DEPARTMENT OF COMPUTER SCIENCE
SCHOOL OF PHYSICAL SCIENCES
ASSAM UNIVERSITY SILCHAR
(A CENTRAL UNIVERSITY CONSTITUTED UNDER ACT XIII OF 1989)
Silchar-788011, Assam, India

Date: …….

DECLARATION

I, Aadil Ahmad Lawaye, bearing Registration No. Ph. D/ 1782/11 Dated 22.09.2011, hereby declare that the subject matter of the thesis entitled "Design and Implementation of Part of Speech Tagger for Kashmiri" is the record of works done by me and that the contents of the thesis did not form the basis for award of any other degree to me or to anybody else to the best of my knowledge. The thesis has not been submitted in any other University / Institute.

This thesis is being submitted to Assam University for the degree of Doctor of Philosophy in Computer Science.

(Aadil Ahmad Lawaye)
Research Scholar

Place:
Date:

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor, guide and mentor Prof. Bipul Syam Purkayastha for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my Ph.D study. I am also thankful to other faculty members and non-teaching staff of the Department of Computer Sciences for their support and cooperation.

I convey my sincere gratitude to Prof. Pushpak Bhattacharyya of IIT Bombay and Prof. Aadil Amin Kak of Kashmir University for their precious help and encouragement.

Credit goes to my family members especially to my beloved Aba, Naana, Daddy and Baiji for constant support and encouragement. I am also thankful to Didi and Rani di for believing in me and nudging me on. I should also mention Sadaf, Choti, Aifa, Insha, Uzma, Qamran, Arfat and Shailah for always being there for me.

I am extremely thankful to my friend's Aabid, Aashiq, Nazima, Sehar, Mansoor, Aijaz, Huma, Samee, Imran and many others whose names might have escaped me this time for their love and support.

And, last and definitely not the least, it was the Grace of the Almighty that made it possible, and many many thanks to Him.

Date: (AADIL AHMAD LAWAYE)

CONTENTS

List of Tables and Figures

List of Tables

List of Figures

Kashmiri Consonants and Vowels

Manner of Articulation	State of Glottis	Place of Articulation							
		Bilabials	**Labio-dentals**	**Dentals**	**Alveolars**	**Palato-alveolars**	**Palatals**	**Velars**	**Glottals**
Stops	**voiceless**	p, ph		t, th	T, Th			k, kh	
	voiced	b		d	D			g	
Fricatives	**voiceless**				s	s			h
	voiced		v		z				
Affricates	**voiceless**				ts, tsh	c, ch			
	voiced					j			
Nasals	**voiced**	m			n			N	
Laterals	**voiced**				l				
Trills	**voiced**				r				
Semivowels	**voiced**	w					y		

Chart of Kashmiri Consonants

		Front	**Central**	**Back**
High	Short	i	I	u
	Long	i:	I:	u:
Mid	Short	e	ə	o
	Long	e:	ə:	o:
Low	Short	ɛ	a	ɔ
	Long	ɛ:	a:	ɔ:

Chart of Kashmiri Vowels

List of Abbreviations

M	Masculine
F	Feminine
S	Singular
P	Plural
nom	Nominative
dat	Dative
abl	Ablative
erg	Ergative
gen	Genitive
Ip	First person
IIp	Second person
IIIp	Third person
prt	Present
pst	Past
fut	Future
prog	Progressive
POS	Part-of-Speech
ILMT	Indian Language Machine Translation
IL-POST	Indian Language Part of Speech Tagging
TnT	Trigrams'n'Tags
TBL	Transformation Based Learning
HMM	Hidden Morkov Model
BNC	British National Corpus
TOSCA	Tools for Syntactic Corpus Analysis