.

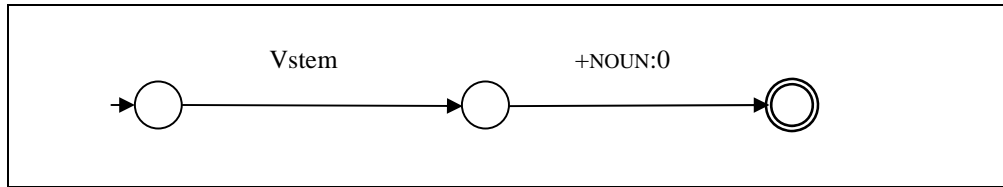


Figure 6.14: A finite state transducer for verb to adverb derivation

The phonological rule involved in this conversion is listed in PR 6.10; it is compiled and composed with finite state transducer illustrated in Figure 6.14.

Phonological rule

PR 6.10

- i. Halanta ् at the end of verb stem is removed.

Regular Expression: ् -> [] || __.#.

6.2.11 Verb to adjective/noun conversion

Some verb stems alter between verb and noun or adjective forms. Some examples of such stems are listed in Table 6.15. Phonologically they are same but orthographically differ by halanta.

Table 6.15: Verb to Adjective/Noun Conversion

Base verb Stem	Gloss	Derived adjective	Gloss
ठग् t ^h Λg	cheat	ठग् t ^h Λg	cheat
चोर tsor	steal	चोर tsor	thief
थप् t ^h Λp	add	थप् t ^h Λp	additional

The finite-state transducer illustrated in Figure 6.14 encodes stems listed in Table 6.15 and it is capable of analyzing and generating the base stems and derived words.

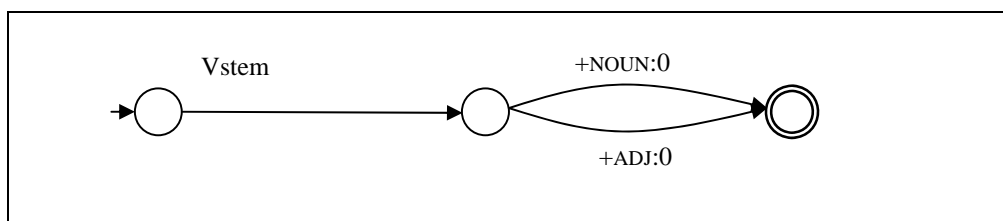


Figure 6.15: A finite state transducer for verb to adverb derivation

The phonological rule involved in this conversion is listed in PR 6.11. It is compiled and composed with the finite state transducer illustrated in Figure 6.15.

Phonological rule

PR 6.11

- i. Halanta ् at the end of verb stem is removed.

Regular expression: ् -> [] || __.#.

6.2.12 Verb to noun derivation (vowel insertion)

Some verb stems change from verb form to noun forms by inserting vowel अ between consonants in the stem. Some examples of such stems are listed in Table 6.16.

Table 6.16: Verb to noun (vowel insertion)

Base verb stem	Gloss	Derived adjective	Gloss
चम्क tsAmkA	shine	चमक tsAmAk	shining
सम्झ sAmDZA	remember	समझ sAmAdZ	understanding
टल्क tAlkA	shine	टलक tAlAk	shining

The finite-state transducer illustrated in Figure 6.16 encodes stems listed in Table 6.16 and it is capable of analyzing and generating the derived words.

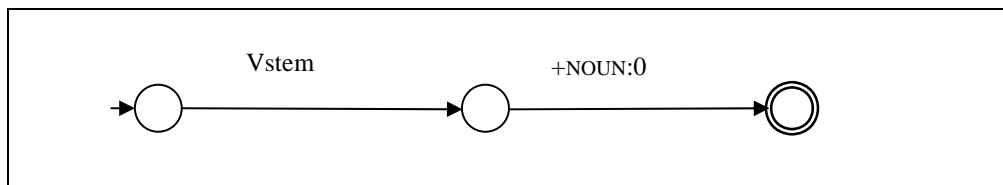


Figure 6.16: A finite state transducer for verb to adverb derivation

The phonological rule involved in this derivation is listed in PR 6.12; it is compiled and composed with the finite state transducer illustrated in Figure 6.16.

Phonological rule

PR 6.12

- i. Halant ɔ̃ between the consonants of verb stem is removed.

Regular expression: ɔ̃ -> [] || cons__ cons;

6.3 Summary

In this chapter, we presented the derivation process in Nepali. The various derivation such as noun to noun, noun to adjective, noun to adverb and adjective to adjective are the former types and noun to noun, noun to adjective, noun to noun/adjective, adjective to noun, adjective/noun to noun, verb to noun, verb to adjective, verb to adverb are the latter types. In addition, there are two kinds of conversions: verb to noun and verb to adjective/noun. And verb to noun derivation due to vowel insertion is also included. Each prefix and suffix has its own set of words from which derivation takes place. The derivation process in Nepali is not as productive and regular as inflectional process in Nepali. However there exists a quite good number of derived words. Two major types of derivation, prefixation and suffixation, are discussed and implemented.

CHAPTER 7

IMPLEMENTATION

7.0 Outline

This chapter presents the implementation of morphological categories and phonological rules analyzed in the earlier chapters to design a computational model using the Xerox finite state toolkit. It consists of four sections. Section 7.1 presents the morphotactics, i.e. syntax of morphemes. The morphological categories and grammatical categories have been separated based on the earlier analysis. Section 7.2 presents the lexicon grammar for nouns, pronouns, adjectives, verbs, numerals, classifiers, adverbs, postpositions, conjunctions, particles, interjections and derivation. Section 7.3 deals with the realization, i.e. rules for alternation using xfst interface for each category. Finally, section 7.4 summarizes the chapter.

7.1 Morphotactics: syntax of morphemes

7.1.1 Morphological categories

As discussed and analyzed in chapters (3-6), two major categories are identified, open word class and closed word class. Table 7.1 shows four open word classes and their corresponding morphological tags used in the morphological analyzer. Table 7.2 shows seven closed word classes with their corresponding morphological tags used in the morphological analyzer.

Table 7.1: The open word classes

S.N.	Morphological Categories	Tags
1.	Nouns	+NOUN
2.	Adjectives	+ADJ
3.	Verbs	+VERB
4.	Adverbs	+ADV

Table 7.2: The closed word classes

S.N.	Morphological Categories	Tags
1.	Pronouns	+PRON
2.	Numeral	+NUM
3.	Classifier	+CLF
4.	Postpositions	+POSTP
5.	Conjunctions	+CCONJ, +SCONJ
6.	Particles	+PARTICLE
7.	Interjections	+INTERJ

7.1.2 Grammatical categories

Altogether 45 grammatical categories are identified in all open and closed word classes. Features in adverbs from 37 to 42 in Table 7.3 and two features in nouns in 44 and 45 of Table 7.3 are semantic whereas the rest of the features in other categories are formal. Table 7.3 lists all the grammatical categories and their corresponding morphological tags used in the morphological analyzer. The redundant features such as augmentative, non-causative, active, direct form have not been incorporated into the analyzer.

Table 7.3: The grammatical categories and features

S.N.	Grammatical categories	Tags
1.	Number	+SG, +PL
2.	Gender	+MASC, +FEM
3.	Form	+DIRT, +OBL
4.	Honorificity	+NHON, +HON, +HHON, +RHON
5.	Evaluation	+AUG, +DIM
6.	Persons	1, 2, 3
7.	Cases	+ERG, +INST, +DAT, +ABL, +LOC, +COM, +GEN, +VOC, +ALL
8.	Distal	+DIST
9.	Proximate	+PROX
10.	Reflexive	+REFL
11.	Demonstrative	+DEM
12.	Relative	+REL
13.	Interrogative	+INTERRO
14.	Indefinite	+INDEF
15.	Definite	+DEF
16.	Reciprocal	+RECIP
17.	Degree	+POSIT, +COMP, +SUPER
18.	Cardinal	+CARD
19.	Ordinal	+ORD
20.	Frequency	+FREQ
21.	Portion	+PORT

22.	Voice	+PASS
23.	Causative	+CAUSE
24.	Existential	+EXIST
25.	Identificational	+ID
26.	Tenses	+NPST, +PST
27.	Aspects	+PERF, +IMPERF, +INFER, +HAB
28.	Moods	+IMP, +OPT, +POT
29.	Absolutive	+ABS
30.	Infinitive	+INF
31.	Purposive	+PURP
32.	Prospective	+PROS
33.	Durative	+DUR
34.	Conjunctive	+CONJUNCT
35.	Conditional	+COND
36.	Perfective	+PERFT
37.	Temporal	+TEMP
38.	Spatial	+SPAC
39.	Amount	+AMOUNT
40.	Manner	+MANNER
41.	Reason	+REASON
42.	Sentential	+SENT
43.	Emphasis	+EMPH
44.	Proper Name	+PROPER
45.	Place Name	+PLACE

There are a number of arbitrary tags used in the lexc file to restrict the scope of replace rules. Finally these tags are removed for the transducer after the application of the replace rules. Table 7.3 lists a sample of such tags.

Table 7.3: The arbitrary tags

S.N.	Purpose of arbitrary tag	Tags
1.	<i>O</i> -ending nouns for plural, honorific, oblique and vocative	^MP
2.	<i>O</i> -ending nouns for feminine and diminutive	^FE
3.	Non-honorific imperative	^IMPsg
4.	Honorific imperative	^IMPhon
5.	Plural imperative	^IMPpl
6.	Noun to adjective derivation	^NA
7.	Noun to adverb derivation	^NADV
8.	Adjective to verb derivation	^ADJV
9.	Verb to noun conversion	^R
10.	Insertion in verb to noun derivation	^a

7.2 Lexc grammar

Nouns, pronouns, adjectives, numerals, classifiers, verbs, adverbs, postpositions and particles and interjections are encoded in lexc files. The main morphological forms are in Devanagari script and the morphological tags are in Roman script using UTF-8 character encoding. The lexc files begin with Multichar_Symbols and lexicon follows it.

7.2.1 Nouns

The nouns discussed and analyzed in (3.1) are implemented in a lexc file named nouns.txt which includes 14 classes of nouns. The upper language contains stems and sequence of morphological tags and the lower language contains surface forms. Besides regular morphological tags, some arbitrary tags are used for restricting the application of replace rules. The encoding of the nouns with their morphological tags is as follows. This lexc file accumulates the transducers from Figure 3.1 to Figure 3.13.

```
Multichar_Symbols +NOUN +MASC +FEM +OBL +PL +SG +DIM +VOC +HON ^MP ^FE  
+PLACE +PROPER
```

```
LEXICON ROOT
```

```
  Nouns;
```

```
LEXICON Nouns
```

```
  !! Type 1a Nouns:
```

```
केटो inflection_1a;
```

```
  !!Type 1b Nouns:
```

```
मुसो inflection_1b;
```

```
  !!Type 1c Nouns:
```

```
डालो inflection_1c;
```

```
  !!Type 1d Nouns:
```

```
फोटो inflection_1d;
```

```
  !!Type 21a Nouns:
```

```
काका inflection_21a;
```

```
  !!Type 21b Nouns:
```

```
नाति inflection_21b;
```

```
  !!Type 21c Nouns:
```

```
बाघ inflection_21c;
```

```
  !!Type 21d Nouns:
```

```
बिस्ट inflection_21d;
```

!!Type 22a Nouns:
 दाइ inflection_22a;
 !!Type 22b Nouns:
 दिदि inflection_22b;
 !!Type 22c Nouns:
 राम inflection_22c;
 !!Type 22d Nouns:
 सीता inflection_22d;
 !!Type 22e Nouns:
 खेत inflection_22e;
 !!Type 22f Nouns:
 पोखरा inflection_22f;

LEXICON inflection_1a
 +NOUN+MASC+SG:0 #;
 +NOUN+MASC+PL:^MP #;
 +NOUN+MASC+OBL:^MP #;
 +NOUN+MASC+HON:^MP #;
 +NOUN+MASC+VOC:^MP #;
 +NOUN+FEM:+FE #;
 LEXICON inflection_1b
 +NOUN+MASC+SG:0 #;
 +NOUN+MASC+PL:^MP #;
 +NOUN+MASC+OBL:^MP #;
 +NOUN+FEM:+FE #;
 LEXICON inflection_1c
 +NOUN+SG:0 #;
 +NOUN+PL:^MP #;
 +NOUN+OBL:^MP #;
 +NOUN+DIM:^FE #;
 LEXICON inflection_1d
 +NOUN+SG:0 #;
 +NOUN+PL:^MP #;
 +NOUN+OBL:^MP #;
 LEXICON inflection_21a
 +NOUN+MASC:0 #;
 +NOUN+FEM:ी #;
 LEXICON inflection_21b
 +NOUN+MASC:0 #;
 +NOUN+FEM:नी #;
 LEXICON inflection_21c
 +NOUN+MASC:0 #;
 +NOUN+FEM:िनी #;


```

LEXICON inflection_21d
+NOUN+MASC:0 #;
+NOUN+FEM:ेनी #;
+NOUN+FEM:िनी #;
LEXICON inflection_22a
+NOUN+MASC:0 #;
LEXICON inflection_22b
+NOUN+FEM:0 #;
LEXICON inflection_22c
+NOUN+PROPER+MASC:0 #;
LEXICON inflection_22d
+NOUN+PROPER+FEM:0 #;
LEXICON inflection_22e
+NOUN:0 #;
LEXICON inflection_22f
+NOUN+PLACE:0 #;
END

```

7.2.2 Pronouns

Nepali pronouns are limited in number and more or less idiosyncratic in their forms and functions. Hence, instead of organizing them where rules can be applied to get their surface forms, they are directly encoded uniting all the finite state transducers from Figure 3.14 to Figure 3.33 along with their morphological tags. Therefore, this section does not contain any replace rules.

```

Multichar_Symbols +PRON +1SG +OBL +EMPH +GEN +MASC +FEM +HON
+1PL +2SG +HHON +RHON +3SG +PROX +DIST +REFL +DEM +HUM +NHUM
+DEF +INTERRO +RECIP +PL +SG +REL +INDEF +3PL

```

LEXICON ROOT

```

pronouns;
LEXICON pronouns
!!First person singular pronoun
म+PRON+1SG:म #;
म+PRON+1SG+OBL:मै #;
म+PRON+1SG+EMPH:मै #;
म+PRON+1SG+OBL+GEN+MASC:मेरो #;
म+PRON+1SG+OBL+GEN+FEM:मेरी #;
म+PRON+1SG+OBL+GEN+PL:मेरा #;
म+PRON+1SG+OBL+GEN+HON:मेरा #;
म+PRON+1SG+OBL+GEN+OBL:मेरा #;

```

म+PRON+1SG+OBL+GEN+EMPH:मेरै #;

!!First Person Plural Prouns

हामी+PRON+1PL:हामी #;

हामी+PRON+1PL+OBL+GEN+MASC:हाम्रो #;

हामी+PRON+1PL+OBL+GEN+FEM:हाम्री#;

हामी+PRON+1PL+OBL+GEN+PL:हाम्रा #;

हामी+PRON+1PL+OBL+GEN+HON:हाम्रा #;

हामी+PRON+1PL+OBL+GEN+OBL:हाम्रा#;

हामी+PRON+1PL+OBL+GEN+EMPH:हाम्रै #;

!! Second Person Singular

तँ+PRON+2SG:तँ #;

तँ+PRON+2SG+OBL:तैं #;

तँ+PRON+2SG+EMPH:तैं #;

तँ+PRON+2SG+OBL+GEN+MASC:तेरो#;

तँ+PRON+2SG+OBL+GEN+FEM:तेरी #;

तँ+PRON+2SG+OBL+GEN+PL:तेरा #;

तँ+PRON+2SG+OBL+GEN+HON:तेरा #;

तँ+PRON+2SG+OBL+GEN+OBL:तेरा #;

तँ+PRON+2SG+OBL+GEN+EMPH:तेरै #;

!!Second Person honorific Pronoun

तिमी+PRON+2SG+HON:तिमी #;

तिमी+PRON+2SG+OBL+HON+GEN+MASC:तिम्रो#;

तिमी+PRON+2SG+OBL+HON+GEN+FEM:तिम्री #;

तिमी+PRON+2SG+OBL+HON+GEN+PL:तिम्रा #;

तिमी+PRON+2SG+OBL+HON+GEN+HON:तिम्रा #;

तिमी+PRON+2SG+OBL+HON+GEN+OBL:तिम्रा #;

तिमी+PRON+2SG+OBL+HON+GEN+EMPH:तिम्रै#;

!! Second person high honorific pronouns

तपाईं+ PRON+2SG+HHON:तपाईं #;

यहाँ+PRON+2SG+HHON:यहाँ #;

उहाँ+PRON+2SG+HHON:उहाँ#;

वहाँ+PRON+2SG+HHON:वहाँ #;
हजुर+PRON+2SG+HHON:हजुर #;

!! Second person royal honorific pronoun
मौसुफ+PRON+2SG+RHON:मौसुफ #;

!!Third Person Singular Pronoun ऊ

ऊ+PRON+3SG:ऊ #;
ऊ+PRON+3SG+EMPH:उही #;
ऊ+PRON+3SG+OBL:उस #;
ऊ+PRON+3SG+OBL+EMPH:उसै #;
ऊ+PRON+3SG+HON:उनी #;
ऊ+PRON+3SG+HON+OBL:उन #;
ऊ+PRON+3SG+HON+OBL+EMPH:उनै #;
ऊ+PRON+3SG+HON:उहाँ #;
ऊ+PRON+3SG+HON:वहाँ #;

!! Third person singular pronoun यो

यो+PRON+3SG+PROX:यो #;
यो+PRON+3SG+PROX+EMPH:यही #;
यो+PRON+3SG+OBL+PROX:यस #;
यो+PRON+3SG+OBL+PROX+EMPH:यसै #;
यो+PRON+3SG+PROX+HON:यी #;
यो+PRON+3PL+PROX:यी #;
यो+PRON+3SG+PROX+HON:यिनी #;
यो+PRON+3SG+PROX+OBL+HON:यिन #;
यो+PRON+3SG+PROX+OBL+HON+EMPH:यिनै #;

!!!Third person singular pronoun त्यो and ती

त्यो+PRON+3SG+DIST:त्यो #;
त्यो+PRON+3SG+DIST+EMPH:त्यही #;
त्यो+PRON+3SG+OBL:त्यस #;
त्यो+PRON+3SG+OBL+EMPH:त्यसै #;
ती+PRON+3SG+HON+DIST:ती #;
ती+PRON+3PL+DIST:ती #;

ती+PRON+3SG+HON+DIST:तिनी #;
ती+PRON+3SG+OBL+HON+DIST:तिन #;
ती+PRON+3SG+OBL+HON+DIST+EMPH:तिनै #;

!!Reflexive pronoun

आफू+PRON+REFL:आफू #;
आफू+PRON+REFL+OBL+EMPH:आफै #;
आफू+PRON+REFL+OBL+EMPH:आफैं #;
आफू+PRON+REFL+OBL+GEN+SG:आफ्नो #;
आफू+PRON+REFL+OBL+GEN+PL:आफ्ना#;
आफू+PRON+REFL+OBL+GEN+HON:आफ्ना #;
आफू+PRON+REFL+OBL+GEN+OBL:आफ्ना #;
आफू+PRON+REFL+OBL+GEN+FEM:आफ्नी #;
आफू+PRON+REFL+OBL+GEN+EMPH:आफ्नै #;

!! Demonstrative pronouns यो

यो+PRON+DEM+PROX:यो #;
यो+PRON+DEM+PROX+EMPH:यही #;
यो+PRON+DEM+PROX:यी #;
यो+PRON+DEM+PROX+HON:यिनी #;
यो+PRON+DEM+PROX+OBL:यिन #;
यो+PRON+DEM+PROX+OBL+EMPH:यिनै #;
यो+PRON+DEM+PROX:यहाँ #;

!!Demonstrative pronoun त्यो and ती

त्यो+PRON+DEM+DIST:त्यो #;
त्यो+PRON+DEM+DIST+EMPH:त्यही #;
ती+PRON+DEM+DIST:ती #;
ती+PRON+DEM+DIST+OBL+HON:तिनी #;
ती+PRON+DEM+DIST+OBL:तिन #;
ती+PRON+DEM+DIST+OBL+EMPH:तिनै #;

!!Demonstrative pronoun ऊ

ऊ+PRON+DEM+DIST:ऊ #;

ऊ+PRON+DEM+DIST+EMPH:उही #;
ऊ+PRON+DEM+DIST+HON:उनी #;
ऊ+PRON+DEM+DIST+OBL:उन#;
ऊ+PRON+DEM+DIST+OBL+EMPH:उनै #;
ऊ+PRON+DEM+DIST+HON:उहाँ #;
ऊ+PRON+DEM+DIST+HON:वहाँ #;

!!Other Demonstrative pronouns

सो+PRON+DEM+DIST:सो #;
सो+PRON+DEM+DIST+EMPH:सोही #;
निज+PRON+DEM+PROX:निज #;
निज+PRON+DEM+PROX+EMPH:निजै #;
उक्त+PRON+DEM+PROX:उक्त #;

!! Relative pronouns

जो+PRON+REL+HUM:जो #;
जो+PRON+REL+OBL+HUM:जस #;
जो+PRON+REL+OBL+HUM+EMPH:जसै #;
जे+PRON+REL+NHUM:जे #;
जुन+PRON+REL:जुन #;
जुन+PRON+REL+EMPH:जुनै #;

!! Interrogative pronouns

को+PRON+INTERRO+HUM:को#;
को+PRON+INTERRO+HUM+OBL:कस #;
को+PRON+INTERRO+HUM+OBL+EMPH:कसै #;
के+PRON+INTERRO+NHUM:के #;
कुन+PRON+INTERRO:कुन #;
किन+PRON+INTERRO:किन #;
कसरी+PRON+INTERRO:कसरी#;

!! Indefinite pronouns

को+PRON+INDEF+HUM:कोही #;
के+PRON+INDEF+NHUM:केही #;
कुन+PRON+INDEF: कुनै #;
जो+PRON+INDEF+HUM:जोसुकै #;

जे+PRON+INDEF+NHUM:जेसुकै #;
जुन+PRON+INDEF:जुनसुकै #;

!! Definite pronouns

प्रत्येक+PRON+DEF:प्रत्येक #;
हरेक+PRON+DEF:हरेक #;
सबै+PRON+DEF:सबै #;
अर्को+PRON+DEF+SG:अर्को #;
अर्को+PRON+DEF+PL:अर्का #;
अर्को+PRON+DEF+HON:अर्का #;
अर्को+PRON+DEF+OBL:अर्का #;
अर्को+PRON+DEF+FEM:अर्की #;
अर्को+PRON+DEF+EMPH:अर्कै #;
अरू+PRON+DEF:अरू #;

!! Reciprocal pronouns

एकअर्को+PRON+RECIP:एकअर्को #;
एकअर्को+PRON+RECIP+OBL:एकअर्का #;
एकअर्को+PRON+RECIP+HON:एकअर्का #;
एकअर्को+PRON+RECIP+PL:एकअर्का #;
एकअर्को+PRON+RECIP+FEM:एकअर्की #;
एकआपस+PRON+RECIP:एकआपस #;
आपस+PRON+RECIP:आपस #;
आफू+PRON+RECIP:आआफू #;

END

7.2.3 Verbs

The verb stems analyzed and classified in (4.4); and auxiliary verbs and inflections analyzed in (4.5) are implemented in a single lexc file. The stems and inflections are concatenated in same lexc file with the help of continuation lexicons. Two flag diacritics @U.NEG.PRESENT@ and @U.NEG.ABSENT@ are defined and implemented to restrict negative prefix. This lexc file includes the tranducers from Figure 4.1 to Figure 4.36.

MULTICHAR_SYMBOLS ^IMP +1PL +1SG +2PL +2SG +3PL +3SG +ABS
+COND +CONJ +DUR +EMPH +FEM +HAB +HON +IMP +IMPERF +INF
+MASC +NEG +NPST +OBL +OPT +PST +PERF +PERFT +PL +POT +PROSP

+PURP +SG +UNA +EXIST +IDEN NEG+ @U.NEG.PRESENT@
@U.NEG.ABSENT@ +VERB +PASS +CAUSE

LEXICON ROOT

!!Auxiliary verbs

छ+EXIST:छ auxcha;

हो+IDEN:हो auxho;

थि+EXIST:थि past;

!!=====Main Verb Stems =====!!

NEG+@U.NEG.PRESENT@:न@U.NEG.PRESENT@ Verbs;

Verbs;

LEXICON Verbs

!! Verb Type 1a

अघा Type1a;

!! Verb Type 1b

चोखि Type1b;

!!Verb Type1c

उक्लि Type1c;

!!Type verb1d

हाँस् Type1d;

!!Verb Type1e

बस् Type1e;

!! Verb Type2a

उचाल् Type2a;

!! Verb Type 2b

पक्रि Type2b;

!!Verb Type2c

नाच् Type2c;

!!Type verb2d

किन् Type2d;

!! Irregular verbs

आउनु+VERB:आ Group;

आउनु+VERB+PASS:आइ intGroup;

खानु+VERB:खा Group;

खानु+VERB+PASS:खाइ Group;

खानु+VERB+CAUSE:ख्या Group;

खानु+VERB+CAUSE+PASS:ख्याइ Group;

बस्नु+VERB+CAUSE:बसाल् Group;

बस्नु+VERB+CAUSE+PASS:बसालि Group;
 बस्नु+VERB+CAUSE:बसाल् Group;
 बस्नु+VERB+CAUSE+PASS:बसालि Group;
 LEXICON Type1a
 उनु+VERB:0 Group;
 उनु+VERB+PASS:इ intGroup;
 LEXICON Type1b
 नु+VERB:0 Group;
 नु+VERB+PASS:इ intGroup;
 नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 LEXICON Type1c
 नु+VERB:0 Group;
 नु+VERB+PASS:इ intGroup;
 नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 LEXICON Type1d
 नु+VERB:0 Group;
 नु+VERB+PASS:इ intGroup;
 नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 LEXICON Type1e
 नु+VERB:0 Group;
 नु+VERB+PASS:इ intGroup;
 नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 LEXICON Type2a
 नु+VERB:0 Group;
 नु+VERB+PASS:इ Group;
 LEXICON Type2b
 नु+VERB:0 Group;
 नु+VERB+PASS:इ Group;
 नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 LEXICON Type2c
 नु+VERB:0 Group;
 नु+VERB+PASS:इ Group;

नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 LEXICON Type2d
 नु+VERB:0 Group;
 नु+VERB+PASS:इ Group;
 नु+VERB+CAUSE:आ Group;
 नु+VERB+CAUSE+PASS:आइ Group;
 !!!=====Verbs end

LEXICON intGroup
 @U.NEG.ABSENT@ intGroup1;
 intGroup2;
 LEXICON intGroup1
 +NPST+3SG+MASC:छ #;
 +NPST+NEG+3SG+MASC:दैन #;
 +NPST+NEG+3SG:न #;
 +PST+3SG+MASC:यो #;
 +PST+3SG+MASC:एन #;
 +PST+HAB+3SG+MASC:थ्यो #;
 +PST+NEG+HAB+3SG+MASC:दैनथ्यो #;
 +PST+INFER+3SG+MASC:एछ #;
 +PST+INFER+NEG+3SG+MASC:एनछ #;

LEXICON intGroup2
 +PERF+SG+MASC:एको #;
 +IMPERF:दै #;
 +OPT+3SG:ओस् #;
 +POT+3SG+MASC:ला #;
 +INF:नु #;
 +INF+OBL:ना #;
 +PURP:न #;
 +PROSP:ने #;
 +DUR+EMPH:दै #;
 +CONJ:एर #;
 +CONJ:इकन #;
 +COND:ए #;

+PERFT:ए #;

LEXICON Group
@U.NEG.ABSENT@ Group1;
Group2;

LEXICON Group1
NonpastAffirmative;
NonpastNegative1;
NonpastNegative2;
PastAffirmative;
PastNegative;
HabitualAspectAffirmative;
HabitualAspectNegative;
InferentialAffirmative;
InferentialNegative;

LEXICON Group2
PerfectAspect;
ImperfectAspect;
Imperative;
Optative;
Potential;
Participles;

LEXICON past
PastAffirmative;
PastNegative;

!Inflection for non-past existential verb छे chz 'be' (Affirmative)

LEXICON auxcha

+NPST+1SG:ः #;
+NPST+1PL:ौँ #;
+NPST+2SG+MASC:स् #;
+NPST+2SG+FEM:ेस् #;
+NPST+2SG+MASC+HON:ौ #;
+NPST+2SG+FEM+HON:्यौ #;
+NPST+2PL:ौ #;
+NPST+3SG+MASC:0 #;
+NPST+3SG+FEM:े #;
+NPST+3SG+MASC+HON:न् #;
+NPST+3SG+FEM+HON:िन् #;
+NPST+3PL:न् #;

!Inflection for non-past existential verb छे chz 'be' (Negative)

+NPST+NEG+1SG:ेँ #;
+NPST+NEG+1PL:ेँौ #;
+NPST+NEG+2SG:ेँस् #;
+NPST+NEG+2SG+HON:ेँौ #;
+NPST+NEG+2PL:ेँौ #;
+NPST+NEG+3SG:ेँ #;
+NPST+NEG+3SG+HON:ेँन् #;
+NPST+NEG+3PL:ेँन् #;
!! end of छे

!Inflections for non-past identificational verb हो ho 'be' (affirmative)

LEXICON auxho

+NPST+1SG:उँ #;
+NPST+1PL:औँ #;
+NPST+2SG:स् #;
+NPST+2SG+HON:औ #;
+NPST+2PL:औ #;
+NPST+3SG:0 #;
+NPST+3SG+HON:न् #;
+NPST+3PL:न् #;

!Inflection for non-past identificational verb हो ho 'be' (Negative)

+NPST+NEG+1SG:इँ #;
+NPST+NEG+1PL:इँौ #;
+NPST+NEG+2SG:इँस् #;
+NPST+NEG+2SG+HON:इँौ #;
+NPST+NEG+2PL:इँौ #;
+NPST+NEG+3SG:इँ #;
+NPST+NEG+3SG+HON:इँन् #;
+NPST+NEG+3PL:इँन् #;

!!!===Tense, aspect and mood =====

!Inflections for non-past tense (affirmative)

LEXICON NonpastAffirmative

+NPST+1SG:छु #;
+NPST+1PL:छौँ #;

+NPST+2SG+MASC:छस् #;
 +NPST+2SG+FEM:छेस् #;
 +NPST+2SG+MASC+HON:छौ #;
 +NPST+2SG+FEM+HON:छ्यौ #;
 +NPST+2PL:छौ #;
 +NPST+3SG+MASC:छ #;
 +NPST+3SG+FEM:छे #;
 +NPST+3SG+MASC+HON:छन् #;
 +NPST+3SG+FEM+HON:छिन् #;
 +NPST+3PL:छन् #;

!Inflections for non-past tense negative 1
 LEXICON NonpastNegative1

+NPST+NEG+1SG:दिनै #;
 +NPST+NEG+1PL:दिनौ #;
 +NPST+NEG+2SG+MASC:दिनस् #;
 +NPST+NEG+2SG+FEM:दिनस् #;
 +NPST+NEG+2SG+MASC+HON:दिनौ #;
 +NPST+NEG+2SG+FEM+HON:दिनौ #;
 +NPST+NEG+2PL:दिनौ #;
 +NPST+NEG+3SG+MASC:दिन #;
 +NPST+NEG+3SG+FEM:दिन #;
 +NPST+NEG+3SG+MASC+HON:दिनन् #;
 +NPST+NEG+3SG+FEM+HON:दिनन् #;
 +NPST+NEG+3PL:दिनन् #;

!Inflections for non-past tense negative 2
 LEXICON NonpastNegative2

+NPST+NEG+1SG:नै #;
 +NPST+NEG+1PL:नौ #;
 +NPST+NEG+2SG:नस् #;
 +NPST+NEG+2SG+HON:नौ #;
 +NPST+NEG+2PL:नौ #;
 +NPST+NEG+3SG:न #;
 +NPST+NEG+3SG+HON:नन् #;
 +NPST+NEG+3PL:नन् #;

!Inflections for past tense (affirmative)

LEXICON PastAffirmative

- +PST+1SG:एँ #;
- +PST+1PL:यौँ #;
- +PST+2SG:इस् #;
- +PST+2SG+HON:यौ #;
- +PST+2PL:यौ #;
- +PST+3SG+MASC:यो #;
- +PST+3SG+FEM:ई #;
- +PST+3SG+MASC+HON:ए #;
- +PST+3SG+FEM+HON:इन् #;
- +PST+3PL:ए #;

!Inflections for past tense (negative)

LEXICON PastNegative

- +PST+NEG+1SG:इँ #;
- +PST+NEG+1PL:एँ #;
- +PST+NEG+2SG:इनस् #;
- +PST+NEG+2SG+MASC+HON:एँ #;
- +PST+NEG+2SG+FEM+HON:इँ #;
- +PST+NEG+2PL:एँ #;
- +PST+NEG+3SG+MASC:एन #;
- +PST+NEG+3SG+FEM:इन #;
- +PST+NEG+3SG+MASC+HON:एनन् #;
- +PST+NEG+3SG+FEM+HON:इनन् #;
- +PST+NEG+3PL:एनन् #;

!Inflections for perfect aspect

LEXICON PerfectAspect

- +PERF+SG+MASC:एको #;
- +PERF+PL:एका #;
- +PERF+SG+FEM:एकी #;
- +PERF+EMPH:एकै #;

!Inflections for imperfect aspect

LEXICON ImperfectAspect

- +IMPERF+SG+MASC:दो #;

+IMPERF+SG+FEM:दी #;
+IMPERF+PL:दा #;
+IMPERF:दै #;

! Inflections for habitual aspect (affirmative)
LEXICON HabitualAspectAffirmative

+PST+HAB+1SG:थै #;
+PST+HAB+1PL:थ्यौ #;
+PST+HAB+2SG:थिस् #;
+PST+HAB+2SG+HON:थ्यौ #;
+PST+HAB+2PL:थ्यौ #;
+PST+HAB+3SG+MASC:थ्यो #;
+PST+HAB+3SG+FEM:थि #;
+PST+HAB+3SG+MASC+HON:थे #;
+PST+HAB+3SG+FEM+HON:थिन् #;
+PST+HAB+3PL:थे #;

! Inflections for habitual aspect (negative)
LEXICON HabitualAspectNegative

+PST+NEG+HAB+1SG:दैनथै #;
+PST+NEG+HAB+1PL:दैनथ्यौ #;
+PST+NEG+HAB+2SG:दैनथिस् #;
+PST+NEG+HAB+2SG+HON:दैनथ्यौ #;
+PST+NEG+HAB+2PL:दैनथ्यौ #;
+PST+NEG+HAB+2SG+FEM:दिनथिस् #;
+PST+NEG+HAB+2SG+FEM+HON:दिनथ्यौ #;
+PST+NEG+HAB+3SG+MASC:दैनथ्यो #;
+PST+NEG+HAB+3SG+FEM:दिनथिस् #;
+PST+NEG+HAB+3SG+MASC+HON:दिनथे #;
+PST+NEG+HAB+3PL:दैनथे#;

! Inflections for Inferential aspect (affirmative)
LEXICON InferentialAffirmative

+PST+INFER+1SG:एछु #;
+PST+INFER+1PL:एछौ #;
+PST+INFER+2SG+MASC:एछस् #;
+PST+INFER+2SG+FEM:इछस् #;

+PST+INFER+2SG+MASC+HON:एछौ #;
 +PST+INFER+2SG+FEM+HON:इछौ #;
 +PST+INFER+2PL:एछौ #;
 +PST+INFER+3SG+MASC:एछ #;
 +PST+INFER+3SG+FEM:इछ #;
 +PST+INFER+3SG+MASC+HON:एछन् #;
 +PST+INFER+3SG+FEM+HON:इछन् #;
 +PST+INFER+3PL:एछन् #;

!Inflections for Inferential aspect (negative)
 LEXICON InferentialNegative

+PST+INFER+NEG+1SG:एनछु #;
 +PST+INFER+NEG+1PL:एनछौ #;
 +PST+INFER+NEG+2SG+MASC:एनछस् #;
 +PST+INFER+NEG+2SG+FEM:इनछेस् #;
 +PST+INFER+NEG+2SG+MASC+HON:एनछौ #;
 +PST+INFER+NEG+2SG+FEM+HON:इनछौ #;
 +PST+INFER+NEG+2PL:एनछौ #;
 +PST+INFER+NEG+3SG+MASC:एनछ #;
 +PST+INFER+NEG+3SG+FEM:इनछ #;
 +PST+INFER+NEG+3SG+HON:एनछन् #;
 +PST+INFER+NEG+3SG+FEM+HON:इनछन् #;
 +PST+INFER+NEG+3PL:एनछन् #;

!Inflection for imperative mood
 LEXICON Imperative

+IMP+2SG:^IMPsg #; !-/-ई
 +IMP+2SG+HON:^IMPPhon #; !-अ/-ऊ
 +IMP+2PL:^IMPpl #; !-अ/-ओ

! Inflections for optative mood (affirmative)
 LEXICON Optative

+OPT+1SG:ऊँ #;
 +OPT+1PL:औँ#;
 +OPT+2SG:एस् #;
 +OPT+2SG+HON:ए #;
 +OPT+2PL:ए #;

+OPT+3SG:ओस् #;
+OPT+3SG+HON:ऊन् #;
+OPT+3PL:ऊन् #;

! Inflections for potential mood (affirmative)

LEXICON Potential

+POT+1SG:उँला #;
+POT+1PL:औँला #;
+POT+2SG+MASC:लास् #;
+POT+2SG+FEM:लिस् #;
+POT+2SG+MASC+HON:औँला #;
+POT+2SG+FEM+HON:औली #;
+POT+2PL:औँला #;
+POT+3SG+MASC:ला #;
+POT+3SG+FEM:ली #;
+POT+3SG+MASC+HON:लान् #;
+POT+3SG+FEM+HON:लिन् #;
+POT+3PL:लान् #;

!Inflection for participles

LEXICON Participles

+ABS:ई #;
+INF:नु #;
+INF+OBL:ना #;
+INF+EMPH:नै #;
+PURP:न #;
+PURP+EMPH:नै #;
+PROSP:ने #;
+DUR:दा #;
+DUR+EMPH:दै #;
+CONJ:एर #;
+CONJ+EMPH:एरै #;
+CONJ:इकन #;
+CONJ+EMPH:इकनै #;
+COND:ए #;
+PERFT:ए #;

END

7.2.4 Adjectives

Adjectives described, analyzed and classified in (3.4) are implemented in a lexc file. The adjectives are classified into four groups. This lexc file contains the transducers from Figure 3.34 to Figure 3.37.

Multichar_Symbols +ADJ +SG +PL +OBL +HON +FEM +POSIT +COMP +SUPER

LEXICON ROOT

Adjectives;

LEXICON Adjectives

!!O-ending Adjectives

राम्रो inflection_o_ending;

!!Non-o-ending Adjective Type 1

चतुर inflection_non_o_ending1;

!!Non-o-ending Adjective Type 2

न्यून inflection_non_o_ending2;

!!Unmarked adjectives

असल inflection_unmarked;

LEXICON inflection_o_ending

+ADJ+SG:0 #;

+ADJ+PL:+MP #;

+ADJ+OBL:+MP #;

+ADJ+HON:+MP #;

+ADJ+FEM:+FE #;

LEXICON inflection_non_o_ending1

+ADJ+SG:0 #;

+ADJ+FEM:नी+FE #;

LEXICON inflection_non_o_ending2

+ADJ+POSIT:0 #;

+ADJ+COMP:तर #;

+ADJ+SUPER:तम #;

LEXICON inflection_unmarked

+ADJ:0 #;

END

7.2.5 Numerals and classifiers

The numerals analyzed and described in (3.5) and classifiers described and analyzed in (3.6) are implemented in a lexc file. Irregular ordinal numerals are directly encoded and this lexc file contains the transducers from Figure 3.38 to Figure 3.45.

Multichar_Symbols +CARD +ORD +NUM +CLF +PORT +FREQ +FEM +PL +HON +MASC
+HUM +NHUM +OBL +CL +MP +SG

LEXICON ROOT

Numbers;

LEXICON Numbers

!!Cardinal Numbers

पाँच CardOrd;

सात CardOrd;

सय CardOrd;

हजार CardOrd;

लाख CardOrd;

करोड CardOrd;

अरब CardOrd;

खरब CardOrd;

नील CardOrd;

शङ्ख CardOrd;

पद्म CardOrd;

!!Classifier like items

!!O-ending classifiers

कोसो ctag1;

दानो ctag1;

!!Non-o-ending classifiers

पोटी ctag2;

थुन ctag2;

!!Exceptional numbers

एक+NUM+CARD:एक #;

दुई+NUM+CARD:दुई#;

तीन+NUM+CARD:तीन #;

चार+NUM+CARD:चार #;
छ+NUM+CARD:छ #;
नौ+NUM+CARD:नौ #;

!!Exceptional ordinal numerals

एक+NUM+ORD+MASC:पहिलो #;
एक+NUM+ORD+PL:पहिला #;
एक+NUM+ORD+OBL:पहिला #;
एक+NUM+ORD+HON:पहिला #;
एक+NUM+ORD+FEM:पहिली #;
दुई+NUM+ORD+MASC:दोस्रो #;
दुई+NUM+ORD+PL:दोस्रा #;
दुई+NUM+ORD+OBL:दोस्रा #;
दुई+NUM+ORD+HON:दोस्रा #;
दुई+NUM+ORD+FEM:दोस्री #;
तीन+NUM+ORD+MASC:तेस्रो #;
तीन+NUM+ORD+PL:तेस्रा #;
तीन+NUM+ORD+OBL:तेस्रा #;
तीन+NUM+ORD+HON:तेस्रा #;
तीन+NUM+ORD+FEM:तेस्री #;
चार+NUM+ORD+MASC:चौथो #;
चार+NUM+ORD+PL:चौथा #;
चार+NUM+ORD+OBL:चौथा #;
चार+NUM+ORD+HON:चौथा #;
चार+NUM+ORD+FEM:चौथी #;
एक+NUM+ORD:प्रथम #;
दुई+NUM+ORD:द्वितीय #;
तीन+NUM+ORD:तृतीय #;
चार+NUM+ORD:चतुर्थ #;
पाँच+NUM+ORD:पञ्चम #;
छैटौं+NUM+ORD:छैटौं #;
नवौं+NUM+ORD:नवौं #;

!! Frequency numerals

एक+NUM+FREQ:एकोहोरो #;

दुई+NUM+FREQ:दोहोरो #;
 तीन+NUM+FREQ:तेहोरो #;
 एक+NUM+FREQ:एकसरो #;
 दुई+NUM+FREQ:दुईसरो #;
 तीन+NUM+FREQ:तीनसरो #;
 दुई+NUM+FREQ:दोबर #;
 तीन+NUM+FREQ:तेबर #;
 चार+NUM+FREQ:चौबर #;
 दुई+NUM+FREQ:दुईगुना #;
 तीन+NUM+FREQ:तीनगुना #;
 चार+NUM+FREQ:चौगुना #;

!! Portion Numerals

आधा+NUM+PORT:आधा #;
 पौने+NUM+PORT:पौने #;
 सवा+NUM+PORT:सवा #;
 डेढ+NUM+PORT:डेढ #;
 साढे+NUM+PORT:साढे #;
 अढाइ+NUM+PORT:अढाइ #;
 चौथाइ+NUM+PORT:चौथाइ #;

!! Classifiers

जना+CLF+HUM:जना #;
 वटा+CLF+NHUM:वटा #;
 ओटा+CLF+NHUM:ओटा #;
 वटी+CLF+FEM:वटी #;
 ओटी+CLF+FEM:ओटी #;
 LEXICON CardOrd
 +NUM+CARD:0 #;
 +NUM+ORD:औँ #;
 LEXICON ctag1
 +CL+SG:0 #;
 +CL+PL:+MP #;
 LEXICON ctag2
 +CL:0 #;
 END

7.2.6 Adverbs

The adverbs described, analyzed and classified in (5.1) are implemented in a lexc file. Since the adverbs do not inflect, they are classified into semantic classes. This lexc file contains the transducers from Figure 5.1 to Figure 5.7.

Multichar_Symbols +ADV +TEMP +SPAC +AMOUNT +MANNER +FREQ
+REASON +SENT

LEXICON Root

!!! Temporal adverbs

अहिले AdvT;

हिजो AdvT;

!!! Spatial Adverbs

तल AdvS;

त्यहाँ AdvS;

!!! Amount adverbs

धेरै AdvAm;

अझ AdvAm;

!!! Manner adverbs

सुस्तरी AdvMa;

फटाफट AdvMa;

!!! Frequency adverbs

बारम्बार AdvFr;

निरन्तर AdvFr;

!!! Reason adverbs

त्यसकारण AdvRe;

तसर्थ AdvRe;

!!! Sentential adverbs

साँच्चै AdvSe;

स्वाभावतः AdvSe;

LEXICON AdvT

+ADV+TEMP:0 #;

LEXICON AdvS

+ADV+SPAC:0 #;

LEXICON AdvAm

+ADV+AMOUNT:0 #;

LEXICON AdvMa

+ADV+MANNER:0 #;

```

LEXICON AdvFr
+ADV+FREQ:0    #;
LEXICON AdvRe
+ADV+REASON:0  #;
LEXICON AdvSe
+ADV+SENT:0    #;

```

END

7.2.7 Postpositions

Postpositions discussed and analyzed in (5.3) are implemented in a lexc file. Plural marker and case markers are directly encoded whereas adverbial postpositions are implemented through continuations lexicons. This lexc file contains transducers from Figure 5.10 to Figure 5.12.

```

Multichar_Symbols +POSTP +EMPH +ERG +INST +DAT +ABL +LOC +COM
+GEN +DIR +SG +PL +FEM

```

LEXICON ROOT

!!Case Markers which do not take emphatic marker

```

+ERG:ले    #;
+INST:ले   #;
+DAT:लाई #;
+ABL:देखि  #;

```

!!Case marker which take emphatic marker also

```

+ABL:बाट   #;
+ABL+EMPH:बाटै #;
+LOC:मा    #;
+LOC+EMPH:मै #;
+COM:सँग  #;
+COM+EMPH:सँगै #;
+COM:सित  #;
+COM+EMPH:सितै #;
+GEN+SG:को #;
+GEN+PL:का #;
+GEN+FEM:की #;

```

+GEN+EMPH:कै #;
+ALL:तिर #;
+ALL+EMPH:तिरै #;
!!Plural/collective marker
+PL:हरू #;

!!Adverbial postpositions which do not take emphatic marker

माथि tag1;

कहाँ tag1;

!!Adverbial Postpositions which take emphatic marker

सहित tag2;

साथ tag2;

अनुसार tag2;

बाहेक tag2;

LEXICON tag1

+POSTP:0 #;

LEXICON tag2

+POSTP:0 #;

+POSTP+EMPH:कै #;

END

7.2.8 Conjunctions, particles and interjections

The conjunctions analyzed in (5.2) and particles and interjections analyzed in (5.4) are implemented in a lexc file. This lexc file contains the transducers from Figure 5.8 to Figure 5.9 and Figure 5.13 to Figure 5.14.

Multichar_Symbols +PART +INTERJ +CCONJ +SCONJ +PARTICLE

LEXICON Root

!!!समपदिक संयोजकहरू

र Coordinate;

वा Coordinate;

अथवा Coordinate;

या Coordinate;

कि Coordinate;

नकि Coordinate;

अनि Coordinate;

पनि Coordinate;
तथा Coordinate;
एवं Coordinate;
तर Coordinate;
किन्तु Coordinate;
परन्तु Coordinate;

!!!विषमपदिक संयोजकहरू

भन्ने Subordinate;
भनेर Subordinate;
भने Subordinate;
कि Subordinate;
किनभने Subordinate;
किनकि Subordinate;
यसकारण Subordinate;

!!!Particles निपातहरू

नै Particle;
मात्र Particle;
केवल Particle;
चाहिँ Particle;
पनि Particle;
ल Particle;
है Particle;
न Particle;
नि Particle;
त Particle;
पो Particle;
क्या Particle;
के Particle;
कि Particle;
रे Particle;
क्यारे Particle;
हँ Particle;
हगि Particle;

खै	Particle;
लौ	Particle;
हौ	Particle;
क्यार	Particle;
ब्यारे	Particle;
क्या	Particle;
है	Particle;
झैँ	Particle;

!!!Interjections विस्मायादिबोधकहरू

अहा	Interjection;
अहो	Interjection;
ओहो	Interjection;
उहु	Interjection;
उफ	Interjection;
आत्था	Interjection;
आत्थु	Interjection;
आच्छु	Interjection;
छि	Interjection;
धत्	Interjection;
धत्तेरि	Interjection;
थुक्क	Interjection;
थुइक्क	Interjection;
बःड	Interjection;
हाय	Interjection;
कठै	Interjection;
हरे	Interjection;
शिव	Interjection;
च्च	Interjection;
बरा	Interjection;
बिचरा	Interjection;
उस्	Interjection;
हाहा	Interjection;
हिहि	Interjection;

ज्या	Interjection;
ए	Interjection;
ऐ	Interjection;
औ	Interjection;
हौ	Interjection;
ऐय्या	Interjection;
ल	Interjection;
हवस्	Interjection;
अँ	Interjection;
ज्यू	Interjection;
हजुर	Interjection;
हँ	Interjection;
अहँ	Interjection;
नाइँ	Interjection;
कुन्नि	Interjection;
सत्ते	Interjection;
साँच्ची	Interjection;
धरोधर्म	Interjection;
भो	Interjection;
ई	Interjection;
ऊ	Interjection;
वाह	Interjection;
स्याबास	Interjection;
अबुइ	Interjection;
आप्पै	Interjection;
उफ्	Interjection;
ओ	Interjection;
चै	Interjection;
ओइ	Interjection;
एइ	Interjection;
ज्याहै	Interjection;
हि	Interjection;
हा	Interjection;
बःड	Interjection;

```

LEXICON Coordinate
+CCONJ:0 #;
LEXICON Subordinate
+SCONJ:0 #;
LEXICON Particle
+PARTICLE:0 #;
LEXICON Interjection
+INTERJ:0 #;
END

```

7.2.9 Derivations

The derivational process prefixation described and analyzed in (6.1) and suffixation described and analyzed in (6.2) are implemented in lexc file. This lexc file contains the transducers from Figure 6.1 to Figure 6.16.

```

Multichar_Symbols +NOUN PFX+ +ADJ +SFX ^R ^a +ADV

```

```

LEXICON ROOT
!Lexcons for prefixation
  PNtoN;
  PNtoAdj;
  PNtoAdv;
  PAdjtoAdj;
!Lexicons for suffixation
  SNtoN;
  SNtoAdj;
  SNtoNAdj;
  SAdjtoN;
  SAdjNtoN;
  SVtoN;
  SVtoAdj;
  SVtoAdv;
  SAdvtoAdj;
  ConVtoN;
  ConVtoNAdj;
  InsVtoN;
LEXICON PNtoN
PFX+:ऋ PNtoN1;
PFX+:परऱ PNtoN2;
PFX+:अपऱ PNtoN3;
PFX+:सम्ऱ PNtoN4;
PFX+:अनुऱ PNtoN5;
PFX+:अवऱ PNtoN6;

```

PFX+:दुस्	PNtoN7;
PFX+:दुर्	PNtoN8;
PFX+:वि	PNtoN9;
PFX+:अधि	PNtoN10;
PFX+:अति	PNtoN11;
PFX+:अभि	PNtoN12;
PFX+:प्रति	PNtoN13;
PFX+:परि	PNtoN14;
PFX+:उप	PNtoN15;
PFX+:सह	PNtoN16;
PFX+:स	PNtoN17;
PFX+:कु	PNtoN18;
PFX+:अ	PNtoN19;
PFX+:अन्	PNtoN20;
PFX+:बे	PNtoN21;
PFX+:बद	PNtoN22;
PFX+:ला	PNtoN23;
PFX+:सु	PNtoN24;

!!Lexicon of underived nouns -----!!

LEXICON PNtoN1

चलन PNtoNtag;

LEXICON PNtoN2

जय PNtoNtag;

LEXICON PNtoN3

शब्द PNtoNtag;

LEXICON PNtoN4

मान PNtoNtag;

LEXICON PNtoN5

शासन PNtoNtag;

LEXICON PNtoN6

गुण PNtoNtag;

LEXICON PNtoN7

परिणामPNtoNtag;

LEXICON PNtoN8

घटना PNtoNtag;

LEXICON PNtoN9

नाश PNtoNtag;

LEXICON PNtoN10

राज्य PNtoNtag;

LEXICON PNtoN11

वृष्टि PNtoNtag;

LEXICON PNtoN12

रुचि PNtoNtag;

LEXICON PNtoN13

ध्वनि PNtoNtag;

LEXICON PNtoN14

योजना PNtoNtag;

LEXICON PNtoN15

ग्रह PNtoNtag;

LEXICON PNtoN16

कार्य PNtoNtag;

LEXICON PNtoN17

परिवार PNtoNtag;

LEXICON PNtoN18

पुत्र PNtoNtag;

LEXICON PNtoN19

ज्ञान PNtoNtag;

LEXICON PNtoN20

आस्था PNtoNtag;

LEXICON PNtoN21

इज्जत PNtoNtag;

LEXICON PNtoN22

नाम PNtoNtag;

LEXICON PNtoN23

वारिस PNtoNtag;

LEXICON PNtoN24

समाचारPNtoNtag;

!Lexicon for common tag

LEXICON PNtoNtag

+NOUN:0 #;

!!---Noun to adjective derivation -----!!

LEXICON PNtoAdj

PFX+:निर् PNtoAdj1;

PFX+:निः PNtoAdj2;

PFX+:नि PNtoAdj3;

PFX+:वि PNtoAdj4;

PFX+:निस् PNtoAdj5;
PFX+:स PNtoAdj6;
PFX+:बे PNtoAdj7;
PFX+:अ PNtoAdj8;
PFX+:अन PNtoAdj9;

!!Lexicon of underived nouns

LEXICON PNtoAdj1

दोष PNtoAdjtag;

LEXICON PNtoAdj2

स्वार्थ PNtoAdjtag;

LEXICON PNtoAdj3

डर PNtoAdjtag;

LEXICON PNtoAdj4

मुख PNtoAdjtag;

LEXICON PNtoAdj5

फल PNtoAdjtag;

LEXICON PNtoAdj6

बल PNtoAdjtag;

LEXICON PNtoAdj7

घर PNtoAdjtag;

LEXICON PNtoAdj8

मूल्य PNtoAdjtag;

LEXICON PNtoAdj9

मोल PNtoAdjtag;

!!Lexicon for common tag

LEXICON PNtoAdjtag

+ADJ:0 #;

!!-----Noun to adverb derivation -----!!

LEXICON PNtoAdv

PFX+:आ PNtoAdv1;

PFX+:स PNtoAdv2;

PFX+:निर् PNtoAdv3;

PFX+:प्रति PNtoAdv4;

!!Lexicon of underived nouns

LEXICON PNtoAdv1

मरण PNtoAdvtag;

LEXICON PNtoAdv2

हर्ष PNtoAdvtag;
LEXICON PNtoAdv3
घात PNtoAdvtag;
LEXICON PNtoAdv4
हसा PNtoAdvtag;

!!Lexicon for common tag
LEXICON PNtoAdvtag
+ADV:0 #;

!!-----Adjective to adjective derivation -----!!
LEXICON PAdjtoAdj
PFX+:सम् PAdjtoAdj1;
PFX+:वि PAdjtoAdj2;
PFX+:दुर् PAdjtoAdj3;
PFX+:उन् PAdjtoAdj4;
PFX+:सु PAdjtoAdj5;
PFX+:परि PAdjtoAdj6;

!!Lexicon of underived nouns
LEXICON PAdjtoAdj1
पूर्ण PAdjtoAdjtag;
LEXICON PAdjtoAdj2
शुद्ध PAdjtoAdjtag;
LEXICON PAdjtoAdj3
भेद्य PAdjtoAdjtag;
LEXICON PAdjtoAdj4
मुक्त PAdjtoAdjtag;
LEXICON PAdjtoAdj5
शिक्षित PAdjtoAdjtag;
LEXICON PAdjtoAdj6
पूर्ण PAdjtoAdjtag;

!!Lexicon for common tag
LEXICON PAdjtoAdjtag
+ADJ:0 #;

!! Suffixation
!!Noun to Noun Derivation
LEXICON SNtoN
!!Nountype1
सुन SNtoN1;
!!Nountpe2

घाँस SNtoN2;
LEXICON SNtoN1
+SFX:आर SNtoNtag;
LEXICON SNtoN2
+SFX:ई SNtoNtag;
LEXICON SNtoNtag
+NOUN:0 #;

!! Noun to adjective derivation

LEXICON SNtoAdj

!!Nountype1

दया SNtoAdj1;

!!Nountpe2

लाभ SNtoAdj2;

!!Nountpe3

सेवा SNtoAdj3;

!!Nountpe4

मुगल SNtoAdj4;

!!Nountpe5

लिम्बु SNtoAdj5;

!!Nountpe6

दान SNtoAdj6;

!!Nountpe7

खर्च SNtoAdj7;

!!Nountpe8

भिर SNtoAdj8;

!!Nountpe9

रिस SNtoAdj9;

!!Nountpe10

शहर SNtoAdj10;

!!Nountpe11

होस SNtoAdj11;

LEXICON SNtoAdj1

+SFX:अनीय SNtoAdjtag;

LEXICON SNtoAdj2

+SFX:अक SNtoAdjtag;

LEXICON SNtoAdj3

+SFX:इका SNtoAdjtag;

LEXICON SNtoAdj4

+SFX:आन SNtoAdjtag;

LEXICON SNtoAdj5

+SFX:वान SNtoAdjtag;
 LEXICON SNtoAdj6
 +SFX:ई SNtoAdjtag;
 LEXICON SNtoAdj7
 +SFX:आलु SNtoAdjtag;
 LEXICON SNtoAdj8
 +SFX:आलो SNtoAdjtag;
 LEXICON SNtoAdj9
 +SFX:आहा SNtoAdjtag;
 LEXICON SNtoAdj10
 +SFX:इया SNtoAdjtag;
 LEXICON SNtoAdj11
 +SFX:इयार SNtoAdjtag;
 LEXICON SNtoAdjtag
 +ADJ:0 #;

!!----Noun to noun/adjective derivation ----!!

LEXICON SNtoNAdj
 !!Nountype1
 झापा SNtoNAdj1;
 !!Nountpe2
 गुल्मी SNtoNAdj2;
 !!Nountpe3
 इलाम SNtoNAdj3;
 !!Nountpe4
 गाउँ SNtoNAdj4;
 !!Nountpe5
 नेपाल SNtoNAdj5;
 LEXICON SNtoNAdj1
 +SFX:ली SNtoNAdjtag1;
 LEXICON SNtoNAdj2
 +SFX:एली SNtoNAdjtag1;
 LEXICON SNtoNAdj3
 +SFX:ए SNtoNAdjtag1;
 LEXICON SNtoNAdj4
 +SFX:ले SNtoNAdjtag1;
 LEXICON SNtoNAdj5
 +SFX:ई SNtoNAdjtag1;
 LEXICON SNtoNAdjtag
 +NOUN:0 #;
 LEXICON SNtoNAdjtag1
 +NOUN:0 #;

+ADJ:0 #;

!!-----Adjective to noun derivation -----!!

LEXICON SAdjtoN

!!Nountype1

लामो SAdjtoN1;

छोटो SAdjtoN1;

!!Nountpe2

!!xy SAdjtoN2;

LEXICON SAdjtoN1

+SFX:आइ SAdjtoNtag;

LEXICON SAdjtoNtag

+NOUN:0 #;

!!----- Adjective/noun to noun derivation -----!!

LEXICON SAdjNtoN

!!Nountype1

गरिब SAdjNtoN1;

LEXICON SAdjNtoN1

+SFX:ई SAdjNtoNtag1;

LEXICON SAdjNtoNtag

+NOUN:0 #;

+ADJ:0 #;

LEXICON SAdjNtoNtag1

+NOUN:0 #;

!!----- Verb to noun derivation -----!!

LEXICON SVtoN

!!verbtype1

चुन् SVtoN1;

!verbtype2

चुन् SVtoN2;

!!verbtype3

किट् SVtoN3;

!!verbtype4

किट् SVtoN4;

!!verbtype5

ढाक् SVtoN5;

!!verbtype6

जल् SVtoN6;

!!verbtype7

चोर् SVtoN7;

!!verbtype8

हाँस् SVtoN8;

!!verbtype9
पढ् SVtoN9;
!!verbtype10
थाक् SVtoN10;
!!verbtype11
छाप् SVtoN11;
!!verbtype12
छान् SVtoN12;
!!verbtype13
चिच्या SVtoN13;
!!verbtype14
झर् SVtoN14;
!!verbtype15
ढोग् SVtoN15;
!!verbtype16
राख् SVtoN16;
!!verbtype17
दाब् SVtoN17;
!!verbtype18
बच् SVtoN18;
!!verbtype19
सङ् SVtoN19;
!!verbtype20
रोप् SVtoN20;
!!verbtype21
छेक् SVtoN21;
!!verbtype22
चिर् SVtoN22;
!!verbtype23
बढ् SVtoN23;
!!verbtype24
सर् SVtoN24;
!!verbtype25
उर् SVtoN25;
!!verbtype26
चाल् SVtoN26;
!!verbtype27
बेर् SVtoN27;
!!verbtype28
गा SVtoN28;
!!verbtype29

भिङ् SVtoN29;
 !!verbtype30
 जित् SVtoN30;
 !!verbtype31
 कोर् SVtoN31;
 !!verbtype32
 खुल् SVtoN32;
 LEXICON SVtoN1
 +SFX:आउ SVtoNtag;
 LEXICON SVtoN2
 +SFX:आब SVtoNtag;
 LEXICON SVtoN3
 +SFX:आनी SVtoNtag;
 LEXICON SVtoN4
 +SFX:आनी SVtoNtag;
 LEXICON SVtoN5
 +SFX:अनी SVtoNtag;
 LEXICON SVtoN6
 +SFX:अन SVtoNtag;
 LEXICON SVtoN7
 +SFX:ई SVtoNtag;
 LEXICON SVtoN8
 +SFX:ओ SVtoNtag;
 LEXICON SVtoN9
 +SFX:आइ SVtoNtag;
 LEXICON SVtoN10
 +SFX:आवट SVtoNtag;
 LEXICON SVtoN11
 +SFX:आ SVtoNtag;
 LEXICON SVtoN12
 +SFX:ओट SVtoNtag;
 LEXICON SVtoN13
 +SFX:हट SVtoNtag;
 LEXICON SVtoN14
 +SFX:अना SVtoNtag;
 LEXICON SVtoN15
 +SFX:आउनी SVtoNtag;
 LEXICON SVtoN16
 +SFX:आलो SVtoNtag;
 LEXICON SVtoN17

+SFX:आब SVtoNtag;
 LEXICON SVtoN18
 +SFX:अत SVtoNtag;
 LEXICON SVtoN19
 +SFX:अल SVtoNtag;
 LEXICON SVtoN20
 +SFX:आई SVtoNtag;
 LEXICON SVtoN21
 +SFX:आरो SVtoNtag;
 LEXICON SVtoN22
 +SFX:औटो SVtoNtag;
 LEXICON SVtoN23
 +SFX:औती SVtoNtag;
 LEXICON SVtoN24
 +SFX:उवा SVtoNtag;
 LEXICON SVtoN25
 +SFX:ती SVtoNtag;
 LEXICON SVtoN26
 +SFX:नी SVtoNtag;
 LEXICON SVtoN27
 +SFX:नो SVtoNtag;
 LEXICON SVtoN28
 +SFX:ना SVtoNtag;
 LEXICON SVtoN29
 +SFX:अन्त SVtoNtag;
 LEXICON SVtoN30
 +SFX:औरी SVtoNtag;
 LEXICON SVtoN31
 +SFX:एसो SVtoNtag;
 LEXICON SVtoN32
 +SFX:अस्त SVtoNtag;
 LEXICON SVtoNtag
 +NOUN:0 #;

!!---- Verb to adjective derivation ----!!
 LEXICON SVtoAdj
 !!verdtype1
 मिच् SVtoAdj1;
 !!verdtype2
 भुल् SVtoAdj2;
 !!verdtype3

पोस् SVtoAdj3;
 !!verbtype4
 घुम् SVtoAdj4;
 !!verbtype5
 घुम् SVtoAdj5;
 !!verbtype6
 खप् SVtoAdj6;
 !!verbtype7
 पढ् SVtoAdj7;
 !!verbtype8
 छाड् SVtoAdj8;
 !!verbtype9
 रोप् SVtoAdj9;
 !!verbtype10
 सिक् SVtoAdj10;
 !!verbtype11
 बिक् SVtoAdj11;
 !!verbtype12
 भाग् SVtoAdj12;
 !!verbtype13
 छेर SVtoAdj13;
 !!verbtype14
 लाग् SVtoAdj14;
 LEXICON SVtoAdj1
 +SFX:आहा SVtoAdjtag;
 LEXICON SVtoAdj2
 +SFX:अक्कड SVtoAdjtag;
 LEXICON SVtoAdj3
 +SFX:इलो SVtoAdjtag;
 LEXICON SVtoAdj4
 +SFX:अन्ते SVtoAdjtag;
 LEXICON SVtoAdj5
 +SFX:अन्ता SVtoAdjtag;
 LEXICON SVtoAdj6
 +SFX:आलु SVtoAdjtag;
 LEXICON SVtoAdj7
 +SFX:ऐया SVtoAdjtag;
 LEXICON SVtoAdj8
 +SFX:आ SVtoAdjtag;
 LEXICON SVtoAdj9

+SFX:अर SVtoAdjtag;
LEXICON SVtoAdj10
+SFX:अरु SVtoAdjtag;
LEXICON SVtoAdj11
+SFX:आउ SVtoAdjtag;
LEXICON SVtoAdj12
+SFX:औटो SVtoAdjtag;
LEXICON SVtoAdj13
+SFX:औटी SVtoAdjtag;
LEXICON SVtoAdj14
+SFX:उ SVtoAdjtag;
LEXICON SVtoAdjtag
+ADJ:0 #;

!!----- Verb to adverb derivation -----!!
LEXICON SVtoAdv
!!verbtype1
गर SVtoAdv1;
LEXICON SVtoAdv1
+SFX:उन्जेल SVtoAdvtag;
+SFX:इन्जेल SVtoAdvtag;
LEXICON SVtoAdvtag
+ADV:0 #;

!!----- Adverb to adjective derivation -----!!
LEXICON SAdvtoAdj
भिन्न SAdvtoAdj1;
LEXICON SAdvtoAdj1
+SFX:ई SAdvtoAdjtag;
LEXICON SAdvtoAdjtag
+ADJ:0 #;

!!----- Verb to noun conversion -----!!
LEXICON ConVtoN
!!verbtype1
खेल् ConVtoNtag;
खोज् ConVtoNtag;
LEXICON ConVtoNtag
+NOUN:^R #;

!!----- Verb to noun/adjective conversion -----!!
LEXICON ConVtoNAdj
!!verbtype1
ठग् ConVtoNAdjtag;

```

चोर् ConVtoNAdjtag;
थप् ConVtoNAdjtag;
LEXICON ConVtoNAdjtag
+NOUN:^R #;
+ADJ:^R #;

```

```

!!----- Verb to noun derivation (by vowel insertion) -----!!
LEXICON InsVtoN
!!verbtype1
चम्क InsVtoNtag;
सम्झ InsVtoNtag;
टल्क InsVtoNtag;
LEXICON InsVtoNtag
+NOUN:^a #;
END

```

7.3 Realization: rules of alternations

When the lexc files are compiled into an finite state transducer, the upper language contains sequence of stem (citation form) and morphological tags and the lower language contains actual spelling of the stems and affixes. At the same time, there may be some of the arbitrary tags used for creating some sorts of environment for the application of rules. The rules demonstrated right after each figure of finite state transducer are respectively collected and put them into certain order. The required variables are defined and these rules are composed into a single finite state transducer. Finally the rules are applied to the lower language of the lexicon finite state transducer. Each category of the word class is treated separately in the subsequent sections.

7.3.1 Phonological rules for nouns

clear

```
define cons |गाघड|चाछ|जझ|जटाठाडाढाणतथदाधानापाफबाभामायरलावासाषशह;
```

```
define liquids रल;
```

```
define change [[ो %+MP -> ा || _ .#.]
```

.o.

```
[ो %+FE -> ी || _ .#.]
```

.o.

[ा -> [] || _ ?* %^b ि न ी .#.]

.o.

[ा -> [] || _ े न ी .#.]

.o.

[ा -> [] || _ ि न ी .#.]

.o.

[ा -> [] || _ ी .#.]

.o.

[ी -> ् || liquids _ न ी .#.]

.o.

[ी -> ि || _ न ी .#.]

.o.

[[. .] -> ् || cons _ न ी .#.]

.o.

[य ा -> [] || _ न ी .#.]

.o.

[ी -> [] || _ ि न ी .#.]

.o.

[ू -> ु || _ ?* %^b ि न ी .#.]

.o.

[ी -> [] || _ %^b ि न ी .#.]

.o.

[%^b -> []]

.o.

[ि -> [] || ु _ न ी .#.]

];¹

¹ When more nouns are included, new rules, if any, can be added to this set of rules.

```

read lexc <nouns.txt
define nouns;
read regex nouns .o. change;
save stack nouns.fst

```

7.3.2 Phonological rules for pronouns

Since the pronouns are limited in their numbers and idiosyncratic in their behaviors, they are directly encoded in the lexc file. Because of this reason there are no orthographic rules for pronouns. The script file contains only commands to compile the file of pronouns as follows.

```

clear
read lexc <pronouns.txt;
save stack pronouns.fst

```

7.3.3 Phonological rules for verbs

```

clear
define imperfect {दो}{दी}{दा}{दै};
define habinfl {थै}{थ्यौ}{थिस}{थ्यौ}{थ्यो}{थि}{थे}{थिन्};
define habneginfl {दैनथै}{दैनथ्यौ}{दैनथिस}{दैनथ्यौ}{दिनथिस}{दैनथ्यो}{दिनथे};
define npastinfl {छु}{छौ}{छस}{छेस}{छौ}{छ्यौ}{छा}{छे}{छन्}{छिन्};
define npastneginfl {दिनँ}{दैनौँ}{दैनस}{दैनौ}{दिनौ}{दैन}{दिन}{दैनन्}{दिनन्};
define cons कागघडाचछजझञटठडढणताथदधानपाफाबाभामायारलावासाषशह;
define change [[ि -> ् य || _ %^ia ]
.o.
[् -> ा || cons _ cons ?* %^a ]
.o.
[ि %^a -> ् ]
.o.
[ा -> [ ] || cons _ ?* %^ia ]

```

.o.

[ा -> [] || cons _ ?* %^ta]

.o.

[ँ %^ia -> ा]

.o.

[ँ %^ta -> ा]

.o.

[%^ia -> ा]

.o.

[ि %^ta -> ा]

.o.

[%^IMPsg -> [], %^IMPpl -> ऊ, %^IMPhon -> ओ || ा _.#.]

.o.

[ँ -> [] || cons _ %^IMPhon]

.o.

[ँ -> [] || cons _ %^IMPpl]

.o.

[ि -> [] || cons _ %^IMPhon]

.o.

[ि -> [] || cons _ %^IMPpl]

.o.

[%^IMPsg -> [] || ँ _.#.]

.o.

[%^IMPsg -> [] || _.#.]

.o.

[%^IMPpl -> [] || _.#.]

.o.

[%^IMPhon -> [] || _.#.]

.0.

[् आ -> ा || cons _]

.0.

[् इ -> ि || cons _]

.0.

[् ई -> ी || cons _]

.0.

[् उ -> ु || cons _]

.0.

[् ऊ -> ू || cons _]

.0.

[् ए -> े || cons _]

.0.

[् ओ -> ो || cons _]

.0.

[् औ -> ौ || cons _]

.0.

[इ इ -> इ]

.0.

[इ ई -> ई]

.0.

[[. .] -> उ ँ || ा _ npastinfl|npastneginfl1|habinfl|habneginfl|imperfect .#.]

.0.

[[. .] -> न् || िइ _ npastinfl|habinfl|imperfect .#.]

.0.

[i -> इ]

.0.

```
[ ् इ -> ि || cons _ ]
```

```
.o.
```

```
[इ इ -> इ ]
```

```
];2
```

```
read lexc <verbinflexions.txt
```

```
eliminate flag NEG
```

```
define verbs;
```

```
read regex verbs .o. change;
```

```
save stack verbs.fst
```

7.3.4 Phonological rules for adjectives

```
clear
```

```
define cons कगघङचछजझञटठडढणतथादधनपफबभमयारलावसषशह;
```

```
define liquids रल;
```

```
define change [[ो %+MP -> ा || _ .#.]
```

```
.o.
```

```
[ो %+FE -> ी || _ .#.]
```

```
.o.
```

```
[[. .] -> ् || liquids _ न ी .#.]
```

```
.o.
```

```
[य ा -> [ ] || _ न ी .#.]
```

```
.o.
```

```
[ा -> [ ] || _ ?* न ी .#.]
```

```
.o.
```

```
[[. .] -> ि || cons _ न ी .#.]
```

```
.o.
```

```
[ि -> [ ] || ध _ न ी .#.]
```

```
];3
```

² More rules may come up when more verbs will be added.

³ When more adjectives are added to lexc file, it may require some other rules.

```

read lexc <adjectives.txt
define aadjectives;
read regex aadjectives .o. change;
save stack adjectives.fst

```

7.3.5 Phonological rules for adverbs

Nepali adverbs do not inflect therefore there is no orthographic rules involved. Therefore, this file contains only commands to compile the adverb lexc file into a network.

```

clear
read lexc <adverbs.txt
save stack adverbs.fst

```

7.3.6 Phonological rules for postpositions

```

clear
define change [[[ॐ -> [] || _ ै .#.]
                .o.
                [॑ -> [] || _ ै .#.]
                .o.
                [॒ -> [] || _ ै .#.]
                .o.
                [॒ -> [] || _ ै .#.]
                ];
read lexc <postpositions.txt
define postpositions;
read regex postpositions .o. change;
save stack postpositions.fst

```

7.3.7 Phonological rules for particles and interjections

Nepali particles and interjections also do not inflect for anything; therefore, there is no orthographic rule involved. Therefore, the script file contains only commands to compile the file into a network.

```
clear
read lexc <conjunctions.txt;
save stack conjunctions.fst
read lexc <particlesinterjections.txt;
save stack particlesinterjections.fst
```

7.3.8 Phonological rules for numerals and classifiers

```
clear
define cons काखागाघडाचछ्जझञटठडढणातथादधानपाफबभमयारलावासाषशह;
define change      [[औ -> ौ || cons _ ]
                    .o.
                    [ो %+MP -> ा || _ .#.]
                    ];
read lexc <numerals.txt
define numerals;
read regex numerals .o. change;
save stack numeralclassifiers.fst
```

7.3.9 Phonological rules for derivations

```
clear
define con
काखागाघडाचछ्जझञटठडढणातथादधानपाफबभमयारलावासाषशह;
define change [[् आ -> ा]
              .o.
              [् ओ -> ो]
              .o.
              [् इ -> ि]
              .o.
```

```

[् ई -> ी]
.o.
[् उ ->ु]
.o.
[् ऊ ->ू]
.o.
[् ए -> े]
.o.
[् ऐ -> ै]
.o.
[् औ -> ौ]
.o.
[् अ -> []]
.o.
[अ -> [] || con _]
.o.
[ए -> े, इ -> ि || con _]
.o.
[ी ए -> े || _ ल ी]
.o.
[आ -> ा || con _]
.o.
[ो आ -> ा || con _]
.o.
[ई -> ी || con _]
.o.
[ा इ -> ि || con _ con]
.o.
[ा अ -> [] || con _]
.o.
[ा -> [] || .#. con _?* इ .#.]
.o.
[् -> [] || _?* %^a]
.o.
[् %^R -> []]
.o.
[%^a -> []]
];

```

```

read lexc <Derivation.txt
define derivation;
read regex derivation .o. change;
save stack derivation.fst

```


Each lexicon implemented in (7.2) is compiled into a network and orthographic rules are applied by composing lexicon transducer with rule transducer. The finite state transducers obtained from (7.3) are combined together to make a single lexicon transducer which is ready to be used as morphological analyzer. While applying the orthographic rules, xfst script files are used in each case and the corresponding lexicon is saved in a .fst file. And finally, all the lexicons are unioned through script file as follows.

```
clear
read regex    nouns
              |pronouns
              |verbs
              |adjectives
              |adverbs
              |postpositions
              |numeralclassifiers
              |conjunctions
              |particlesinterjections
              |derivation;
save stack nepali-lexicon.fst
```

7.4 Summary

In this chapter, the implementations of analyzed morphological categories and grammatical categories have been performed. The categories have been implemented in lexc file and phonological rules have been implemented in rule file. Both of them have been separately compiled and composed into a single lexicon file. Therefore, the final product of the implementation has been saved in a binary file 'nepal-lexicon.fst' and it can be used as morphological analyzer and generator for Nepali.

CHAPTER 8

SUMMARY AND CONCLUSION

This study is the computational analysis of morphology in Nepali and development of a computational model for natural language processing. In this chapter, we try to summarize the findings of the study and recommend some potential area where further research can be carried out.

In chapter 1, we discussed the objectives, methodology, significance, limitations and general concept of the study.

In chapter 2, we presented the theoretical framework that has been employed for the implementation. The finite state transducer is the main tool for the development of morphological analyzer and the implementation has been based on Beesley and Karttunen (2003) in particular.

In chapter 3, we analyzed nouns, pronouns, adjectives, numerals and classifiers in Nepali with reference to their corresponding characteristic features. The nouns have been described with features such as significant stem finals, number, gender, form, honorificity and augmentative/diminutive. On the basis of these features they have been classified into fourteen different groups.

Pronouns in Nepali have been discussed and described with reference to features such as person, number, form and honorificity. On the basis of the features, the pronouns are grouped into eight groups.

Adjectives in Nepali have been analyzed with reference to the characteristic features such as significant stem finals, number, gender, form, honorificity and degree. On the basis of features that they possess, the adjectives are classified into four groups.

Numerals in Nepali basically are of two kinds: cardinals and ordinals. The numerals are grouped into three groups. The classifiers in Nepali are found to be of two types: true classifiers and quasi-classifiers. Former class differentiates human vs. non-human and also has feminine form. But each classifier in latter group classifies a small set of nouns and these classifiers inflect for number.

In chapter 4, we analyzed the verbs in Nepali. The verb stems are analyzed and discussed with reference to various characteristic features such as stem finals, transitivity, syllabicity, sound *a*, causativization/transitivization, and negativization. On the basis of stem final segments, the verb stems in Nepali are vowel final stems and consonant final stems. In terms of transitivity, there are intransitive verb stems and transitive/ditransitive verb stems. From syllabicity point of view, the verb stem in Nepali are of two types: monosyllabic and polysyllabic. Sound *a* in Nepal verb stem is significant in causativization. Causativization in Nepali verb stems is a productive process, except few verb stems. Passivization is the most regular morphological process in Nepali verb stems. Both transitive and intransitive verb stems can be passivized. Negativization occurs by two ways: prefixation and suffixation. On the basis of these features, the verb stems in Nepali are grouped into ten groups. Inflectional paradigms for tenses, aspects and moods are analyzed for both affirmation and negation.

In chapter 5, we discussed adverbs, conjunctions, postpositions and particles. These categories do not inflect for any grammatical features, except emphasis in some cases. Therefore, they are simply grouped into various classes for computational purpose. The adverbs are grouped into seven groups, conjunctions into two groups, postpositions into three groups and particles into one group.

In chapter 6, we have presented the derivational process of morphology in Nepali. The derivation in Nepali is primarily of two kinds: prefixational and suffixational. The prefixational derivations are noun to noun, noun to adjective, noun to adverb and adjective to adjective derivations. The suffixational derivations are noun to noun, noun to adjective, noun to noun/adjective, adjective to noun, adjective/noun to noun, verb to noun, verb to adjective, verb to adverb and adverb to adjective derivations. In addition, there are some cases where verb to noun and verb to adjective derivations take place through conversion and verb to noun derivation by vowel insertion.

We presented the implementation of analyzed categories into a finite state transducer using Xerox Finite State Toolkit in chapter 7. First, all forms of nouns, pronouns, adjectives, numerals, classifiers, adverbs, postpositions, conjunctions, particles and interjections have been implemented in separate lexc files. The rules identified have been implemented in xfst script files respectively. The finite state transducers of each category and finite state transducers of rules for respective categories are composed

separately. All the finite state transducers have been unioned together resulting into a single lexicon finite state transducer which can be used as morphological analyzer and generator.

At the end, we like to recommend that the compounding and reduplication processes can also be analyzed and implement using the same techniques. Furthermore, this study is confined only to a model, thus, one can investigate Nepali morphology for designing a complete morphological analyzer.

ANNEXES

Annex –1: Devanagari – IPA

The subset of Devanagari scripts used in Nepali writing system and IPA equivalents

S.N.	Devnagari	IPA	Remarks
1.	अ	ʌ	
2.	आ ा	a	
3.	इ ि	i	
4.	ई ी	i:	
5.	उ उ	u	
6.	ऊ ू	u:	
7.	ए े	e	
8.	ऐ ै	ʌi	
9.	ओ ो	o	
10.	औ ौ	ʌu	
11.	क्	k	
12.	ख्	k ^h	
13.	ग्	g	
14.	घ्	g ^h	
15.	ङ्	ŋ	
16.	च्	ts	
17.	छ्	ts ^h	
18.	ज्	dz	
19.	झ्	dz ^h	
20.	ञ्	ɲ	
21.	ट्	t̪	
22.	ठ्	t̪ ^h	
23.	ड्	d̪	
24.	ढ्	d̪ ^h	
25.	ण्	ɳ	
26.	त्	t	
27.	थ्	t ^h	
28.	द्	d	

29.	ध्	d ^h	
30.	न्	n	
31.	प्	p	
32.	फ्	p ^h	
33.	ब्	b	
34.	भ्	b ^h	
35.	म्	m	
36.	य्	j	
37.	र्	r	
38.	ल्	l	
39.	व्	w	
40.	श्	ʃ	
41.	ष्	ʂ	
42.	स्	s	
43.	ह्	h	

Annex -2

Nepali Nouns: Sample

Nouns

Type 1a Nouns:

केटो	<i>keṭo</i>	'boy'
चेलो	<i>tseḷo</i>	'disciple'
छोरो	<i>tsʰoro</i>	'son'
छौँडो	<i>tsʰũḍo</i>	'derogatory word for child'
नाठो	<i>naṭʰo</i>	'boy friend [soically unacceptable]'
बच्चो	<i>baṭtso</i>	'child'
बेहुलो	<i>beḥulo</i>	'bridegroom'
बोक्सो	<i>bokso</i>	'male witch'
रण्डो	<i>raṇḍo</i>	'male prostitute'
रौँडो	<i>rũḍo</i>	'widower'

==End of Type 1a Nouns==

Type 1b Nouns:

मुसो	<i>muso</i>	'rat, mouse'
कोरलो	<i>korɽo</i>	'adult calf'
घोडो	<i>gʰoḍo</i>	'horse'
च्याङ्ग्रो	<i>tsjaŋgro</i>	'a variety of sheep'
छाउरो	<i>tsʰauro</i>	'male puppy'
छुचुन्द्रो	<i>tsʰuṣundro</i>	'mole'
छेपारो	<i>tsʰeparo</i>	'lizard'
पाडो	<i>paḍo</i>	'buffalo calf'
बाख्रो	<i>baḵʰro</i>	'goat'
बाछ्रो	<i>baṣʰo</i>	'he-calf'
बिरालो	<i>biralo</i>	'cat'
भँगरो	<i>bʰŋgero</i>	'sparrow'
भेडो	<i>bʰeḍo</i>	'sheep'
भ्यागुतो	<i>bʰjaɣuto</i>	'frog'
भमरो	<i>bʰɽmaɽo</i>	'bumble-bee'

==End of Type 1b Nouns==

Type 1c Nouns:

डालो	<i>ḍalo</i>	'basket'
कुचो	<i>kutso</i>	'broom'
कोठो	<i>koṭʰo</i>	'room'

कोदालो	<i>kodalo</i>	'spade'
खोलसो	<i>k^holso</i>	'rivulet'
गाग्रो	<i>gagro</i>	'water vase'
घुम्टो	<i>g^humṭo</i>	'veil'
घैटो	<i>g^hḷṭo</i>	'earthen water-pot'
चक्को	<i>tsakko</i>	'wheel'
चम्चो	<i>tsamtso</i>	'spoon'
चालनो	<i>tsalno</i>	'sieve'
चिम्टो	<i>tsimṭo</i>	'tongs'
चुङ्को	<i>tsuḍko</i>	'a brief time'
चुरो	<i>tsuro</i>	'bangle'
चुल्ठो	<i>tsult^ho</i>	'plait'
चुल्हो	<i>tsulho</i>	'stove'
जिभ्रो	<i>dzib^hro</i>	'tongue'
जिलो	<i>dzilo</i>	'seed money'
टपरो	<i>ṭaparo</i>	'leaf-plate'
थैलो	<i>t^hailo</i>	'bag'
नाभो	<i>nab^ho</i>	'navel'

==End of Type1c Nouns==

Type 1d Nouns:

फोटो	<i>p^hoto</i>	'photo'
अंगालो	<i>ḷgalo</i>	'arms'
अगुल्ठो	<i>ḷgulṭo</i>	'a piece of burning firewood'
अगेनो	<i>ḷgeno</i>	'fireplace'
अचानो	<i>ḷtsano</i>	'a piece of wood to cut things'
अछेतो	<i>ḷts^heto</i>	'auspicious rice-grain'
अण्डो	<i>ḷndo</i>	'egg'
बोको	<i>boko</i>	'he-goat'
अम्रिसो	<i>ḷmriso</i>	'a plant to make brooms'
आँकुरो	<i>ḷkuro</i>	'new shoot'
आगो	<i>ago</i>	'fire'
आटो	<i>aṭo</i>	'maize rice'
उकेरो	<i>ukero</i>	'heaping of earth around a plant'
ओथारो	<i>ot^haro</i>	'hatching'
कटेरो	<i>kḷtero</i>	'cattle-shad'
कटौरो	<i>kḷṭauro</i>	'bowl'

कमिलो	<i>kʌmilo</i>	'ant'
कमेरो	<i>kʌmero</i>	'white clay'
कर्कलो	<i>kʌrkʌlo</i>	'arum leaf'
काँक्रो	<i>kãkro</i>	'cucumber'
काँडो	<i>kãdo</i>	'thorn'
कात्रो	<i>katro</i>	'clothe to cover corpse'
कान्लो	<i>kanlo</i>	'slopy land'
किलो	<i>kilo</i>	'nail'
कीरो	<i>ki:ro</i>	'insect'
कुइनो	<i>kuino</i>	'elbow'
कुइरो	<i>kuiro</i>	'cloud'
कुटो	<i>kuʈo</i>	'a small hoe'
कुट्को	<i>kuʈko</i>	'a piece'
कुनो	<i>kuno</i>	'corner'
कुरो	<i>kuro</i>	'matter'
कुलो	<i>kulo</i>	'canal'
केरो	<i>kerō</i>	'banana'
कोइलो	<i>koilo</i>	'charcoal'
कोक्रो	<i>kokro</i>	'cradle'
कोदो	<i>kodo</i>	'millet'
कोपिलो	<i>kopilo</i>	'flower bud'
कोयो	<i>kojo</i>	'stone of mango'
कोसो	<i>koso</i>	'pod'
खजुरो	<i>kʰʌdzuro</i>	'centipade'
खबटो	<i>kʰʌbʌʈo</i>	'broken piece of earthen pot'
खाल्डो	<i>kʰʌldo</i>	'ditch'
खास्टो	<i>kʰʌstō</i>	'shawl'
खुट्टो	<i>kʰuʈʈo</i>	'leg'
खोटो	<i>kʰoʈo</i>	'pine resin'
खोप्रो	<i>kʰopro</i>	'hut'
खोयो	<i>kʰojo</i>	'maize cob'
गँगटो	<i>gãgʌʈo</i>	'crab'
गँड्यौलो	<i>gãdʒʌulo</i>	'earthworm'
गमलो	<i>gʌmʌlo</i>	'flower vase'
खोलो	<i>kʰolo</i>	'river'
गारो	<i>garo</i>	'wall'
गालो	<i>galo</i>	'face skin'

गुच्छो	<i>guts^ho</i>	'bunch'
गुनासो	<i>gunaso</i>	'complaint'
गुफो	<i>gup^ho</i>	'cave'
गेडो	<i>gedo</i>	'grain of corn'
घट्टो	<i>g^hʌt̥to</i>	'water-mill'
घारो	<i>g^haro</i>	'hive'
घुँडो	<i>g^hũdo</i>	'knee'
घैलो	<i>g^hʌilo</i>	'water pot'
चमेरो	<i>tsʌmero</i>	'bat'
चिउँडो	<i>tsiũdo</i>	'chin'
चिनो	<i>tsino</i>	'token of love'
चुर्नो	<i>tsurno</i>	'ring worm'
चोलो	<i>tsolo</i>	'blouse'
चौतारो	<i>tsʌutaro</i>	'resting place'
छहरो	<i>ts^hʌfiʌro</i>	'water fall'
छातो	<i>ts^hato</i>	'umbrella'
छानो	<i>ts^hano</i>	'roof'
छाप्रा	<i>ts^hapro</i>	'hut'
छेस्को	<i>ts^hesko</i>	'piece of stick'
जुम्नो	<i>dzumro</i>	'louse'
जरो	<i>dzʌro</i>	'root'
जाँतो	<i>dzãto</i>	'grindstone'
जुको	<i>dzuko</i>	'leech'
झिँगो	<i>dz^hĩgo</i>	'housefly'
झिक्रो	<i>dz^hikro</i>	'dried twig'
झिल्को	<i>dz^hilko</i>	'sparkle'
झुरो	<i>dz^huro</i>	'fine and dried fibres to make fire'
टाउको	<i>ʈauko</i>	'head'
टाटो	<i>ʈaʈo</i>	'scar'
टालो	<i>ʈalo</i>	'piece of clothe'
टुक्रो	<i>ʈukro</i>	'piece'
टुप्पो	<i>ʈuppa</i>	'tip'
ठोसो	<i>ʈ^hoso</i>	'piece of wood'
डढेलो	<i>ɖʌɖ^helo</i>	'forest fire'
डण्डो	<i>ɖʌɳdo</i>	'large stick'
डल्लो	<i>ɖʌllo</i>	'clod'
डिको	<i>ɖiko</i>	'line over letter'

डेरो	<i>dero</i>	'rented room'
ढिस्को	<i>d^hisko</i>	'hillock'
ढुङ्गो	<i>d^hungo</i>	'stone'
ढुटो	<i>d^hu^o</i>	'flour'
तलो	<i>talo</i>	'house-story'
ताँबो	<i>tābo</i>	'copper'
तिघ्रो	<i>tig^hro</i>	'thigh'
तुवाँलो	<i>tuwālo</i>	'fog'
थाप्लो	<i>t^haplo</i>	'upper part of head'
थिग्रो	<i>t^higro</i>	'settled mud'
थुँगो	<i>t^hũgo</i>	'flower'
थुतुनो	<i>t^hutuno</i>	'snout'
थैलो	<i>t^halo</i>	'sack'
थोपो	<i>t^hpo</i>	'drop'
दाउरो	<i>dauro</i>	'firewood'
दाबिलो	<i>dabilo</i>	'ladder'
दियालो	<i>dijalo</i>	'pine stick used as lamp'
ढुनो	<i>duno</i>	'leaf-cup'
ढुलो	<i>dulo</i>	'hole'
ढैलो	<i>dalo</i>	'door'
धमिरो	<i>d^hamiro</i>	'termite'
धागो	<i>d^hago</i>	'thread'
धारो	<i>d^haro</i>	'water tap'
नसो	<i>naso</i>	'vein'
नाङ्गलो	<i>nanlo</i>	'winnow'
नातो	<i>nato</i>	'relation'
निम्तो	<i>nimto</i>	'invitation'
पखेटो	<i>pak^heto</i>	'wing'
पाङ्गो	<i>pangro</i>	'wheel'
पातो	<i>pato</i>	'sheet of metal'
पात्रो	<i>patro</i>	'calendar'
पासो	<i>paso</i>	'trap'
पिठो	<i>pit^ho</i>	'flour'
पैतालो	<i>paitalo</i>	'foot-sole'
पोथ्रो	<i>pot^hro</i>	'bush'
फट्याङ्गो	<i>p^hatjangro</i>	'grasshopper'
फाँचो	<i>p^hātso</i>	'udder'

फियो	<i>p^hijo</i>	'spleen'
फेटो	<i>p^heto</i>	'turban'
फेरो	<i>p^hero</i>	'turn'
फोक्सो	<i>p^hokso</i>	'lungs'
बकुल्लो	<i>bAkullo</i>	'crane-bird'
बटुको	<i>bAtuko</i>	'bowl'
बन्चरो	<i>bAntsaro</i>	'axe'
बाटो	<i>baTo</i>	'path'
बिटो	<i>biTo</i>	'bundle'
बिक्रो	<i>birko</i>	'lid'
बोक्रो	<i>bokro</i>	'bark'
बोरो	<i>boro</i>	'sack'
बोसो	<i>boso</i>	'fat'
ब्वौसो	<i>bwāso</i>	'fox'
भकुण्डो	<i>b^hakundo</i>	'ball'
भाँडो	<i>b^hādo</i>	'utensil'
भित्तो	<i>b^hitto</i>	'side of wall'
भुत्तलो	<i>b^hutlo</i>	'hair'
भुसुनो	<i>b^husuno</i>	'gnat'
भेट्नो	<i>b^hetno</i>	'plant stalk'
मखुन्डो	<i>mAk^hundo</i>	'mask'
माकुरो	<i>makuro</i>	'spider'
माछो	<i>mats^ho</i>	'fish'
माटो	<i>mato</i>	'soil'
मिगौँलो	<i>mirgAulo</i>	'kidney'
मुन्टो	<i>munTo</i>	'tip'
मोसो	<i>moso</i>	'soot'
राँको	<i>rāko</i>	'torch'
राँगो	<i>rāgo</i>	'he-buffalo'
लहरो	<i>lAġaro</i>	'creeper'
लिस्नो	<i>lisno</i>	'ladder'
लुगो	<i>lugo</i>	'clothe'
लुतो	<i>luto</i>	'scabies'
लेदो	<i>ledo</i>	'slurry'
लौको	<i>luko</i>	'gourd'
लौरो	<i>luro</i>	'walking stick'
सनासो	<i>sAnaso</i>	'plier'

सलेदो	<i>sAledo</i>	'wick'
सल्लो	<i>sAllo</i>	'pine'
ससुरो	<i>sAsuro</i>	'father-in-law'
साइनो	<i>saino</i>	'relation'
साङ्लो	<i>saŋlo</i>	'cocroach'
सालो	<i>salo</i>	'wife's younger brother'
सितो	<i>sito</i>	'one grain of boiled-rice'
सिनो	<i>sino</i>	'flesh of dead animal'
सियो	<i>sijo</i>	'needle'
सुइरो	<i>suiro</i>	'big needle'
सुस्केरो	<i>suskerō</i>	'whistle'
हत्केलो	<i>hAtkelo</i>	'palm'
हलो	<i>hAllo</i>	'plough'
हाँगो	<i>hāgo</i>	'branch'
हिलो	<i>hilo</i>	'mud'

Type 21a Nouns:

काका	<i>kaka</i>	'uncle'
कुमार	<i>kumar</i>	'lad'

Type 21b Nouns:

नाति	<i>nati</i>	'grandson'
पनाति	<i>pAnati</i>	'great grandson'
जोगी	<i>dzoji:</i>	'saint'
सम्धी	<i>sAmD^{hi}:</i>	'father-in-law of offspring'

Type 21c Nouns:

बाघ	<i>bag^h</i>	'tiger'
थापा	<i>t^hapa</i>	'Thapa'
लामा	<i>lama</i>	'Lama'
कामी	<i>kami:</i>	'blacksmith'

Type 21d Nouns:

बिस्ट	<i>bist</i>	'Bista'
खस	<i>k^has</i>	'Khas'
मीत	<i>mi:t</i>	'bond-friend'
भूत	<i>b^hu:t</i>	'ghost'

Type 22a Nouns

दाइ	<i>dai</i>	'elder brother'
भाइ	<i>b^hai</i>	'younger brother'
बुबा	<i>buba</i>	'father'

Type 22b Nouns

दिदि	<i>didī</i>	'elder sister'
बहिनी	<i>b^hini:</i>	'younger sister'
आमा	<i>ama</i>	'mother'

Type 22c Nouns

राम	<i>ram</i>	'Ram'
श्याम	<i>sjam</i>	'Shyam'
हरि	<i>h^hari</i>	'Hari'

Type 22d Nouns

सीता	<i>si:ta</i>	'Sita'
गीता	<i>gi:ta</i>	'Gita'
सुनिता	<i>sunita</i>	'Sunita'

Type 22e Nouns:

खेत	<i>k^het</i>	'land'
अंगार	<i>ãgar</i>	'charcoal'
अंगुठी	<i>ãgut^hi</i>	'finger-ring'
अंश	<i>ãsA</i>	'share'
अकास	<i>ãkas</i>	'sky'
अक्षर	<i>ãksar</i>	'letter'
अखबार	<i>ãk^hãbar</i>	'newspaper'
अग्नि	<i>ãgni</i>	'fire'
अचार	<i>ãtsar</i>	'pickle'
अड्डा	<i>ãdãda</i>	'office'
अदालत	<i>ãdalat</i>	'court'
अदुवा	<i>ãduwa</i>	'ginger'
अनुहार	<i>ãnuhar</i>	'face'
अन्त्य	<i>ãntjã</i>	'end'
अन्न	<i>ãnnã</i>	'crops'
अपराध	<i>ãpãradh</i>	'crime'

अभाव	<i>Ab^hawA</i>	'scarcity'
अम्बक	<i>AmBAk</i>	'guava'
अवसर	<i>AWASAR</i>	'opportunity'
असर	<i>ASAR</i>	'effect'
असिना	<i>ASina</i>	'hailstone'
आँगन	<i>ãgAn</i>	'cortyard'
आँप	<i>ãp</i>	'mango'
आदर	<i>adAr</i>	'respect'
आन्द्रा	<i>andra</i>	'intestine'
आलु	<i>alu</i>	'potato'
आस्था	<i>ast^ha</i>	'belief'
इँटा	<i>ĩta</i>	'brick'
इच्छा	<i>itsts^ha</i>	'interest'
इतिहास	<i>itihAs</i>	'history'
इनार	<i>inar</i>	'well'
उखु	<i>uk^hu</i>	'sugarcane'
उत्तर	<i>uttAr</i>	'answer'
उत्पत्ति	<i>utpAti</i>	'origin'
उदाहरण	<i>udahArAn</i>	'example'
ऊन	<i>u:n</i>	'wool'
उपियाँ	<i>upijã</i>	'flea'
ऊँट	<i>ũt</i>	'camel'
ऐना	<i>Aina</i>	'mirror'
ओठ	<i>oṭ^h</i>	'lip'
औजार	<i>Audzar</i>	'device'
औषधी	<i>ASAdhi:</i>	'medicine'
कक्षा	<i>kAkṣa</i>	'class'
कटहर	<i>kAṭAhAr</i>	'jackfruit'
कथा	<i>kAt^ha</i>	'story'
कपाल	<i>kApal</i>	'hair'
कपास	<i>kApas</i>	'cotton'
कमर	<i>kAMAr</i>	'waist'
कमल	<i>kAMAl</i>	'lotus'
कम्पनी	<i>kAMpAni:</i>	'company'
कर	<i>kAr</i>	'tax'
करार	<i>kArar</i>	'contract'
कर्जा	<i>kArDza</i>	'loan'

कलम	<i>kʌlʌm</i>	'pen'
कला	<i>kʌlʌ</i>	'art'
काख	<i>kak^h</i>	'lap'
काठ	<i>kat^h</i>	'wood'
काम	<i>kam</i>	'work'
किताब	<i>kitab</i>	'book'
किसान	<i>kisan</i>	'farmer'
कुम	<i>kum</i>	'shoulder'
कृषि	<i>kriʃi</i>	'agriculture'
केराउ	<i>kerau</i>	'pea'
कैंची	<i>kʌiʃi:</i>	'scissors'
कोसिस	<i>kosis</i>	'attempt'
क्रम	<i>krʌm</i>	'sequence'
क्रोध	<i>krodh</i>	'anger'
क्षण	<i>kʃʌŋ</i>	'moment'
क्षेत्र	<i>kʃetrʌ</i>	'area'
खकार	<i>k^hʌkar</i>	'cough'
खत	<i>k^hʌt</i>	'scar'
खबर	<i>k^hʌbʌr</i>	'news'
खरायो	<i>k^hʌrajo</i>	'rabbit'
खल्ली	<i>k^hʌlti:</i>	'pocket'
खाडल	<i>k^hʌdʌl</i>	'ditch'
खाना	<i>k^hʌna</i>	'food item'
खिया	<i>k^hiʃa</i>	'rust'
खुर्सानी	<i>k^hʌrsani:</i>	'chilli'
खेत	<i>k^het</i>	'farm land'
खेल	<i>k^hel</i>	'game'
खोजी	<i>k^hodzi:</i>	'rearch'
ख्याति	<i>k^hiʃti</i>	'fame'
गणित	<i>gʌŋit</i>	'mathematics'
गधा	<i>gʌdha</i>	'donkey'
गफ	<i>gʌp^h</i>	'talk'
गलैँचा	<i>gʌlʌiʃa</i>	'carpet'
गहना	<i>gʌfiʌna</i>	'ornament'
गिजा	<i>gidza</i>	'gum'
गिलास	<i>gilas</i>	'tumbler'
गुन्द्री	<i>gundri:</i>	'mat'

गुरु	<i>guru</i>	'teacher'
गोबर	<i>gobar</i>	'dung'
गोही	<i>gofii</i>	'crocodile'
ग्रह	<i>grʌfiʌ</i>	'planet'
घड़ी	<i>g^hʌdi:</i>	'wrist watch'
घन	<i>g^hʌn</i>	'hammer'
घर	<i>g^hʌr</i>	'house'
घाम	<i>g^hʌm</i>	'sun'
चक़्कु	<i>tsʌkku</i>	'knife'
चट्टान	<i>tsʌttʌn</i>	'rock'
चना	<i>tsʌnʌ</i>	'gram'
चरन	<i>tsʌrʌn</i>	'grazing'
चरित्र	<i>tsʌritrʌ</i>	'character'
चामल	<i>tsʌmʌl</i>	'rice-grain'
चिउरा	<i>tsiura</i>	'beaten-rice'
चितुवा	<i>tsituwa</i>	'leopard'
चित्र	<i>tsitrʌ</i>	'picture'
चिनी	<i>tsini:</i>	'sugar'
चिया	<i>tsija</i>	'tea'
चुनाव	<i>tsunaw</i>	'election'
चेत	<i>tset</i>	'consciousness'
चोर	<i>tsor</i>	'thief'
चौकी	<i>tsʌuki:</i>	'police-post'
छड़ी	<i>ts^hʌdi:</i>	'stick'
छलफल	<i>ts^hʌlp^hʌl</i>	'discussion'
छाती	<i>ts^hʌti:</i>	'chest'
छायँ	<i>ts^hʌjã</i>	'shadow'
छिमेक	<i>ts^himek</i>	'neighbour'
जग	<i>dzʌg</i>	'foundation'
जगगा	<i>dzʌgga</i>	'land'
जन्तु	<i>dzʌntu</i>	'animal'
जन्म	<i>dzʌnma</i>	'birth'
जवाफ	<i>dzʌwap^h</i>	'answer'
जाँड	<i>dzãḍ</i>	'liquor'
जादु	<i>dzadu</i>	'magic'
जाल	<i>dzal</i>	'net'
जिउ	<i>dziu</i>	'body'

जिल्ला	<i>dzilla</i>	'district'
जीवन	<i>dzi:wan</i>	'life'
जुन	<i>dzun</i>	'moon'
जुवा	<i>dzuwa</i>	'gamble'
जेल	<i>dzɛl</i>	'prison'
जोस	<i>dzos</i>	'courage'
ज्याला	<i>dzjala</i>	'wage'
ज्वाला	<i>dzwala</i>	'flame'
झगडा	<i>dzʰʌgʌɖa</i>	'wage'
झण्डा	<i>dzʰʌɳɖa</i>	'flag'
झरना	<i>dzʰʌrʌna</i>	'water fall'
झाडी	<i>dzʰaɖi:</i>	'bush'
झार	<i>dzʰar</i>	'grass'
झील	<i>dzʰi:l</i>	'lake'
झुपडी	<i>dzʰupʌɖi:</i>	'hut'
झुल	<i>dzʰul</i>	'mosquito net'
झोल	<i>dzʰol</i>	'soup'
झ्याल	<i>dzʰjal</i>	'window'
टमाटर	<i>ʈʌmaʈʌr</i>	'tomato'
टाँक	<i>ʈãk</i>	'botton'
टापु	<i>ʈapu</i>	'island'
टिपोट	<i>ʈipot</i>	'note'
टुकी	<i>ʈuki:</i>	'oil lamp'
टेबुल	<i>ʈebul</i>	'table'
टेवा	<i>ʈewa</i>	'support'
टोपी	<i>ʈopi:</i>	'cap'
ठग	<i>ʈʰʌg</i>	'cheat'
ठट्टा	<i>ʈʰʌʈʈa</i>	'joke'
ठाउँ	<i>ʈʰaũ</i>	'place'
ठाना	<i>ʈʰana</i>	'police post'
ठेक्का	<i>ʈʰekka</i>	'contract'
ठोकर	<i>ʈʰokʌr</i>	'collision'
डकैत	<i>ɖʌkʌit</i>	'bandit'
डर	<i>ɖʌr</i>	'fear'
डाँठ	<i>ɖãʈʰ</i>	'stalk of plant'
डाडु	<i>ɖaɖu</i>	'ladle'
डिल	<i>ɖil</i>	'edge'

डेक्ची	<i>deksi:</i>	'a variety of cooking utensil'
डोली	<i>doli:</i>	'palanquin'
ढक	<i>dʰʌk</i>	'weighing weight'
ढाँचा	<i>dʰatsa</i>	'frame'
ढाड	<i>dʰaḍ</i>	'backbone'
ढाल	<i>dʰal</i>	'shield'
ढुकुटी	<i>dʰukuṭi:</i>	'storage'
ढुकुर	<i>dʰukur</i>	'dove'
ढुसी	<i>dʰusi:</i>	'fungus'
ढेडु	<i>dʰeḍu</i>	'ape'
ढोड	<i>dʰoḍ</i>	'maize stalk'
तत्त्व	<i>tʌttwʌ</i>	'element'
तथ्य	<i>tʌtʰjʌ</i>	'fact'
तन्तु	<i>tʌntu</i>	'tissue'
तरकारी	<i>tʌrkari:</i>	'vegetable'
तरिका	<i>tʌrika</i>	'method'
तलास	<i>tʌlas</i>	'research'
ताँती	<i>tāti:</i>	'long queue'
तागत	<i>tagʌt</i>	'energy'
ताप	<i>tap</i>	'heat'
तार	<i>tar</i>	'wire'
तिर्खा	<i>tirkʰa</i>	'thirst'
तिहुन	<i>tifiun</i>	'curry'
तुलना	<i>tulʌna</i>	'comparision'
तेल	<i>tel</i>	'oil'
तोरी	<i>tori:</i>	'mustard'
त्रास	<i>tras</i>	'fear'
त्रुटि	<i>truṭi</i>	'mistake'
थपड	<i>tʰʌpʌḍ</i>	'slap'
थर	<i>tʰʌr</i>	'surname'
थाल	<i>tʰal</i>	'plate'
थिति	<i>tʰiti</i>	'norms'
थुन	<i>tʰun</i>	'teat'
थुन्से	<i>tʰunse</i>	'back basket'
दक्षिण	<i>dʌkʃiŋ</i>	'south'
दण्ड	<i>dʌṇḍ</i>	'punishment'
दवाई	<i>dʌbai:</i>	'medicine'

दया	<i>dʌja</i>	'kindness'
दर	<i>dʌr</i>	'rate'
दराज	<i>dʌradz</i>	'cupboard'
दलाल	<i>dʌlʌl</i>	'broker'
दाँत	<i>dā̃t</i>	'tooth'
दाउरा	<i>daura</i>	'firewood'
दाग	<i>dag</i>	'spot'
दाम	<i>dam</i>	'cost'
दिन	<i>din</i>	'day'
दिशा	<i>difa</i>	'direction'
दुरी	<i>duri:</i>	'distance'
दुध	<i>dud^h</i>	'milk'
देउता	<i>deuta</i>	'god'
देश	<i>def</i>	'country'
दैत्य	<i>dʌĩtjʌ</i>	'titan'
दोकान	<i>dokan</i>	'shop'
दोसल्ला	<i>dosʌlla</i>	'shawl'
दौड	<i>dʌuḍ</i>	'race'
द्वार	<i>dwar</i>	'entrance'
धन	<i>d^hʌn</i>	'wealth'
धनु	<i>d^hʌnu</i>	'bow'
धरती	<i>d^hʌrʌti:</i>	'earth'
धरहरा	<i>d^hʌrʌhʌra</i>	'tower'
धराप	<i>d^hʌrap</i>	'trap'
धर्म	<i>d^hʌrmʌ</i>	'religion'
धातु	<i>d^hʌtu</i>	'metal'
धारणा	<i>d^hʌrʌṇa</i>	'concept'
धितो	<i>d^hito</i>	'deposit'
धुवाँ	<i>d^huwā̃</i>	'smoke'
धोबी	<i>d^hobi:</i>	'washerman'
ध्यान	<i>d^hjan</i>	'concentration'
ध्वनि	<i>d^hwʌni</i>	'sound'
नक्कल	<i>nʌkkʌl</i>	'imitation'
नक्सा	<i>nʌksa</i>	'map'
नगद	<i>nʌgʌt</i>	'cash'
नङ	<i>nʌṅ</i>	'nail'
नतिजा	<i>nʌtidza</i>	'result'

नरिवल	<i>nΛriwΛl</i>	'coconut'
नहर	<i>nΛfiΛr</i>	'canal'
नाइके	<i>naike</i>	'leader'
नाक	<i>nak</i>	'nose'
नाच	<i>nats</i>	'dance'
नाफा	<i>nap^ha</i>	'profit'
निधन	<i>nid^hΛn</i>	'death'
निबन्ध	<i>nibΛnd^h</i>	'essay'
नियम	<i>nijΛm</i>	'rule'
निर्माता	<i>nirmata</i>	'producer'
निवेदन	<i>niwedΛn</i>	'application'
नेता	<i>neta</i>	'leader'
नोकर	<i>nokΛr</i>	'servant'
नोकसान	<i>noksan</i>	'loss'
नोट	<i>not</i>	'paper money'
नौनी	<i>nΛuni:</i>	'butter'
न्याय	<i>njaj</i>	'justice'
पक्षी	<i>pΛkʃi:</i>	'bird'
पच्छिम	<i>pΛts^him</i>	'west'
पतन	<i>pΛtΛn</i>	'downfall'
पति	<i>pΛti</i>	'husband'
पत्थर	<i>pΛtt^hΛr</i>	'stone'
पत्नी	<i>pΛtni:</i>	'wife'
पदार्थ	<i>pΛdart^h</i>	'matter'
परम्परा	<i>pΛrΛmpΛrΛ</i>	'tradition'
परिचय	<i>pΛritsΛjΛ</i>	'introduction'
परिवार	<i>pΛriwar</i>	'family'
पशु	<i>pΛʃu</i>	'animal'
पाउ	<i>pau</i>	'foot'
पाचन	<i>patsΛn</i>	'digestion'
पात	<i>pat</i>	'leaf'
पानी	<i>pani:</i>	'water'
पालिस	<i>palis</i>	'polish'
पालुवा	<i>paluwa</i>	'new shoot'
पिँजडा	<i>pĩdzΛdΛ</i>	'bird's cage'
पिङ	<i>piŋ</i>	'swing'
पिता	<i>pita</i>	'father'

पिसाब	<i>pisab</i>	'urine'
पुच्छर	<i>putsts^hʌr</i>	'tail'
पुतली	<i>putʌli:</i>	'butterfly'
पुर्खा	<i>purk^ha</i>	'ancestor'
पुल	<i>pul</i>	'bridge'
पृथ्वी	<i>prit^hwi:</i>	'earth'
पेट	<i>peʈ</i>	'stomach'
पेसा	<i>pesa</i>	'profession'
पेस्की	<i>peski:</i>	'advance money'
पैनी	<i>pʌini:</i>	'canal'
पैसा	<i>pʌisa</i>	'money'
पोइ	<i>poi</i>	'husband'
पोखरी	<i>pok^hʌri:</i>	'pond'
पोथी	<i>pot^hi</i>	'female'
पोसाक	<i>posak</i>	'dress'
पौडी	<i>pʌuʈi:</i>	'swimming'
प्याज	<i>pjadz</i>	'onion'
प्यार	<i>pjar</i>	'love'
प्रकार	<i>prʌkar</i>	'type'
प्रदर्शन	<i>prʌdʌrʃʌn</i>	'exhibition'
प्रश्न	<i>prʌʃnʌ</i>	'question'
प्वौख	<i>pwāk^h</i>	'feather'
फरक	<i>p^hʌrʌk</i>	'different'
फर्सी	<i>p^hʌrsi</i>	'pumpkin'
फल	<i>p^hʌl</i>	'fruit'
फलाम	<i>p^hʌlam</i>	'iron'
फलेक	<i>p^hʌlek</i>	'plank'
फाइदा	<i>p^hʌida</i>	'profit'
फापर	<i>p^hʌpʌr</i>	'buck wheat'
फिँज	<i>p^hʌdz</i>	'foam'
फित्ता	<i>p^hʌitta</i>	'tape'
फूल	<i>p^hʌ:l</i>	'flower'
फोहोर	<i>p^hʌfior</i>	'dirt'
फौज	<i>p^hʌudz</i>	'soldier'
बँदेल	<i>bʌ̃del</i>	'wild boar'
बखान	<i>bʌk^hʌn</i>	'description'
बग्गी	<i>bʌggi:</i>	'carriage'

बजार	<i>bʌdzar</i>	'market'
बतास	<i>bʌtas</i>	'wind'
बन	<i>bʌn</i>	'forest'
बन्दुक	<i>bʌnduk</i>	'gun'
बरफ	<i>bʌrʌpʰ</i>	'ice'
बहस	<i>bʌɦʌs</i>	'discussion'
बहिनी	<i>bʌɦini:</i>	'sister'
बाँस	<i>bʌ̃s</i>	'bamboo'
बाजा	<i>badza</i>	'musical instrument'
बादल	<i>badʌl</i>	'cloud'
बानी	<i>bani:</i>	'habit'
बाफ	<i>bapʰ</i>	'vapour'
बालुवा	<i>baluwa</i>	'sand'
बिउ	<i>biu</i>	'seed'
बिक्री	<i>bikri:</i>	'sale'
बिजुली	<i>bidzuli</i>	'electricity'
बुँदा	<i>būda</i>	'point'
बुबा	<i>buba</i>	'father'
बुरुस	<i>burus</i>	'brush'
बेत	<i>bet</i>	'cane'
बेला	<i>bela</i>	'time'
बैगुन	<i>bʌigun</i>	'brinjal'
बोतल	<i>botʌl</i>	'bottle'
बगैँचा	<i>bʌgʌĩtsa</i>	'garden'
बाँसुरी	<i>bʌ̃suri:</i>	'flute'
ब्याज	<i>bjadz</i>	'interest'
भगवान	<i>bʰʌgʌwan</i>	'God'
भटमास	<i>bʰʌtmas</i>	'soyabean'
भतिज	<i>bʰʌtidz</i>	'nephew'
भर्ना	<i>bʰʌrna</i>	'admission'
भन्‍याङ	<i>bʰʌrjaŋ</i>	'ladder'
भवन	<i>bʰʌwan</i>	'building'
भविष्य	<i>bʰʌwiʃjʌ</i>	'future'
भाग	<i>bʰʌg</i>	'share'
भाग्य	<i>bʰʌgʃʌ</i>	'luck'
भात	<i>bʰʌt</i>	'rice'
भाला	<i>bʰʌla</i>	'spear'

भालु	<i>b^halu</i>	'bear'
भाषा	<i>b^haṣa</i>	'language'
भिखारी	<i>b^hik^hari:</i>	'beggar'
भिर	<i>b^hir</i>	'cliff'
भुँडी	<i>b^hũdi:</i>	'stomach'
भुकम्प	<i>b^hukʌmpʌ</i>	'earthquake'
भुत	<i>b^hut</i>	'ghost'
भुस	<i>b^hus</i>	'husk'
भूमि	<i>b^hu:mi</i>	'land'
भेला	<i>b^hela</i>	'gathering'
भोज	<i>b^hodz</i>	'feast'
भ्रम	<i>b^hrʌm</i>	'illusion'
मकै	<i>mʌkʌi</i>	'maize'
मगज	<i>mʌgʌdz</i>	'brain'
मजदुर	<i>mʌdzdur</i>	'labour'
मट्टितेल	<i>mʌt̪titel</i>	'kerosine'
मन्जन	<i>mʌndʒʌn</i>	'tooth paste'
मन्त्री	<i>mʌntri:</i>	'minister'
मरुभूमि	<i>mʌrʌb^hu:mi</i>	'desert'
मल	<i>mʌl</i>	'fertilizer'
मसला	<i>mʌsʌlʌ</i>	'spices'
मह	<i>mʌɦʌ</i>	'honey'
महिना	<i>mʌɦina</i>	'month'
माध्यम	<i>mʌd^hjʌm</i>	'medium'
मान्छे	<i>mʌnts^he</i>	'man'
माला	<i>mʌlʌ</i>	'garland'
मासु	<i>mʌsu</i>	'meat'
मिति	<i>mʌiti</i>	'date'
मुकुट	<i>mʌkuʈ</i>	'crown'
मुख	<i>mʌk^h</i>	'mouth'
मुट्टु	<i>mʌtu</i>	'heart'
मुर्ति	<i>mʌrti</i>	'idol'
मुला	<i>mʌlʌ</i>	'raddish'
मृग	<i>mʌriɡʌ</i>	'beer'
मेवा	<i>mʌwʌ</i>	'papaya'
मैदान	<i>mʌidʌn</i>	'ground'
मैन	<i>mʌin</i>	'wax'

मोजा	<i>modza</i>	'sock'
मोती	<i>moti:</i>	'pearl'
मौका	<i>mΛuka</i>	'occasion'
मौरी	<i>mΛuri:</i>	'bees'
म्वाइँ	<i>mwaĩ</i>	'kiss'
यन्त्र	<i>jantra</i>	'machine'
यातायात	<i>jatajat</i>	'transport'
युद्ध	<i>judd^hΛ</i>	'war'
योजना	<i>jodzΛna</i>	'plan'
रक्षा	<i>rΛkṣa</i>	'security'
रक्सी	<i>rΛksi:</i>	'alcohol'
रगत	<i>rΛgAt</i>	'blood'
रङ	<i>rΛŋ</i>	'colour'
रस	<i>rAs</i>	'juice'
रसायन	<i>rAsajΛn</i>	'chemical'
राजा	<i>raja</i>	'king'
राज्य	<i>radzjΛ</i>	'state'
राहदानी	<i>raɦΛdani:</i>	'passport'
रिस	<i>ris</i>	'anger'
रुमाल	<i>rumal</i>	'handkerchief'
रेडियो	<i>redijo</i>	'radio'
रोग	<i>rog</i>	'disease'
रौँ	<i>rΛũ</i>	'hair'
र्याल	<i>rjal</i>	'saliva'
लक्ष्य	<i>lΛkṣjΛ</i>	'objective'
लगन	<i>lAgΛn</i>	'auspicious time'
लडाइँ	<i>lΛdaĩ</i>	'battle'
लडाकु	<i>lΛdaku</i>	'warrior'
लसुन	<i>lAsun</i>	'garlic'
लाज	<i>lAdz</i>	'shame'
लाम	<i>lam</i>	'queue'
लाश	<i>laf</i>	'corpse'
लिलाम	<i>lilam</i>	'auction'
लुगा	<i>luga</i>	'cloth'
लोभ	<i>lob^h</i>	'greed'
वकिल	<i>wΛkil</i>	'lawyer'
वचत	<i>wΛtsAt</i>	'saving'

वर्ग	<i>wAɾgA</i>	'class'
वस्तु	<i>wAstu</i>	'matter'
वाक्य	<i>wakjA</i>	'sentence'
विज्ञान	<i>widzjan</i>	'science'
विभाग	<i>wib^hag</i>	'department'
शब्द	<i>ʃAbdA</i>	'word'
शहर	<i>ʃAɦAɾ</i>	'city'
शोषण	<i>ʃoʃAŋ</i>	'exploitation'
श्राप	<i>ʃɾap</i>	'curse'
श्रीपेच	<i>ʃɾipets</i>	'crown'
सँस्कार	<i>sãskar</i>	'ritual'
संझना	<i>sAMdz^hANA</i>	'remembrance'
संरचना	<i>sAMɾAɾsANA</i>	'structure'
संस्था	<i>sãst^ha</i>	'organization'
सङ्गीत	<i>sAŋgi:t</i>	'music'
सजाय	<i>sAdzajA</i>	'punishment'
सतह	<i>sAtAɦA</i>	'surface'
सदस्य	<i>sAdAsjA</i>	'member'
सन्तुलन	<i>sAtulAn</i>	'balance'
सपना	<i>sApANA</i>	'dream'
समय	<i>sAMAJA</i>	'time'
समाचार	<i>sAMatsar</i>	'news'
समाज	<i>sAMadz</i>	'society'
समुन्द्र	<i>sAMundrA</i>	'sea'
सम्बन्ध	<i>sAMBAND^hA</i>	'relation'
सरकार	<i>sArkar</i>	'government'
सर्प	<i>sArpA</i>	'snake'
सलाई	<i>sAlai:</i>	'match box'
सहमति	<i>sAɦAMATI</i>	'agreement'
साग	<i>sag</i>	'leafy vegetable'
साबुन	<i>sabun</i>	'soap'
सारस	<i>sArAs</i>	'crane'
सालिक	<i>salik</i>	'statue'
सास	<i>sas</i>	'breath'
सिक्री	<i>sikri:</i>	'chain'
सिपाही	<i>sipahi:</i>	'army'
सिरक	<i>sirAk</i>	'quilt'

सिर्जना	<i>sirdzΛna</i>	'creation'
सुँगुर	<i>sũgur</i>	'pig'
सुगा	<i>suga</i>	'parrot'
सुती	<i>suti:</i>	'cotton'
सुन	<i>sun</i>	'gold'
सूची	<i>sutsi:</i>	'list'
सूर्य	<i>surjΛ</i>	'sun'
सेना	<i>sena</i>	'soldier'
सेवा	<i>sewa</i>	'service'
स्तन	<i>stΛn</i>	'breast'
स्याउ	<i>sjau</i>	'apple'
स्वर्ग	<i>swΛrgΛ</i>	'heaven'
हँसिया	<i>hĩsija</i>	'sickle'
हड्डी	<i>hΛd̪di:</i>	'bone'
हतियार	<i>hΛtijaɾ</i>	'weapon'
हमला	<i>hΛmΛla</i>	'attack'
हाँस	<i>hĩs</i>	'duck'
हाट	<i>hΛt</i>	'local market'
हात	<i>hΛt</i>	'hand'
हावा	<i>hΛwΛ</i>	'air'
हिउँ	<i>hiũ</i>	'snow'
हिसाब	<i>hisab</i>	'maths'
हुलाक	<i>hulak</i>	'post office'
हौसला	<i>hΛusΛla</i>	'encouragement'

Type 22f Nouns:

पोखरा	<i>pok^hΛrΛ</i>	'Pokhara'
काठमान्डौ	<i>kat^hmandΛu</i>	'Kathmandu'
ललतिपुर	<i>lΛlitpur</i>	'Lalitpur'

Annex -3

Nepali Pronouns

Pronouns

First Person Singular Pronouns

म	<i>ma</i>	'1SG'
मै	<i>mai</i>	'1SG.OBL'
मै	<i>mai</i>	'1SG.EMPH'
मेरो	<i>mero</i>	'1SG.OBL.GEN.MASC'
मेरी	<i>meri:</i>	'1SG.OBL.GEN.FEM'
मेरा	<i>mera</i>	'1SG.OBL.GEN.PL'
मेरा	<i>mera</i>	'1SG.OBL.GEN.HON'
मेरा	<i>mera</i>	'1SG.OBL.GEN.OBL'
मेरै	<i>merai</i>	'1SG.OBL.GEN.EMPH'

!!First Person Plural Pronouns

हामी	<i>fami:</i>	'1PL'
हाम्रो	<i>famro</i>	'1PL.OBL.GEN.MASC'
हाम्री	<i>famri</i>	'1PL.OBL.GEN.FEM'
हाम्रा	<i>famra</i>	'1PL.OBL.GEN.PL'
हाम्रा	<i>famra</i>	'1PL.OBL.GEN.HON'
हाम्रा	<i>famra</i>	'1PL.OBL.GEN.OBL'
हाम्रै	<i>famai</i>	'1PL.OBL.GEN.EMPH'

!! Second Person Singular Pronouns

तँ	<i>ta</i>	'2SG'
तैं	<i>tai</i>	'2SG.OBL'
तैं	<i>tai</i>	'2SG.EMPH'
तेरो	<i>tero</i>	'2SG.OBL.GEN.MASC'
तेरी	<i>teri:</i>	'2SG.OBL.GEN.FEM'
तेरा	<i>tera</i>	'2SG.OBL.GEN.PL'
तेरा	<i>tera</i>	'2SG.OBL.GEN.HON'
तेरा	<i>tera</i>	'2SG.OBL.GEN.OBL'
तेरै	<i>terai</i>	'2SG.OBL.GEN.EMPH'

!!Second Person honorific Pronouns

तिमी	<i>timi:</i>	'2SG.HON'
तिम्रो	<i>timro</i>	'2SG.OBL.HON.GEN.MASC'

तिम्री	<i>timri:</i>	'2SG.OBL.HON.GEN.FEM'
तिम्रा	<i>timra</i>	'2SG.OBL.HON.GEN.PL'
तिम्रा	<i>timra</i>	'2SG.OBL.HON.GEN.HON'
तिम्रा	<i>timra</i>	'2SG.OBL.HON.GEN.OBL'
तिम्रै	<i>timrai</i>	'2SG.OBL.HON.GEN.EMPH'

!! Second Person High Honorific Pronouns

तपाईं	<i>tapaĩ</i>	'2SG.HHON'
यहाँ	<i>jaɦã</i>	'2SG.HHON'
उहाँ	<i>uɦã</i>	'2SG.HHON'
वहाँ	<i>waɦã</i>	'2SG.HHON'
हजुर	<i>ɦadzur</i>	'2SG.HHON'

!! Second Person Royal Honorific Pronoun

मौसुफ	<i>mausup^h</i>	'2SG+RHON'
-------	---------------------------	------------

!! Third Person Singular Pronouns ऊ u:

ऊ	<i>u:</i>	'3SG'
उही	<i>ufi:</i>	'3SG.EMPH'
उस	<i>us</i>	'3SG.OBL'
उसै	<i>usai</i>	'3SG.OBL.EMPH'
उनी	<i>uni:</i>	'3SG.HON'
उन	<i>un</i>	'3SG.HON.OBL'
उनै	<i>unai</i>	'3SG.HON.OBL.EMPH'
उहाँ	<i>uɦã</i>	'3SG.HON'
वहाँ	<i>waɦã</i>	'3SG.HON'

!! Third Person singular Pronouns यो

यो	<i>jo</i>	'3SG.PROX'
यही	<i>jaɦi:</i>	'3SG.PROX.EMPH'
यस	<i>jas</i>	'3SG.OBL.PROX'
यसै	<i>jasai</i>	'3SG.OBL.PROX.EMPH'
यी	<i>ji:</i>	'3SG.PROX.HON'
यी	<i>ji:</i>	'3PL.PROX'
यिनी	<i>jini:</i>	'3SG.PROX.HON'
यिन	<i>jin</i>	'3SG.PROX.OBL.HON'
यिनै	<i>jinai</i>	'3SG.PROX.OBL.HON.EMPH'

!!!Third Person singular Pronouns त्यो and ती

त्यो	<i>tjo</i>	'3SG.DIST'
त्यही	<i>tjΛhi:</i>	'3SG.DIST.EMPH'
त्यस	<i>tjas</i>	'3SG.OBL'
त्यसै	<i>tjasai</i>	'3SG.OBL.EMPH'
ती	<i>ti:</i>	'3SG.HON.DIST'
ती	<i>ti:</i>	'3PL.DIST'
तिनी	<i>tini:</i>	'3SG.HON.DIST'
तिन	<i>tin</i>	'3SG.OBL.HON.DIST'
तिनै	<i>tinai</i>	'3SG.OBL.HON.DIST.EMPH'

!!Reflexive Pronoun

आफू	<i>ap^hu:</i>	'REFL'
आफै	<i>ap^hai</i>	'REFL.OBL.EMPH'
आफै	<i>ap^haĩ</i>	'REFL.OBL.EMPH'
आफ्नो	<i>ap^hno</i>	'REFL.OBL.GEN.SG'
आफना	<i>ap^hna</i>	'REFL.OBL.GEN.PL'
आफना	<i>ap^hna</i>	'REFL.OBL.GEN.HON'
आफना	<i>ap^hna</i>	'REFL.OBL.GEN.OBL'
आफनी	<i>ap^hna:</i>	'REFL.OBL.GEN.FEM'
आफ्नै	<i>ap^hna</i>	'REFL.OBL.GEN.EMPH'

!! Demonstrative Pronouns यो

यो	<i>jo</i>	'DEM.PROX'
यही	<i>jΛhi:</i>	'DEM.PROX.EMPH'
यी	<i>ji:</i>	'DEM.PROX'
यिनी	<i>jini:</i>	'DEM.PROX.HON'
यिन	<i>jin</i>	'DEM.PROX.OBL'
यिनै	<i>jinai</i>	'DEM.PROX.OBL.EMPH'
यहाँ	<i>jΛhĩ</i>	'DEM.PROX'

!!Demonstrative pronoun त्यो and ती

त्यो	<i>tjo</i>	'DEM.DIST'
त्यही	<i>tjΛhi:</i>	'DEM.DIST.EMPH'
ती	<i>ti:</i>	'DEM.DIST'
तिनी	<i>tini:</i>	'DEM.DIST.OBL.HON'
तिन	<i>tin</i>	'DEM.DIST.OBL'
तिनै	<i>tinai</i>	'DEM.DIST.OBL.EMPH'

!!Demonstrative pronoun ऊ

ऊ	<i>u</i>	'DEM.DIST'
उही	<i>ufii</i>	'DEM.DIST.EMPH'
उनी	<i>uni</i>	'DEM.DIST.HON'
उन	<i>un</i>	'DEM.DIST.OBL'
उनै	<i>unAi</i>	'DEM.DIST.OBL.EMPH'
उहाँ	<i>ufiã</i>	'DEM.DIST.HON'
वहाँ	<i>wafiã</i>	'DEM.DIST.HON'

!!Other Demonstrative Pronouns

सो	<i>so</i>	'DEM.DIST'
सोही	<i>sofi</i>	'DEM.DIST.EMPH'
निज	<i>nidz</i>	'DEM.PROX'
निजै	<i>nidzAi</i>	'DEM.PROX.EMPH'
उक्त	<i>ukta</i>	'DEM.PROX'

!! Relative Pronouns

जो	<i>dzo</i>	'REL.HUM'
जस	<i>dzAs</i>	'REL.OBL.HUM'
जसै	<i>dzAsAi</i>	'REL.OBL.HUM.EMPH'
जे	<i>dze</i>	'REL.NHUM'
जुन	<i>dzun</i>	'REL'
जुनै	<i>dzunAi</i>	'REL.EMPH'

!! Interrogative Pronouns

को	<i>ko</i>	'INTERRO.HUM'
कस	<i>kaS</i>	'INTERRO.HUM.OBL'
कसै	<i>kaSAi</i>	'INTERRO.HUM.OBL.EMPH'
के	<i>ke</i>	'INTERRO.NHUM'
कुन	<i>kun</i>	'INTERRO'
किन	<i>kiNa</i>	'INTERRO'
कसरी	<i>kaSAri</i>	'INTERRO'

!! Indefinite Pronouns

कोही	<i>kofii</i>	'INDEF.HUM'
केही	<i>kefi</i>	'INDEF.NHUM'
कुनै	<i>kunAi</i>	'INDEF'
जोसुकै	<i>dzosukAi</i>	'INDEF.HUM'

जेसुकै *dzesukai* 'INDEF.NHUM'
जुनसुकै *dzunsukai* 'INDEF'

!! Definite Pronouns

प्रत्येक *pratjek* 'DEF'
हरेक *harek* 'DEF'
सबै *sabai* 'DEF'
अको *ako* 'DEF.SG'
अका *aka* 'DEF.PL'
अका *aka* 'DEF.HON'
अका *aka* 'DEF.OBL'
अकी *aki:* 'DEF.FEM'
अकै *akai* 'DEF.EMPH'
अरू *aru* 'DEF'

!! Reciprocal pronouns

एकअको *ekako* 'RECIP'
एकअका *ekaka* 'RECIP.OBL'
एकअका *ekaka* 'RECIP.HON'
एकअका *ekaka* 'RECIP.PL'
एकअकी *ekaki:* 'RECIP.FEM'
एकआपस *ekapas* 'RECIP'
आपस *apas* 'RECIP'
आआफू *aap^hu* 'RECIP'

END

Annex -4

Adjectives in Nepali

Adjectives:

!!O-ending Adjectives

राम्रो	<i>ramro</i>	'good'
कालो	<i>kalo</i>	'black'
खस्रो	<i>k^hAsro</i>	'coarse'
बुढो	<i>bud^ho</i>	'old'
अनौठो	<i>anA^uto</i>	'strange'
अन्धो	<i>and^ho</i>	'blind'
अप्टेरो	<i>apt^hero</i>	'difficult'
अमिलो	<i>amilo</i>	'sour'
लाटो	<i>lato</i>	'dumb'
सानो	<i>sano</i>	'small'

!!Non-o-ending Adjective Type 1

चतुर	<i>tsatur</i>	'clever'
धुर्त	<i>d^hurta</i>	'cunning'
पुर्विया	<i>purwija</i>	'of east'
पाखे	<i>pak^he</i>	'rural'

!!Non-o-ending Adjective Type 2

न्यून	<i>nju:nA</i>	'less'
महत	<i>mAhAt</i>	'big'
निम्न	<i>nimnA</i>	'low'
गहन	<i>gAhAn</i>	'intensive'
लघु	<i>lag^hu</i>	'small'

!!Unmarked Adjectives

असल	<i>asAl</i>	'nice'
खराब	<i>k^hArab</i>	'bad'
नयाँ	<i>nAjã</i>	'new'
धनी	<i>d^hAni:</i>	'rich'
अमर	<i>amar</i>	'live-long'
अल्छी	<i>alts^hi</i>	'lazy'
आन्तरिक	<i>antarik</i>	'internal'
उन्नत	<i>unnAt</i>	'improved'
संभव	<i>sAmb^hAV</i>	'possible'

Annex - 5

Numerals and Classifiers in Nepali

Numbers

!!Cardinal Numbers

पाँच	<i>pāts</i>	'five'
सात	<i>sat</i>	'seven'
आठ	<i>aṭ</i>	'eight'
दस	<i>das</i>	'ten'
एधार	<i>eg^harA</i>	'ten'
बाह	<i>bafirA</i>	'twelve'
तेह	<i>tefirA</i>	'thirteen'
चौध	<i>tsAud^h</i>	'fourteen'
पन्ध्र	<i>pAnd^hrA</i>	'fifteen'
सोह	<i>sofirA</i>	'sixteen'
सत्र	<i>sAt^rA</i>	'seventeen'
अठार	<i>Aṭ^harA</i>	'eighteen'
उन्नाइस	<i>unnais</i>	'nineteen'
बीस	<i>bi:s</i>	'twenty'
एक्काईस	<i>ekkai:s</i>	'twenty one'
बाईस	<i>bai:s</i>	'twenty two'
तेईस	<i>tei:s</i>	'twenty three'
चौबीस	<i>tsAubi:s</i>	'twenty four'
पच्चीस	<i>pAtsi:s</i>	'twenty five'
छब्बीस	<i>ts^hAbi:s</i>	'twenty six'
सत्ताईस	<i>sAttai:s</i>	'twenty seven'
अठ्ठाईस	<i>Aṭ^hṭ^hai:s</i>	'twenty eight'
उनन्तीस	<i>unAntis</i>	'twenty nine'
तीस	<i>ti:s</i>	'thirty'
एकतीस	<i>ekti:s</i>	'thirty one'
बत्तीस	<i>bAtti:s</i>	'thirty two'
तेत्तीस	<i>tetti:s</i>	'thirty three'
चौत्तीस	<i>tAṭṭi:s</i>	'thirty four'
पैंतीस	<i>pAṭṭi:s</i>	'thirty five'
छत्तीस	<i>tsAtti:s</i>	'thirty six'
सैत्तीस	<i>sAṭṭi:s</i>	'thirty seven'
अठ्तीस	<i>Aṭ^hti:s</i>	'thirty eight'

उन्चालीस	<i>untsali:s</i>	'thirty nine'
चालीस	<i>tsali:s</i>	'forty'
एकचालीस	<i>ektsali:s</i>	'forty one'
बयालीस	<i>bʌjali:s</i>	'forty two'
त्रिचालीस	<i>tritsali:s</i>	'forty three'
चौवालीस	<i>tsʌuwali:s</i>	'forty four'
पैंतालीस	<i>pʌĩtali:s</i>	'forty five'
छयालीस	<i>tsʌjali:s</i>	'forty six'
सठ्चालीस	<i>sʌtʰtsali:s</i>	'forty seven'
अठ्चालीस	<i>ʌtʰtsali:s</i>	'forty eight'
उन्चास	<i>untsas</i>	'forty nine'
पचास	<i>pʌtsas</i>	'fifty'
एकाउन्न	<i>ekaunna</i>	'fifty one'
बाउन्न	<i>bʌunna</i>	'fifty two'
त्रिपन्न	<i>tripʌnna</i>	'fifty three'
चौवन्न	<i>tsʌunna</i>	'fifty four'
पचपन्न	<i>pʌtsapʌnna</i>	'fifty five'
छपन्न	<i>tsapʌnna</i>	'fifty six'
सन्ताउन्न	<i>sʌntaunna</i>	'fifty seven'
अन्ठाउन्न	<i>ʌntʰaunna</i>	'fifty eight'
उन्साठी	<i>unsaʰti:</i>	'fifty nine'
साठी	<i>saʰti:</i>	'sixty'
एकसठ्ठी	<i>eksʌtʰti:</i>	'sixty one'
बैसठ्ठी	<i>bʌisʌtʰti:</i>	'sixty two'
त्रिसठ्ठी	<i>trisʌtʰti:</i>	'sixty three'
चौसठ्ठी	<i>tsʌusʌtʰti:</i>	'sixty four'
पैंसठ्ठी	<i>pʌĩsʌtʰti:</i>	'sixty five'
छैसठ्ठी	<i>tsʌisʌtʰti:</i>	'sixty six'
सत्सठ्ठी	<i>sʌtsʌtʰti:</i>	'sixty seven'
अठ्सठ्ठी	<i>ʌtʰsʌtʰti:</i>	'sixty eight'
उन्सत्तरी	<i>unʌnsattari:</i>	'sixty nine'
सत्तरी	<i>sattari:</i>	'seventy'
एकहत्तर	<i>ekʰattar</i>	'seventy one'
बहत्तर	<i>bʌʰattar</i>	'seventy two'
त्रिहत्तर	<i>triʰattar</i>	'seventy three'
चौरहत्तर	<i>tsʌurʌʰattar</i>	'seventy four'
पचहत्तर	<i>pʌtsʌʰattar</i>	'seventy five'

छयत्तर	<i>ts^hʌjattʌr</i>	'seventy six'
सतहत्तर	<i>sʌtʌfʌttʌr</i>	'seventy seven'
अठहत्तर	<i>ʌt^hʌfʌttʌr</i>	'seventy eight'
उनासी	<i>unasi:</i>	'seventy nine'
असी	<i>ʌsi:</i>	'eighty'
एकासी	<i>ekasi:</i>	'eighty one'
बयासी	<i>bʌjasi:</i>	'eighty two'
त्रियासी	<i>trijasi:</i>	'eighty three'
चौरासी	<i>tsʌrasi:</i>	'eighty four'
पचासी	<i>pʌtsasi:</i>	'eighty five'
छयासी	<i>ts^hʌjasi:</i>	'eighty six'
सतासी	<i>sʌtasi:</i>	'eighty seven'
अठासी	<i>ʌt^hʌsi:</i>	'eighty eight'
उनानब्बे	<i>unanʌbbe</i>	'eighty nine'
नब्बे	<i>nʌbbe</i>	'ninety'
एकानब्बे	<i>ekanʌbbe</i>	'ninety one'
बयानब्बे	<i>bʌjanʌbbe</i>	'ninety two'
त्रियानब्बे	<i>trijanʌbbe</i>	'ninety three'
चौरानब्बे	<i>tsʌuranʌbbe</i>	'ninety four'
पन्चानब्बे	<i>pʌntsanʌbbe</i>	'ninety five'
छयानब्बे	<i>tsʌjanʌbbe</i>	'ninety six'
सन्तानब्बे	<i>sʌntanʌbbe</i>	'ninety seven'
अन्ठानब्बे	<i>ʌnt^hʌnʌbbe</i>	'ninety eight'
उनान्सय	<i>unansʌjʌ:</i>	'ninty nine'
सय	<i>sʌjʌ</i>	'hundred'
हजार	<i>fʌdzar</i>	'thousand'
लाख	<i>lak^h</i>	'hundred thousand'
करोड	<i>kʌroɖ</i>	'ten million'
अरब	<i>ʌrʌb</i>	'ten billion'
खरब	<i>k^hʌrʌb</i>	'ten trillion'

!!Classifier Like items

!!O-ending classifiers

कोसो	<i>koso</i>	'pod'
दानो	<i>dano</i>	'bead'
डाँकलो	<i>ɖʌklo</i>	'stalk'
सिको	<i>sirko</i>	'stream of (milk)'
तुको	<i>turko</i>	'a small quantity of liquid'

पासो	<i>paso</i>	'a dice ??'
थोपो	<i>t^hopo</i>	'drop'
छिटो	<i>ts^hiʔo</i>	'drop'
आँठो	<i>ãt^ho</i>	'sheaf'
चुल्लो	<i>tsult^ho</i>	'plaited hair'
चोक्ठो	<i>tsokʔo</i>	'fragment'
पानो	<i>pano</i>	'sheet'
सियो	<i>sijo</i>	'needle'
सितो	<i>sito</i>	'a grain of boiled rice'
भक्कानो	<i>b^hʌkkano</i>	'clod of'
ढिक्को	<i>d^hkkō</i>	'lump of'
धरो	<i>d^hʌro</i>	'a small piece of cloth'
कोइलो	<i>koilo</i>	'a charcoal'
झिल्लको	<i>dz^hiʎko</i>	'sparkle'
पित्को	<i>pitko</i>	'small amount'
लस्को	<i>ʌsko</i>	'certain amount'
डल्लो	<i>dʌllo</i>	'a round lump'
टस्को	<i>tʌsko</i>	'barest greeting'
झल्को	<i>dz^hʌlko</i>	'glance, flash'
मुसलो	<i>muslo</i>	'rapids in stream'
सिन्को	<i>sinko</i>	'a small piece of stick or twig'
ठुन्को	<i>t^hunko</i>	'fragment, slice'
काँगियो	<i>kãgijo</i>	'bunch of bananas'
गेडो	<i>geɖo</i>	'grain of corn'
घनो	<i>g^hʌno</i>	'whole bamboo'
केस्रो	<i>kesro</i>	'a quarter of an orange'
बिजुलो	<i>bidzulo</i>	'juice sacs inside orange'
कोयो	<i>kojo</i>	'stone of a fruit'
फब्ल्याँटो	<i>p^hʌbljãʔo</i>	'a piece of firewood'
चोइटो	<i>tsoiʔo</i>	'small cutting'
चिरो	<i>tsiro</i>	'a slice, a splinter'
कुङ्को	<i>kuɖko</i>	'a piece'
टोप्रो	<i>ʔopro</i>	'stalk of a fruit'
ठेउको	<i>t^heuko</i>	'a piece of wood'
पाखो	<i>pak^ho</i>	'slant'
लाँक्रो	<i>lãkro</i>	'whole sugarcane'
आँख्लो	<i>ãk^hʎo</i>	'joint in plants'

लत्को	<i>lʌtko</i>	'hanging mass'
लट्टो	<i>lʌʈʈo</i>	'cluster'
जुरो	<i>dzuro</i>	'crest, hood, hump'
चोइलो	<i>tsoilo</i>	'a slice'

!!Non-o-ending classifiers

पोटी	<i>poʈi:</i>	'bulb of garlic'
थुन	<i>t^hun</i>	'teat, udder'
पन्ना	<i>pʌnna</i>	'sheet of paper'
जुवा	<i>dzuwa</i>	'pole of a cart'
गाँस	<i>gãs</i>	'mouthful'
हल	<i>fiʌl</i>	'pair of bullocks'
झर	<i>dz^hʌr</i>	'moderate rain'
सल	<i>sʌl</i>	'pile'
चक्की	<i>tsʌkki:</i>	'small wheel'
तार	<i>tar</i>	'wire'
झप्की	<i>dz^hʌpki:</i>	'pounce'
माउ	<i>mau</i>	'mother of'
घरी	<i>g^hʌri:</i>	'a bunch of plaintains'
फ्याक	<i>pjak</i>	'division'
हार	<i>fiar</i>	'line, row'
खात	<i>k^hat</i>	'heap'
ढिँडी	<i>d^hĩdi:</i>	'pod of a chilli'
खिल्ली	<i>k^hilli</i>	'stick of ciggaret, roll of betel'
फूल	<i>p^hu:l</i>	'flower'

!!Exceptional numbers

एक	<i>ek</i>	'one'
दुई	<i>dui:</i>	'two'
तीन	<i>ti:n</i>	'three'
चार	<i>tsar</i>	'four'
छ	<i>tsʌ</i>	'six'
नौ	<i>nʌu</i>	'nine'

!!Exceptional ordinal numerals

पहिलो	<i>pʌfiʌlo</i>	'first'
पहिला	<i>pʌfiʌla</i>	'first.PL'
पहिला	<i>pʌfiʌla</i>	'first.OBL'

पहिला	<i>pʌfila</i>	'first.HON'
पहिली	<i>pʌfili:</i>	'first.FEM'
दोस्रो	<i>dosro</i>	'second.MASC'
दोस्रा	<i>dosra</i>	'second.PL'
दोस्रा	<i>dosra</i>	'second.OBL'
दोस्रा	<i>dosra</i>	'second.HON'
दोस्री	<i>dosri:</i>	'second.FEM'
तेस्रो	<i>tesro</i>	'third.MASC'
तेस्रा	<i>tesra</i>	'third.PL'
तेस्रा	<i>tesra</i>	'third.OBL'
तेस्रा	<i>tesra</i>	'third.HON'
तेस्री	<i>tesri:</i>	'third.FEM'
चौथो	<i>tsʌut^ho</i>	'fourth.MASC'
चौथा	<i>tsʌut^ha</i>	'fourth.PL'
चौथा	<i>tsʌut^ha</i>	'fourth.OBL'
चौथा	<i>tsʌut^ho</i>	'fourth.HON'
चौथी	<i>tsʌut^hi:</i>	'fourth.FEM'
प्रथम	<i>prʌt^hʌm</i>	'first'
द्वितीय	<i>dwiti:jʌ</i>	'second'
तृतीय	<i>triti:jʌ</i>	'third'
चतुर्थ	<i>tsʌturt^hʌ</i>	'fourth'
पञ्चम	<i>pʌntsʌm</i>	'fifth'
छैटौं	<i>ts^hʌiʈʌũ</i>	'sixth'
नवौं	<i>nʌwʌũ</i>	'nineth'

!! Frequency Numerals

एकोहोरो	<i>ekohoro</i>	'once'
दोहोरो	<i>dohoro</i>	'double'
तेहोरो	<i>tehoro</i>	'thrice'
एकसरो	<i>eksaro</i>	'single'
दुईसरो	<i>dui:saro</i>	'two fold'
तीनसरो	<i>tinsaro</i>	'three fold'
दोबर	<i>dobar</i>	'double'
तेबर	<i>tebar</i>	'thrice'
चौबर	<i>tsaubar</i>	'four times'
दुईगुना	<i>dui:guna</i>	'two times.'
तीनगुना	<i>tinguna</i>	'three times'
चौगुना	<i>tsauguna</i>	'four times'

!! Portion Numerals

आधा	<i>ad^ha</i>	'half'
पौने	<i>pʌune</i>	'a quarter less'
सवा	<i>sʌwa</i>	'a quarter more'
डेढ	<i>d̪eḍ^h</i>	'one and half'
साढे	<i>saḍ^he</i>	'a half more'
अढाइ	<i>ʌḍ^hai</i>	'two and half'
चौथाइ	<i>tsʌut^hai</i>	'a quarter'

!! Classifiers

जना	<i>dzʌna</i>	'CLF.HUM'
वटा	<i>wʌṭa</i>	'CLF.NHUM'
ओटा	<i>oṭa</i>	'CLF.NHUM'
वटी	<i>wʌṭi:</i>	'CLF.FEM'
ओटी	<i>wʌṭi:</i>	'CLF.FEM'

Annex – 6

Adverbs in Nepali

!!! Temporal Adverbs

अहिले	<i>ʌfile</i>	'now'
हिजो	<i>fidzo</i>	'yesterday'
उहिले	<i>ufile</i>	'long time ago'
पहिले	<i>pʌfile</i>	'before'
जहिले	<i>dzʌfile</i>	'when'
जहिलेसुकै	<i>dzʌfilesukʌi</i>	'any time'
जब	<i>dzʌbʌ</i>	'when'
अब	<i>ʌbʌ</i>	'now'
तब	<i>tʌbʌ</i>	'then'
आज	<i>adzʌ</i>	'today'
भोलि	<i>b^holi</i>	'tomorrow'
पर्सि	<i>pʌrsi</i>	'day after tomorrow'
निपर्सि	<i>nipʌrsi</i>	'two days after tomorrow'
अस्ति	<i>ʌsti</i>	'last time'
पोहोर	<i>poʌior</i>	'last year'
परार	<i>pʌrar</i>	'year before last year'
आघौं	<i>ag^hʌũ</i>	'next year'
पराघौं	<i>pʌrag^hʌũ</i>	'year after next year'
भरे	<i>b^hʌre</i>	'later'
दिउँसो	<i>diũso</i>	'afternoon'
बेर	<i>ber</i>	'late'
अबेर	<i>ʌber</i>	'late'
अघि	<i>ʌg^hi</i>	'before'
पछि	<i>pʌts^hi</i>	'later'
भरखर	<i>b^hʌrk^hʌr</i>	'just now'
तुरुन्त	<i>turuntʌ</i>	'recently'
कतिखेर	<i>kʌtik^herʌ</i>	'any time'
यतिखेर	<i>jʌtik^herʌ</i>	'this time'
त्यतिखेर	<i>tjʌtik^herʌ</i>	'that time'
उतिखेर	<i>utik^herʌ</i>	'that time'
जतिखेर	<i>dzʌtik^herʌ</i>	'same time'

कतिबेला	<i>kʌtibela</i>	'what time'
यतिबेला	<i>jʌtibela</i>	'this time'
त्यसबेला	<i>tjʌsbela</i>	'that time'
कतिन्जेल	<i>kʌtindzel</i>	'till what time'
उतिन्जेल	<i>utindzel</i>	'till that time'
यतिन्जेल	<i>jʌtindzel</i>	'till this time'
त्यतिन्जेल	<i>tjʌtindzel</i>	'till that time'
जतिन्जेल	<i>dzʌtindzel</i>	'till some time'
उसबेला	<i>usbela</i>	'that time'
अचेल	<i>ʌtsel</i>	'nowadays'
हिजोआज	<i>fidzoadzʌ</i>	'nowadays'
आजभोलि	<i>adzʌb^holi</i>	'nowadays'
आइतबार	<i>aitʌbar</i>	'Sunday'
सोमबार	<i>sombar</i>	'Monday'
मङ्गलबार	<i>mʌŋgʌlbar</i>	'Tuesday'
बुधवार	<i>bud^hʌbar</i>	'Wednesday'
बिहिबार	<i>bifiibar</i>	'Thursday'
शुक्रबार	<i>fukrʌbar</i>	'Friday'
शनिबार	<i>ʃʌnibar</i>	'Saturday'
बैशाख	<i>bʌisak^h</i>	'Nepali first month (April-May)'
जेठ	<i>dzet^h</i>	'Nepali second month (May-June)'
जेष्ठ	<i>dzest^h</i>	'Nepali second month (May-June)'
असार	<i>ʌsar</i>	'Nepali third month (June-July)'
आषाढ	<i>ʌʃad^h</i>	'Nepali third month (June-July)'
श्रावण	<i>ʃrawʌŋ</i>	'Nepali fourth month (July-August)'
भाद्र	<i>b^hadrʌ</i>	'Nepali fifth month (August-September)'
आश्विन	<i>aʃwin</i>	'Nepali sixth month (September-October)'
कार्तिक	<i>kartik</i>	'Nepali seventh month (October-November)'
मार्ग	<i>margʌ</i>	'Nepali eighth month (November-December)'
पौष	<i>pʌuʃ</i>	'Nepali ninth month (December-January)'
माघ	<i>marg^h</i>	'Nepali tenth month (January-February)'
फल्गुन	<i>p^halgun</i>	'Nepali eleventh month (February-March)'
चैत्र	<i>tsʌitrʌ</i>	'Nepali twelveth month (March-April)'
साउन	<i>saun</i>	'Nepali fourth month (July-August)'

भदौ	<i>b^hadΛu</i>	'Nepali fifth month (August-September)'
असोज	<i>Λsodz</i>	'Nepali sixth month (September-October)'
बसन्त	<i>bΛsAntΛ</i>	'Spring season'
गृष्म	<i>griṣmΛ</i>	'Summer season'
शिशिर	<i>ʃiʃir</i>	'Autumn season'
हेमन्त	<i>fiemAnt</i>	'Winter season'
हिउँद	<i>fiũd</i>	'winter'
बर्खा	<i>bArk^ha</i>	'rainy season'
कात्तिक	<i>kattik</i>	'Nepali seventh month (October-November)'
मङ्सिर	<i>mΛṣir</i>	'Nepali eighth month (November-December)'
पुस	<i>pus</i>	'Nepali ninth month (December-January)'
फागुन	<i>p^hagun</i>	'Nepali eleventh month (February-March)'
चैत	<i>tʃaitΛ</i>	'Nepali twelveth month (March-April)'

!!! Spatial Adverbs

तल	<i>tΛΛ</i>	'below'
त्यहाँ	<i>tjΛfiã</i>	'there'
यहाँ	<i>jΛfiã</i>	'here'
उहाँ	<i>ufiã</i>	'there'
जहाँ	<i>dzΛfiã</i>	'where'
कहाँ	<i>kΛfiã</i>	'where'
यता	<i>jΛta</i>	'here'
त्यता	<i>tjΛta</i>	'there'
कता	<i>kΛta</i>	'where'
उता	<i>uta</i>	'there'
जता	<i>dzΛta</i>	'where'
कता	<i>kΛta</i>	'where'
अगाडि	<i>Λgaḍi</i>	'in front'
अघि	<i>Λg^hi</i>	'in front'
पछाडि	<i>pΛts^haḍi</i>	'behind'
पछि	<i>pΛts^hi</i>	'behind'
भिन्न	<i>b^hitrΛ</i>	'inside'
बाहिर	<i>bahirΛ</i>	'outside'
माथि	<i>mat^hi</i>	'above'
तल	<i>tΛΛ</i>	'below'

मुनि	<i>muni</i>	'under'
वर	<i>wΛrΛ</i>	'near'
पर	<i>pΛrΛ</i>	'far'
उँभो	<i>ũb^ho</i>	'up'
उँधो	<i>ũd^ho</i>	'down'
वारि	<i>wari</i>	'this side'
पारि	<i>pari</i>	'other side'
सामु	<i>samu</i>	'in front of'
सामुन्ने	<i>samunne</i>	'in front of'
नजिक	<i>nΛdzik</i>	'near'
जहाँ	<i>dzΛhã</i>	'where'
जहाँसुकै	<i>dzΛhãsukΛi</i>	'wherever'
जतासुकै	<i>dzΛtasukΛi</i>	'wherever'
जहाँतहाँ	<i>dzΛhãtΛhã</i>	'every where'
जताततै	<i>dzΛtatΛtΛi</i>	'every where'
जहींतहीं	<i>dzΛhĩtΛhĩ</i>	'every where'
वरपर	<i>wΛrΛpΛrΛ</i>	'here and there'
वारिपारि	<i>waripari</i>	'this side and that side'
भिन्नबाहिर	<i>b^hitrΛbahirΛ</i>	'inside outside'
अघिपछि	<i>Λg^hipΛts^hi</i>	'before and after'
तलमाथि	<i>tΛΛmat^hi</i>	'up and down'
आमनेसामने	<i>amnesamne</i>	'face to face'
सँगसँगै	<i>sãgAsãgai</i>	'together'

!!! Amount Adverbs

धेरै	<i>d^herΛi</i>	'more'
एकदम	<i>ekdΛm</i>	'very much'
अझ	<i>Λdz^hΛ</i>	'still more'
झन	<i>dz^hΛnΛ</i>	'still more'
अलि	<i>Λli</i>	'little more'
थोरै	<i>t^horΛi</i>	'less'
अलिकति	<i>ΛlikΛti</i>	'little more'
प्रशस्त	<i>prAsAstΛ</i>	'enough'
प्रशस्तै	<i>prAsAstΛi</i>	'enough.EMPH'
बहुत	<i>bΛhuta</i>	'much more'

निकै	<i>nikai</i>	'more'
साहै	<i>safirai</i>	'too much'
ज्यादा	<i>dzjada</i>	'more than enough'
असाध्य	<i>ʌsad^hʌ</i>	'too much'
अति	<i>ʌsad^hjʌ</i>	'too much'
बढि	<i>bʌd^hi</i>	'more'
खुब	<i>k^hubʌ</i>	'plenty'
बिलकुल	<i>bilkul</i>	'absolutely'
बिछट्ट	<i>bits^hʌttʌ</i>	'extremely'
औधि	<i>ʌud^hi</i>	'more'
कम	<i>kʌm</i>	'less'
अलिक	<i>ʌlik</i>	'more'
अलिअलि	<i>ʌliʌli</i>	'little be little'
पूर्णतः	<i>pu:rtʌtʌ</i>	'completely'
पुरै	<i>purai</i>	'completely'
पुरापुर	<i>purapur</i>	'completely'
यति	<i>jʌti</i>	'this much'
त्यति	<i>tjʌti</i>	'that much'
उति	<i>uti</i>	'that much'
जति	<i>dzʌti</i>	'so much'
कति	<i>kʌti</i>	'how much'

!!! Manner Adverbs

सुस्तरी	<i>sustari:</i>	'slowly'
फटाफट	<i>p^hʌtʌp^hʌt</i>	'instantly'
राम्ररी	<i>ramrari:</i>	'nicely'
बल्ल	<i>bʌllʌ</i>	'at last'
यसरी	<i>jʌsari:</i>	'this way'
त्यसरी	<i>tjʌsari:</i>	'that way'
उसरी	<i>usari:</i>	'that way'
जसरी	<i>dzʌsari:</i>	'any how'
कसरी	<i>kʌsari:</i>	'how'
यसो	<i>jʌso</i>	'this way'
उसो	<i>uso</i>	'that way'
जसो	<i>dzʌso</i>	'some way'

कसो	<i>kʌso</i>	'what way'
बेसरी	<i>besʌri:</i>	'very much'
सुटुक्क	<i>sutukkʌ</i>	'quietly'
झलमल्ल	<i>dzʌʌmʌʌʌ</i>	'brightly'
प्वाक्क	<i>pwakkʌ</i>	'manner of saying instantly'
फुत्त	<i>p^huttʌ</i>	'abruptly'
सरसर्ति	<i>sʌrsʌrti</i>	'manner of doing things lightly'
फरफर	<i>p^hʌrp^hʌr</i>	'manner ...'
इवाट्ट	<i>dz^hwattʌ</i>	'suddenly'
अकस्मात्	<i>ʌkʌsmat</i>	'suddenly'
छिटो	<i>ts^hiʈo</i>	'quickly'
चाँडो	<i>tsãdo</i>	'quickly'
सितै	<i>sitʌi</i>	'how much'
स्वतः	<i>swʌtʌ</i>	'automatically'

!!! Frequency Adverbs

कहिलेकाहीं	<i>kʌfiʌkafiĩ:</i>	'sometimes'
जहिलेतहिले	<i>dzʌfiʌetʌfiʌe</i>	'anytime'
बारम्बार	<i>barʌmbar</i>	'frequently'
निरन्तर	<i>nirʌntar</i>	'continuously'
सधैँ	<i>sʌd^hʌĩ</i>	'always'
सदा	<i>sʌda</i>	'always'
धेरैजसो	<i>d^herʌidzʌso</i>	'more and like'
प्रायः	<i>prajʌ</i>	'often'
सायदै	<i>sajʌdʌi</i>	'perhaps'
बिरलै	<i>birʌʌi</i>	'rarely'
कमै	<i>kʌmʌi</i>	'less'
यदाकदा	<i>jʌdakʌda</i>	'sometimes'
प्रत्येकपल्ट	<i>prʌtekpʌʌtʌ</i>	'each time'
धेरैचोटि	<i>d^herʌitsotʌi</i>	'many times'
बातैपिच्छे	<i>batʌipitsts^he</i>	'each times'

!!! Reason Adverbs

कारण	<i>karʌŋ</i>	'reason'
त्यसकारण	<i>tjʌskarʌŋ</i>	'therefore'

यसकारण	<i>jʌskarʌŋ</i>	'therefore'
फलस्वरूप	<i>p^hʌlswʌru:p</i>	'as a result'
यसर्थ	<i>jʌsʌrt^hʌ</i>	'therefore'
तसर्थ	<i>tʌsʌrt^hʌ</i>	'therefore'
परिणामस्वरूप	<i>pʌriŋʌmswʌru:p</i>	'as a result'

!!! Sentential Adverbs

सायद	<i>sajʌd</i>	'perhaps'
सामान्यतः	<i>samanjʌtʌ</i>	'normally'
साँच्चै	<i>sātstsʌi</i>	'truly'
साँच्चिकै	<i>sātstsikʌi</i>	'really'
सत्ते	<i>sʌtte</i>	'truly'
सके	<i>sʌke</i>	'possibly'
अवश्य	<i>ʌwʌʃjʌ</i>	'definitely'
जरुर	<i>dzʌrur</i>	'surely'
वस्तुतः	<i>wʌstutʌ</i>	'as a fact'
स्वाभावतः	<i>swab^hʌwʌtʌ</i>	'naturally'

Annex -7

Verbs in Nepali

!!Verbs

!! Verb Type 1a

अघा	<i>ʌg^ha</i>	'be satisfied'
करा	<i>kʌra</i>	'shout'
निदा	<i>nida</i>	'sleep'
बहुला	<i>bʌfiula</i>	'be mad'
मुस्कुरा	<i>muskura</i>	'smile'
लजा	<i>ʌdza</i>	'shy'
टुसा	<i>tusa</i>	'sprout'

!! Verb Type 1b

चोखि	<i>tsok^hi</i>	'purify'
गुम्सि	<i>gumsi</i>	'suffocate'
घोप्टि	<i>g^hopti</i>	'spill'
टुक्रि	<i>tukri</i>	'break'

!!Verb Type1c

उक्लि	<i>ukli</i>	'climb'
उघ्रि	<i>ug^hri</i>	'open'
उफ्रि	<i>up^hri</i>	'jump'
घस्रि	<i>g^hʌsri</i>	'crawl'
थुप्रि	<i>t^hupri</i>	'pile'
निखि	<i>nik^hri</i>	'come to an end'
पग्लि	<i>pʌgli</i>	'melt'
सप्रि	<i>sʌpri</i>	'grow well'
सुघ्रि	<i>sud^hri</i>	'improve'

!!!Type verb1d

काँप्	<i>kāp</i>	'shiver'
हाँस्	<i>hā̃s</i>	'laugh'

!!Verb Type1e

बस्	<i>bʌs</i>	'sit'
खस्	<i>k^hʌs</i>	'drop'

छिन्	<i>ts^hin</i>	'cut right through'
मर्	<i>maɾ</i>	'die'
गल्	<i>gaɪ</i>	'tire'
चल्	<i>tsaɪ</i>	'move'
जँच्	<i>dzãts</i>	'be examined'
झर्	<i>dz^har</i>	'drop'
टर्	<i>taɾ</i>	'pass by'
सर्	<i>saɾ</i>	'shift'

!! Verb Type2a

उचाल्	<i>utsal</i>	'lift'
अर्जाप्	<i>ɹdzap</i>	'sharpen'
झपार्	<i>dz^hapar</i>	'scold'
पछार्	<i>pats^har</i>	'topple'
सराप्	<i>sarap</i>	'curse'
मार	<i>maɾ</i>	'kill'
गाल्	<i>gal</i>	'melt'
चाल्	<i>tsal</i>	'move'
जाँच्	<i>dzãts</i>	'examine'
झार्	<i>dz^har</i>	'drop'
टार्	<i>taɾ</i>	'avoid'
सार	<i>sar</i>	'move'

!! Verb Type 2b

पक्रि	<i>pakri</i>	'arrest'
पर्खि	<i>parki</i>	'wait'
बिसि	<i>birsi</i>	'forget'
मन्सि	<i>mansi</i>	'ward off'
सम्झि	<i>samdz^hi</i>	'remember'
कुल्चि	<i>kultsi</i>	'stamp out'
उइँटि	<i>uĩti</i>	'card cotton'

!!Verb Type2c

खाप्	<i>k^hap</i>	'add on a pile'
गाड्	<i>gad</i>	'bury'

छाप्	<i>ts^hap</i>	'print'
टाँस्	<i>tãs</i>	'stick'
तान्	<i>tan</i>	'pull'
नाच्	<i>nats</i>	'dance'
बाँच्	<i>bãts</i>	'survive'
हाँक्	<i>fiãk</i>	'drive'

!!!Type verb2d

पढ्	<i>pʌd^h</i>	'read'
किन्	<i>kin</i>	'buy'
जोत्	<i>dzot</i>	'plough'
घस्	<i>g^hʌs</i>	'rub'

!! Irregular verbs

आ	<i>a</i>	'come'
जा	<i>dza</i>	'go'
रो	<i>ro</i>	'cry'
खा	<i>k^ha</i>	'eat'
पा	<i>pa</i>	'get'
दि	<i>di</i>	'give'
लि	<i>li</i>	'take'
धो	<i>d^ho</i>	'wash'

!! Auxiliary verbs

छ	<i>ts^hʌ</i>	'be.EXIST'
हो	<i>fiʌ</i>	'be.IDN'
थि	<i>t^hi</i>	'be.P'

Annex -8

Verbal inflections in Nepali

!! Inflections for non-Past existential verb छे *cha* 'be' (affirmative)

उ	-u	'NP.1SG'
औँ	-Au	'NP.1PL'
स्	-s	'NP.2SG.MASC'
एस्	-es	'NP.2SG.FEM'
औ	-Au	'NP.2SG.MASC.HON'
यौ	<i>ja</i> u	'NP.2SG.FEM.HON'
औ	-Au	'NP.2PL'
	- ϕ	'NP.3SG.MASC'
ए	-e	'NP.3SG.FEM'
न्	-n	'NP.3SG.MASC.HON'
इन्	-in	'NP.3SG.FEM.HON'
न्	-n	'NP.3PL'

!! Inflection for non-past existential verb छे *cha* 'be' (Negative)

इनँ	-in \tilde{a}	'NP.NEG.1SG'
इनौँ	-inAu	'NP.NEG.1PL'
इनस्	-inAs	'NP.NEG.2SG'
इनौ	-inAu	'NP.NEG.2SG.HON'
इनौ	-inAu	'NP.NEG.2PL'
इन	-inA	'NP.NEG.3SG'
इनन्	-inAn	'NP.NEG.3SG.HON'
इनन्	-inAn	'NP.NEG.3PL'

!! Inflections for non-past identificational verb हो *ho* 'be' (affirmative)

उँ	- \tilde{u}	'NP.1SG'
औँ	-Au	'NP.1PL'
स्	-s	'NP.2SG'
औ	-Au	'NP.2SG.HON'
औ	-Au	'NP.2PL'
	- ϕ	'NP.3SG'
न्	-nA	'NP.3SG.HON'
न्	-nA	'NP.3PL'

!! Inflection for non-past identificational verb हो *ho* 'be' (Negative)

इँ	- <i>inā</i>	'NP.NEG.1SG'
इँ	- <i>ināū</i>	'NP.NEG.1PL'
इँस्	- <i>inās</i>	'NP.NEG.2SG'
इँ	- <i>ināu</i>	'NP.NEG.2SG.HON'
इँ	- <i>ināu</i>	'NP.NEG.2PL'
इँ	- <i>inā</i>	'NP.NEG.3SG'
इँन्	- <i>inān</i>	'NP.NEG.3SG.HON'
इँन्	- <i>inān</i>	'NP.NEG.3PL'

!! Inflections for non-past tense (affirmative)

छु	- <i>ts^hu</i>	'NP.1SG'
छँ	- <i>ts^hāū</i>	'NP.1PL'
छस्	- <i>ts^hās</i>	'NP.2SG.MASC'
छेस्	- <i>ts^hes</i>	'NP.2SG.FEM'
छँ	- <i>ts^hāu</i>	'NP.2SG.MASC.HON'
छँ	- <i>ts^hjāu</i>	'NP.2SG.FEM.HON'
छँ	- <i>ts^hāu</i>	'NP.2PL'
छ	- <i>ts^hā</i>	'NP.3SG.MASC'
छे	- <i>ts^he</i>	'NP.3SG.FEM'
छँन्	- <i>ts^hān</i>	'NP.3SG.MASC.HON'
छँन्	- <i>ts^hin</i>	'NP.3SG.FEM.HON'
छँन्	- <i>ts^hān</i>	'NP.3PL'

!!Table 4.46: Inflections for non-past tense negative 1

दिँ	- <i>dinā</i>	'NP.NEG.1SG'
दिँ	- <i>dināū</i>	'NP.NEG.1PL'
दिँस्	- <i>dinās</i>	'NP.NEG.2SG.MASC'
दिँस्	- <i>dinās</i>	'NP.NEG.2SG.FEM'
दिँ	- <i>dināu</i>	'NP.NEG.2SG.MASC.HON'
दिँ	- <i>dināu</i>	'NP.NEG.2SG.FEM.HON'
दिँ	- <i>dināu</i>	'NP.NEG.2PL'
दिँ	- <i>dinā</i>	'NP.NEG.3SG.MASC'
दिँ	- <i>dinā</i>	'NP.NEG.3SG.FEM'
दिँन्	- <i>dinān</i>	'NP.NEG.3SG.MASC.HON'
दिँन्	- <i>dinān</i>	'NP.NEG.3SG.FEM.HON'
दिँन्	- <i>dinān</i>	'NP.NEG.3PL'

!! Inflections for non-past tense negative 2

नँ	-nã	'NP.NEG.1SG'
नौँ	-nãũ	'NP.NEG.1PL'
नस्	-nas	'NP.NEG.2SG'
नौ	-nau	'NP.NEG.2SG.HON'
नौ	-nau	'NP.NEG.2PL'
न	-na	'NP.NEG.3SG'
नन्	-nan	'NP.NEG.3SG.HON'
नन्	-nan	'NP.NEG.3PL'

!! Inflections for past tense (affirmative)

एँ	-ẽ	'P.1SG'
यौँ	-jãũ	'P.1PL'
इस्	-is	'P.2SG'
यौ	-jau	'P.2SG.HON'
यौ	-jau	'P.2PL'
यो	-jo	'P.3SG.MASC'
ई	-i:	'P.3SG.FEM'
ए	-e	'P.3SG.MASC.HON'
इन्	-in	'P.3SG.FEM.HON'
ए	-e	'P.3PL'

!! Inflections for past tense (negative)

इनँ	-inã	'P.NEG.1SG'
एनौँ	-enãũ	'P.NEG.1PL'
इनस्	-inas	'P.NEG.2SG'
एनौ	-enau	'P.NEG.2SG.HON'
एनौ	-enau	'P.NEG.2PL'
एन	-ena	'P.NEG.3SG.MASC'
इन	-ina	'P.NEG.3SG.FEM'
एनन्	-enan	'P.NEG.3SG.MASC.HON'
इनन्	-inan	'P.NEG.3SG.FEM.HON'
एनन्	-enan	'P.NEG.3PL'

!! Inflections for perfect aspect

एको	-eko	'PERF.SG.MASC'
एका	-eka	'PERF.PL'
एकी	-eki	'PERF.SG.FEM'
एकै	-ekai	'PERF.SG.FEM.EMPH'

!! Inflections for imperfect aspect

दो	-do	'IMPERF.SG.MASC'
दी	-di	'IMPERF.SG.FEM'
दा	-da	'IMPERF.PL'
दैं	-dΛi	'IMPERF'

!! Inflections for habitual aspect (Affirmative)

थैं	-t ^h ẽ	'P.HAB.1SG'
थ्यौँ	-t ^h jΛũ	'P.HAB.1PL'
थिस्	-t ^h is	'P.HAB.2SG'
थ्यौ	-t ^h jΛu	'P.HAB.2SG.HON'
थ्यौ	-t ^h jΛu	'P.HAB.2PL'
थ्यो	-t ^h jo	'P.HAB.3SG.MASC'
थि	-t ^h i	'P.HAB.3SG.FEM'
थे	-t ^h e	'P.HAB.3SG.MASC.HON'
थिन्	-t ^h in	'P.HAB.3SG.FEM.HON'
थे	-t ^h e	'P.HAB.3PL'

!! Inflections for habitual aspect (Negative)

दैनथैं	-dΛinΛt ^h ẽ	'P.NEG.HAB.1SG'
दैनथ्यौँ	-dΛinΛt ^h jΛũ	'P.NEG.HAB.1PL'
दैनथिस्	-dΛinΛt ^h is	'P.NEG.HAB.2SG'
दैनथ्यौ	-dΛinΛt ^h jΛu	'P.NEG.HAB.2SG.HON'
दैनथ्यौ	-dΛinΛt ^h Λu	'P.NEG.HAB.2PL'
दिनथिस्	-dΛinΛt ^h is	'P.NEG.HAB.2SG.FEM'
दिनथ्यौ	-dΛinΛt ^h Λu	'P.NEG.HAB.2SG.FEM.HON'
दैनथ्यो	-dΛinΛt ^h jo	'P.NEG.HAB.3SG.MASC'
दिनथिस्	-dinΛt ^h is	'P.NEG.HAB.3SG.FEM'
दिनथे	-dinΛt ^h e	'P.NEG.HAB.3SG.MASC.HON'
दैनथे	-dΛinΛt ^h e	'P.NEG.HAB.3PL'

!! Inflections for Inferential aspect (Affirmative)

एछु	-ets ^h u	'P.INFER.1SG'
एछौँ	-ets ^h Λũ	'P.INFER.1PL'
एछस्	-ets ^h Λs	'P.INFER.2SG.MASC'
इछस्	-its ^h Λs	'P.INFER.2SG.FEM'
एछौ	-ets ^h Λu	'P.INFER.2SG.MASC.HON'
इछौ	-its ^h Λu	'P.INFER.2SG.FEM.HON'

एछौ	-ets ^h Λu	'P.INFER.2PL'
एछ	-ets ^h Λ	'P.INFER.3SG.MASC'
इछ	-its ^h Λ	'P.INFER.3SG.FEM'
एछन्	-ets ^h Λn	'P.INFER.3SG.MASC.HON'
इछन्	-its ^h Λn	'P.INFER.3SG.FEM.HON'
एछन्	-ets ^h Λn	'P.INFER.3PL'

!! Inflections for Inferential aspect (Negative)

एनछु	-enΛts ^h u	'P.INFER.NEG.1SG'
एनछौ	-enΛts ^h Λũ	'P.INFER.NEG.1PL'
एनछस्	-enΛts ^h Λs	'P.INFER.NEG.2SG.MASC'
इनछेस्	-inΛts ^h es	'P.INFER.NEG.2SG.FEM'
एनछौ	-enΛts ^h Λu	'P.INFER.NEG.2SG.MASC.HON'
इनछौ	-inΛts ^h Λu	'P.INFER.NEG.2SG.FEM.HON'
एनछौ	-enΛts ^h Λu	'P.INFER.NEG.2PL'
एनछ	-enΛts ^h Λ	'P.INFER.NEG.3SG.MASC'
इनछ	-inΛts ^h Λ	'P.INFER.NEG.3SG.FEM'
एनछन्	-enΛts ^h Λn	'P.INFER.NEG.3SG.HON'
इनछन्	-inΛts ^h Λn	'P.INFER.NEG.3SG.FEM.HON'
एनछन्	-enΛts ^h Λn	'P.INFER.NEG.3PL'

!! Inflection for imperative mood

	φ	'IMP.2SG'
ई	-i'	'IMP.2SG'
अ	-A	'IMP.2SG.HON'
ऊ	-u'	'IMP.2SG.HON'
अ	-A	'IMP.2PL'
ओ	-o	'IMP.2PL'

!! Inflections for optative mood (Affirmative)

ऊँ	-ũ	'OPT.1SG'
औँ	-Λũ	'OPT.1PL'
एस्	-es	'OPT.2SG'
ए	-e	'OPT.2SG.HON'
ए	-e	'OPT.2PL'
ओस्	-os	'OPT.3SG'
ऊन्	-u:n	'OPT.3SG.HON'
ऊन्	-u:n	'OPT.3PL'

!! Inflections for potential mood (affirmative)

उँला	- <i>ũla</i>	'POT.1SG'
औँला	- <i>āũla</i>	'POT.1PL'
लास्	- <i>las</i>	'POT.2SG.MASC'
लिस्	- <i>lis</i>	'POT.2SG.FEM'
औला	- <i>āula</i>	'POT.2SG.MASC.HON'
औली	- <i>āuli:</i>	'POT.2SG.FEM.HON'
औला	- <i>āula</i>	'POT.2PL'
ला	- <i>la</i>	'POT.3SG.MASC'
ली	- <i>li:</i>	'POT.3SG.FEM'
लान्	- <i>lan</i>	'POT.3SG.MASC.HON'
लिन्	- <i>lin</i>	'POT.3SG.FEM.HON'
लान्	- <i>lan</i>	'POT.3PL'

!! Inflection for participles

ई	- <i>i:</i>	'ABS'
नु	- <i>nu</i>	'INF'
ना	- <i>na</i>	'INF.OBL'
नै	- <i>nai</i>	'INF.EMPH'
न	- <i>n</i>	'PURP'
नै	- <i>nai</i>	'PURP.EMPH'
ने	- <i>ne</i>	'PROSP'
दा	- <i>da</i>	'DUR'
दै	- <i>dai</i>	'DUR.EMPH'
एर	- <i>er</i>	'CONJ'
एरै	- <i>erai</i>	'CONJ.EMPH'
इकन	- <i>ikn</i>	'CONJ'
इकनै	- <i>iknai</i>	'CONJ.EMPH'
ए	- <i>e</i>	'COND'
ए	- <i>e</i>	'PERFT'

Annex - 9

Conjunctions and Particles in Nepali

!!Conjunctions

!!!समपदिक संयोजकहरू

र	<i>ra</i>	'and'
वा	<i>wa</i>	'or'
अथवा	<i>at^hawa</i>	'or'
या	<i>ja</i>	'or'
कि	<i>ki</i>	'or'
नकि	<i>naki</i>	'neither??'
अनि	<i>ani</i>	'and then'
पनि	<i>p^ani</i>	'also'
तथा	<i>ta^ha</i>	'and'
एवं	<i>ewam</i>	'and'
तर	<i>ta^a</i>	'but'
किन्तु	<i>kintu</i>	'but'
परन्तु	<i>pa^antu</i>	'but'

!!!विषमपदिक संयोजकहरू

भन्ने	<i>b^hanne</i>	'saying'
भनेर	<i>b^hanera</i>	'said that'
भने	<i>b^hene</i>	'then'
कि	<i>ki</i>	'that'
किनभने	<i>kinab^hane</i>	'because'
किनकि	<i>kinaki</i>	'because'
यसकारण	<i>jas^akarana</i>	'therefore'

!!! Particles निपातहरू

नै	<i>nai</i>	'Particle'
मात्र	<i>matra</i>	'Particle'
केवल	<i>kewal</i>	'Particle'
चाहिँ	<i>tsahi</i>	'Particle'
पनि	<i>p^ani</i>	'Particle'
ल	<i>la</i>	'Particle'
है	<i>hai</i>	'Particle'
न	<i>na</i>	'Particle'
नि	<i>ni</i>	'Particle'
त	<i>ta</i>	'Particle'

पो	<i>po</i>	'Particle'
क्या	<i>kja</i>	'Particle'
के	<i>ke</i>	'Particle'
कि	<i>ki</i>	'Particle'
रे	<i>re</i>	'Particle'
क्यारे	<i>kjare</i>	'Particle'
हाँ	<i>fiḥ</i>	'Particle'
हगि	<i>fiḷgi</i>	'Particle'
खै	<i>k^hḷi</i>	'Particle'
लौ	<i>lɒu</i>	'Particle'
हौ	<i>fiḷu</i>	'Particle'
क्यार	<i>kjarɒ</i>	'Particle'
ब्यारे	<i>bjare</i>	'Particle'
झै	<i>dz^hḷi</i>	'Particle'

!!!विस्मायादि बोधकहरू Interjections

अहा	<i>ḷfiɒ</i>	'Interjection'
अहो	<i>ḷfio</i>	'Interjection'
ओहो	<i>ofo</i>	'Interjection'
उहु	<i>ufiu</i>	'Interjection'
उफ	<i>up^h</i>	'Interjection'
आत्था	<i>att^ha</i>	'Interjection'
आत्थु	<i>att^hu</i>	'Interjection'
आच्छु	<i>atsts^hu</i>	'Interjection'
छि	<i>ts^hi</i>	'Interjection'
धत्	<i>d^hḷt</i>	'Interjection'
धत्तेरि	<i>d^hḷtteri</i>	'Interjection'
थुक्क	<i>t^hukkɒ</i>	'Interjection'
थुइक्क	<i>t^huikkɒ</i>	'Interjection'
बड	<i>bḷḍɒ</i>	'Interjection'
हाय	<i>fiaj</i>	'Interjection'
कठै	<i>kḷt^hḷi</i>	'Interjection'
हरे	<i>fiḷre</i>	'Interjection'
शिव	<i>fiwɒ</i>	'Interjection'
च्च	<i>tstsɒ</i>	'Interjection'
बरा	<i>bḷrɒ</i>	'Interjection'
बिचरा	<i>bitsɒrɒ</i>	'Interjection'
उस्	<i>us</i>	'Interjection'

हाहा	<i>fɪfɪ</i>	'Interjection'
हिहि	<i>fɪfɪ</i>	'Interjection'
ज्या	<i>dzja</i>	'Interjection'
ए	<i>e</i>	'Interjection'
ऐ	<i>ɛi</i>	'Interjection'
औ	<i>ɛu</i>	'Interjection'
हौ	<i>fɪɛu</i>	'Interjection'
ऐय्या	<i>ɛijja</i>	'Interjection'
ल	<i>lɛ</i>	'Interjection'
हवस्	<i>fɪɛwɛs</i>	'Interjection'
अँ	<i>ɛ̃</i>	'Interjection'
ज्यू	<i>dzju:</i>	'Interjection'
हजुर	<i>fɪɛdzur</i>	'Interjection'
हाँ	<i>fɪɛ̃</i>	'Interjection'
अहाँ	<i>ɛ̃fɪɛ̃</i>	'Interjection'
नाइँ	<i>naĩ</i>	'Interjection'
कुन्नि	<i>kunni</i>	'Interjection'
सत्ते	<i>sɛtte</i>	'Interjection'
साँच्ची	<i>sɛ̃tstsi:</i>	'Interjection'
धरोधर्म	<i>dʰɛrodʰɛrma</i>	'Interjection'
भो	<i>bʰo</i>	'Interjection'
ई	<i>i:</i>	'Interjection'
ऊ	<i>u:</i>	'Interjection'
वाह	<i>wahi</i>	'Interjection'
स्याबास	<i>sjabas</i>	'Interjection'
अबुइ	<i>abui</i>	'Interjection'
आप्पै	<i>appɛi</i>	'Interjection'
ओ	<i>o</i>	'Interjection'
चै	<i>tɛɛi</i>	'Interjection'
ओइ	<i>oi</i>	'Interjection'
एइ	<i>ei</i>	'Interjection'
ज्याहै	<i>dzjafɛi</i>	'Interjection'
हि	<i>fɪi</i>	'Interjection'
हा	<i>fɪa</i>	'Interjection'

Annex - 10

Postpositions in Nepali

!! Postpositions-case markers plural marker adverbial postpositions

!!Case Markers which do not take emphatic marker

ले	- <i>le</i>	'ERG'
ले	- <i>le</i>	'INST'
लाई	- <i>lai:</i>	'DAT'
देखि	- <i>dek^{hi}</i>	'ABL'

!!Case Marker which take emphatic marker also

बाट	- <i>batA</i>	'ABL'
बाटै	- <i>batA</i>	'ABL+EMPH'
मा	- <i>ma</i>	'LOC'
मै	- <i>mai</i>	'LOC+EMPH'
सँग	- <i>sāgA</i>	'COM'
सँगै	- <i>sāgAi</i>	'COM+EMPH'
सित	- <i>sitA</i>	'COM'
सितै	- <i>sitAi</i>	'COM+EMPH'
को	- <i>ko</i>	'GEN+SG'
का	- <i>ka</i>	'GEN+PL'
की	- <i>ki:</i>	'GEN+FEM'
कै	- <i>kAi</i>	'GEN+EMPH'
तिर	- <i>tirA</i>	'DIR'
तिरै	- <i>tirAi</i>	'DIR+EMPH'

!!Plural/collective marker

हरू	- <i>haru:</i>	'PL'
-----	----------------	------

!!Adverbial Postpositions which do not take emphatic marker

माथि	- <i>mat^{hi}</i>	'above'
कहाँ	- <i>kahā</i>	'in'
मुनि	- <i>muni</i>	'under'
वारि	- <i>wari</i>	'this side'
पारि	- <i>pari</i>	'that side'
वरि	- <i>wari</i>	'this side'
परि	- <i>pari</i>	'that side'
पट्टि	- <i>patṭi</i>	'towards'
प्रति	- <i>prati</i>	'for'
पालि	- <i>pali</i>	'time'

खेरि	-k ^h eri	'time'
पालि	-pali	'time'
छेउ	-ts ^h eu	'edge'
पछि	-pʌts ^h i	'later'
पछाडि	-pʌts ^h adi	'behind'
अघि	-ʌg ^h i	'before'
अगाडि	-ʌgadi	'in front of'
भरि	-b ^h ʌri	'full of'
पहिले	-pʌhile	'before'
निम्ति	-nimti	'for'
लागि	-lagi	'for the sake of'
बारे	-bare	'about'
सामु	-samu	'in front of'
सरि	-sʌri	'equivalent to'
मध्ये	-mʌd ^h je	'among'
जति	-dzʌti	'about'
पिच्छे	-pitsts ^h e	'each of'

!! Adverbial Postpositions which take emphatic marker

सहित	-sʌhit	'along with'
साथ	-sat ^h	'with'
सम्म	-sʌmma	'till'
बाहिर	-ba ^h iirʌ	'outside'
वार	-war	'this side'
पार	-par	'that side'
वर	-wʌrʌ	'this side'
पर	-pʌrʌ	'that side'
उँभो	-ũb ^h o	'up'
उँधो	-ũd ^h o	'down'
तर्फ	-tʌrp ^h ʌ	'toward'
नेर	-nerʌ	'near'
निर	-nirʌ	'near'
समक्ष	-sʌmʌkʂʌ	'in front of'
पर्यन्त	-pʌrjʌntʌ	'up to now'
खेर	-k ^h erʌ	'time'
उप्रान्त	-uprʌntʌ	'then after'
वित्तिकै	-bittikʌi	'as soon as'
साथ	-sat ^h	'with'

पख	-pAk ^h Λ	'time'
ताक	-takΛ	'time'
ताका	-taka	'time'
पाला	-pala	'time'
पालि	-pali	'time'
पटक	-pAtak	'time'
पलट	-pAlʈΛ	'time'
पश्चात्	-pAstsat	'after'
छेक	-ts ^h ekΛ	'time'
भिन्न	-b ^h itrΛ	'inside'
नजिक	-nΛdzik	'near'
भर	-b ^h Ar	'whole'
तक	-tAk	'upto'
यता	-jAta	'here'
उता	-uta	'there'
बीच	-bitsΛ	'between'
निमित्त	-nimittΛ	'for the sake of'
खातिर	-k ^h atir	'for the sake of'
अन्तर्गत	-Antarɡat	'within'
बमोजिम	-bΛmodzim	'according to'
माफिक	-map ^h ik	'in accordance with'
मुताबिक	-mutabik	'according to'
अनुसार	-Anusar	'according to'
उपर	-upAr	'on top of'
मार्फत	-marp ^h At	'via'
अलावा	-Alawa	'other than'
अतिरिक्त	-Atirikta	'in addition'
बाहेक	-bahek	'besides'
सरह	-sarAhΛ	'equivalent to'
बाबजुद	-babΛdzud ^h	'in spite of'
विरुद्ध	-wirudd ^h Λ	'against'
बापत	-bapat	'in return'
सट्टा	-sAtta	'in exchange'
बदला	-bAdAla	'in exchange'
लेखा	-lek ^h a	'number'
सुद्ध	-suddΛ	'even??'
समेत	-sAmet	'along with'

Annex - 11

Words and Affixes for Derivation in Nepali

!!---- Derivation ----- Prefixes Noun to Noun !!

प्र	<i>pra-</i>	'PFX'/1
परा	<i>para-</i>	'PFX'/2
अप	<i>apa-</i>	'PFX'/3
सम्	<i>sam-</i>	'PFX'/4
अनु	<i>anu-</i>	'PFX'/5
अव	<i>ava-</i>	'PFX'/6
दुस्	<i>dus-</i>	'PFX'/7
दुर्	<i>dur-</i>	'PFX'/8
वि	<i>vi-</i>	'PFX'/9
अधि	<i>adhi-</i>	'PFX'/10
अति	<i>ati-</i>	'PFX'/11
अभि	<i>abhi-</i>	'PFX'/12
प्रति	<i>prati-</i>	'PFX'/13
परि	<i>pari-</i>	'PFX'/14
उप	<i>upa-</i>	'PFX'/15
सह	<i>saha-</i>	'PFX'/16
स	<i>sa-</i>	'PFX'/17
कु	<i>ku-</i>	'PFX'/18
अ	<i>a-</i>	'PFX'/19
अन्	<i>an-</i>	'PFX'/20
बे	<i>be-</i>	'PFX'/21
बद	<i>bada-</i>	'PFX'/22
ला	<i>la-</i>	'PFX'/23
सु	<i>su-</i>	'PFX'/24

!! Lexicon of underived nouns for Noun to Noun derivation!

चलन	<i>tsalan</i>	'tradition'/N1
जय	<i>dzaja</i>	'victory'/N2
शब्द	<i>shabda</i>	'word'/N3
मान	<i>man</i>	'respect'/N4
शासन	<i>shasan</i>	'governance'/N5
गुण	<i>gun</i>	'attribute'/N6
परिणाम	<i>parinam</i>	'result'/N7
घटना	<i>ghatana</i>	'incident'/N8

नाश	<i>naf</i>	'damage'/N9
राज्य	<i>radzjʌ</i>	'state'/N10
वृष्टि	<i>wriʃti</i>	'rain'/N11
रुचि	<i>rutsi</i>	'interest'/N12
ध्वनि	<i>dʰwʌni</i>	'sound'/N13
योजना	<i>jodzʌna</i>	'plan'/N14
ग्रह	<i>grʌɦʌ</i>	'planet'/N15
कार्य	<i>karjʌ</i>	'work'/N16
परिवार	<i>pʌriwar</i>	'family'/N17
पुत्र	<i>putra</i>	'son'/N18
ज्ञान	<i>gjan</i>	'knowledge'/N19
आस्था	<i>ʌnastʰa</i>	'belief'/N20
इज्जत	<i>idzdzʌt</i>	'respect'/N21
नाम	<i>nam</i>	'name'/N22
वारिस	<i>waris</i>	'care'/N23
समाचार	<i>sʌmatsar</i>	'news'/N24

!!---Noun to Adjective Derivation ---Prefixes--!!

निर्	<i>nir-</i>	'PFX'/1
निः	<i>ni:</i>	'PFX'/2
नि	<i>ni</i>	'PFX'/3
वि	<i>wi</i>	'PFX'/4
निस्	<i>nis</i>	'PFX'/5
स	<i>sʌ</i>	'PFX'/6
बे	<i>be</i>	'PFX'/7
अ	<i>ʌ</i>	'PFX'/8
अन	<i>ʌnʌ</i>	'PFX'/9

!!Lexicon of underived nouns for Noun to Adjective derivation

दोष	<i>doʃ</i>	'blame'/NA1
स्वार्थ	<i>swartʰʌ</i>	'self-interest'/NA2
डर	<i>dʌr</i>	'fear'/NA3
मुख	<i>mukʰ</i>	'mouth'/NA4
फल	<i>pʰʌl</i>	'fruit'/NA5
बल	<i>bʌl</i>	'strength'/NA6
घर	<i>gʰʌr</i>	'house'/NA7
मूल्य	<i>muljʌ</i>	'cost'/NA8
मोल	<i>mol</i>	'price'/NA9

!!-----Noun to Adverb Derivation ---Prefixes--!!

आ	<i>a</i>	'PFX'/1
स	<i>sA</i>	'PFX'/2
निर	<i>nir</i>	'PFX'/3
प्रति	<i>prAti</i>	'PFX'/4

!!Lexicon of underived nouns

LEXICON PNtoAdv1

मरण	<i>mArAṇ</i>	'death'/NAdv1
हर्ष	<i>hArṣA</i>	'happiness'/NAdv2
घात	<i>g^hat</i>	'decieve'/NAdv3
हसा	<i>hApta</i>	'week'/NAdv4

!!-----Adjective to Adjective Derivation ---Prefixes--!!

सम्	<i>sAm</i>	'PFX'/1
वि	<i>wi</i>	'PFX'/2
दुर	<i>dur</i>	'PFX'/3
उन्	<i>un</i>	'PFX'/4
सु	<i>su</i>	'PFX'/5
परि	<i>pAri</i>	'PFX'/6

!!Lexicon of underived nouns

पूर्ण	<i>pu:rṇA</i>	'full'/AA1
शुद्ध	<i>śudd^hA</i>	'pure'/AA2
भेद्य	<i>b^hedjA</i>	'penetratable'/AA3
मुक्त	<i>muktA</i>	'free'/AA4
शिक्षित	<i>śikṣit</i>	'educated'/AA5
पूर्ण	<i>pu:rṇA</i>	'full'/AA6

!! Derivation by Suffixation

!!Nouns for Noun to Noun Derivation

सुन	<i>sun</i>	'gold'/N1
घाँस	<i>g^hās</i>	'grass'/N2

!! Suffixes for Noun to Noun Derivation

आर	<i>-ar</i>	'SFX'/1
ई	<i>i:</i>	'SFX'/2

!! Noun to Adjective Derivation

दया	<i>dAja</i>	'affection'/NA1
लाभ	<i>lab^h</i>	'profit'/NA2

सेवा	<i>sewa</i>	'service'/NA3
मुगल	<i>mugaɫ</i>	'Mugal'/NA4
लिम्बु	<i>limbu</i>	'Limbu'/NA5
दान	<i>dan</i>	'donation'/NA6
खर्च	<i>k^hArtsA</i>	'expense'/NA7
भिर	<i>b^hir</i>	'cliff'/NA8
रिस	<i>ris</i>	'anger'/NA9
शहर	<i>ʃʌɦʌr</i>	'city'/NA10
होस	<i>fios</i>	'consciousness'/NA11

!! Suffixes for Noun to Adjective Derivation-

अनीय	<i>-Ani:jA</i>	'SFX'/1
अक	<i>-AkA</i>	'SFX'/2
इका	<i>-ika</i>	'SFX'/3
आन	<i>-an</i>	'SFX'/4
वान	<i>-wan</i>	'SFX'/5
ई	<i>-i:</i>	'SFX'/6
आलु	<i>-alu</i>	'SFX'/7
आलो	<i>-alo</i>	'SFX'/8
आहा	<i>-afia</i>	'SFX'/9
इया	<i>-ija</i>	'SFX'/10
इयार	<i>-ijar</i>	'SFX'/11

!!-----Noun to Noun/Adjective Derivation -----!!

झापा	<i>dz^hapa</i>	'Jhapa'/NNA1
गुल्मी	<i>gulmi</i>	'Gulmi'/NNA2
इलाम	<i>ilam</i>	'Ilam'/NNA3
गाउँ	<i>gaũ</i>	'village'/NNA4
नेपाल	<i>nepal</i>	'Nepal'/NNA5

!-Suffixes for Noun to Noun/Adjective Derivation

ली	<i>-li:</i>	'SFX'/1
एली	<i>-eli:</i>	'SFX'/2
ए	<i>-e</i>	'SFX'/3
ले	<i>-le</i>	'SFX'/4
ई	<i>-i:</i>	'SFX'/5

!!-----Adjective to Noun Derivation -----!!

लामो	<i>lamo</i>	'long'/AN1
------	-------------	------------

छोटो	<i>ts^hoʈo</i>	'short'/AN1
xy	<i>-xxx</i>	'xxx'/AN2

!—Suffix for Adjective to Noun Derivation

आइ	<i>-ai</i>	'SFX'/1
अक	<i>-AkA</i>	'SFX'/2

!!----- Adjective/Noun to Noun Derivation -----!!

गरिब	<i>gArīb</i>	'poor'/ANN1
xy	<i>-xxx</i>	'xxx'/ANN2

!--Suffixes for Adjective/Noun to Noun Derivation -----

ई	<i>-i:</i>	'SFX'/1
sd	<i>-sss</i>	'SFX'/2

!!----- Verb to Noun Derivation -----!!

चुन्	<i>tsun</i>	'elect'/VN1
चुन्	<i>tsun</i>	'elect'/VN2
किट्	<i>kiṭ</i>	'fix?'/VN3
किट्	<i>kiṭ</i>	'fix?'/VN4
ढाक्	<i>ḍ^hak</i>	'cover'/VN5
जल्	<i>dzAl</i>	'burn'/VN6
चोर्	<i>tsor</i>	'steal'/VN7
हाँस्	<i>hã̃s</i>	'laugh'/VN8
पढ्	<i>pAd^h</i>	'read'/VN9
थाक्	<i>t^hak</i>	'tire'/VN10
छाप्	<i>ts^hap</i>	'print'/VN11
छान्	<i>ts^han</i>	'choose'/VN12
चिच्या	<i>tsitsja</i>	'shout'/VN13
झर्	<i>dz^hAr</i>	'drop'/VN14
ढोग्	<i>ḍ^hog</i>	'bow to feet'/VN15
राख्	<i>raK^h</i>	'keep'/VN16
दाब्	<i>dab</i>	'press'/VN17
बच्	<i>bAts</i>	'survive'/VN18
सड्	<i>sAd</i>	'decay'/VN19
रोप्	<i>rop</i>	'plant'/VN20
छेक्	<i>ts^hek</i>	'block'/VN21
चिर्	<i>tsir</i>	'saw'/VN22
बढ्	<i>bAd</i>	'grow'/VN23
सर	<i>sAr</i>	'move'/VN24

उठ्	<i>ut^h</i>	'get up'/VN25
चाल्	<i>tsal</i>	'sieve'/VN26
बेर	<i>ber</i>	'wrap'/VN27
गा	<i>ga</i>	'sing'/VN28
भिड्	<i>b^hid</i>	'fight'/VN29
जित्	<i>dzit</i>	'win'/VN30
कोर्	<i>kor</i>	'scratch'/VN31
खुल्	<i>k^hul</i>	'open'/VN32

!---Suffixes for Verb to Noun Derivation-----

आउ	<i>-au</i>	'SFX'/1
आब	<i>-ab</i>	'SFX'/2
आनी	<i>-ani:</i>	'SFX'/3
आनी	<i>-ani:</i>	'SFX'/4
अनी	<i>-Ani:</i>	'SFX'/5
अन	<i>-An</i>	'SFX'/6
ई	<i>-i:</i>	'SFX'/7
ओ	<i>-o</i>	'SFX'/8
आइ	<i>-ai</i>	'SFX'/9
आवट	<i>-awAt</i>	'SFX'/10
आ	<i>-a</i>	'SFX'/11
ओट	<i>-ot</i>	'SFX'/12
हट	<i>-hAt</i>	'SFX'/13
अना	<i>-Ana</i>	'SFX'/14
आउनी	<i>-auni:</i>	'SFX'/15
आलो	<i>-alo</i>	'SFX'/16
आब	<i>-ab</i>	'SFX'/17
अत	<i>-At</i>	'SFX'/18
अल	<i>-Al</i>	'SFX'/19
आइँ	<i>-aĩ</i>	'SFX'/20
आरो	<i>-aro</i>	'SFX'/21
औटो	<i>-Auto</i>	'SFX'/22
औती	<i>-Auti:</i>	'SFX'/23
उवा	<i>-uwa</i>	'SFX'/24
ती	<i>-ti:</i>	'SFX'/25
नी	<i>-ni:</i>	'SFX'/26
नो	<i>-no</i>	'SFX'/27
ना	<i>-na</i>	'SFX'/28

अन्त	- <i>AntA</i>	'SFX'/29
औरी	- <i>Auri:</i>	'SFX'/30
एसो	- <i>eso</i>	'SFX'/31
अस्त	- <i>AstA</i>	'SFX'/32

!!----- Verb to Adjective Derivation -----!!

मिच्	<i>mits</i>	'squeeze'/VA1
भुल्	<i>b^hul</i>	'forget'/VA2
पोस्	<i>pos</i>	'feed'/VA3
घुम्	<i>g^hum</i>	'roam'/VA4
घुम्	<i>g^hum</i>	'roam'/VA5
खप्	<i>k^hAp</i>	'bear'/VA6
पढ्	<i>pAd^h</i>	'read'/VA7
छाड्	<i>ts^had</i>	'leave'/VA8
रोप्	<i>rop</i>	'plant'/VA9
सिक्	<i>sik</i>	'learn'/VA10
बिक्	<i>bik</i>	'sell'/VA11
भाग्	<i>b^hag</i>	'flee'/VA12
छेर	<i>ts^her</i>	'pass stool'/VA13
लाग्	<i>lag</i>	'attach'/VA14

!--Suffixes for Verb to Adjective derivation-----

आहा	- <i>afia</i>	'SFX'/1
अक्कड	- <i>AkkAd</i>	'SFX'/2
इलो	- <i>ilo</i>	'SFX'/3
अन्ते	- <i>Ante</i>	'SFX'/4
अन्ता	- <i>Anta</i>	'SFX'/5
आलु	- <i>alu</i>	'SFX'/6
ऐया	- <i>Aija</i>	'SFX'/7
आ	- <i>a</i>	'SFX'/8
आर	- <i>ar</i>	'SFX'/9
आरु	- <i>aru</i>	'SFX'/10
आउ	- <i>au</i>	'SFX'/11
औटो	- <i>Au^o</i>	'SFX'/12
औटी	- <i>Auⁱ</i>	'SFX'/13
उ	- <i>u</i>	'SFX'/14

!!----- Verb to Adverb Derivation -----!!

गर्	<i>gar</i>	'do'/VAdv
-----	------------	-----------

!---Suffixes for Verb to Adverb Derivation----

उन्जेल *-undzel* 'SFX'

इन्जेल *-indzel* 'SFX'

!!----- Adverb to Adjective Derivation -----!!

भिन्न *b^hitrΛ* 'inside'

!----Suffix for Adverb to Adjective Derivation----

LEXICON SAdvtoAdj1

ई *-i:* 'SFX'

!!----- Verb to Noun conversion -----!!

खेल् *k^hel* 'play'

खोज् *k^hodz* 'search'

!!----- Verb to Noun/Adjective conversion -----!!

ठग् *t^hΛg* 'cheat'

चोर् *tsor* 'steal'

थप् *t^hΛp* 'add'

!!----- Verb to Noun Derivation (by vowel insertion) -----!!

चम्क *tsΛmkΛ* 'shine'

सम्झ *sΛmdz^hΛ* 'remember'

टल्क *tΛlkΛ* 'shine'

REFEENCES

- Abdulky, V.J. 1969. *Case grammar and Nepalese language*. MA Thesis. Cornell University.
- Acharya, Jayraj. 1991. *A descriptive grammar of Nepali and an analyzed corpus*. Washinton: Georgetown University Press.
- Acharya, Vratraj. 2058 VS. *Adharbhut Nepali vyakaran tatha rachana* (Basic Nepali grammar and composition). Lalitpur: Sajha Prakashan.
- Adhikari, Hamanga. 1993. *Samasaamayik Nepali vyakaran* (Contemporary Nepali grammar). Kathmandu: Vidhyarthi Pustak Bhandar.
- Adhikari, Hamanga. 2062VS. *Samasaamayik Nepali vyakaran* (Contemporary Nepali grammar). 3rd edn. Kathmandu: Vidhyarthi Pustak Bhandar
- Adhikari, Hemang Raj. 2052VS. *Nepali karak vyakaran* (Nepali case grammar). Kathmandu: Royal Nepal Academy.
- Aryal, Sudip, Nabin Karmacharya and Kalyan K Ghimire. 2006. 'Computational disambiguation for Nepali language parsing'. In *Nepalese linguistics*. Vol.22.
- Aryal, Sudip, Nabin Karmacharya and Kalyan K. Ghimire. 2007. 'Disambiguation based on syntactic constraints over lexical tag sequences and possible semantic constraints for Nepali language parsing'. In *Recent studies in Nepalese linguistics*.
- Awasthi, Jai Raj, Ganga Ram Gautam and Bhim Narayan Regmi (eds.). 2008. *Nepalese Linguistics*, Vol. 23, Kathmandu: Linguistic Society of Nepal.
- Ayton, J.A. 1820. *A grammar of the Nepalese language*. Calculta: Philip Pereira.
- Bal, Bal Krishna and Prajol Shrestha. 2007a. 'Nepali spellchecker'. In *PAN localization working papers*.
- Bal, Bal Krishna and Prajol Shrestha. 2007b. 'Stemmer for Nepali'. In Rai et al. (eds.) *Recent studies in Nepalese linguistics*.
- Bal, Bal Krishna and Prajol Shrestha. 2007c. 'A morphological analyzer and a stemmer for Nepali. In *PAN localization working papers*.
- Bal, Bal Krishna. 2007. 'Structure of Nepali grammar'. In *PAN localization working papers*.
- Bandhu, Churamani. 1971. A computer concordance of Nepali spoken text. [http://cqpweb.lancs.ac.uk/bandhu/index.php]

- Bandhu, Churamani. 1973. 'Clause Patterns in Nepali'. In Davis, Irvine (ed.) *Clause, sentence, and discourse patterns in selected languages of Nepal*. Part II. Summer Institute of Linguistics. Norman: University of Oklahoma.
- Beesley, Kenneth R. and Lauri Karttunen. 2003. *Finite state morphology*. CSLI Publications.
- Bharati, Akshar and Amba Kulkarni. 2007. *Sanskrit and computational linguistics*. Department of Sanskrit Studies. University of Hyderabad. [www.sanskrit.uohyd.ernet.in].
- Bista, Sanat Kumar, Birendra Keshari, Laxmi Prasad Khatiwada, Pawan Chitrakar and Srishtee Gurung. 2007. 'Nepali lexicon development'. In *PAN localization working papers*. 2004-2007.
- Bloomfield, L. 1933. *Language*. New York: Holt.
- Bögel, Tina, Mariam Butt, Annette Hauti and Sebastian Sulger (2007) *Developing a finite-state morphological analyzer for Urdu and Hindi*. . In T. Hanneforth and K.-M. W`urzner (eds.) *Finite-state methods and natural Language processing, revised papers of the sixth international workshop on finite-state methods and natural language processing*, 86–96. Potsdam University Press.
- Chapagain, Narendra. 2046VS. *Bhashatattwa* (Elements of language). Kathmandu: Ratna Pustak Bhandar.
- Chomsky, Noam. 1957. *Syntactic structures*. Mouton. The Hague.
- Dasgupta, Sajib, Naira Khan, Asif Iqbal Sarkar, Dewan Shahriar Hossain Pavel and Mumit Khan. 2007. 'Morphological analysis of inflecting compound words in Bangla'. In *PAN localization working papers*.
- Dhahal, Ballabh Mani. 1974. *A description of Nepali: literary and colloquial*. Ph.D. thesis. University of Poona (Faculty of Arts).
- Dhakal Santi Prasad. 2058VS. *Nepali sankhyabachak sabda: bikas prakriya ra bhasabaigyanik bislesan* (Nepali numerical word: evolution process and linguistic analysis). Itahari: Ms. Anjala Dhakal.
- Grierson, G. A. 1927. *Linguistic survey of India*. Vol IX. Part IV. Delhi: Motilal Banarasidass.
- Gurung, Sreshtee and Laxmi Prasad Khatiwada. 2007. 'Collation sequence in Nepali'. In *PAN localization working papers*.
- Gurung, Sristhtee and Ishwor Thapa. 2007. 'Building Nepali text-to-speech'. In Rai et al. (eds.) *Recent studies in Nepalese linguistics*.

- Hardie, A, Lohani, RR, Regmi, BN, and Yadava, YP. 2009. 'A morphosyntactic categorisation scheme for the automated analysis of Nepali'. In Singh, R. (ed.) *Annual review of South Asian languages and linguistics*. Mouton de Gruyter.
- Hardie, A. 2008. 'A collocation-based approach to Nepali postpositions'. In *Corpus linguistics and linguistic theory* 4(1): 19-62.
- Hussain, Sara. 2004. *Finite-state morphological analyzer for Urdu*. MS Thesis. National University of Computer and Emerging Sciences. Lahor.
- Hyoju, Dibyendra and Vijay Kumar Shrestha. 2007. 'Dictionary software for contemporary Nepali dictionary. In Rai et al. (eds.) *Recent studies in Nepalese linguistics*.
- Islam, Md. Zahural. 2007. 'Research report on Bangla verb and noun morphological analysis. In *PAN localization working papers*.
- Johson, C. Douglas. 1972. *Formal aspects of phonological description*. Mouton.
- Jurafsky, Daniel and James H. Martin. 2000. *Speech and language processing: an introduction to natural language processing, computational linguistics and speech recognition*. Pearson Education.
- Karttunen Lauri. 2000. 'Application of finite-state transducers in natural language processing'. In *the proceedings of CIIA-2000*. Lecture notes in computer science. Springer Verlag.
- Keshari, Birendra, Jaganath Bhatta and Sanat Kumar Bista. 2005. 'Nepali part-of-speech guesser and its application in lexicon building'. In *Proceedings of international conference on natural language processing-ICON*. IIT Kanpur. India.
- Khan, Abid Khan and Fatima Tuz Zuhra. 2007. 'The computational morphology of Pashto noun'. In *South Asian language review*. Vol.XVII. No. 1. January.
- Khatiwada, Laxmi Prasad and Sreshtee Gurung. 2007. 'Nepali lexicon'. In *PAN localization working papers*.
- Koskeniemmi, Kimmo. 1983. *Two-level morphology: a general computational model for word-form recognition and production*. Ph.D. thesis. University of Helsinki.
- Lohani, Ram Raj. 1999. *Complex predicates in Nepali*. MA Thesis. Tribhuvan University. Kathmandu.

- Makedonski, Philip. 2005. *Finite state morphology: the Turkish nominal paradigm*. BA thesis. Eberhard Karls Universität Tübingen. Germany.
- Mathews, P.H. 1991. *Morphology*. 2nd edn. Cambridge: Cambridge University Press.
- Megerdoomian. 2003. *Persian computational morphology: a unification-based approach*. Computing research laboratory. New Mexico State University. New Mexico.
- PAN localization working papers*. 2007. PAN Localization Project Teams. IDRC and CRDI. Canada. www.nu.edu.pk and www.idrc.ca.
- Pandey, Shailesh Bdr., Krishna Sapkota and Prajol Shrestha. 'Stochastic and rule based part of speech tagger for Nepali Language'. [www.geocities.com/shaileshpandeynec/skima/MPPnec.ppt]
- Pandit, G.H. 2051 VS (1969 VS). *Chandrikaa: gorkhaabhaasaa-vyaakaran* (The grammar of Gorkha Language). Kathmandu: Sajha Prakasan.
- Parajuli, Krishna Prasad. 2001. *Ramro rachana mitho Nepali* (A grammar of Nepali). Kathamndu: Sahayogi Press.
- Paudel, Subhash, Sudeep Ghimire, Suresh Kumar Bhattarai and Uttam Shrestha. 2006. 'Morphological analyzer in Nepali'. Department of Electronics and Computer Engineering, Institute of Engineering. Tribhuvan University.
- Payne, Thomas E. 1997. *Describing morphosyntax: a guide for field linguists*. Cambridge: Cambridge University Press.
- Pokharel Madhav Prasad. 2054VS. *Nepali vaakarana* (Nepali Syntax). Kathmandu: Royal Nepal Academy.
- Pokharel, Madhap P. 2010a. 'Noun class agreement in Napali'. In *Kobe papers in linguistics*. pp. 40-59.
- Pokharel, Madhav Prasad. 1991. 'Compound verbs in Nepali'. In *Contribution to Nepalese Studies*, Vol.18. No.2. pp.149-173.
- Pokharel, Madhav Prasad. 2010b. 'A strategic search for the general formula of verb root derivation in Nepali. In *Nepalese linguistics*. Vol.25.
- Pokharel, Madhav Prasad. 2053VS. *Kyampus stariya Nepali vakya-vyakaran* (Campus level Nepali Syntax). Kathmandu: Navin Prakashan.
- Pradhan, K.L.B. 1982. *The structure of the simple clause in Nepali*. Ph.D. Thesis. University of Wisconsin- Madison.
- Pradhan, P. 1944. *Nepali vyaakarana* (Nepali Grammar). 4th ed. Calcutta: Macmillan and Co. Ltd.

- Prasain, Balam. 2008. 'Computational analysis of Nepali basic verbs (written form)'. In *Nepalese linguistics*. Vol. 23.
- Prasain, Balam. 2010. 'A computational analysis of Nepali basic nouns: a finite state approach'. In *Nepalese linguistics*. Vol. 25.
- Rai, Novel Kisor, Yogendra P Yadava, Bhim Narayan Regmi and Balam Prasain (eds.). 2007. *Recent studies in Nepalese linguistics*. Kathmandu: Linguistics Society of Nepal
- Sharma, Mohan Raj. 2056VS. *Sabdarachana ra varnabinyas: vakyatattwa, abhivyakti ra pathaharu* (Word formation, phonology: Syntax, expression and texts). Kathmandu: Navin Prakashan.
- Sharma, Tara Nath. 1980. *The Auxiliary in Nepali*. Ph.D. Thesis. Graduate School of the University of Wisconsin-Madison.
- Shrestha, Niraj. 2008. *Nepali word sense disambiguation using Lesk algorithm*. MA thesis. Department of Computer Science and Engineering, School of Engineering. Kathmandu University.
- Shrivastava, Manish, Nitin Agrawal, Bibhuti Mahapatra, Smriti Singh and Pushpak Bhattacharya. 'Morphology based natural language processing tools for Indian languages'. [http://www.cse.iitk.ac.in/users/iriss05/m_shrivastava.pdf]
- Siddiqui, Tanveer and U.S. Tiwari. 2008. *Natural language processing and information retrieval*. Oxford: Oxford University Press
- Sigdel, Somnath. 2055VS (1976VS). *Madhyachandrikaa* (Nepali Grammar). Kathmandu: Sajha Prakasan.
- Stump, Gregory T. 2001. 'Inflectional morphology: a theory of paradigm structure'. *Cambridge studies in linguistics*. Vol. 93. Cambridge University Press.
- Tiwari, Gopal Prasad. 2030VS. *Nepali bhaasaako banot* (The structure of Nepali language). Kathmandu: Royal Nepal Academy.
- Turnbull, R. A. 1887. *A Nepali grammar and English-Nepali and Nepali-English vocabulary*. Darjeeling: NL Roy.
- Upadhyaya, Bhim Prasad, Birendra Keshari and Sanat Kumar Bista. 2005. 'Morphological analyzer for Nepali language'. In *Research KU-CSE-1-2005*. Department of Computer Science and Engineering. Kathmandu University.
- Wallace, William David. 1985. *Subjects and subjecthood in Nepali: an analysis of Nepali clause structure and its challenges to relational grammar and*

government and binding. Ph.D. Thesis. Graduate College of the University of Illinois at Urbana-Champaign.

- Yadava, Y.P., Hardie, A., Lohani R.R., Regmi B.N., Gurung, S., Gurung, A., McEnery, T., Allwood, J., and Hall, P. 2008. 'Construction and annotation of a corpus of contemporary Nepali'. In *Corpora* 3(2): 213-225.
- Yadava, Yogendra P. 2003. 'Language'. in *Population monograph of Nepal*, Vol.1. Ch.4. Central Bureau of Statistics. Nepal.
- Ziai, Ramon. 2006. *Finite state methods applied to verbal inflection in Persian*. A BA thesis. Eberhard-Karls-Universität Tübingen.