

मशीन अनुवाद हेतु-कॉर्पोरा आधारित हिंदी-अंग्रेजी संज्ञापद
सम्मिलन

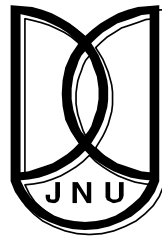
Corpora Based Hindi-English Noun Phrase Mapping for Machine
Translation

जवाहरलाल नेहरू विश्वविद्यालय में एम.फिल. (हिंदी अनुवाद) उपाधि हेतु प्रस्तुत
लघु शोध-प्रबंध

शोध निर्देशक
प्रो. चमन लाल

सह-शोध निर्देशक
डॉ. गिरीश नाथ झा

शोधार्थी
सुमेध खुशालराव हाडके



भारतीय भाषा केंद्र
भाषा, साहित्य एवं संस्कृति अध्ययन संस्थान
जवाहरलाल नेहरू विश्वविद्यालय
नई दिल्ली-110067

2012



जवाहरलाल नेहरू विश्वविद्यालय
JAWAHARLAL NEHRU UNIVERSITY
Centre of Indian Languages
School of Language, Literature & Culture Studies,
New Delhi – 110067, India

Dated : 26th July, 2012

DECLARATION

I declare that the work done in this dissertation entitled in Hindi Transcription 'MASHIN ANUVAD HETU- KARPORA ADHARIT HINDI – ANGREZI SANJNAPAD SAMMILAN' (**Corpora Based Hindi-hrase English Noun P Mapping for Machine Translation**) submitted by me for the award of the degree of **Master of Philosophy**, is an original research work and has not been previously submitted for any other degree or diploma in any other university or institution.

Sumedh khushalrao Hadke
(**Research Scholar**)

Prof. Chaman Lal
(**SUPERVISOR**)
CIL / SLL&CS

Prof. Ram Bux
(**CHAIRPERSON**)
CIL / SLL&CS

Dr. Ranjeet Kr. Saha
(**CO-SUPERVISOR**)
CIL / SLL&CS

Dr. Girish Nath Jha
(**CO-SUPERVISOR**)
Special Centre for Sanskrit Studies

डॉ. भीमराव रामजी आंबेडकर
को समर्पित.....!

आभार

में, उन सभी के प्रति आभार व्यक्त करता हूँ, जिन्होंने इस शोध कार्य के लिए प्रत्यक्ष और अप्रत्यक्ष रूप से सहायता की।

सर्वप्रथम मैं प्रो.चमन लाल और डॉ. रणजीत साहा सर का आभार व्यक्त करना चाहता हूँ, जिन्होंने शोध विषय के चयन से लेकर इस शोध को सफल बनाने तक उचित सुझाव और मार्गदर्शन दिया।

डॉ. गिरीश नाथ झा सर (वि.स.अ.के., जे.एन.यू.) का मैं विशेष आभार व्यक्त करना चाहता हूँ, जिन्होंने मुझे विषय चयन से लेकर इस विषय के लिए 'भारतीय भाषा कॉर्पोरा उपक्रम' परियोजना की सामग्री (कॉर्पोरा) पर अध्ययन करने के लिए प्रेरित किया। साथ ही, इस परियोजना में कार्य करने के लिए मुझे अवसर प्रदान किया। यह कार्य करते समय ही मुझे इस शोध विषय की संपूर्ण जानकारी मिली। इस शोध के दौरान प्रत्येक अध्याय के लिए सर के साथ चर्चा कर सुझाव मिलता रहा। इसके साथ ही मशीनी अनुवाद और कॉर्पोरा अध्ययन की बारीकियों को समझने के लिए हर समय प्रेरित करते रहे।

अपने वरिष्ठ मित्र स्वाती ठाकुर और नरेन्द्र कुमार का आभार व्यक्त करना मेरे लिए औपचारिकता निभाने जैसे होगा। उन्होंने जे.एन.यू. की प्रवेश परीक्षा से लेकर इस शोध कार्य की प्रक्रिया तक समय-समय पर मुझे सहायता और मार्गदर्शन करते रहे, इसलिए मैं उनका भी आभार व्यक्त करता हूँ।

सारदा प्रसाद सर का सहयोग भी मेरे लिए महत्वपूर्ण रहा है, जिनके सहयोग के बिना यह शोध कार्य सफल नहीं हो पाता। उन्होंने मुझे हमेशा शोध कार्य को बेहतर बनाने के लिए प्रेरित किया और शोध कार्य पूर्ण होने तक मार्गदर्शन और सुझाव देते रहे।

दिवाकर मिश्रा सर का भी विशेष आभार व्यक्त करना चाहूँगा, जिन्होंने इस शोध-प्रबंध की प्रस्तुति को बारीकियों से देखा और इस शोध-प्रबंध में उचित सुझाव के साथ इसे नया रूप दिया। साथ ही, पिंकी नैनवानी का भी जितना आभार व्यक्त करूँ, उतना कम ही होगा। उनका सहयोग मेरे लिए बहुत उपयोगी रहा है। उनसे मुझे 'कंप्यूटेशनल भाषाविज्ञान' और 'हिंदी-अंग्रेजी व्याकरण' के बारे में विविध जानकारी मिली तथा इस शोध के चतुर्थ अध्याय में 'स्थानांतरण व्याकरण के नियम' बनाने में सहयोग भी मिला।

नृपेन्द्र पाठक को भी धन्यवाद देना चाहूँगा, जिनसे इस शोध विषय पर चर्चा करने तथा उनके शोध-प्रबंध का आधार लेने से विशेष सहयोग मिला।

इस शोध के दौरान कई मित्रों और वरिष्ठों ने मुझे सहायता की उनमें प्रकाश कांबले, नितिन गायकवाड़, एशा बैनर्जी, आत्मप्रकाश, रंजीत, डॉ. दिलीप दास, धीरज, गजेन्द्र, सतरुद्र किरन कुंभारे आदि का आभार व्यक्त करना चाहूँगा। अपने सहपाठी मित्रों में अजय, मुन्नी, लखिमा, चोरोल, रमेश, दर्शनी, मीना, सीमा सभी का मैं आभार व्यक्त करना चाहता हूँ, जिनका सहयोग इस शोध के दौरान हमेशा मुझे मिलता रहा है।

माता-पिता और भाई-बहन का स्नेह और सहयोग हमेशा मुझे मिलता रहा है, जिससे यह शोध कार्य पूर्ण हुआ।

भारतीय भाषा केंद्र के रावतजी, रमेश भाई का भी आभार व्यक्त करता हूँ, जो हमेशा इस शोध के दौरान महत्वपूर्ण जानकारी देते रहे।

वर्धा के महात्मा गांधी विश्वविद्यालय और जे.एन.यू. के ग्रंथालय से शोध संबंधी आवश्यक पुस्तकें कर्मचारियों ने उपलब्ध कराईं, इसलिए उन्हें भी धन्यवाद देना चाहता हूँ।

इस शोध कार्य में कितनी सफलता प्राप्त हुई, यह मार्गदर्शक, परीक्षक, अध्येता तथा शोधार्थी ही जान पाएँगे। अगर इस शोध से 'मशीनी अनुवाद और कॉर्पोरा अध्ययन' के विषय क्षेत्र पर कार्य करनेवाले शोधार्थियों, पाठकों, या विद्वज्जनों को थोड़ी भी सहायता मिलती है, तो मुझे इस कार्य की सार्थकता और सफलता पर प्रसन्नता होगी।

अतः मैं सभी को धन्यवाद देकर उनके प्रति आदर प्रस्तुत करता हूँ।

- सुमेध खुशालराव हाडके

संक्षिप्ताक्षर
Abbreviations

स.प./सं.प.	संज्ञा पदबंध
क्रि.प.	क्रिया पदबंध
म.अनु.	मशीनी अनुवाद
अ.भा.	अभिकलनात्मक भाषाविज्ञान
प्रा.भा.स.	प्राकृतिक भाषा संसाधन
कृ.बु.	कृत्रिम बुद्धि
स्रो.भा.	स्रोत भाषा
ल.भा.	लक्ष्य भाषा
भा.भा.प्रौ.वि.	भारतीय भाषाओं के लिए प्रौद्योगिकी विकास
अं.हि.म.अ.तंत्र	अंग्रेजी-हिंदी मशीनी अनुवाद तंत्र
भा.भा.कॉ.उ.	भारतीय भाषा कॉर्पोरा उपक्रम
श.मा.प्रा.	शक्ति मानक प्रारूप
नि.	निर्धारक
विशे.पद.	विशेषण पदबंध
संबंध. उप.	संबंधसूचक उपवाक्य
पूर्.विशे.	पूर्व विशेषक
प.विशे.	पश्च विशेषक
अभि.	अभिकर्ता
स्था.व्या.नि.	स्थानांतरण व्याकरण नियम
IL-1	एक भारतीय भाषा
IL-2	अन्य भारतीय भाषा

M.T.	Machine Translation
ILMT	Indian Language Machine Translation
C.L.	Computational Linguistics
NLP	Natural Language Processing
A.A.I.	Applied Artificial Intelligence
S.L.	Source Language
T.L	Target Language
NP	Noun Phrase
VP	Verb Phrase
POS	Part of Speech
QUANT	Quantifiers
DET	Determiners
EHMT	English Hindi Machine Translation
MAHT	Machine Aided Human Translation
HAMT	Human Aided Machine Translation
SMTS	Statistical Machine Translation System
RBMT	Rule Based Machine Translation
TDIL	Technical Development for Indian Languages
HEBMT	Hybrid Example Based Machine Translation
TAG	Tree Adjoining Grammar
NCFST	National Centre for Software Technology
UNL	Universal Networking Language
IIIT	International Institute of Information Technology
IISc	Indian Institute of Science
C-DAC	Centre for Development of Advance Computing
SMTS	Sampark Machine Translation System
ILCI	Indian Language Corpora Initiative
BIS	Bureau of Indian Standards
PENN.	University of Pennsylvania (Tagset)
CPG	Computational Paninian Grammar

भारतीय भाषा कॉर्पोरा अंकन के लिए भाषाई मानक शब्दवर्ग श्रेणियाँ और लेबल

Linguistic Standards for Indian Language (IL) Corpora Annotation

POS Categories and Labels

(June 12, 2010)

POS for Hindi (हिंदी के लिए प्रयुक्त शब्दवर्ग अंकन)

Sl. No	Category		Label	Annotation Convention	Examples	Remarks
	Top level	Subtype (level 1)				
1	Noun ()		N	N	ladakaa, raajaa, kitaaba	
1.1		Common	NN	N__NN	kitaab, kalam, cashmaa	
1.2		Proper	NNP	N__NNP	Mohan, Ravi, Rashmi	
1.3		Nloc (Non-locative)	NST	N__NST	(KE)upara, niice, aage, piiche, bad, subah-sham,	
2	Pronoun		PR	PR	Yaha, vaha, jo	
2.1		Personal	PRP	PR__PRP	Vaha, main, tuma, ve,vahi, esake, aapka,aap	
2.2		Reflexive	PRF	PR__PRF	Apanaa, hum swayam, khuda	
2.3		Relative	PRL	PR__PRL	Jo, jis, jab, jahaaM, usaka esase,Unme,eise	
2.4		Reciprocal	PRC	PR__PRC	Paraspara, aapasa	
2.5		Wh-word	PRQ	PR__PRQ	Kauna, kab, kahaaM	
		Indefinite	PRI	PR__PRI	Koii, kis	
3	Demonstrative		DM	DM	Vaha, jo, yaha,	
3.1		Deictic	DMD	DM__DM D	Vaha, yaha,en, YE, esaki,esme ,eske	
3.2		Relative	DMR	DM__DM R	jo, jis	
3.3		Wh-word	DMQ	DM__DM Q	kis, kaun	
		Indefinite	DMI	DM__DMI	Koi, kisi	

4	Verb		V	V	giraa, gayaa, sonaa, haMstaa, hai, rahaa	
4.1		Main	VM	V__VM	giraa, gayaa, sonaa, haMstaa,	
4.2		Auxiliary	VAUX	V__VAUX	hai, rahaa, huaa,	
5	Adjective		JJ	JJ	sundara, acchaa, baRaa, kam	
6	Adverb		RB	RB	jaldii, teza,	
7	Postposition		PSP	PSP	ne, ko, se, mein, ka, par	
8	Conjunction		CC	CC	aur, agar, tathaa, kyonki	
8.1		Co-ordinator	CCD	CC__CCD	aur, balki, parantu, ya, va	
8.2		Subordinator	CCS	CC__CCS	agar, kyonki, to, ki, lekin, balki, jab, yadi, khastour, aamtour	
9	Particles		RP	RP	to, bhii, hii	
9.1		Default	RPD	RP__RPD	to, bhii, hii	
9.2		Interjection	INJ	RP__INJ	are, he, o	
9.3		Intensifier	INTF	RP__INTF	bahut, behada, adhik, sabse, atyant	
9.4		Negation	NEG	RP__NEG	nahiin, mat, binaa	
10	Quantifiers		QT	QT	thoRaa, bahut, kuch, ek, pahalaa	
10.1		General	QTF	QT__QTF	thoRaa, bahut, kuch, sabhi khub, har, kafi	
10.2		Cardinals	QTC	QT__QTC	eka, do, tiina,	
10.3		Ordinals	QTO	QT__QTO	pahalaa, duusaraa, pratham	
11	Residuals		RD	RD		
11.1		Foreign word	RDF	RD__RDF		A word written in script other than the script of the original text
11.2		Symbol	SYM	RD__SYM	\$, &, *, (,)	For symbols such as \$, & etc

11.3		Punctuation	PUNC	RD__PUN C	., : ; ?	Only for punctuatio ns
11.4		Unknown	UNK	RD__UNK		
11.5		Echowords	ECH	RD__ECH	(Paanii-) vaanii, (khaanaa-) vaanaa	

The annotation is to be done using the lowest level tag of the type hierarchy. Once the lower level tag is selected, the higher level tags should be stored automatically.

This POS tag set for Hindi Document was :

Prepared by: Dipti Misra Sharma, LTRC, IIIT-H, Date: 04-06-2010

Revised by: Dipti Misra Sharma Date : 22-03-2010

<http://www.ldcil.org/up/morph.html>

Penn Treebank Tagset for English

(अंग्रेज़ी के लिए प्रयुक्त पेन्सिलवेनिया विश्वविद्यालय के पेन टैगसेट के लेबल)¹

POS Tag	Description	Example
CC	coordinating conjunction	and
CD	cardinal number	1, third
DT	determiner	the
EX	existential there	<i>there is</i>
FW	foreign word	d'hoevre
IN	preposition/subordinating conjunction	in, of, like
JJ	adjective	green
JJR	adjective, comparative	greener
JJS	adjective, superlative	greenest
LS	list marker	1)
MD	modal	could, will
NN	noun, singular or mass	table
NNS	noun plural	tables
NNP	proper noun, singular	John
NNPS	proper noun, plural	Vikings
PDT	predeterminer	<i>both</i> the boys
POS	possessive ending	friend's

¹ M. Marcus, Beatrice Santorini and M.A. Marcinkiewicz: Building a large annotated corpus of English: The Penn Treebank. In *Computational Linguistics*, volume 19, number 2, pp.313-330.

PRP	personal pronoun	I, he, it
PRP\$	possessive pronoun	my, his
RB	Adverb	however, usually, naturally, here, good
RBR	adverb, comparative	better
RBS	adverb, superlative	best
RP	Particle	give <i>up</i>
TO	To	<i>to go, to him</i>
UH	Interjection	uhhuhhuhh
VB	verb, base form	take
VBD	verb, past tense	took
VBG	verb, gerund/present participle	taking
VBN	verb, past participle	taken
VBP	verb, sing. present, non-3d	take
VBZ	verb, 3rd person sing. present	takes
WDT	wh-determiner	which
WP	wh-pronoun	who, what
WP\$	possessive wh-pronoun	whose
WRB	wh-abverb	where, when

अनुक्रमणिका

(INDEX)

▪ समर्पण	
▪ आभार	iv
▪ संक्षिप्ताक्षर	vi
▪ हिंदी-अंग्रेजी कॉर्पोरा के लिए प्रयुक्त शब्दवर्ग अंकन लेबल (Pos Tagset labels)	viii
▪ आकृतियों, टेबल, सारणियों और फलक चित्रों (Screen shots) की सूची	xvii
▪ प्रस्तावना	xviii
अध्याय - प्रथम : मशीनी अनुवाद और त्रुटि विश्लेषण	1-40
1.1 मशीनी अनुवाद का परिचय	1
1.1.1 मशीनी अनुवाद के प्रकार	3
1.1.2 मशीनी अनुवाद की प्रक्रिया	3
1.2 विश्व स्तर पर मशीनी अनुवाद सिस्टम का विकास	5
1.2.1 विश्व में मशीन अनुवाद का संक्षिप्त इतिहास	5
1.2.2 गूगल अनुवाद	7
1.2.2.1 गूगल अनुवाद प्रक्रिया	8
1.2.2.2 गूगल के संस्करण	9
1.2.2.2.1 Android संस्करण	9
1.2.2.2.2 IOS संस्करण	10
1.2.3 माइक्रोसॉफ्ट अनुवादक	13
1.2.3.1 बिंग अनुवाद तंत्र की विशेषताएं	13
1.2.3.2 बिंग अनुवाद तंत्र का प्रयोग	15
1.2.4 याहू बेबेल फिश	17
1.3 भारत के प्रमुख मशीनी अनुवाद तंत्र	19
1.3.1 आंग्ल-भारती	19
1.3.2 आंग्ल-भारती-II	20
1.3.3 आंग्ल हिन्दी	21
1.3.4 अनुसारक	22
1.3.5 अणुभारती	23
1.3.6 मंत्र-राजभाषा	24
1.3.7 मात्रा	24
1.3.8 शक्ति और शिवा	25
1.3.9 अनुवादक	26

1.3.10	UCSG (MAT) - आधारित अंग्रेजी-कन्नड़ मशीनी अनुवाद तंत्र	26
1.3.11	UNL - आधारित अंग्रेजी, हिंदी एवं मराठी भाषा का मशीन अनुवाद तंत्र	26
1.3.12	तमिल-हिंदी अनुसारक एवं अंग्रेजी-तमिल मशीन अनुवाद तंत्र	27
1.3.13	जादवपुर विश्वविद्यालय का अंग्रेजी-हिंदी मैट (MAT) और 'अनुवाद' तंत्र	27
1.3.14	असमिया मशीनी अनुवाद तंत्र	28
1.3.15	ओडिसी मशीनी अनुवाद तंत्र	28
1.3.16	आई.बी.एम (IBM) अंग्रेजी से हिंदी मशीनी अनुवाद तंत्र	28
1.3.17	लीला तंत्र	29
1.3.18	संपर्क	29
1.3.19	हिंग्लिश	29
1.3.20	अन्य मशीनी अनुवाद तंत्र	29
1.4	त्रुटि विश्लेषण	30
1.5	संबंधित शोध कार्य	37
1.6	उपलब्ध अन्य शोध कार्यो से वर्तमान शोध की भिन्नता	38
1.7	शोध प्रविधि	39
1.8	निष्कर्ष	40

अध्याय - द्वितीय : हिंदी और अंग्रेजी की वाक्यसंरचना और संज्ञापदों का व्यतिरेकी विश्लेषण 41-65

2.1	हिंदी-अंग्रेजी भाषाओं की संरचना	42
2.1.1	भाषाओं के वाक्य का स्वरूप	43
2.1.2	वाक्य संरचना के आधार	44
2.1.2.1	शब्दक्रम या पदक्रम	44
2.1.2.2	अन्विति	45
2.1.2.3	कारक संबंध	46
2.1.2.4	अर्थ-संगति	46
2.1.2.5	अध्याहार या लोप	47
2.1.3	वाक्य के प्रकार	47
2.1.3.1	सरल वाक्य	47
2.1.3.2	उपवाक्य	48
2.1.3.3	संयुक्त वाक्य	48
	क) संयोजक संबंध	49
	ख) विभाजक संबंध	49
	ग) विरोधवाची संबंध	49
	घ) परिणामवाची संबंध	49

2.1.3.4 मिश्र वाक्य	50
क) संज्ञा उपवाक्य	50
ख) विशेषण उपवाक्य	50
ग) क्रिया विशेषण उपवाक्य	51
2.2 हिंदी संज्ञा पदबंध	51
2.2.1 निर्धारक (Determiners)	53
2.2.2 पूर्व विशेषक (Pre-modifiers)	53
2.2.3 पश्च विशेषक (Post modifiers)	54
2.3 अंग्रेजी संज्ञा पदबंध	54
2.3.1 निर्धारक	54
2.3.2 पूर्व विशेषक	55
2.3.3 पश्च विशेषक	55
2.4 हिंदी और अंग्रेजी संज्ञा पदबंध घटकों की तुलना	56
2.5 संज्ञा पदबंध के प्रकार	57
2.5.1 सामान्य संज्ञा पदबंध	57
2.5.2 संयुक्त संज्ञा पदबंध (Compound Noun Phrase)	58
2.5.3 मिश्र संज्ञा पदबंध (Complex Noun Phrase)	58
क) पूरक उपवाक्य (Complement Clause)	58
ख) संज्ञा उपवाक्य (Noun clause)	58
ग) संबंधवाचक उपवाक्य (Relative Clause)	59
घ) कृदंत/वाला विशेषक (Participial Modifiers)	59
2.6 हिंदी अंग्रेजी संज्ञापदों का व्यतिरेकी विश्लेषण	59-65
अध्याय - तृतीय : मशीनी अनुवाद में स्थानांतरण व्याकरण की भूमिका	66-80
3.1 स्थानांतरण व्याकरण का परिचय	66
3.2 स्थानांतरण व्याकरण (TG) का मशीनी अनुवाद में प्रयोग	69
3.3 संपर्क मशीनी अनुवाद तंत्र में स्थानांतरण व्याकरण नियम का प्रयोग	70
3.4 तेलुगु-हिंदी भाषा में स्थानांतरण से होनेवाला बदलाव और स्थानांतरण व्याकरण नियम	72
3.5 शक्ति मशीनी अनुवाद तंत्र में स्थानांतरण व्याकरण का प्रयोग	76-80
अध्याय - चतुर्थ : कार्पोरा आधारित हिंदी-अंग्रेजी संज्ञापद का सम्मिलन	81-107
4.1 सम्मिलन का परिचय	81
4.2 हिंदी - अंग्रेजी संज्ञापद सम्मिलन के प्रमुख घटक	82
4.3 अंग्रेजी प्रजनक	87
4.4 स्थानांतरण व्याकरण	88

4.5	कारक (परसर्ग-पूर्वसर्गों) का सम्मिलन और स्थानांतरण व्याकरण नियम	88
4.5.1	कर्ता कारक (agentive) का विभाजन और स्थानांतरण व्याकरण नियम	89
4.5.2	कर्म कारक (Accusative/objective) का विभाजन और स्थानांतरण व्याकरण नियम	91
4.5.3	करण कारक (Instrumental) का विभाजन और स्थानांतरण व्याकरण नियम	92
4.5.4	संप्रदान कारक (Beneficiary) का विभाजन और स्थानांतरण व्याकरण नियम	93
4.5.5	अपादान कारक (Ablative) का विभाजन और स्थानांतरण व्याकरण नियम	94
4.5.6	अधिकरण कारक (Locative) का विभाजन और स्थानांतरण व्याकरण नियम	95
4.5.7	संबंध कारक (Genitive) का विभाजन और स्थानांतरण व्याकरण नियम	96
4.5.8	संबोधन कारक (Vocative) का विभाजन और स्थानांतरण व्याकरण नियम	97
4.6	विशेषण का सम्मिलन (Adjective Mapping) और स्थानांतरण व्याकरण नियम	98
4.6.1	संख्यावाचक विशेषण का सम्मिलन	101
4.7	संयुक्त और मिश्र वाक्यसंरचना में समुच्चयबोधक (Conjunctive) का सम्मिलन	103
4.8	व्याकरण कोटियों का सम्मिलन (Mapping of grammatical Categories)	104
4.9	शब्दक्रम सम्मिलन (Word order Mapping)	106
	▪ उपसंहार	108-111
	▪ संदर्भ ग्रंथ-सूची	112-118
	▪ परिशिष्ट	119-128

आकृतियों, टेबल और फलक चित्रों (Screen Shots) की सूची

आकृतियाँ	Page No.
1. मशीनी अनुवाद (अंतरण) की प्रक्रिया के घटक	4
2. हिंदी-अंग्रेजी वाक्य की व्यतिरेकी संरचना और शब्दक्रम	68
3. मशीनी अनुवाद की संरचना और स्थानांतरण व्याकरण की प्रक्रिया	70
4. शक्ति मशीनी अनुवाद तंत्र में प्रयुक्त स्थानांतरण व्याकरण नियम	76
टेबल गूगल अनुवाद भाषाओं के संस्करण	10
1. गूगल और बिंग मशीनी अनुवाद तंत्र का निर्गत अनुवाद	35
2. हिंदी- अंग्रेजी भाषा की तुलनात्मक संरचना	43
3. एकवचन और बहुवचन के प्रत्यय	52
4. हिंदी-अंग्रेजी संज्ञा पदबंध घटकों की तुलनात्मक संरचना	56
5. हिंदी-अंग्रेजी कारक परसर्ग और पूर्वसर्गों की भिन्नता	89
फलक - चित्र (Screen Shots)	
1. Languages of Google Translation	8
2. Google Translation Process 1	9
3. Languages of Microsoft Bing Translator	14
4. Microsoft Translator Hub Architecture	16
5. Languages of Babel fish system	18
6. Google Translation Process 2	31
7. Microsoft Bing Translation Process	33

प्रस्तावना

समस्त विश्व में अनुवाद का महत्त्व दिनों दिन बढ़ता ही जा रहा है। पहले अनुवाद मानव के द्वारा किया जाता था, लेकिन आज सहायता स्वरूप और समय की बचत करने के लिए मशीन (कम्प्यूटर) के द्वारा भी अनुवाद कार्य किया जा रहा है, जिसमें मशीन के माध्यम से एक भाषा से दूसरी भाषा में अनुवाद किया जाता है। मशीनी अनुवाद की यह प्रक्रिया अनुवाद की तरह विश्लेषण-अंतरण-प्रजनन पर आधारित है। मशीन अनुवाद 'अभिकलनात्मक भाषाविज्ञान' (Computational Linguistics) का महत्त्वपूर्ण क्षेत्र माना जाता है। अभिकलनात्मक भाषाविज्ञान ने प्राकृतिक भाषा (Natural Language) का वैज्ञानिक अध्ययन करके भाषा संबंधी अनुसंधानों को तथा अनुवाद के क्षेत्र में सुगमता प्रदान की है।

मशीन अनुवाद का कार्य केवल अनुवाद करना ही नहीं बल्कि भाषा के अनुवाद में आनेवाली समस्याओं पर विचार करना भी होता है। उन समस्याओं का निवारण करने के लिए तथा मशीनी अनुवाद को विकसित और संचालित करने के लिए देश-विदेश में बहुत प्रयास हो रहे हैं फिर भी मशीनी अनुवाद का प्रयोग कार्यालय, विज्ञान, व्यापार, वाणिज्य, खेल, मनोरंजन, आदि विभिन्न क्षेत्रों में बेहतर अनुवाद करने के लिए किया जा रहा है।

मशीनी अनुवाद के द्वारा किस भाषा से किस भाषा में अनुवाद किया जा रहा है, यह अनुवाद करने के लिए उस भाषा की संरचना क्या है, और अनुवाद के स्वरूप दो भाषाओं में भिन्नता कौन सी है- यह जानना अनुवाद के लिए महत्त्वपूर्ण है। इसके माध्यम से अनुवाद करने के लिए मशीन में कुछ व्याकरणिक नियम और भाषाई सामग्री का समावेश किया जा सकता है। यह शोध भी भाषाई सामग्री (कॉर्पोरा) पर आधारित है। यह कॉर्पोरा किसी भी भाषा की सामग्री के प्रायोगिक स्वरूप का संचयन तथा संकलन होता है, जिसका प्रयोग किसी भाषा के अध्ययन और विश्लेषण के लिए किया जाता है। प्रस्तुत शोध में कॉर्पोरा के हिंदी अंग्रेजी वाक्यों के आधार पर संज्ञापद का अध्ययन किया है। इस अध्ययन में दोनों भाषाओं के अनुवाद में पाई जाने वाली

भिन्नता का सम्मिलन किया गया है और मशीनी अनुवाद में सुधार हेतु सैद्धांतिक रूप से स्थानांतरण व्याकरण के नियमों का निर्माण किया गया है।

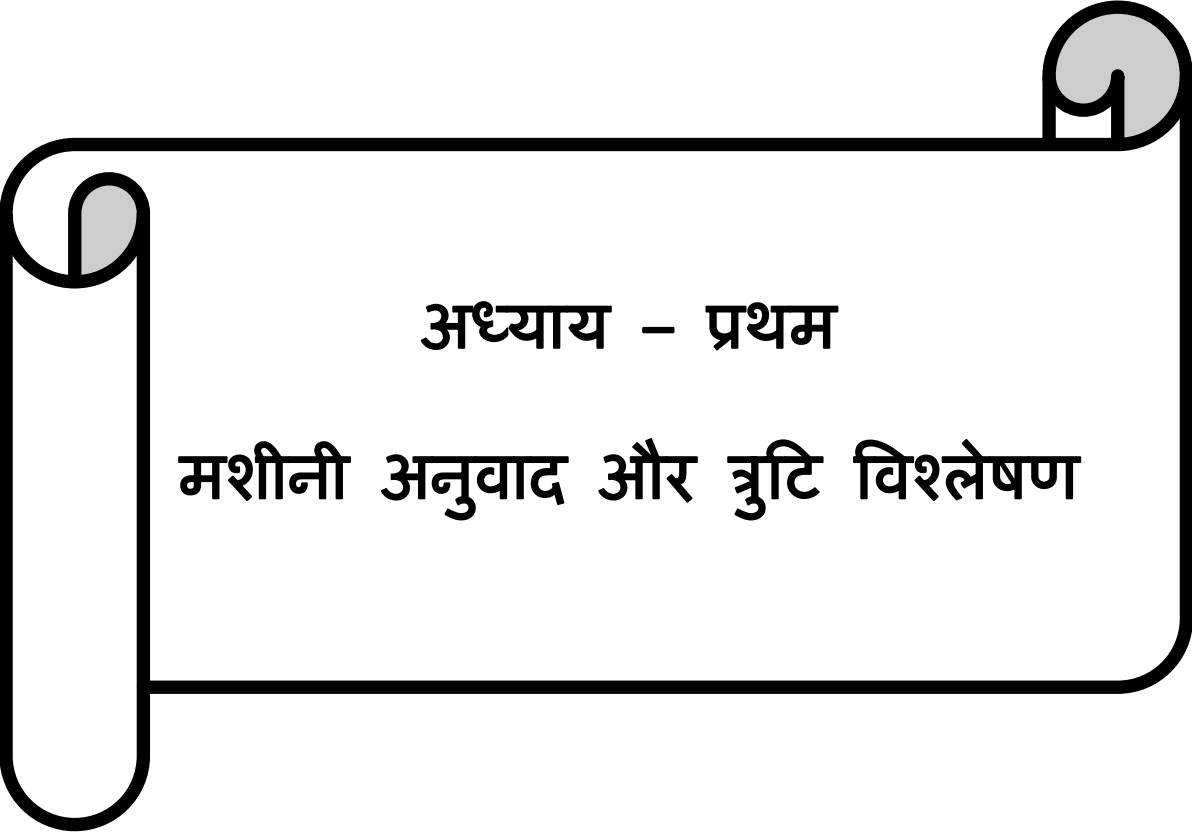
इस शोध के प्रथम अध्याय में मशीनी अनुवाद के परिचय के साथ विश्व (world) और भारत के मशीनी अनुवाद तंत्रों के विकास का अध्ययन किया गया है। ये अनुवाद तंत्र सीमित क्षेत्रों में सक्षम रूप से अनुवाद कार्य कर रहे हैं। साथ ही अन्य, क्षेत्रों में अनुवाद करने के लिए तंत्रों का विकास कई संस्थानों में किया जा रहा है। दूसरी ओर, इस अध्याय में, मशीनी अनुवाद द्वारा प्राप्त त्रुटियों का अध्ययन-विश्लेषण किया गया है, जिसमें विश्व स्तर के भारतीय भाषाओं के लिए अनुवाद करने वाले मशीनी अनुवाद तंत्रों (गूगल और बिंग अनुवादक) से प्राप्त अनुवाद का परीक्षण किया गया है और पाया जानेवाला परिणाम त्रुटि विश्लेषण में स्पष्ट किया गया है।

द्वितीय अध्याय में हिंदी और अंग्रेजी भाषाओं की संरचना का अध्ययन किया गया है। इन दोनों भाषाओं की संरचना में वाक्य के प्रमुख घटकों के साथ वाक्य के विभिन्न प्रकारों को स्पष्ट किया गया है। अध्याय में दूसरी ओर संज्ञापद के साथ प्रयुक्त होनेवाले घटक निर्धारक, पूर्व विशेषक और पश्च विशेषक की तुलना दोनों भाषाओं में की गई है और व्यतिरेकी विश्लेषण को स्पष्ट करने के लिए वाक्यों के उदाहरण के द्वारा संज्ञापद के साथ होनेवाले बदलाव को व्यतिरेकी विश्लेषण में रेखांकित कर उसे स्पष्ट किया गया है।

तृतीय अध्याय में, मशीनी अनुवाद में स्थानांतरण व्याकरण की भूमिका की जानकारी दी गई है। मशीनी अनुवाद में किसी भी भाषा का अनुवाद तीन प्रक्रियाओं से गुजरकर होता है। यह प्रक्रिया विश्लेषण (Analysis), अंतरण (Transfer), प्रजनन (Generation) पर आधारित है। इस प्रक्रिया के द्वारा अनुवाद के लिए दूसरा चरण महत्वपूर्ण है जिसमें दोनों भाषाओं के अनुवाद स्वरूप शब्दकोश और व्याकरणिक नियमों का समावेश मशीन में किया जाता है, जिसके द्वारा मशीनी अनुवाद में सुधार किया जा सकता है। इसलिए इन नियमों का निर्माण कर मशीन में संचालित करना महत्वपूर्ण है।

चतुर्थ अध्याय में हिंदी और अंग्रेजी दोनों भाषाओं के कॉर्पोरा के वाक्यों का अध्ययन किया गया है। इस अध्ययन में दोनों भाषाओं के अनुवाद हेतु संज्ञापद का सम्मिलन किया गया है। इस सम्मिलन में संज्ञापद के साथ प्रयोग किए जानेवाले घटकों का मशीनी अनुवाद की दृष्टि से सम्मिलन किया गया है और हिंदी से अंग्रेजी में भाषा के अनुवाद में सुधार हेतु मशीन के लिए स्थानांतरण व्याकरण नियमों का निर्माण किया गया है। यह नियम आई. आई. टी. हैदराबाद के मशीनी अनुवाद तंत्र पर आधारित है। ये नियम सैद्धांतिक रूप से निर्मित किए गए हैं, जिनका प्रायोगिक रूप में प्रयोग के लिए अल्गोरिदम बनाकर इसे मशीन में समावेशित तथा संचालित किया जा सकता है।

इन उपर्युक्त अध्यायों में दो भाषाओं के अनुवाद स्वरूप विभिन्न घटकों के अध्ययन विश्लेषण में संज्ञापद का विशेष अध्ययन किया गया है और मशीनी अनुवाद के संदर्भ में दोनों भाषाओं के अनुवाद में होने वाले बदलाव को सुधारने के लिए व्याकरणिक नियम बताएँ गए हैं, जिनका समावेश मशीन में करके मशीनी अनुवाद हेतु सुधार किया जा सकता है और इन्हीं नियमों के आधार पर आज विभिन्न भाषाओं के लिए मशीनी अनुवाद तंत्रों का विकास 'भारतीय भाषा के लिए प्रौद्योगिकी विकास' (TDIL) संस्थान, भारत में संचालित कर रहा है।



अध्याय – प्रथम

मशीनी अनुवाद और त्रुटि विश्लेषण

अध्याय - प्रथम
मशीनी अनुवाद और त्रुटि विश्लेषण
(Machine translation and Error Analysis)

1.1. मशीनी अनुवाद का परिचय (Introduction of Machine Translation)

भारत एक बहुभाषिक देश है। यहाँ विभिन्न भाषाओं का परस्पर संप्रेषण होता रहता है, इसलिए यहाँ भिन्न-भिन्न भाषाओं की कला, उनके साहित्य, संस्कृति को जानने के लिए पहले से अनुवाद होता रहा है लेकिन समय के साथ अनुवाद केवल मानव आधारित अनुवाद तक ही सीमित नहीं रहा, बल्कि आज उसका संबंध मशीन के साथ भी जुड़ गया है और 'मशीनी अनुवाद' अनुवाद का एक नया प्रकार बन गया है। यह भाषिक अध्ययन और विश्लेषण के फलस्वरूप भाषा का वैज्ञानिक अध्ययन प्रस्तुत करनेवाला शास्त्र भाषाविज्ञान के अनुप्रायोगिक क्षेत्र 'अभिकलनात्मक भाषाविज्ञान' (Computational Linguistics) में विकसित हुआ है। 'अभिकलनात्मक भाषाविज्ञान' का मुख्य उद्देश्य प्राकृतिक भाषा (Natural Language) को समझ कर, ऐसे व्यापक मॉडल (Model) और सिद्धांतों का विकास करना है, जिनकी सहायता से मानव और मशीन के बीच संवाद स्थापित हो सके। इसमें कृत्रिम बुद्धि (Artificial Intelligence) का समावेश किया गया है, इस कृत्रिम बुद्धि ने शब्दों और अर्थों का विश्लेषण करके मशीनी अनुवाद में भाषाई पक्ष को मजबूत किया है।

एक स्रोत भाषा (SL) की पाठ्यसामग्री को दूसरी लक्ष्य भाषा (TL) में प्रस्तुत करने की प्रक्रिया अनुवाद है। अनुवाद की इस प्रक्रिया में मशीन (कंप्यूटर) का उपयोग करना मशीनी अनुवाद है, जिससे स्रोत भाषा (SL) की पाठ्य सामग्री का अनुवाद लक्ष्य भाषा (TL) में स्वतः हो जाता है। एक भाषा की सामग्री का दूसरी भाषा में अनूदित करने का तात्पर्य केवल शब्दों का स्थानान्तरण नहीं है, बल्कि अनुवाद एक विचार प्रक्रिया है। जब अनुवादक किसी एक भाषा का दूसरी भाषा में अनुवाद करता है, तो वह दोनों भाषाओं के साहित्य, भाव, उनकी संस्कृति और संवेदना को मूल भाषा की तरह बनाए रखने का प्रयास करता है लेकिन मशीन के द्वारा यह अनुवाद त्रुटिग्रस्त प्राप्त होता है, क्योंकि, मानव अपने मस्तिष्क, बुद्धि, और उससे जुड़े हुए संदर्भ की जानकारी के आधार पर भाषा की अभिव्यक्तियों (Expressions) का सही निर्वचन (Interpretation) कर उनके सही अर्थों को ग्रहण करता है। उसके पास भाषा की समझ, और सांसारिक ज्ञान (World Knowledge) पहले से होता है, लेकिन कंप्यूटर केवल शब्द का वही अर्थ ग्रहण करता है, जो उसकी स्मृति (Memory) में संचित रहता है और उसे प्रोग्रामिंग के द्वारा

संचालित किया जाता है। इस मशीनी अनुवाद को समझने के लिए इसकी परिभाषा को जानना महत्वपूर्ण है।

सूरजभान सिंह (2003) के अनुसार “मशीनी अनुवाद अनुवाद की ऐसी प्रक्रिया है, जिसमें कंप्यूटर प्रणाली (System) के जरिए एक भाषा से दूसरी भाषा में अपने आप अनुवाद की जाने वाली सामग्री (Text) को आगत शब्द (Input Word) के रूप में देते हैं। कंप्यूटर की भीतरी प्रणाली (System), जिसमें दोनों भाषाओं के शब्दों, मुहावरों और व्याकरणिक नियमों का ज्ञान संचित रहता है, जो अपने आप उस सामग्री का दूसरी भाषा में अनुवाद करता है और कुछ ही क्षणों में निर्गत पाठ (Output) के रूप में अनूदित सामग्री प्राप्त हो जाती है।”¹

विशाल गोयल (2010) के अनुसार, स्वतःअनुवाद (अर्थात् मशीनी अनुवाद) का लक्ष्य, एक मानव भाषा के पाठ का दूसरी भाषा में कंप्यूटर के द्वारा अनुवाद करना है। “*The goal of automatic translation (also called Machine Translation, or MT) is to translate text from one human language into other using computers.*”²

इन परिभाषाओं से यह स्पष्ट होता है कि एक प्राकृतिक भाषा का कंप्यूटर के द्वारा दूसरी भाषाओं में अनुवाद करना ही मशीनी अनुवाद है।

भारत में मशीनी अनुवाद पर किए जा रहे प्रयासों से लगता है कि यह कार्य कठिन होने पर भी बहुत से क्षेत्रों में सुचारु रूप से सफलता हासिल की जा रही है। वास्तव में म.अनु. का उद्देश्य शीघ्र गति से अनुवाद प्राप्त कर समय और श्रम की बचत करना है और उसका प्रयोग सामान्य प्रयोक्ता के लिए भी साधक बनाना है। लेकिन भाषा वैज्ञानिकों के लिए आज भी यह चुनौती है कि वह मानव सहायता रहित म.अनु. का तंत्र को विकसित करे। इसलिए म.अनु. के लिए प्राकृतिक भाषा (NL) को विकसित कर कृत्रिम बुद्धि (Artificial Inteligence) का समावेश किया गया है। इस कृत्रिम बुद्धि के निर्माण द्वारा मशीन को मानव बुद्धि की तरह ही सृजनशील बनाने का प्रयास किया जा रहा है, जिससे स्वचालित मशीनी अनुवाद का उद्देश्य सफल हो सके। इस दिशा में, भारत में किए गए कार्यों में आई.आई.टी. हैदराबाद द्वारा निर्मित अनुसारक तथा सी-डैक पुणे द्वारा निर्मित मंत्र-राजभाषा अंग्रेजी-हिंदी मशीनी अनुवाद तंत्र (Hindi-English Machine Translation system) काफ़ी हद तक कारगर साबित हुआ है। हमें मानना पड़ेगा कि कंप्यूटर से मशीनी अनुवाद 100 प्रतिशत सही नहीं हो रहा है लेकिन 80 से 90 प्रतिशत का अनुवाद सही मात्रा में पाया जा सकता है। लेकिन आज इसके लिए मानव की सहायता बहुत ही महत्वपूर्ण है।

¹ सिंह सूरजभान (2003) : अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ.252

² Goyal, Vishal (2010) : *Development of a Hindi to Punjabi Machine Translation System* (Ph.D Thesis), p.18

1.1.1. मशीनी अनुवाद के प्रकार: (मानव सहभागिता के आधार पर)

1. विशुद्ध मानव अनुवाद (Human Translation)
2. मशीन साधित मानव अनुवाद (Machine Aided Human Translation)
3. मानव साधित मशीन अनुवाद (Human Aided Machine Translation)
4. मशीनी अनुवाद (Machine Translation)³

विशुद्ध मानव अनुवाद की प्रक्रिया में मानव स्वतः अनुवाद करता है। वह किसी भी विषय क्षेत्र के पाठ का अनुवाद करता है। साहित्यिक अनुवाद के पाठ में संस्कृति, भाव, संवेदना, लक्षणा एवं व्यंजना रहती है, जिसे मानव अपने सहज ग्रहण की हुई मस्तिष्क क्षमता के आधार पर यह अनुवाद प्रदान करता है, जिसमें मशीन की कोई सहायता नहीं ली जाती ।

मशीन साधित मानव अनुवाद की प्रक्रिया में कंप्यूटर के स्मृति कोश (Memory) में संग्रहीत उपयुक्त शब्दावली को उपलब्ध कराके अनुवादक की सहायता करता है। यह तकनीकी और वैज्ञानिक अनुवाद के लिए ज्यादा उपयुक्त है, क्योंकि इस अनुवाद में शब्द प्रति शब्द अनुवाद किया जाता है।

मानव साधित मशीनी अनुवाद में अनुवादक ही पूर्व संपादन करता है और वह पाठ कंप्यूटर को देता है। इसमें अनुवादक आवश्यकता पड़ने पर वाक्यों की संदिग्धता को स्पष्ट करके वाक्य को सरल बनाता है और अनुवाद पूर्ण होने के बाद संपादन करता है ।

यह म. अनु के लिए मशीनी अनुवाद तंत्रों का विकास होना आवश्यक है, जिसमें प्रमुखतः स्रोत और लक्ष्य भाषा की व्याकरणिक संरचना, पूर्व संसाधक, पद निरूपक (Parser) और शब्द विश्लेषक (Word analyzer), शब्द संचयन, आदि का होना आवश्यक है। म.अनु. के प्रकारों से यह स्पष्ट होता है कि अभी तक के मशीन अनुवाद में मानव की सहायता महत्वपूर्ण रही है लेकिन अभी मशीन के माध्यम से पूर्ण अनुवाद प्राप्त करने का प्रयास देश और विदेश में जोरों पर चल रहा है।

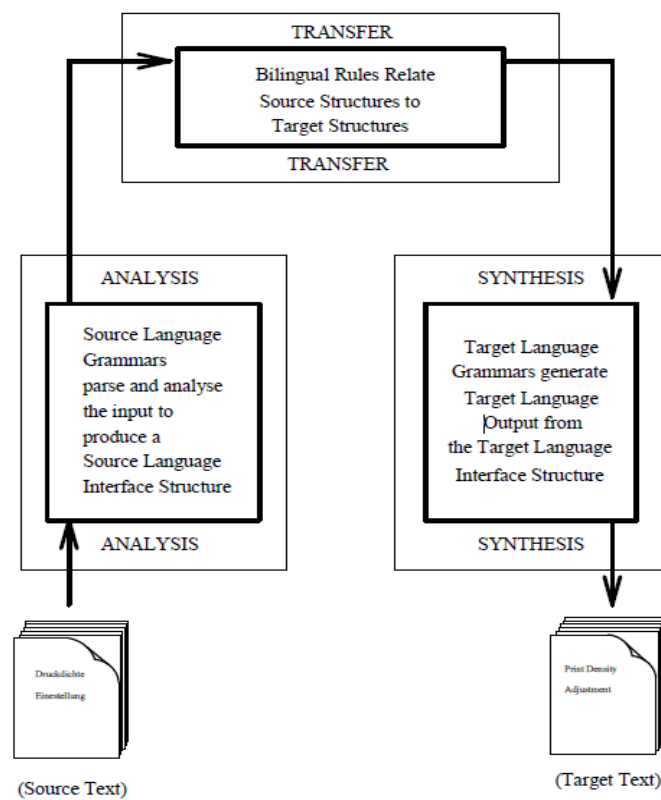
1.1.2. मशीनी अनुवाद की प्रक्रिया

अनुवाद एक भाषा से दूसरी भाषा में करना केवल शब्दों का स्थानांतरण नहीं है। यह एक बौद्धिक प्रक्रिया है। अनुवाद की प्रक्रिया में जिस तरह विश्लेषण, अंतरण और पुनर्गठन तीन चरण होते हैं, उसी तरह मशीनी अनुवाद में भी भाषा संबंधी विश्लेषण, अंतरण और संश्लेषण के पक्ष

³ Karunkaran, k. and Jaykumar (ed.) (1988) : Translation As Synthesis “Human Aided Machine Translation :Problems and Perspective” Vasu R. (Article), p. 183

समाहित होते हैं। “मशीनी अनुवाद में प्रथम कार्य को पद-अन्वय (Parsing) कहा जाता है, जिसमें पाठ के अंतर्गत वाक्य के घटकों (Components) के मध्य कार्यरत प्रकार्यात्मक संबंधों की पहचान कर वाक्य को पद विच्छेदित किया जाता है और अंतरण (Transfer) के चरण में यह कार्य ‘लेक्सिकन ट्रान्सफर’ (LT) के द्वारा किया जाता है। यह स्रोत भाषा पाठ के लिए आगत (Input) शब्दों एवं अभिव्यक्तियों के लिए समानार्थी शब्द प्रदान करता है।”⁴ अनुवाद में पुनर्गठन की तरह जेनरेटर लक्ष्य भाषा के अनुवाद को लिपि, शब्द, वाक्य, शैली तथा रूप को प्रस्तुत करता है अर्थात् पार्सर से आए शब्दों द्वारा लक्ष्य भाषा की संरचनानुसार वाक्य गठन का कार्य प्रजनक (Generator) करता है। इस म.अनु. की प्रक्रिया को हम इस आरेख के द्वारा समझ सकते हैं -

MACHINE TRANSLATION ENGINES (PROCESS)



आकृति 1 : म.अनु. अंतरण प्रक्रिया के घटक⁵

मशीनी अनुवाद प्रक्रिया के लिए आवश्यक संसाधन के घटक पूर्ण रूप में विकसित न होने की वजह से भारत में म.अनु. तंत्र (MT System) से प्राप्त अनुवाद तथा तंत्र निर्माण में विभिन्न समस्याएँ पायी जाती हैं। वे समस्याएँ निम्न हैं - 1. भाषिक संरचना, 2. वाक्यगत भाषिक अस्पष्टता, 3. संदर्भ परक अस्पष्टता, 4. मुहावरों एवं कहावतों के अनुवाद की समस्या, 5.

⁴ रस्तोगी, कविता (2010) : समसामयिक अनुप्रयुक्त भाषाविज्ञान, पृ. 101

⁵ Arnold, Balkan, Humphreys, Meijer, Sadler (1994) : “Machine Translation: an Introductory Guide”, p.68

शब्दगत संबंध की समस्या, 6. विकसित नए शब्दों के अनुवाद की समस्या आदि। इन समस्याओं का मूल कारण है भाषिक संसाधनों का पूर्ण विकास न होना। इसलिए म. अनु. तंत्र को उन्नत या विकसित करने की आवश्यकता है। मशीनी अनुवाद तंत्र के विकास की दिशा में विश्व स्तर पर और भारत में चल रहे प्रयासों को इस अध्याय में नीचे निर्दिष्ट किया गया है।

1.2 विश्व स्तर पर मशीनी अनुवाद सिस्टम का विकास

(Development of Machine Translation System in World)

आज विश्व में मशीनी अनुवाद पर काफी कार्य हो रहे हैं। इसका महत्वपूर्ण उद्देश्य है कि, एक भाषा के माध्यम से दूसरी भाषा में मानव के लिए सहज अनुवाद प्रदान करना। अगर यह अनुवाद बेहतर होता है, तो इसका उपयोग विभिन्न कार्यों के लिए कर सकते हैं। फिर वह क्षेत्र, उद्योग व्यापार के साथ सूचना प्रौद्योगिकी का हो, या विज्ञान एवं तकनीकी के साथ प्रशासन का हो या और भी अनुवाद के क्षेत्र हो सकते हैं। इस सह-अध्याय में विश्व के म.अनु. तंत्रों का अध्ययन किया जाएगा।

1.2.1. विश्व में मशीन अनुवाद का संक्षिप्त इतिहास

विश्व में मशीन अनुवाद का संक्षिप्त इतिहास अगर देखा जाए तो यह ज्ञात होता है, कि म.अनु. की दिशा में प्रयास 1946 के बाद वारेन विवर के भाषा विश्लेषण और अनुवाद संबंधी विचारों से होता है। लेकिन यह प्रयास स्रोत भाषा से लक्ष्य भाषा के शब्द प्रतिस्थापन पर आधारित था। म.अनु. तंत्र के निर्माण का पहला प्रयास 1956 में जार्ज टाउन विश्वविद्यालय में किया गया था, जिस पर 1964 में अमरीकी प्रशासन द्वारा गठित एल्पेक (ऑटोमेटिक लैंग्वेज प्रोसेसिंग एडवाइज़री कमेटी) ने 1966 में प्रस्तुत रिपोर्ट में अनुवाद तंत्र को विकसित करने के लिए मूलभूत शोध करने के लिए कहा। इस रिपोर्ट की मान्यताओं को अस्वीकार कर अमरीकी सरकार ने अनुवाद संबंधी शोध के लिए वित्तीय सहायता बंद कर दी। इससे अमेरीका में अनुवाद संबंधी शोध कार्य काफी समय तक बंद रहा, लेकिन यूरोप के अन्य देशों में तथा सोवियत संघ में अनुवाद कार्य जारी रहा। उसके बाद जापान में भी इस दिशा में प्रयास शुरू किए गए। तो कुछ निजी कंपनियों ने भी भारी मात्रा में वित्तीय सहायता प्राप्त कर म. अनु. के लिए प्रयास शुरू किया।⁶

⁶ शर्मा, प्रमोद (1999) : "मशीनी अनुवाद: स्थिति, सीमाएँ और संभावनाएँ", अनुवाद शतक विशेषांक (100-101) संपा. - नीता गुप्ता, (जुलाई-दिसंबर), पृ. 229

1970 के बाद से 1980 तक विश्वभर में म.अनु में रुचि बढ़ी, और मशीनी अनुवाद तथा मशीन साधित अनुवाद प्रणालियों का प्रयोग सरकारी प्रशासन, व्यापार और उद्योग में किया जाने लगा। उसके बाद से 60 के दशक के पूर्वार्ध तक मशीनी अनुवाद संबंधी प्रयासों का तीन चरणों में अनुवाद की विधियों के रूप में प्रारंभ हुआ। उसमें प्रत्यक्ष विधि का प्रयोग स्रोत भाषा की सामग्री को विश्लेषित करने के पश्चात् उसे स्रोत भाषा के शब्दों के स्थान पर लक्ष्य भाषा के शब्दों में प्रतिस्थापित कर लक्ष्य भाषा के वाक्य विन्यास के अनुरूप संयोजित कर दिया जाता था। इस विधि के फलस्वरूप 1964 में जॉर्ज टाउन विश्वविद्यालय में रूसी से अंग्रेजी अनुवाद तंत्र GAT विकसित किया गया और 1970 में इसे SYSTRAN के रूप में परिवर्तित एवं परिमार्जित किया गया। इसकी सहायता से NASA ने अपोलो सोयूज सहयोग से संबंधित सामग्री का अनुवाद किया। यह तंत्र विश्व का प्रथम व्यावसायिक म.अनु. तंत्र था। इस विधि के आधार पर हांगकांग के चीनी विश्वविद्यालय ने CULT (1968) नामक म.अनु. तंत्र का प्रयोग कर गणित और भौतिकी विषयों से संबंधित शोध पत्रों का अंग्रेजी में अनुवाद किया था। अन्य सफल प्रणालियों में TITUS प्रणाली तथा LOGOS आदि उल्लेखनीय हैं।⁷

कनाडा में अंग्रेजी के साथ फ्रेंच के राजभाषा बनने के बाद मॉंट्रियल विश्वविद्यालय में मौसम संबंधी जानकारी का अंग्रेजी से फ्रेंच में अनुवाद करने के लिए एक पूर्णस्वचालित मशीनी अनुवाद तंत्र TAUM METEO को विकसित किया गया। यह अनुवाद तंत्र आज भी कार्यरत है। अमरिका में टेक्सस यूनिवर्सिटी में व्यावहारिक म.अनु. प्रणाली के विकास के लिए METAL प्रोजेक्ट चल रहा है, जिसे जर्मनी की सीमेन्स कंपनी वित्तीय अनुदान प्रदान करती है। 1984 में विकसित SUSY प्रणाली का विकास रूसी जर्मन प्रोटोटाइप सिस्टम द्वारा अंतरण विधि के अंतर्गत किया गया, तो ग्रेनेल विश्वविद्यालय, फ्रांस ने इस विधि के अंतर्गत GETA प्रणाली का सफलतापूर्वक विकास किया। यूरोप में यूरोपियन आर्थिक समुदाय (European Economic Community) में व्यवहार हेतु नौ भाषाएँ स्वीकृत हैं। 1986 में यूरोपीय संसद ने म. अनु. के विकास पर काफ़ी धन खर्च किया था। इसके परिणामस्वरूप आज म. अनु. पर यूरोप में कई परियोजनाओं पर काम चल रहा है। EUROTRA और GETA वहाँ सरकारी अनुदान से पोषित प्रमुख परियोजनाएँ हैं। म.अनु. में जापान ने भी विकसित टेक्नोलॉजी का प्रयोग कर 1980 में अनुवाद की दिशा में जापानी से अंग्रेजी में अनुवाद के लिए परियोजना शुरू की है, और 1988 से CICC (Centre for International Cooperation for Computerization) ने बहुभाषिक म. अनु. प्रणाली के विकास पर काम शुरू कर इसमें अंग्रेजी, जापानी के अतिरिक्त थाई, मलय, इंडोनेशियाई, चीनी

⁷ शर्मा, राजमणि (2004) : अनुवाद विज्ञान, पृ. 184

भाषाओं का मशीनी अनुवाद के लिए शामिल किया गया है। इसके साथ ही ATLAS-I, ATLAS-II, TAURUS, MU, PIVOT ये विश्व के कुछ प्रमुख व्यावहारिक म.अनु. तंत्र हैं।⁸

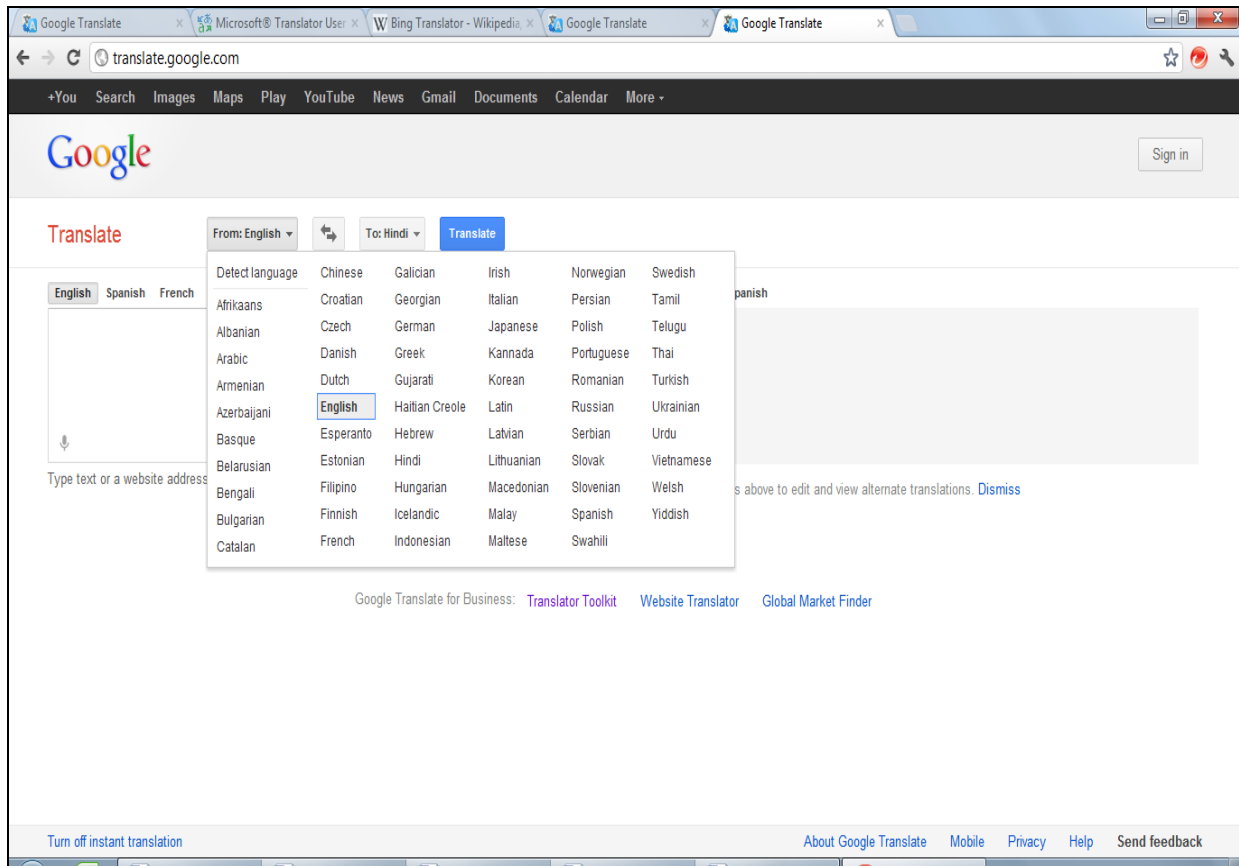
उक्त म. अनु. तंत्रों की तरह विश्व में अनेक परियोजनाएँ, इन तंत्रों को विकसित करने में लगी है। उसमें से कुछ महत्वपूर्ण सिस्टम आज भी काम कर रहे हैं। विश्व में आज कुछ मशीनी अनुवाद के तंत्र सार्वभौमिक रूप में प्रयोग किए जा रहे हैं वे हैं - गूगल अनुवाद (Google Translate) माइक्रोसाफ्ट ट्रान्सलेटर (Microsoft Translator) और बेबेल फिश (Babel Fish) जिनका इस शोध में अनुवाद और त्रुटि विश्लेषण के लिए प्रयोग किया गया है।

1.2.2 गूगल अनुवाद⁹ (Google Translate 2007)

गूगल विश्व की 65 भाषाओं में अनुवाद प्रदान करता है, उसमें भारत की प्रमुख 7 भाषाएँ और 1 सह-राजभाषा (अंग्रेजी) शामिल है। गूगल अनुवाद (Google Translate) सांख्यिकीय मशीनी अनुवाद (Statistical Machine Translation Based) पर आधारित अनुवादक साफ्टवेयर (software) है। यह गूगल इन्कॉर्पोरेशन (Google Incorporation) द्वारा विकसित एवं परिचालित है। यह तंत्र एक भाषा के पाठ (Text) या वेबपेज (webpage) का दूसरी भाषा में अनुवाद करता है। गूगल ने यह अनुवाद सेवा अरबी भाषा के लिए 28 अप्रैल, 2006 को शुरू की थी, फिर अक्टूबर 2007 में अरबी से अन्य भाषाएँ चीनी और रूसी के लिए भी शुरुआत की गई। गूगल में सिस्ट्रान (SYSTRAN) आधारित अनुवादक तंत्र का इस्तेमाल किया गया है। (सिस्ट्रान यह मशीनी अनुवाद की पुरानी कंपनी है जिसे डॉ. पीटर टोमा (Peter Toma) ने 1968 में स्थापित किया गया था, सिस्ट्रान ने संयुक्त राष्ट्र अमेरिका के रक्षा विभाग और यूरोपीय आयोग के लिए व्यापक अनुवाद कार्य किया है।) इसके साथ ही याहू बेबेल फिश (Yahoo Babel Fish), अमेरिकन आनलाइन (AOL) अन्य अनुवाद-तंत्र सेवाएँ भी आज सिस्ट्रान का प्रयोग कर रही हैं। आरंभ में गूगल ने अंग्रेजी से अरबी में अनुवाद की शुरुआत की थी और आज विश्व की अनेक भाषाओं में अनुवाद प्रदान कर रहा है। गूगल विश्व की जितनी भाषाओं का अनुवाद प्रदान कर रहा है, वे फलक-चित्र (screen-shot) में दिया गया है।

⁸ ओम विकास (1993) : "मशीनी अनुवाद की समस्याएँ", (संपा. - डॉ. नगेन्द्र), अनुवाद विज्ञान सिध्दांत और अनुप्रयोग पृ.341-342

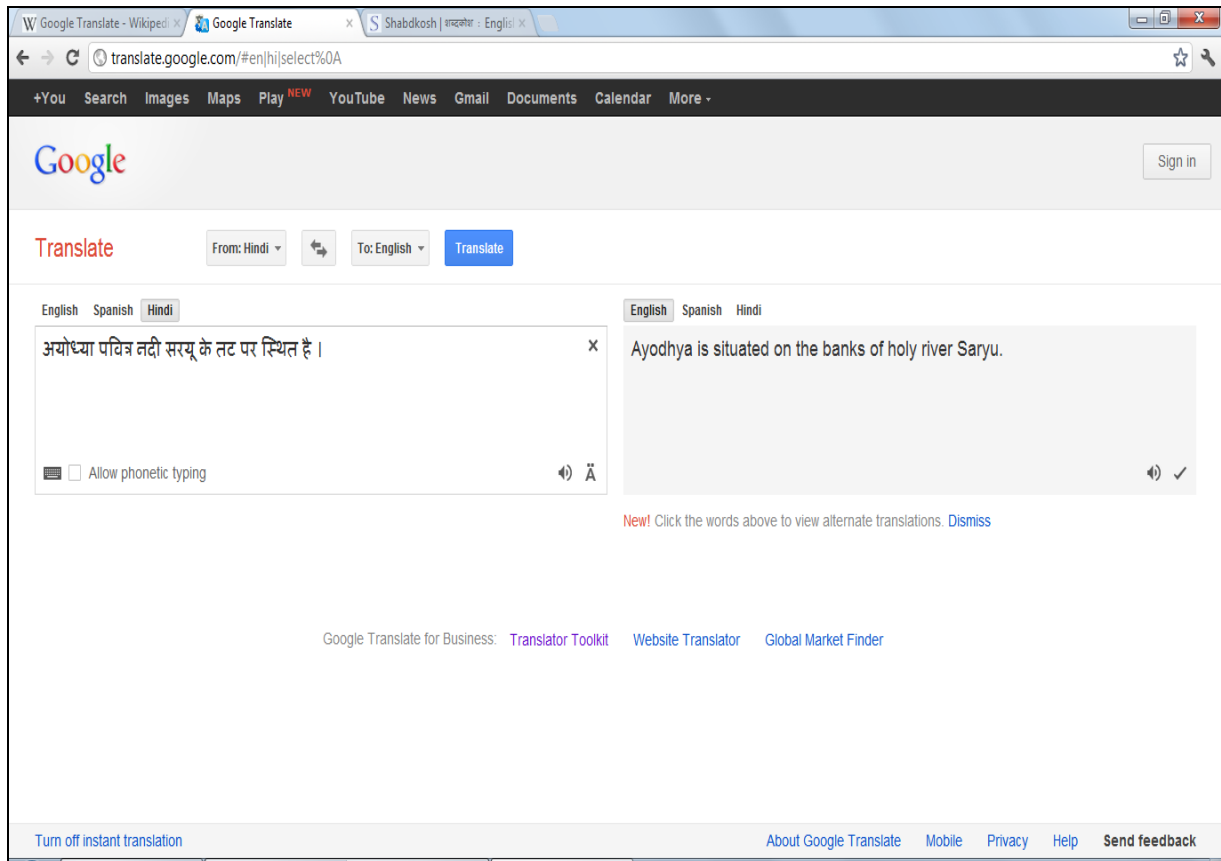
⁹ http://en.wikipedia.org/wiki/Google_Translate retrieved on 06.04.2012



Screen Shot 1: Showing Languages of Google Translation, Source taken from <http://translate.google.com/>

1.2.2.1 गूगल अनुवाद प्रक्रिया (Translation Process)

गूगल अनुवाद की प्रक्रिया, बेहतर अनुवाद प्रदान करती है, फिर भी इसके अनुवाद की कुछ सीमाएँ हैं। अगर अनुवाद के लिए सीमित पैराग्राफ की संख्या और तकनीकी शब्द (Technical Terms) दिए जाएंगे तो उसका अनुवाद गूगल द्वारा अच्छा होता है लेकिन हम पूर्णतः सही अनुवाद प्राप्त नहीं कर सकते। कुछ भाषाओं का अनुवाद गूगल द्वारा सही होता है। यह तंत्र 2010 के बाद से फ्रांसीसी से अंग्रेजी तथा चीनी से अंग्रेजी भाषा में बेहतर अनुवाद प्रदान कर रहा है। गूगल अनुवाद सात भारतीय भाषाओं (हिंदी, बांग्ला, गुजराती, कन्नड, तमिल, तेलुगु, उर्दू) और अंग्रेजी के साथ एक दूसरी भाषा में युग्म अनुवाद (Paired Translation) की सुविधा प्रदान कर रहा है। हमें गूगल में जिस भाषा से अनुवाद करना है, उस भाषा का चयन करना होता है और जिस भाषा में अनुवाद प्राप्त करना है, उस भाषा का भी चयन करना होता है। उदाहरण के लिए हिंदी भाषा के किसी वाक्य का अनुवाद अंग्रेजी में करना है तो यह निम्न फलक-चित्र (screen-shot) के द्वारा देख सकते हैं।



Screen Shot 2: Google Translation Process, taken from <http://translate.google.com/>

इसमें लिप्यंतरण की सुविधा भी है। हिंदी वाक्य को हिंदी वाले बक्से में रोमन में टाइप कर उसे देवनागरी में आप प्राप्त कर सकते हैं और इसमें यह सुविधा भी है कि किसी भी वाक्य (Sentence) या पाठ (Text) के अनुवाद को ध्वनि (Speech) में बदलकर सुन सकते हैं।

1.2.2.2 गूगल के संस्करण

1.2.2.2.1 Android संस्करण (Android Version)

गूगल अनुवाद यह फ्री डाउनलोड के रूप में Android ओएस (Operating System for Mobile devices) संस्करण उपयोगकर्ताओं के लिए उपलब्ध है। इसका प्रथम संस्करण जनवरी 2010 में विकसित किया गया था। यह बहुत आसान तरीके से अनुवाद का कार्य करता है। 2011 से इसे वार्तालाप माध्यम (Conversation Mode) जोड़ दिया गया जो अंग्रेजी से स्पेनिश अनुवाद करता है। गूगल ने इसे और परिष्कृत कर अक्टूबर में 14 भाषाओं में अनुवाद के लिए जारी किया है। अब यह तंत्र 53 भाषाओं में अनुवाद के साथ 15 भाषा के लिए स्पीच इनपुट प्रणाली (Speech System) में अनुवाद करता है। यह संस्करण जनवरी 2011 से गूगल ने अद्यतन कर उपलब्ध कराया है।

1.2.2.2.2 IOS संस्करण (Version)

अगस्त, 2008 में गूगल ने गूगल ट्रान्सलेट में HTML5 वेब का अनुप्रयोग कर IOS के iPhone और iPad का उपयोगकर्ताओं के लिए शुरुआत की है। सहकारी रूप से गूगल का अनुप्रयोग IOS के लिए 8 फ़रवरी 2011 को शुरू किया गया था। वह 15 भाषाओं में ध्वनि इनपुट (Input) स्वीकार करता है और शब्द या वाक्य का अनुवाद एक से 50 भाषाओं में करता है, साथ ही यह अनुवाद वाक (Speech) के रूप में 23 भाषाओं में प्राप्त कर सकते हैं।

आरम्भ में गूगल ने अंग्रेजी से अरबी में अनुवाद की शुरुआत की थी। उसके बाद गूगल के संस्करण में बदलाव और भाषाओं में वृद्धि की गई। इसका विवरण निम्नलिखित है-

1. चरण	अंग्रेजी से जर्मन, फ्रेंच से अंग्रेजी, अंग्रेजी से स्पेनिश, जर्मन से अंग्रेजी, स्पेनिश से अंग्रेजी
2. चरण	अंग्रेजी से पोर्तुगाली, पोर्तुगाली से अंग्रेजी, अंग्रेजी से डच, डच से अंग्रेजी
3. चरण	अंग्रेजी से इतालवी, इतालवी से अंग्रेजी
4. चरण	अंग्रेजी से चीनी(सरलीकृत), अंग्रेजी से जापानी, अंग्रेजी से कोरियाई, चीनी(सरलीकृत) से अंग्रेजी, जापानी से अंग्रेजी, कोरियाई से अंग्रेजी
5. चरण (अप्रैल 2006 में शुरुआत)	अंग्रेजी से अरबी, अरबी से अंग्रेजी
6. चरण (दिसंबर 2006 में शुरुआत)	अंग्रेजी से रूसी, रूसी से अंग्रेजी
7. चरण (फ़रवरी, 2007 में शुरुआत)	अंग्रेजी से चीनी (पारंपारिक), चीनी (पारंपारिक) से अंग्रेजी, चीनी (सरलीकृत से पारंपरिक), चीनी (पारंपरिक से सरलीकृत)
8. वाँ चरण (फरवरी 2007)	उपरोक्त सभी 25 भाषा युग्म मशीन अनुवाद प्रणाली में कार्यान्वित की गई।
9. वाँ चरण	अंग्रेजी से हिंदी, हिंदी से अंग्रेजी
10.वाँ चरण (मई-2008)	बल्गेरियाई, क्रोएशियाई, चेक, डेनिश, फ़ि, ग्रीक, नार्वेजियन, पोलिश, रोमानियाई, स्वीडिश. इस चरण में अंग्रेजी भाषा के माध्यम से किसी भी दो भाषाओं में अनुवाद करने के लिए प्रारंभ किया।

11. वाँ चरण (25 सितंबर, 2008)	कैटलन, फिलिपिनो, हिब्रू, इंडोनेशियन, लातवियाई, लिथुअनियन, सर्बियन, स्लोवाक, कैरिथियावासी स्लाव जाति की भाषा, यूक्रेनी, वियतनामी
12. वाँ चरण	अल्बेनियन, एस्तोनियन, गैलिशियन्, हंगेरियन, मालतिश, थाई, तुर्किश
13. वाँ चरण (19 जून, 2009)	पर्शियन
14. वाँ चरण	आफ्रिकन, बेलारूसियन, आइसलैंडिक, आयरिश, मेसीडोनियन, मलय, स्वाहिली, वेल्श, यिद्दीश
15. वाँ चरण (19 नवंबर, 2009 को शुरू)	बीटा प्रणाली समाप्त होने के बाद, चीनी, जापानी, कोरियाई, रूसी, यूक्रेनी, बेलारूसी, बल्गेरियाई, ग्रीक, हिंदी और थाई के लिए रोमन में लिखना शुरू हुआ। अंग्रेजी, इतालवी, फ्रेंच और जर्मन भाषा में पाठ को वाक में प्रोग्राम द्वारा पढ़ा जा सकता है।
16. वाँ चरण (30 जनवरी, 2010)	हार्डटियन क्रियोल
17. वाँ चरण (अप्रैल 2010)	वाक प्रणाली प्रोग्राम (Speech Programme) की शुरुआत हिंदी और स्पेनिश भाषा के लिए हुई।
18 वाँ. चरण (मई, 2010)	वाक प्रणाली प्रोग्राम (Speech Programme) की शुरुआत, अल्बेनियन, कातालान, चीनी (Mandarin), क्रोएशियाई, चेक, डेनिश, डच, फिनिश, ग्रीक, हंगेरियन, आइसलैंडिक, इन्डोनेशियाई, लातीवियाई, मैसेडोनियाई, नॉर्वेजियाई, पोलिश, पुर्तगाली, रोमानियाई, रूसी, सर्बियाई, स्लोवाक, स्वाहिली, स्वीडीश, तुर्किश, वियतनामी और वेल्श आदि. भाषा में हुई।
19. वाँ चरण (13 मई 2010)	आर्मीनियाई, आज़रबाइजानी, बास्क्यु, जार्जियन
20. वाँ चरण (जून, 2010)	अरबी के लिए रोमनीकरण की सुविधा प्रदान की।
21. वाँ चरण (सितम्बर 2010)	अरबी, ग्रीक, हिंदी, फ़ारसी, रूसी, सर्बियाई, और उर्दू लैटिन के लिए ध्वन्यात्मक टाइपिंग की सुविधा प्रदान की।
22.वाँ चरण (दिसम्बर 2010)	अरबी का रोमनीकरण निकालकर वर्तनी जाँचक समावेशित किया और पाठ से वाक में विश्लेषक प्रणाली का समावेश किया।
23. वाँ चरण	एक शब्द के लिए वैकल्पिक अनुवाद प्रस्तुत किया।

(जनवरी,2011)	
24. वाँ चरण (जून, 2011)	बांग्ला, गुजराती, कन्नड, तमिल, तेलुगु - इन 5 नई भारतीय भाषाओं में अल्फा और एक में लिप्यंतरण (Transliteration) की आगत विधि (Input Method) का प्रारंभ किया।
25. वाँ चरण (जुलाई, 2011)	अनुवाद का मूल्यांकन प्रारंभ
26. वाँ चरण(जनवरी 2012)	डच भाषा के पुरुष वाक संश्लेषण (voice synthesizer) को बदलकर महिला की आवाज में रूपांतरण के लिए सुविधा की तथा यीदिश भाषा में लिप्यंतरण का समावेश किया।
27. वाँ चरण (फ़रवरी 2012)	थाई भाषा में वाक प्रोग्राम (Speech Programme) का प्रारंभ किया और गूगल तंत्र ने अनुवाद के लिए एस्पेरन्तो भाषा का समावेश किया।

टेबल 1: गूगल अनुवाद के भाषाओं के संस्करण

गूगल अनुवाद व्याकरणिक नियमों के अनुसार काम नहीं करता। यह पारंपरिक नियम आधारित विश्लेषण के बजाय सांख्यिकीय विश्लेषण के अल्गोरिदम पर आधारित है। सिस्टम के निर्माता फ्रांज जोसेफ ओक (Franz Josef Och) ने पारंपरिक नियम आधारित विश्लेषण की आलोचना कर सांख्यिकीय आधारित मशीनी अनुवाद (Statistical Based Machine Translation) को अधिक महत्त्व दिया।

गूगल एक भाषा से दूसरी भाषा (L1 - L2) में सीधा अनुवाद नहीं करता, वह अक्सर स्रोत भाषा (SL) से पहले अंग्रेजी में अनुवाद करता है, फिर उसे लक्ष्य भाषा (TL) में अनूदित करता है (L1 – EN –L2)। मानव की सभी भाषाएँ अस्पष्ट हैं और संदर्भ के अनुसार अर्थ पर निर्भर करती हैं, इससे अनुवाद में त्रुटियाँ होती हैं। इसलिए गूगल द्वारा किया जानेवाला यह अनुवाद स्रोत भाषा से प्रत्यक्ष अनुवाद न होकर, वह अंग्रेजी भाषा में प्रत्यक्ष अनुवाद होता है। गूगल अनुवादक में विभिन्न भाषा के युग्म के साथ दस लाख शब्दों का द्विभाषिक पाठ कॉर्पोरा (Bilingual Text Corpora) और समानांतर कॉर्पोरा का संकलन (Parallel Collection) है। अन्य दो एक-भाषी कॉर्पोरा एक अरब से अधिक शब्दों से बने हैं। सांख्यिकीय मॉडल (Stastical Model) से इस आधार सामग्री (Data) का अनुवाद के लिए उपयोग किया गया है। यह भाषाई सामग्री संयुक्त राष्ट्र के

कार्यालयों के दस्तावेजों से पाई गयी है, जिनकी छह कार्यालयी भाषाएँ हैं और छह भाषाओं में बृहत कॉर्पोरा प्रकाशित है। गूगल ने जापान से भी द्विभाषिक आधार सामग्री प्राप्त की है।¹⁰

निष्कर्षतः गूगल के अनुवाद में कुछ गलतियाँ और विषमताएँ हैं क्योंकि, गूगल अनुवाद शब्दकोश या व्याकरणिक नियमों के बजाय सांख्यिकीय मॉडल पर आधारित है, लेकिन इसका अनुवाद सामान्य तौर पर अन्य अनुवाद तंत्रों से बेहतर होता है।

1.2.3 माइक्रोसॉफ्ट अनुवादक¹¹ (Microsoft Bing Translator 2009)

स्वचालित अनुवाद (Automatic Translation) के लिए माइक्रोसॉफ्ट का बिंग (Bing) अनुवादक विश्व के म. अनु. तंत्रों में गूगल तंत्र की तरह मशीनी अनुवाद का एक तंत्र है। ये विश्व की 38 भाषाओं में अनुवाद करता है। भारत की हिंदी भाषा का भी समावेश इस तंत्र में किया गया है। माइक्रोसॉफ्ट बिंग के द्वारा एक भाषा के पाठ (Text) को या वेबपेज (webpage) को दूसरी भाषा में अनुवाद किया जाता है। यह अनुवाद के भाषा युग्म (Pairs) माइक्रोसॉफ्ट ट्रांसलेटर टेक्नोलॉजी द्वारा संचालित है, जिसे माइक्रोसॉफ्ट रिसर्च ने अनुवादक तंत्र के रूप में विकसित किया है। चीनी पारंपारिक और चीनी सरलीकृत लिप्यंतरण के इस नए युग्म (Transliteration Pairs) को माइक्रोसॉफ्ट विंडोज अंतरराष्ट्रीय टीम ने विकसित कर इसे अनुवाद हेतु प्रदान किया है। मशीन से संबंधित पाठ का अनुवाद माइक्रोसॉफ्ट वाक्य-विन्यास आधारित सांख्यिकीय मशीनी अनुवाद (Syntax-Based Statistical Machine Translation) प्रणाली से करता है। इस तंत्र से प्रत्येक दिन 10 लाख पृष्ठों का अनुवाद किया जाता है। साथ ही, यह किसी भी दिए गए भाषा के पाठ की जानकारी देता है। हाल ही में माइक्रोसॉफ्ट बिंग ने हमोंग भाषा का समावेश अनुवाद के लिए किया है।

1.2.3.1 बिंग अनुवाद तंत्र की विशेषताएँ :

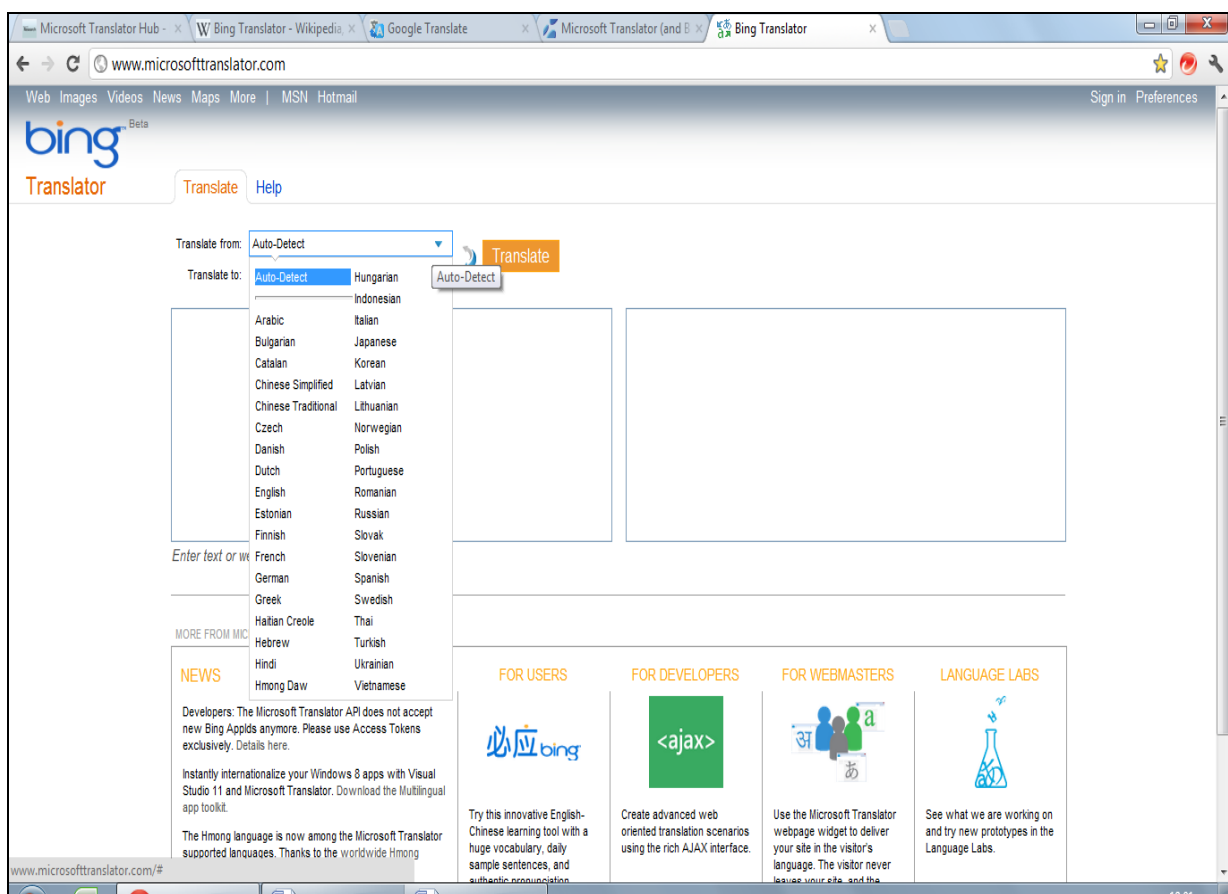
- किन्हीं दो भाषाओं के बीच अनुवाद प्रदान करना।
- माइक्रोसॉफ्ट तंत्र केवल मानक पाठ और वेबपेज के अनुवाद के साथ उपयोगकर्ताओं के लिए द्विभाषिक रूप में अनुवाद करता है।

¹⁰ http://en.wikipedia.org/wiki/Google_Translate retrieved on 10.04.2012 (उपरोक्त जानकारी गूगल ट्रांसलेट विकिपीडिया से ली गई है।)

¹¹ http://en.wikipedia.org/wiki/Bing_Translator retrieved on 19.04.2012 (उपरोक्त जानकारी बिंग ट्रांसलेटर विकिपीडिया से ली गई है।)

- पाठ (Text) से वाक (Speech) प्रणाली के अंतर्गत, निर्गत (Output) अनुवाद को स्पीच (Speech) के रूप में सुन सकते हैं।
- वेबसाइट या ब्लॉग के मालिक अपनी वेबसाइट या ब्लॉग में बिंग अनुवादक को समाहित (Integrate) कर सकते हैं। इस बिंग अनुवादक के माध्यम से आप अपनी पसंद की भाषा में वेबपेज का अनुवाद कर सकते हैं।

वर्तमान में बिंग अनुवादक तंत्र निम्नलिखित भाषाओं के लिए अनुवाद प्रदान कर रहा है – अरबी, बल्गेरियाई, कैटलन, चीनी सरलीकृत, पारंपारिक चीनी, चेक, डेनिश, डच, अंग्रेजी, एस्तोनियन, फिन जातीय, फ्रेंच, जर्मन, ग्रीक, हाईटियन, हिब्रू, हिंदी, हमोंग, इंडोनेशियन, इटालियन, जापानी, कोरियाई, लात्वियन, लिथुनिअन, नार्वेजियन, पोलिश, पुर्तगाली रोमनियाई, रूसी, स्लोवाक, स्लॉवेनियन, स्पेनिश, स्वीडिश, थाई, तुर्की, युक्रेनी, वियतनामी आदि इस तंत्र में समावेशित भाषा को इस निम्न फलक-चित्र (Screen-Shot) के द्वारा देखा जा सकता है।



Screen Shot 3: Showing Languages of Microsoft Bing Translator

1.2.3.2 बिंग अनुवाद तंत्र का प्रयोग¹²

पाठ का अनुवाद : जिस भाषा (मूल भाषा) का जिस भाषा में (लक्ष्य भाषा) अनुवाद करना है, उस भाषा का चयन कर, भाषा के ड्राप बाक्स में पाठ या वेबपृष्ठ का URL दीजिए अथवा पाठ को लिखें या पेस्ट करें और 'अनुवाद' बटन पर क्लिक करें, इससे अनुवाद प्राप्त हो जाएगा।

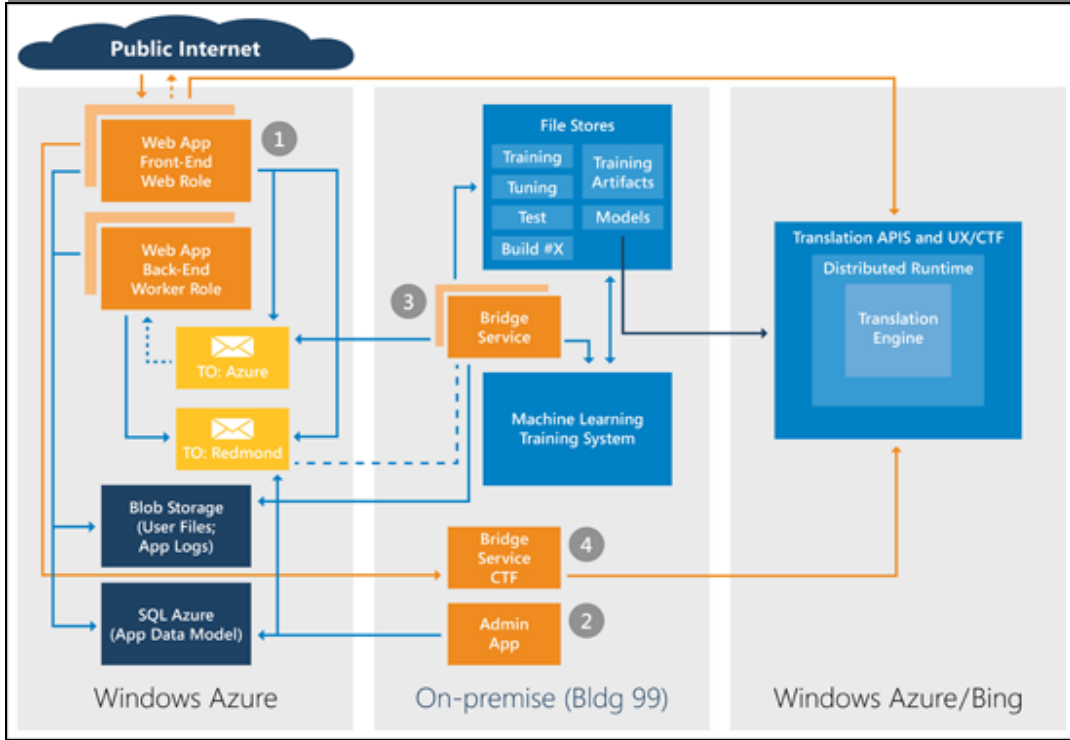
विंडोज़ फ़ोन के लिए नए अनुवाद तंत्र का अनुप्रयोग¹³ : इस तंत्र का अनुप्रयोग यह है कि अगर किसी स्थान पर आप जाते हैं और आप वहाँ की भाषा नहीं जानते हैं, तो इस तंत्र के प्रयोग से आप वहाँ के रास्तों के बोर्ड, पोस्टर, रेस्तरां के मेनू आदि पर, मुद्रित पाठ को फोन के कैमरे में तस्वीर खींचकर इससे स्वचालित अनुवाद (Automatic Translation) प्राप्त कर सकते हैं। माइक्रोसॉफ्ट ने इस अनुवाद तंत्र का निर्माण व्यावसायिकों और समुदायों के लिए स्वचालित अनुवाद प्राप्त करने के लिए किया है, जो भाषा को विकसित करने के साथ सटीक और सही अनुवाद प्रदान करने के लिए प्रयासरत है। स्वचालित भाषा के अनुवाद में कुछ प्रदाताओं का प्रभुत्व है, इसलिए विश्व में 100 से कम भाषाओं का भाषावैज्ञानिक रूप से विकास हुआ जबकि विश्व में 7000 से अधिक भाषाएँ हैं। माइक्रोसॉफ्ट इस उद्देश्य की पूर्ति हेतु स्वदेशी भाषा का विकास करने के लिए वैश्विक ज्ञान का उपयोग (Global Knowledge Base) कर विशाल सूचना प्रदान करता है। माइक्रोसॉफ्ट अनुवाद तंत्र अनुवाद के ऐसे मॉडल का निर्माण कर रहा है जिससे स्वदेशी (indigenous) भाषा को जीवित रहने में सहायता होगी। और भाषा में अनुवाद के लिए किसी एक समुदाय या संघटन के विशिष्ट हितों के लिए परिवहन, कृषि, स्वास्थ्य और प्रौद्योगिकी जैसे क्षेत्रों को विकसित करना माइक्रोसॉफ्ट बिंग अनुवादक का एकमात्र लक्ष्य कहा जाता है।

माइक्रोसॉफ्ट अनुवाद तंत्र से अनुवाद करने में **माइक्रोसॉफ्ट अनुवाद हब¹⁴** (Microsoft Translator Hub) की भूमिका महत्वपूर्ण है। यह हब इस तंत्र को बेहतर सेवाएँ प्रदान करता है और अन्य अनुवाद तंत्रों के निर्माण में सहायता भी करता है। इस हब के जरिए आप अपने तंत्र का निर्माण भी कर सकते हैं और अपने म. अनु. तंत्र का या निजी वेबसाइट के अनूदित दस्तावेजों का संयोजन माइक्रोसॉफ्ट बिंग अनुवादक से करके देख सकते हैं। अगर पाया गया अनुवाद सही है, तो आप अनूदित भाषिक अनुवाद को माइक्रोसॉफ्ट अनु. तंत्र की तरह सार्वजनिक रूप से दुनिया के साथ वेब पर साझा कर सकते हैं।

¹² <http://www.microsofttranslator.com/help/?FORM=R5FD> retrieved on 19.04.2012

¹³ <http://blogs.msdn.com/b/translation/> retrieved on 19.04.2012

¹⁴ <http://research.microsoft.com/en-us/projects/microsofttranslatorhub/> retrieved on 19.04.2012



Screen shot 4: Microsoft Translator Hub Architecture¹⁵

उपर्युक्त फलक-चित्र (screen-shot) में माइक्रोसॉफ्ट अनुवाद हब की अनुवाद प्रक्रिया को दिखाया गया है। प्रथम चरण में समानांतर दस्तावेज का एक सेट है, जो दो भाषाओं में लिखा होता है, जिसे स्रोत भाषा से लक्ष्य भाषा में अनुवाद करना होता है, उन दोनों भाषाओं को समानांतर युग्म के रूप में संचित रखता है। द्वितीय चरण में - समानांतर दस्तावेज के प्रयोग से माइक्रोसॉफ्ट अनुवाद हब सीखने (Learning Process) की प्रक्रिया को दिखाया गया है। शब्दों, वाक्यांशों, और वाक्यों का आमतौर पर कैसे अनुवाद करते हैं, अनुवाद के संदर्भ के अनुसार कैसे उचित संदर्भ का चयन करें, कैसे क्रियाएँ (Verbs) संयुग्मित और संज्ञाएँ विभक्तियुक्त होती हैं, यह संदर्भ पर निर्भर करता है, इन बिंदुओं को इस प्रक्रिया में स्पष्ट किया गया है। तृतीय चरण में - मशीन से निर्गत अनुवाद को गुणवत्तापूर्ण बनाने के लिए विशेषज्ञों से सहायता ली जाती है, और सिस्टम को परिष्कृत कर अनुवाद के लिए तैयार किया जाता है। चतुर्थ चरण - तंत्र का निर्माण होने के बाद वेब सेवाओं के लिए मानक उपकरण Microsoft Translator APIs - HTTP, AJAX, and SOAP interfaces, के साथ ही वेबपेज मशीन (widget) का प्रयोग कर रहे हैं। इस माइक्रोसॉफ्ट अनुवाद हब के चरणों से म.अनु. तंत्र के निर्माण के साथ इसे और परिष्कृत करने के लिए माइक्रोसॉफ्ट कार्यरत है। अनुवाद के लिए यह बहुत महत्वपूर्ण है कि विभिन्न भाषाओं के लिए इसमें कार्पस सामग्री, व्याकरणिक नियमों और उपकरणों का समावेश कर विश्व स्तर की भाषाओं का अनुवाद किया जा रहा है।

¹⁵ <http://research.microsoft.com/en-us/projects/microsofttranslatorhub/> retrieved on 19.04.2012,

बिंग अनुवाद तंत्र से प्राप्त अनुवाद पूर्ण रूप से सही है, यह तो नहीं कह सकते, लेकिन यह 60 प्रतिशत अनुवाद सही करता है तो कुछ अनअनुदित (Untranslated) शब्द को लिप्यंतरित करता है। इस अनुवाद से प्राप्त परिणाम को देखते हुए, इसमें आनेवाले समय में बदलाव के लिए विश्व स्तर पर माइक्रोसॉफ्ट प्रयास कर रहा है।

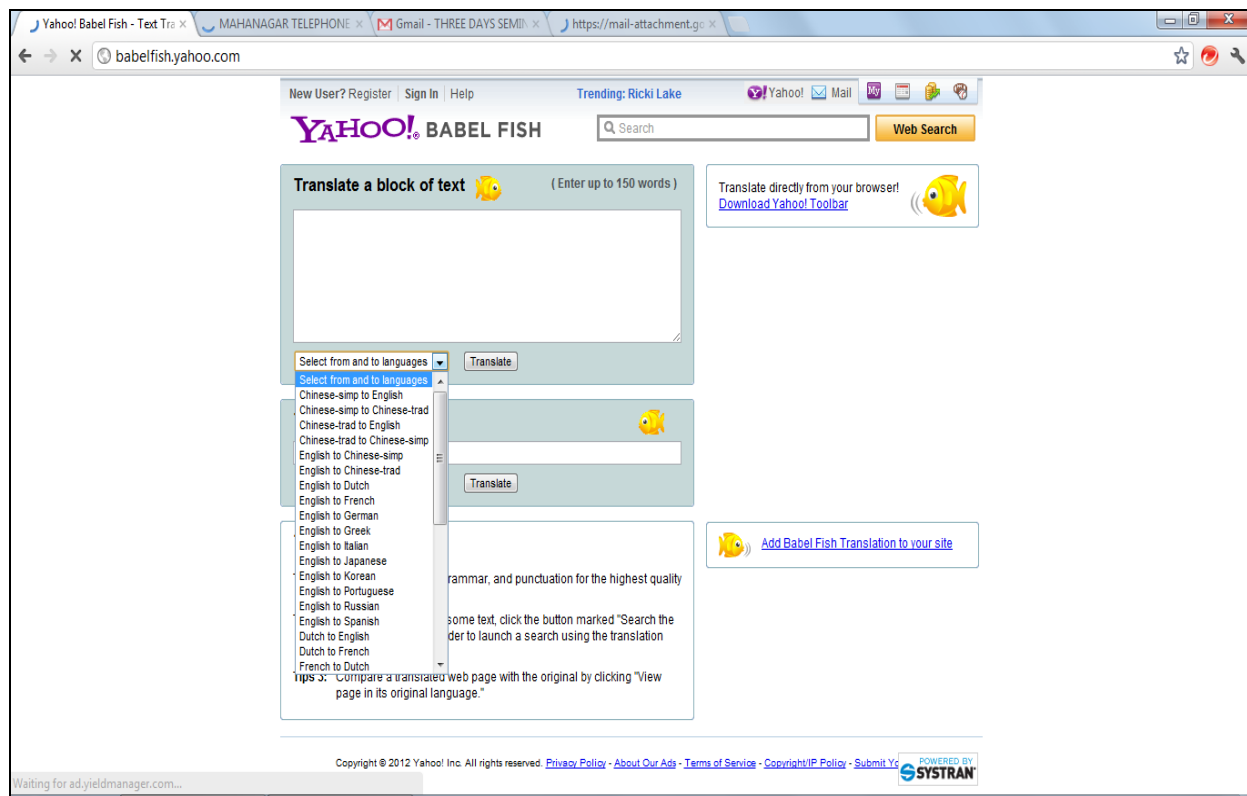
1.2.4 याहू बेबेल फिश (Yahoo Babel Fish 2008)

याहू बेबेल फिश (Babel fish) अल्टा विस्टा (AltaVista) द्वारा विकसित याहू पर एक वेब आधारित अनुप्रयोग (Application) है, जो एक भाषा के पाठ (Text) या वेबपेज (webpage) का किसी दूसरी भाषा में अनुवाद करता है। यह अनुवाद की प्रौद्योगिकी (Translation Technology) बेबेल फिश को सिस्ट्रान (SYSTRAN) द्वारा प्रदान की गयी है जिसने अन्य अनुवाद की सेवाओं को यह प्रौद्योगिकी प्रदान की है। यह विश्व की प्रमुख 14 भाषाओं का युग्म (Pairs) के रूप में अनुवाद प्रदान करता है। इस तंत्र को ऑस्कर ए. जोफ़ (Oscar A Jofre) और उनकी टीम ने जनवरी, 1999 में विकसित किया है। बाद में इसे 'Fictional animal' (कल्पित प्राणी) नाम दिया गया। बेबेल फिश तंत्र (Babel Fish) का प्रयोग डगलस एडम्स (Douglas Adams's) की श्रृंखला 'The Hitchhikers Guide to the Galaxy' के तत्कालिक भाषिक अनुवाद के लिए किया जाता था। फिश नाम बेबेल के साथ लगाया गया है, वह बाइबिल के बेबेल (Babel) शहर के नाम से उद्धृत किया गया है। बेबेल फिश को लंबे समय अंतराल के बाद babelfish.altavista.com से 9 मई, 2008 को babelfish.yahoo.com में तब्दील कर दिया गया है।¹⁶

बेबेल फिश तंत्र में विविध भाषाओं का समावेश किया गया है। वह एक भाषा से दूसरी भाषा में अनुवाद करता है। इस तंत्र में भी विश्व स्तर की भाषाओं के युग्मों (Pairs) का समावेश किया गया है। वे भाषाएँ निम्न हैं- चीनी सरलीकृत से अंग्रेजी, चीनी सरलीकृत से चीनी पारंपरिक, चीनी पारंपरिक से अंग्रेजी, चीनी पारंपरिक से चीनी सरलीकृत, अंग्रेजी से चीनी सरलीकृत, अंग्रेजी से चीनी पारंपरिक, अंग्रेजी से डच, अंग्रेजी से फ्रेंच, अंग्रेजी से जर्मन, अंग्रेजी से यूनानी, अंग्रेजी से इतालवी, अंग्रेजी से जापानी, अंग्रेजी से कोरियन, अंग्रेजी से पुर्तगाली, अंग्रेजी से रूसी, अंग्रेजी से स्पैनिश, डच से अंग्रेजी, डच से फ्रेंच, फ्रेंच से डच, फ्रेंच से अंग्रेजी, फ्रेंच से जर्मन, फ्रेंच से यूनानी, फ्रेंच से इतालवी, फ्रेंच से पुर्तगाली, फ्रेंच से स्पैनिश, जर्मन से अंग्रेजी, जर्मन से फ्रेंच, यूनानी से अंग्रेजी, यूनानी से फ्रेंच, इतालवी से अंग्रेजी, इतालवी से फ्रेंच, जापानी से अंग्रेजी, कोरियन से अंग्रेजी, पुर्तगाली से अंग्रेजी, पुर्तगाली से फ्रेंच, रूसी से अंग्रेजी,

¹⁶ http://en.wikipedia.org/wiki/Yahoo!_Babel_Fish retrieved on 19.04.2012

स्पैनिश से अंग्रेजी, स्पैनिश से फ्रेंच आदि। इस तंत्र की अनुवाद प्रक्रिया को निचे दिए फलक-चित्र (screen-shot) द्वारा देख सकते हैं।¹⁷



Screen shot 5: Languages of Babel fish system¹⁸

उक्त फलक-चित्र (screen-shot) से ज्ञात होता है कि बेबेल फिश तंत्र भाषाओं के युग्म का अनुवाद सहज रूप से करता है। बेबेलफिश अनुवाद तंत्र का प्रयोग कोई भी उपयोक्ता कर सकता है। इस म.अनु. तंत्र को याहू के द्वारा परिष्कृत किया जा रहा है ताकि एक भाषा से किसी भी भाषा में अनुवाद सही और सटीक प्राप्त हो।

वर्तमान में अनुवाद सेवा प्रदाता बेबेलफिश कार्पोरेशन एडमोन्टन (Babel Fish Corporation of Edmonton) ने ओटावा की दो कंपनियों के साथ विलय कर कंपनी के कार्यालय को राजधानी में पुनर्स्थापित किया है। बेबेलफिश कनाडा में बहुभाषिक अनुवाद सेवाएँ प्रदान करनेवाला सबसे बड़ा प्रदाता है, अब कनाडा के बाद इसका विस्तार अमेरिका में करने की योजना है। इस अनुवाद तंत्र को विकसित तथा परिष्कृत करने के लिए कंपनी के पास 7500 अनुवादक, 75 विभिन्न भाषाओं के अध्यापक हैं और 38 कर्मचारी हैं तथा कंपनी के कार्यालय कैलगरी, टोरंटो, मॉन्ट्रियल, वैंकूवर और ग्रेट ब्रिटेन में स्थित हैं।¹⁹

¹⁷ [view-source:http://babelfish.yahoo.com/](http://babelfish.yahoo.com/) retrieved on 19.04.2012

¹⁸ <http://babelfish.yahoo.com/> retrieved on 20.04.2012

¹⁹ <http://www.businessedge.ca/archives/article.cfm/babelfish-on-the-move-3074/> retrieved on 20.04.2012,

1.3 भारत के प्रमुख मशीनी अनुवाद तंत्र (Indian MT Systems)

विश्व स्तर पर जिस तरह से म.अनु. के तंत्र विकसित किए गए और विकसित किए जा रहे हैं उसी तरह भारत भी इस दिशा में प्रगति पर है। भारत एक बहुभाषिक देश है, जिसमें एकाधिक भाषा परिवारों की भाषाएँ हैं और म.अनु. के लिए किसी एक भाषा के क्षेत्र या विषय को लेकर विकास कार्य किया जा सकता है। इसलिए प्राकृतिक भाषा को विकसित करने के लिए भाषा के कॉर्पस (Corpus), द्विभाषिक कोश (Bilingual Dictionary), शब्द संसाधक (Word Processor), विच्छेदक (Parser), रूपवैज्ञानिक विश्लेषक (Morphological Analyzer), प्रजनक (Generator), वर्तनी जाँचक (Spell Checker), वाक्य विश्लेषक (Syntactic analysis), वाक प्रौद्योगिकी (Speech Technology) आदि उपकरणों (Tools) का निर्माण आवश्यक है।

भारत में इस विकास कार्य में भारतीय भाषाओं के लिए प्रौद्योगिकी विकास (TDIL), आई.आई.टी. कानपुर, आई.आई.आई.टी. हैदराबाद, हैदराबाद विश्वविद्यालय, सी-डैक बंगलोर, पुणे, राष्ट्रीय सॉफ्टवेयर प्रौद्योगिकी केंद्र, मुंबई; अन्ना विश्वविद्यालय, चेन्नई; ज. ने. वि.वि.नई दिल्ली, आई.बी.एम.रिसर्च इंडिया, आदि का कार्य विशेष उल्लेखनीय है। भारत में प्रमुख रूप से विकसित किए गए मशीन अनुवाद के तंत्र निम्नलिखित हैं।

1.3.1 आंग्ल-भारती (Anglabharati 1991)

“आंग्ल-भारती योजना का प्रारंभ 1991 में प्रो. आर.एम.के.सिन्हा के निर्देशन में भारतीय प्रौद्योगिकी संस्थान कानपुर (IITK) में हुआ है। इस तंत्र का निर्माण अंग्रेजी (English) से भारतीय भाषा में अनुवाद के लिए किया गया था। प्राथमिक रूप से इसका प्रयोग अंग्रेजी से हिंदी में अनुवाद के लिए किया गया।”²⁰ इस तंत्र में नियम आधारित प्रणाली (Rule Based System) का उपयोग किया गया है। यह विशेष रूप से स्वास्थ्य सेवाओं में प्रयोग की जानेवाली भाषा का अंग्रेजी से हिंदी में अनुवाद करता है। अंग्रेजी भाषा की वाक्य संरचना कर्ता-कर्म-क्रिया (SOV) पर आधारित है; जबकि भारतीय भाषाओं की वाक्यसंरचना शब्दक्रम मुक्त (free word order) और कर्ता-क्रिया-कर्म (SVO) के क्रम पर आधारित है। अंग्रेजी से प्रत्येक भारतीय भाषा में अनुवाद के लिए आंग्ल-भारती ने Pseudo – interlingua (Pseudo Lingua for Indian Languages (PLIL)) प्रणाली का प्रयोग किया है। इस प्रणाली के अंतर्गत यह अंग्रेजी वाक्य का पहले विश्लेषण करती है और फिर भारतीय भाषाओं के लिए मध्यवर्ती संरचना (intermediate structure) की अस्पष्टता को दूर करते हुए लक्ष्य भाषा पाठ का प्रजनन (text generation) करती है। इस तरह

²⁰ Durgesh, Rao (2001) : “Machine Translation in India: Brief survey”, In Proceeding of SCALLA 2001 Conference, Bangalore, India,
<http://www.elda.org/en/proj/scalla/SCALLA2001/SCALLA2001Rao.pdf> retrieved on 07.04. 2012

यह अंग्रेजी वाक्य का विश्लेषण 70 प्रतिशत और भारतीय भाषाओं के पाठ का प्रजनन 30 प्रतिशत करती है। इस प्रणाली के तहत अंग्रेजी से भारतीय भाषा में अनुवाद के लिए अतिरिक्त 30 प्रतिशत प्रयास करने की ज़रूरत है। इससे मशीन के द्वारा 90 प्रतिशत अनुवाद प्राप्त किया जा सकता है और 10 प्रतिशत मानव के द्वारा पश्च संपादन (Post-editing) के द्वारा पूर्ण किया जा सकता है।²¹

आंग्लभारती तंत्र के प्रमुख घटक²²

1. नियम आधारित प्रणाली अंग्रेजी से भारतीय भाषाओं के लिए वाक्य के सम्मिलन संरचनाओं (mapping structure) के लिए नियम प्रदान करता है।
2. तात्पर्य – निर्धारक सेन्स डिसेम्बिग्युएटरर्स (sense disambiguators) स्रोत भाषा के प्रत्येक शब्द के लिए सही अर्थ को चुनता है।
3. लक्ष्य पाठ्य निर्मापक (target text generator) लक्ष्य भाषा के लिए अनुवाद प्रदान करने का काम करता है।
4. लक्ष्य भाषा में अंग्रेजी के प्रत्येक शब्द के लिए बहुभाषिक शब्दकोश (multilingual dictionary) प्रदान करता है जैसे वाक्य की श्रेणियाँ, अर्थ की अस्पष्टता. इ. ।
5. सूत्राधार-अधिग्राहक (Rule-base acquirer) नियम आधारित अधिग्रहण में म.अनु. (Machine Translation) के लिए नियम निर्मित करने का कार्य करता है।

1.3.2 आंग्ल-भारती-II (Anglabharati 2004)

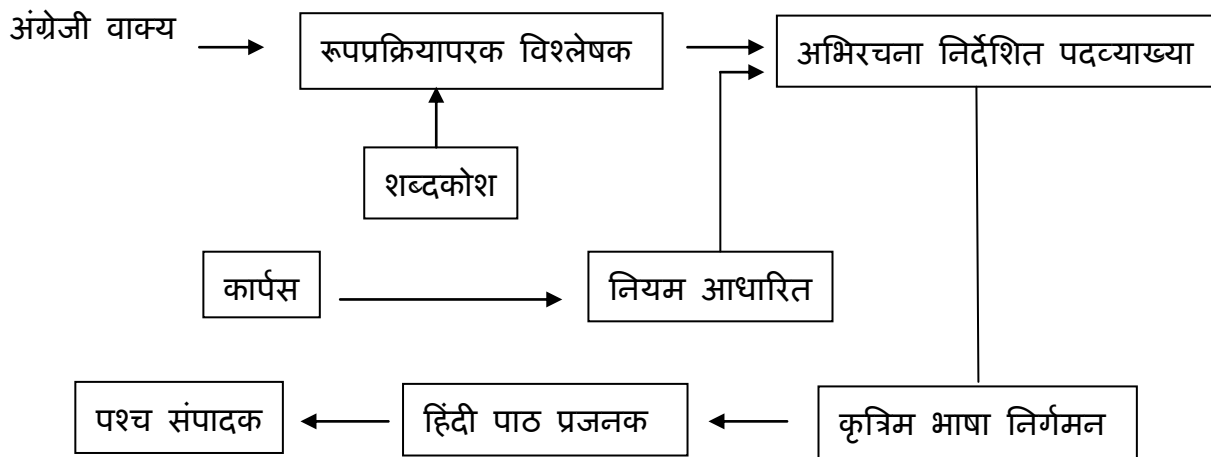
आंग्ल-भारती को विकसित करने के लिए (IIT) आई.आई.टी. कानपुर में प्रो. सिन्हा के निर्देशन में इसका निर्माण किया गया और इसे फिर से अद्यतन नए रूप में ई.आर.एंड डी.सी.आई. (भारतीय इलेक्ट्रॉनिकी अनुसंधान एवं विकास केन्द्र) नोएडा (वर्तमान में सी-डैक नोएडा) ने निर्माण किया जिसे आंग्ल-भारती-II कहते हैं। यह परियोजना कार्य भारतीय भाषा के लिए प्रौद्योगिकी विकास (TDIL) द्वारा वित्तपोषित है। इसमें नियम आधारित प्रणाली (rule-based system) और सामान्य उदाहरण आधारित प्रणाली (example based system) दोनों का संयुक्त रूप से प्रयोग किया गया है। यह अंग्रेजी से हिंदी में लोक स्वास्थ्य सेवा (Public Health Care),

²¹ Sinha, R. M. K., Jain, R. and Jain A. (2001): “Translation from English to Indian languages: ANGLABHARTI Approach”. In *proceedings of Symposium on Translation Support System STRANS 2001*. February 15-17, IIT Kanpur, India. p.167-172
<http://www.cse.iitk.ac.in/users/langtech/anglabharti.htm> retrieved on 07.04. 2012

²² Asamidinova, Ainura (2007): *Knowledge Base for Russian-English Machine Translation Divergences*, (Ph.D.Thesis) p.56

कार्यालयीन पत्राचार (Official Letter) और तकनीकी संदर्भिका (Technical context) आदि विषय सामग्री का अनुवाद करता है। यह स्वचालित रूप से अंग्रेजी वाक्य को पूर्व-संपादित (Pre-editing), रूपप्रक्रियापरक विश्लेषण (Morphological Analysis), हिंदी पाठ प्रजनन (Text Generator) और पश्च संपादन (Post editing) करता है।²³

आंग्ल-भारती-II अनुवाद प्रक्रिया को निम्न अभिकल्प (Design) द्वारा समझा जा सकता है-²⁴



आंग्ल-भारती ने म.अनु. के लिए भारतीय भाषाओं में, अंग्रेजी से भारतीय भाषा, अंग्रेजी से बांग्ला, अंग्रेजी से मलयालम, अंग्रेजी से उर्दू, अंग्रेजी से पंजाबी, संपर्क तंत्र आदि तंत्रों का निर्माण किया है। http://tdil-dc.in/index.php?option=com_content&view=article&id=61&lang=en इस लिंक पर यह तंत्र वेब (web) संस्करण के रूप में उपलब्ध है।

1.3.3 आंग्लहिंदी (Anglahindi)

यह एक Pseudo – Interlingual नियम आधारित अंग्रेजी से हिंदी मशीनी अनुवाद तंत्र है, जिसके निर्माण में आंग्लभारती के प्रतिरूपक (Modules) का प्रयोग किया गया है। यह संज्ञा पदबंध और क्रिया पदबंध को उदाहरण आधारित और सांख्यिकीय विधियों के द्वारा अस्पष्टता को दूर करता है। यह तंत्र 20 शब्दों तक के सरल, जटिल और मिश्रित वाक्य का लगभग 90 प्रतिशत स्वीकार्य अनुवाद कर सकता है। यह तंत्र वर्तमान में दो संगठनों में स्थानांतरित किया

²³ सिंह, सूरजभान (2003) अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ. 259

²⁴ कृष्णकुमार, गोस्वामी (2012). अनुवाद विज्ञान की भूमिका, पृ. 178

<http://www.languageinindia.com/oct2010/vishalthesis.pdf> retrieved on 07.04. 2012

गया है, जिसे हाल में Linux और Windows पर उपलब्ध कराया जा रहा है। यह तंत्र <http://anglahindi.iitk.ac.in/> इस लिंक पर उपलब्ध है।²⁵

1.3.4 अनुसारक²⁶ (Anusaaraka 1995)

आई.आई.टी.कानपुर ने 'अनुसारक' मशीनी अनुवाद तंत्र को भी विकसित किया है। इस परियोजना में कार्य करने वालों के स्थानांतरण के परिणामस्वरूप यह कार्य पहले केंद्रीय हैदराबाद विश्वविद्यालय में तथा अब आई.आई.आई.टी.(IIT) हैदराबाद में स्थानांतरित हो गया। यह परियोजना टी.डी.आई.एल. (TDIL) द्वारा वित्त पोषित थी। इस तंत्र ने पाणिनीय (Paninian) व्याकरण के सिद्धान्तों का प्रयोग कर भारतीय भाषाओं में अनुवाद को उन्नत बनाने का कार्य किया है। यह स्रोत भारतीय भाषा के पाठ्य को लक्ष्य भारतीय भाषा के निकटतम और ग्राह्य रूप में अनूदित करता है। इसकी मूल अवधारणा यह है कि भारतीय भाषाओं की व्याकरणिकता और अव्याकरणिकता को नजर अंदाज करते हुए वाक्य से अर्थ ग्रहण करने पर उसकी बोधगम्यता को आधार बनाया जाए और उसमें मानव की सहायता भी ली जाए। एक तरह से यह तंत्र सूचना स्थानान्तरण का कार्य करता है। यह तंत्र तेलुगु, कन्नड़, बांग्ला, मराठी, और पंजाबी से हिंदी आदि भारतीय भाषाओं में अनुवाद करता है। इस तंत्र का प्रयोग सामान्य अनुवाद और बच्चों की कहानियों का अनुवाद करने में किया जाता है। आई.आई.आई.टी. (IIT) हैदराबाद और हैदराबाद विश्वविद्यालय ने तेलुगु से हिंदी तंत्र, तथा संस्कृत अध्ययन विभाग हैदराबाद विश्वविद्यालय ने अंग्रेजी-हिंदी तंत्र विकसित किया है।²⁷

अनुसारक तंत्र के प्रमुख घटक:²⁸ (The components of the Anusaaraka system)

- रूपवैज्ञानिक विश्लेषक (Morphological Analyzer)
- स्थानीय शब्द समूहक (Local Word Grouper)
- बहु-भाषिक शब्दकोश (Bi-lingual Dictionaries)
- स्रोत भाषा से लक्ष्य भाषा के लिए सम्मिलक (Mapper from Source Language to Target Language)
- शाब्दिक संसाधक (Word Synthesizer)
- पश्च संपादन (Post-editing interface)

²⁵ <http://www.cse.iitk.ac.in/users/rmk/proj/proj.html> retrieved on 08.04.2012

²⁶ <http://anusaaraka.iit.ac.in/> retrieved on 08.04.2012

²⁷ झा गिरीश नाथ, दि.मणि, दि.मिश्र, (अंक 90), अप्रैल-जून (2008): "भारत में भाषा प्रौद्योगिकी: एक सर्वेक्षण", गवेषणा पत्रिका, पृ. 41

²⁸ Banerjee, Esha (2011): *Intra-phrasal agreement within the Hindi Noun Phrase: A Grammar Checker Approach* (M.Phil Dissertation) p.17

यह तंत्र निःशुल्क रूप में <http://trc.iiit.ac.in/showfile.php?filename=downloads/anu/index.htm> इस लिंक पर आनलाइन और डाउनलोड के लिए उपलब्ध है। अनुसारक के माध्यम से अंग्रेजी से हिंदी अनुवाद के लिए <http://anusaaraka.iiit.ac.in/node/1018> यह लिंक देख सकते हैं।

1.3.5 अणुभारती (Anubharati 1995)

आई.आई.टी.कानपुर की दूसरी परियोजना 'अणुभारती' है। प्रो.आर.एम.के सिन्हा के निर्देशन में हिंदी से अंग्रेजी और अन्य भारतीय भाषाओं में अनुवाद करने के लिए, इस तंत्र का विकास 1995 में किया गया था। यह उदाहरण आधारित प्रणाली (Example Base System) पर आधारित है, जो स्रोत भाषा से लक्ष्य भाषा में उदाहरण आधारित कॉर्पोरा के माध्यम से अनुवाद करता है। यह स्रोत भाषा के वाक्य और लक्ष्य भाषा के वाक्य के अनुवाद युग्म (Translation pairs) का शब्द के स्तर पर सम्मिलन (Mapping) करता है।²⁹ इससे "मिश्रित उदाहरण आधारित यांत्रिक अनुवाद प्रणाली (Hybrid Example-based Machine Translation) को नई दिशा प्राप्त हुई। यह तंत्र हिंदी की सजातीय भारतीय भाषाओं पर लागू होता है। इस तंत्र की महत्वपूर्ण विशेषताएँ ये हैं कि निष्पादित एवं संक्षिप्त उदाहरण-आधार का निर्माण, उदाहरण-आधार का विभाजन, तथा उदाहरण-आधार से सुमेलन को खोजता है।"³⁰

इस तंत्र में मिश्रित उदाहरण आधारित मशीनी अनुवाद प्रणाली (HEBMT) के द्वारा स्रोत भाषा हिंदी से किसी भी भारतीय भाषा में अनुवाद के लिए इसे सरलीकृत किया है। उदाहरण आधारित दृष्टिकोण से हिंदी के वाक्य को इनपुट (Input) में देने के बाद उसे पदबंध के स्तर पर विश्लेषित कर शब्दक्रम के अनुसार वाक्य का मानकीकरण किया जाता है और हिंदी के मानकीकृत वाक्य का अनूदित वाक्य के साथ मिलान किया जाता है। इसमें कुछ मिलान नहीं हुआ तो इनपुट वाक्य को Shallow Chunker के द्वारा विच्छेद किया जाता है और फिर से पदानुक्रम उदाहरण के अनुसार मिलान किया जाता है। इस मिलान के बाद मानव द्वारा पश्च संपादन (Post editing) किया जाता है।³¹

²⁹ Vishwabharat, A Ministry of Information Technology publication on Language Technology in India, Issues:

May, 2001, p.6

³⁰ झा गिरीश नाथ, दि.मणि, दि. मिश्र, (अंक 90), अप्रैल-जून (2008): "भारत में भाषा प्रौद्योगिकी: एक सर्वेक्षण", गवेषणा पत्रिका, पृ. 41

³¹ Goyal, Vishal (2010) : *Development of a Hindi to Punjabi Machine Translation System* (Phd Thesis), p. 674

http://www.iitk.ac.in/infocell/Archive/dirn02/techno_machine.html retrieved on 09.04.2012

1.3.6 मंत्र-राजभाषा³² (Mantra-Rajbhasha 1999) :

‘मंत्र’ मशीनी अनुवाद की तरह ही सी.डैक. पुणे के AAI (एप्लाइड आर्टिफिशियल इंटेलिजेन्स) ग्रुप द्वारा मंत्र-राजभाषा को सन 1999 में विकसित किया गया है। यह टी.डी.आई.एल एवं भारत सरकार के राजभाषा विभाग द्वारा वित्तपोषित है। मंत्र-राजभाषा एक मशीन साधित अनुवाद सिस्टम (Machine Assisted Translation) है। इस तंत्र में व्याकरणिक विश्लेषण सिद्धांत के सुप्रसिद्ध कम्प्यूटेशनल भाषाविज्ञानी प्रो.अरविन्द जोशी (पेन्सिलवेनिया यूनिवर्सिटी, अमेरिका से पीएच.डी.) के टैग व्याकरण प्रणाली (Tree Adjoining Grammar) का उपयोग किया गया है।³³ इसका मुख्य उद्देश्य भारत सरकार के गृह मंत्रालय, राजभाषा विभाग के अधीन राजभाषा और प्रशासनिक दस्तावेजों, सरकारी सूचनाओं, संकल्पों, आदेशों आदि कार्यालयीन साहित्य का अंग्रेजी से हिंदी में अनुवाद करना है। आरंभ में यह भारत सरकार के मंत्रालयों, कार्यालयों द्वारा जारी आदेशों, सूचनाओं आदि का अनुवाद करता था, लेकिन इस समय यह राजभाषा के प्रशासनिक, वित्त एवं कृषि क्षेत्र के दस्तावेजों का अंग्रेजी से हिंदी में अनुवाद करता है। “इस अनुवाद तंत्र में निर्विशिष्ट अंग्रेजी वाक्यों का (XTAG) के रूप में विश्लेषण होता है और उसके अनुसार हिंदी में अनुवाद किया जाता है। यह अनुवाद तंत्र प्रशासनिक क्षेत्र में 92 प्रतिशत और इतर क्षेत्र में 60-65 प्रतिशत अनुवाद करने में सक्षम है। यह अब ऑनलाइन और डाउनलोड के लिए उपलब्ध है”।³⁴

1.3.7 मात्रा³⁵ (MaTra 2004)

‘मात्रा’ यह पूर्णतःस्वचालित अंग्रेजी से हिंदी मशीन अनुवाद तंत्र है। यह टी.डी.आई.एल (TDIL) द्वारा वित्तपोषित है। इसे राष्ट्रीय सॉफ्टवेयर प्रौद्योगिकी केंद्र (National Centre for Software Technology) ने विकसित किया है, इस समय सी-डैक मुंबई में इस तंत्र पर विकास कार्य चल रहा है। ‘मात्रा’ ने प्राकृतिक भाषा विश्लेषण की पारंपरिक समस्याओं पर अभिनव दृष्टिकोण को अपना कर नई फ्रेम संरचित प्रणाली का प्रयोग किया है। यह तंत्र वाक्यों का अर्थपरक विश्लेषण करता है और इसके अंतर्गत अंतःक्रियात्मक संरचना शोधक (Interactive Structural Editor) उपलब्ध है, जिसके माध्यम से निवेशित अंग्रेजी वाक्यों का संरचनात्मक प्रतिरूपण (Structural Representation) प्राप्त होता है। जहाँ वाक्य में अस्पष्टता होती है, नियम आधारित प्रणाली का

³² <http://www.mantra-rajbhasha.cdac.in/mantrarajbhasha/> retrieved on 09.04.2012

³³ Rao, Durgesh (2001): “Machine Translation in India: Brief survey”, In Proceeding of SCALLA 2001 Conference, Bangalore, India,

³⁴ http://www.prayaslt.blogspot.in/2009_08_01_archive.html retrieved on 09.04.2012

³⁵ Vishwabharat, A Ministry of Information Technology publication on Language Technology in India, Issues: Oct,2003, p.106

प्रयोग कर उस अस्पष्टता को दूर किया जाता है।³⁶ मात्रा का शब्दकोश (Lexicon) और उपागम सामान्य प्रयोजन के लिए है। वर्तमान समय में इसका प्रयोग सामान्य वाक्य के अनुवाद के लिए हो रहा है। आगे इसे विस्तारित कर मिश्र वाक्य (Complex Sentence) के लिए विकसित किया जाना है। इसका प्रयोग पत्रकारिता, समाचार, वार्षिक रिपोर्ट (annual reports) और तकनीकी पद (Technical Phrase) आदि विषय क्षेत्र पर आधारित है। यह ऑनलाइन अनुवाद के लिए भी उपलब्ध है। इस तंत्र के MaTra Pro- The Professional Translators Tool; और MaTra Lite-automatic On-line Translator- ये दो संस्करण वेबसाइट पर उपलब्ध हैं।

1.3.8 शक्ति और शिवा³⁷ (Shakti 2004)

शक्ति म.अनु. तंत्र का विकास अंग्रेजी से भारतीय भाषाओं में अनुवाद के लिए किया गया है। इस समय यह तंत्र तीन भाषाओं अंग्रेजी से हिंदी, मराठी, और तेलुगु में अनुवाद करता है। इस तंत्र के लिए जो संयुक्त नियम बनाए गए हैं, वे भाषाई प्रकृति (Linguistic Nature) और सांख्यिकीय दृष्टिकोण (Statistical approach) पर आधारित हैं। सांख्यिकीय दृष्टिकोण इससे परिणाम निकालने का या भाषाई सूचना प्राप्त करने का कार्य करता है। तो कुछ मॉड्यूलस का प्रयोग अर्थ सूचना (Semantic Information) प्रदान करने के लिए होता है। यह तंत्र एकाधिक दृष्टिकोणों से संयोजित है, जिसका उद्देश्य भाषाई विश्लेषण करना है। इस तंत्र में 69 मॉड्यूलस का समावेश किया गया है, उसमें से 9 मॉड्यूलस का प्रयोग स्रोत भाषा (अंग्रेजी) विश्लेषण के लिए किया गया है और 24 मॉड्यूलस का प्रयोग द्विभाषी कार्य (Bilingual Tasks) में लक्ष्य भाषा के मूल शब्दों का प्रतिस्थापन और क्रम बदलने के लिए किया गया है और शेष मॉड्यूलस का प्रयोग लक्ष्य भाषा के प्रजनन (Generation) के लिए किया गया है। इस तंत्र की बनावट (Architecture) अत्यंत सरल है। सभी मॉड्यूलस को आधार सामग्री के आधार पर परिचालन किया जाता है, जिसका प्रारूप 'शक्ति मानक प्रारूप (Shakti Standard Format) है। दूसरी ओर "शिवा तंत्र अंग्रेजी से हिंदी में अनुवाद करता है और यह उदाहरण आधारित म.अनु. (EBMT) तंत्र है। इन दोनों तंत्रों का विकास अंतरराष्ट्रीय सूचना प्रौद्योगिकी संस्थान (IIIT) हैदराबाद, कार्नेगी मेलन विश्वविद्यालय, अमेरिका और भारतीय विज्ञान संस्थान (IIS) बेंगलूर के संयुक्त तत्वावधान में हुआ है।"³⁸ यह ऑनलाइन रूप में इस <http://shakti.iiit.ac.in/> (<http://shakti.iiit.net/>) लिंक पर उपलब्ध है।

³⁶ Ananthkarishnan, R, (ed), (2006): "MaTra: A Practical Approach to Fully-Automatic Indicative English-Hindi Machine Translation" CDAC:Mumbai .

<http://www.ncst.ernet.in/matra/> retrieved on 10.04.2012

³⁷ Dwivedi, Sanjay Kumar et. al. (2010): "Machine Translation System in Indian Prospectives", Journal of Computer Science 6 (10): 1111-1116, ISSN 1549-3636 © 2010 Science Publications

³⁸ Naskar, Sudip et. al. (2004): "Use of Machine Translation in India: current Status", Computer Science and Engineering department jadavpur university, Kolkata, p.4

1.3.9 अनुवादक³⁹ (Anuvadak)

नई दिल्ली की एक प्राइवेट संस्था सुपर इन्फोसाफ्ट प्रा.लि. ने अनुवादक 5.0 तंत्र का विकास किया है। इसका प्रयोग सामान्य प्रयोजनार्थ अंग्रेजी के सरल वाक्यों के हिंदी अनुवाद के लिए किया जाता है। इसमें प्रशासन, बैंक, कृषि, आदि विभिन्न क्षेत्रों की अंग्रेजी-हिंदी शब्दावली और पदबंधीय कोश सन्निहित हैं। इसमें सरकारी कार्यालयों, उपक्रमों, प्रतिष्ठानों, बैंकों आदि के अंग्रेजी वाक्यविन्यास का हिंदी में अनुवाद करने की क्षमता है। इससे अंग्रेजी के दस्तावेजों का हिंदी व्याकरण के नियमों का यथासंभव पालन करते हुए हिंदी में अनुवाद प्राप्त किया जा सकता है। यह तंत्र विंडोज के सभी संस्करणों (95/98/2000/NT/XP) में कार्य करता है। इस तंत्र में दिए गए शब्दकोश में नए शब्दों को जोड़ने तथा बदलने एवं शब्दों में संशोधन की सुविधा उपलब्ध है। यदि किसी शब्द के अनेक अर्थ हों तो सभी अर्थों को शब्दकोश में शामिल किया जा सकता है। साथ ही, अंग्रेजी और हिंदी दोनों भाषाओं के लिए वर्तनी जाँचक (Spell Checker) की सुविधा भी इस तंत्र में उपलब्ध है।⁴⁰

1.3.10 UCSG (MAT) - आधारित अंग्रेजी-कन्नड़ मशीनी अनुवाद तंत्र (2002)⁴¹

यह मशीन आश्रित अनुवाद तंत्र (Machine Assisted Translation) है, जिसे हैदराबाद विश्वविद्यालय के कंप्यूटर और सूचना विज्ञान विभाग (The Computer and Information Sciences Department) के डॉ. के.नारायण मूर्ति के निर्देशन में अंग्रेजी से हिंदी में अनुवाद करने के लिए विकसित किया गया है। इस तंत्र में सार्वभौमिक उपवाक्य संरचना व्याकरण (Universal Clause Structure Grammar) का प्रयोग हुआ है, जिससे अंग्रेजी से हिंदी अनुवाद के लिए निवेशित वाक्य (Input Sentences) को UCSG पार्सर के द्वारा विश्लेषित किया जाता है और यह अनुवाद कन्नड़ में द्विभाषी शब्दकोश (Bilingual Dictionary) के माध्यम से किया जाता है। इसमें रूपवैज्ञानिक विश्लेषक (Morphological analyzer) का भी समावेश किया गया है। इसका प्रयोग सरकारी परिपत्रों के अनुवाद के लिए किया जा रहा है। यह तंत्र कर्नाटक सरकार द्वारा वित्तपोषित है।

1.3.11 UNL - आधारित अंग्रेजी, हिंदी एवं मराठी भाषा का मशीन अनुवाद तंत्र (2003)⁴²

UNL- यूनिवर्सल नेटवर्किंग लैंग्वेज यूनाइटेड नेशन युनिवर्सिटी, टोकियो की अंतरराष्ट्रीय परियोजना है, जिसका उद्देश्य अंतरराष्ट्रीय स्तर की सभी मानव भाषाओं को अंतरभाषा

³⁹ <http://www.mysmartschool.com/pls/portal/portal.MSSStatic.ProductAnuvaadak> 11.04.2012

⁴⁰ नैटियाल ज.प्र., (अंक 88), अक्टूबर-दिसंबर (2007) : “भाषा विज्ञान और सूचना प्रौद्योगिकी”, गवेषणा पत्रिका, पृ. 41

⁴¹ Goyal, Vishal (2010) : “Development of a Hindi to Punjabi Machine Translation System” (Ph.D Thesis), p.675

⁴² Rao, Durgesh (2001) : “Machine Translation in India: Brief survey”, In Proceeding of SCALLA 2001 Conference, Bangalore, India.

(Interlingua) के माध्यम से विकसित करना है। UNL के माध्यम से अंग्रेजी से हिंदी को संयुक्त राष्ट्र संघ की भाषाओं से जोड़ने के लिए आई.आई.टी. (IIT) मुंबई ने इस परियोजना में महत्वपूर्ण भूमिका निभाई है और अंग्रेजी हिंदी एवं मराठी भाषाओं में म.अनु. तंत्र का निर्माण किया। इस तंत्र में हिंदी से यू.एन.एल. में रूपांतरण के लिए संपरिवर्तक (enconverter) और यू.एन.एल. से हिंदी में रूपांतरण के लिए विपरिवर्तक (deconverter) का प्रयोग किया है। यह तंत्र ऑनलाइन के लिए <http://www.cfilt.iitb.ac.in/machine-translation/eng-hindi-mt/> इस लिंक पर उपलब्ध है।

1.3.12 तमिल-हिंदी अनुसारक एवं अंग्रेजी-तमिल म. अनु. तंत्र⁴³

इस तंत्र का विकास के.बी चन्द्रशेखर शोध केन्द्र, अन्ना विश्वविद्यालय, चेन्नई के प्रो. सी.एन. कृष्णन के निर्देशन में हुआ है। यह तंत्र अनुसारक के आधार पर कार्य करता है, जिसमें तमिल पाठ (Text) का इनपुट दिया जाता है और उसका हिंदी में अनुवाद प्राप्त किया जाता है। इस तंत्र के तीन महत्वपूर्ण घटक (Major Components) हैं -

- रूपवैज्ञानिक विश्लेषक (Morphological Analyzer)
- सम्मिलक इकाई (Mapping unit)
- प्रजनक (Generator)

यह तंत्र वर्तमान में शब्द प्रति शब्द (Word to word translation) का अनुवाद करता है, लेकिन यह अर्थ विश्लेषण (Semantic analysis) पर आधारित नहीं है। इसी शोध केंद्र ने अंग्रेजी-तमिल मशीन साधित अनुवाद तंत्र को विकसित किया है, जो इसमें सीमित शब्दावली (100-150) के अंतरण नियमों (Transfer rules) का लघु समुच्चय (Small Sets) है। इस तंत्र में संपूर्ण वाक्यात्मक विश्लेषण (Syntactical analysis) का समावेश किया गया है।⁴⁴

1.3.13 जादवपुर विश्वविद्यालय का अंग्रेजी-हिंदी मैट (MAT) और 'अनुवाद' तंत्र (2004)⁴⁵

जादवपुर विश्वविद्यालय ने समाचार के वाक्यों के अनुवाद के लिए उदाहरण आधारित (Example Based) अंग्रेजी से हिंदी मशीन साधित अनुवाद तंत्र को विकसित किया है। यह तंत्र उदाहरण आधारित होने के साथ इसमें अंतरण पद्धति (Transfer approach) का भी प्रयोग किया गया है। यह समाचार के शीर्षकों का अनुवाद ज्ञान आधार (Knowledge bases) और उदाहरण संरचना के आधार पर करता है। उसी तरह इस विश्वविद्यालय ने समाचार के शीर्षकों का अनुवाद

⁴³ http://www.au-kbc.org/research_areas/nlp/projects/anusaaraka.html 12.04.2012

⁴⁴ Dwivedi Sanjay Kumar et. al (2010) : "Machine Translation System in Indian Prospectives" Journal of Computer Science 6 (10): 1111-1116, ISSN 1549-3636 © 2010 Science Publications, p-04

⁴⁵ Asamidinova, Ainura (2007) : *Knowledge Base for Russian-English Machine Translation Divergences*, (Ph.D. Thesis) p.66

करने के लिए अंग्रेजी से बांग्ला उदाहरण आधारित म. अनुवाद तंत्र को विकसित किया है। इसमें नियम और अंतरण प्रणाली का प्रयोग किया है। वर्तमान में इस तंत्र का नया संस्करण वाक्य स्तर पर अनुवाद कर रहा है।

1.3.14 असमिया मशीनी अनुवाद तंत्र

आई.आई.टी. गुवाहटी ने अंग्रेजी से असमिया मशीनी अनुवाद तंत्र को विकसित किया है। यह तंत्र मूलतः नियम आधारित है, जो अंग्रेजी से असमिया द्विभाषिक शब्दकोश के आधार पर अनुवाद करता है। इस तंत्र के शब्दकोश में 5000 मूल शब्दों (Root word) का समावेश किया गया है। यह अंग्रेजी से अनुवाद करते समय इनपुट वाक्य का (रूपवैज्ञानिक, वाक्यीय और अर्थ के स्तर पर) विश्लेषण करता है। इस तंत्र में अनुवाद की प्रक्रिया स्रोत भाषा से लक्ष्य भाषा में शब्द प्रति शब्द और पद से पद द्विभाषी शब्दकोश के माध्यम से होती है। यह तंत्र वर्तमान में अंग्रेजी से असमिया में सरल वाक्यों का अनुवाद करता है।⁴⁶

1.3.15 ओड़िसी मशीनी अनुवाद तंत्र

इस तंत्र को उत्कल विश्वविद्यालय की डॉ. संघमित्रा मोहंती ने विकसित किया है। इसमें स्रोत भाषा अंग्रेजी है और लक्ष्य भाषा ओड़िया है। अनुवाद के लिए विभिन्न उपकरण (Tools) ओड़िया से अंग्रेजी शब्दकोश (e-Dictionary Oriya to English), ओड़िया वर्डनेट (Oriya WorldNet), ओड़िया वर्तनी जाँचक (Oriya Spell checker), ओड़िया रूपवैज्ञानिक विश्लेषक (Morphological Analyzer), ओड़िया शब्द-संसाधक (Oriya Word Processor) का विकास किया जा रहा है।⁴⁷

1.3.16 आई.बी.एम (IBM) अंग्रेजी से हिंदी मशीनी अनुवाद तंत्र (2006)

आई.बी.एम (International Business Machines) रिसर्च इंडिया, नई दिल्ली ने सांख्यिकीय (Statistical) मशीनी अनुवाद तंत्र का विकास किया है। ये तंत्र अंग्रेजी से भारतीय भाषाओं में अनुवाद के लिए महत्वपूर्ण भूमिका निभा रहा है।⁴⁸

⁴⁶ Vishwabharat, A Ministry of Information Technology publication on Language Technology in India, Issues: Oct., 2003, p.108

⁴⁷ Vishwabharat, A Ministry of Information Technology publication on Language Technology in India, Issues: Oct., 2003, p.71

⁴⁸ Sitender, Bawa Seema (2012). "Survey of Indian Machine Translation Systems" JCST Vol.3, Issue 1, Jan-March 2012, ISSN:0976-8491(online)

1.3.17 लीला तंत्र (Leela software)

सी-डैक पुणे के AAI (एप्लाइड आर्टिफिसियल इंटेलिजेन्स) ग्रुप द्वारा लीला तंत्र विकसित किया गया है। Leela (Learning Indian Languages Through Artificial Intelligence) का अर्थ है- कृत्रिम बुद्धि की सहायता से भारतीय भाषाओं का अधिगम कार्यक्रम। भारत सरकार के गृह मंत्रालय के अंतर्गत कार्यरत राजभाषा विभाग द्वारा प्रस्तुत इस तंत्र ने हिंदी पाठ्यक्रम को तीन स्तर 'प्रबोध', 'प्रवीण' और 'प्राज्ञ' पर संयोजित किया है। यह तंत्र विभिन्न सरकारी कर्मचारियों को कक्षा शिक्षण तथा दूरस्थ शिक्षण प्रणाली द्वारा हिंदी शिक्षण प्रदान करता है। यह कार्यक्रम विशेषतः उन कर्मचारियों के लिए उपयोगी है जो नियमित रूप से कक्षा में उपस्थित रहकर हिंदी सीख पाने में असमर्थ हैं। आज यह सॉफ्टवेयर कार्यक्रम सभी कर्मचारियों को ऑनलाइन रूप में शिक्षा दे रहा है।⁴⁹

1.3.18 संपर्क (Sampark 2009) का विकास सभी भारतीय भाषाओं में अनुवाद तंत्र के विकास के लिए किया गया। वर्तमान में पंजाबी से हिंदी, हिंदी से पंजाबी, तेलुगु से तमिल, उर्दू से हिंदी, हिंदी से उर्दू और हिंदी से तेलुगु का अनुवाद करता है⁵⁰।

1.3.19 हिंग्लिश⁵¹ (Hinglish 2004)

यह तंत्र मानक हिंदी से अंग्रेजी में अनुवाद करता है, इसे प्रो. सिन्हा और ठाकुर ने विकसित किया है। विकसित अंग्रेजी से हिंदी आंग्लभारती द्वितीय और हिंदी से अंग्रेजी अनुभारती द्वितीय से अच्छा अनुवाद हिंग्लिश से प्राप्त किया जा सकता है। प्राप्त किया जाने वाला अनुवाद 90 प्रतिशत स्वीकार्य होता है। यह तंत्र केवल बहुअर्थक क्रियापदों के अर्थ को हल करने में केवल सक्षम नहीं है।

1.3.20 अन्य मशीनी अनुवाद तंत्र (Other Machine Translation System)

भारतीय भाषाओं में मशीनी अनुवाद की दिशा में और भी तंत्र कार्य कर रहे हैं और कुछ का प्रयास चल रहा है। बंगाली से आसामी में समाचार अनुवाद के लिए अनु. तंत्र एवं तेलुगु से तमिल म. अनु. तंत्र को हैदराबाद विश्वविद्यालय ने विकसित किया है। पंजाबी से हिंदी में म.अनु. तंत्र का विकास पंजाबी विश्वविद्यालय पटियाला में हुआ है। बंगाली से हिंदी में म.अनु.

⁴⁹ डबीरू, शिरीषा, (अंक 88), अक्टूबर-दिसंबर (2007) : "हिंदी एवं दक्षिण भारतीय भाषाओं में ई-लर्निंग का कार्यान्वयन", गवेषणा पत्रिका, पृ. 92

⁵⁰ <http://sampark.iiit.ac.in/sampark/web/index.php/registration/validuser> retrieved on 15.04.2012

⁵¹ Dwivedi Sanjay Kumar et. al (2010) : "Machine Translation System in Indian Perspectives" Journal of Computer Science 6 (10): 1111-1116, ISSN 1549-3636 © 2010 Science Publications, p. 04

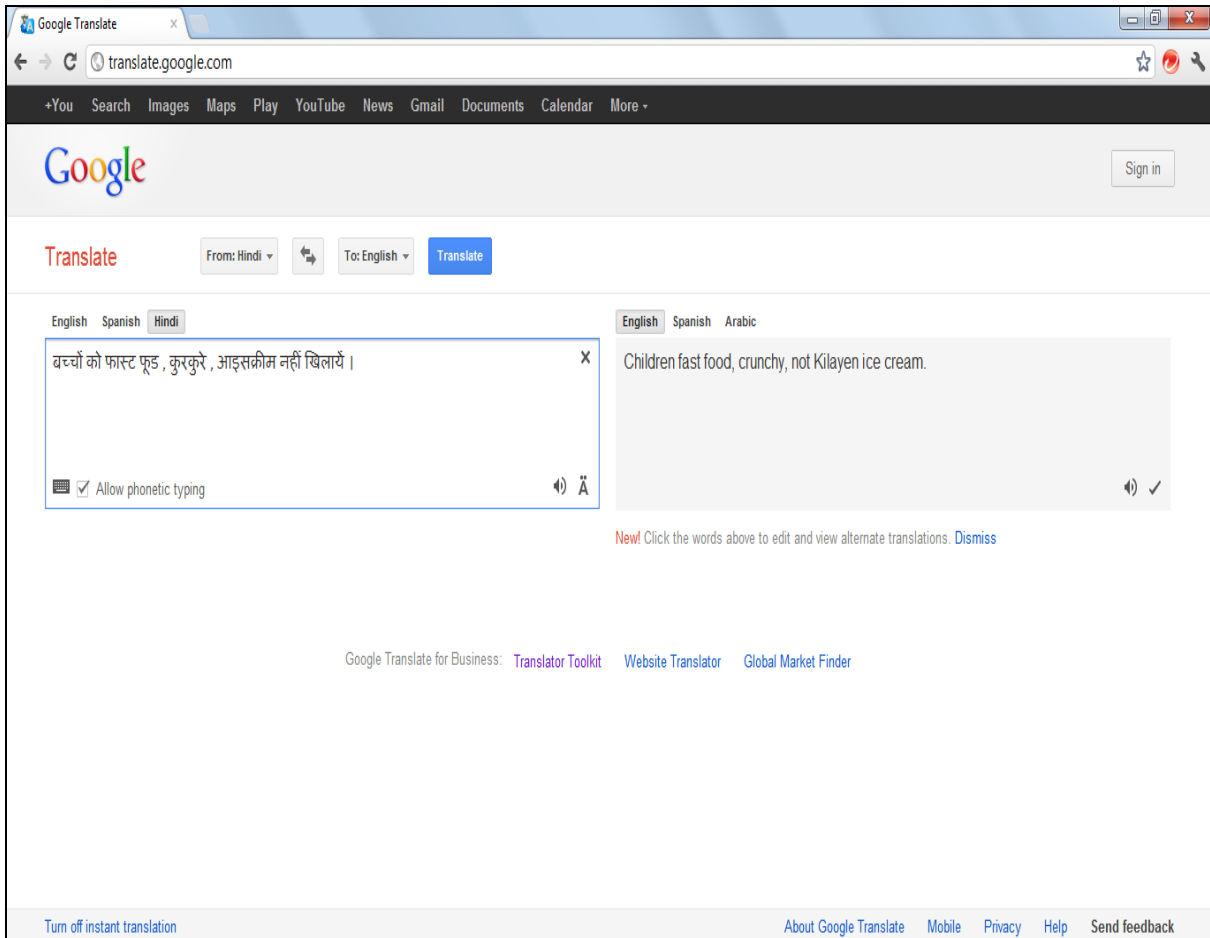
तंत्र आई.आई.टी. खड़गपुर ने विकसित किया है। अंग्रेजी से मणिपुरी में शिवाजी बंधोपाध्याय के निर्देशन में जादवपुर विश्वविद्यालय, कोलकाता में कार्य चल रहा है। हिंदी से पंजाबी म. अनु. तंत्र विशाल गोयल और प्रो. लेहाल ने पंजाबी विश्वविद्यालय, पटियाला में विकसित किया है और 'संस्कृत कारक विश्लेषक' सुधीर मिश्र (जे.एन.यू. नई दिल्ली) ने विकसित किया है। इसी तरह और भी मशीनी अनुवाद तंत्रों का विकास विश्वविद्यालयों, संस्थानों, प्राइवेट कंपनियों के स्तर पर चल रहा है। उपर्युक्त विकसित और प्रयास किए जा रहे मशीनी अनुवाद तंत्रों से यह स्पष्ट होता है कि यह तंत्र का विकास कार्य भारत में अधिक तेज़ गति से हो रहा है और मानव को इससे अनुवाद में सहायता हो रही है। भारत में इस विकास कार्य में भारतीय भाषा के लिए प्रौद्योगिकी विकास (TDIL), आई.आई.टी कानपुर, आई.आई.आई.टी हैदराबाद, हैदराबाद विश्वविद्यालय, सी-डैक, बंगलौर, पुणे., राष्ट्रीय सॉफ्टवेयर प्रौद्योगिकी केंद्र, मुंबई, अन्ना विश्वविद्यालय, चेन्नई, आई.बी. एम. रिसर्च इंडिया, आदि के द्वारा किया गया कार्य का योगदान महत्वपूर्ण है।

1.4 त्रुटि विश्लेषण (Error Analysis)

त्रुटि विश्लेषण संकल्पना का प्रयोग शुरू में भाषा शिक्षण के संदर्भ में हुआ। बाद में इसका प्रयोग अनुवाद के संदर्भ में किया जाने लगा। आज इसका प्रयोग मशीनी अनुवाद से प्राप्त किए जाने वाले अनुवाद के गुण-दोष के विवरण के लिए किया जाता है। मशीनी अनुवाद में ये त्रुटियाँ वाक्य स्तर के साथ शब्द और अर्थ के स्तर पर भी पायी जाती हैं। जब इस संकल्पना का प्रयोग मशीनी अनुवाद के संदर्भ में हो रहा है तो इस अध्याय में कुछ प्रश्नों का उठना स्वाभाविक है। अगर अनुवाद मशीन के द्वारा प्राप्त किया जाता है, तो त्रुटियाँ क्यों होती हैं? त्रुटियाँ हो रही हैं तो उसे विश्लेषित कैसे करें? त्रुटि विश्लेषण की प्रक्रिया क्या है? त्रुटियों की पहचान कैसे करें? त्रुटियों का विवरण और मूल्यांकन कैसे करें? इन प्रश्नों की प्रतिपूर्ति को इस संकल्पना में स्पष्ट किया गया है।

यह त्रुटि विश्लेषण ILCI कॉर्पोरा के हिंदी से अंग्रेजी स्वास्थ्य और पर्यटन के वाक्यों के अनुवाद पर आधारित है। जिसमें 2000 हिंदी के वाक्य मशीनी अनुवाद तंत्र - गूगल, और माइक्रोसाफ्ट बिंग ट्रांसलेटर में अंग्रेजी अनुवाद के लिए दिए गए और उसका अंग्रेजी अनुवाद प्राप्त किया गया। फिर दोनों के वाक्यों में क्या त्रुटियाँ हो रही हैं, यह देखा गया है। इस त्रुटि विश्लेषण को सैद्धांतिक तथा प्रायोगिक रूप में देखा जा सकता है।

प्रथमतः गूगल में अनुवाद के लिए हिंदी के वाक्य दिए गए और उनका अंग्रेजी अनुवाद प्राप्त किया गया है।



Screen Shot 6: Google Translation Process 2

उपर्युक्त फलक-चित्र (screen-shot) में दिखाया गया है कि गूगल कैसे अनुवाद करता है। इस तंत्र में अनुवाद के लिए जो वाक्य दिए गए, उन अधिकांश वाक्य में निम्नलिखित गलतियाँ पायी गयी हैं।

उदाहरण के लिए कॉर्पोरा के वाक्य निम्नलिखित है :

1. हिंदी वाक्य: बच्चों को फास्ट फूड , कुरकुरे , आइसक्रीम नहीं खिलायें।

गूगल अनुवाद: Children fast food, crunchy, not Kilayen ice cream.

2. हिंदी वाक्य: इससे साँसों से बदबू आने लगती है।

गूगल अनुवाद : This leads to bad-smelling breath.

3. हिंदी वाक्य: मुख के कैंसर के चिह्नों के लिए मुख की नियमित रूप से जाँच आवश्यक है ।

गूगल अनुवाद: Mouth to mouth checked regularly for signs of cancer is essential

4. हिंदी वाक्य: शल्य-चिकित्सा कैंसर के चरण पर निर्भर करता है ।

गूगल अनुवाद: Surgery depending on stage of cancer.

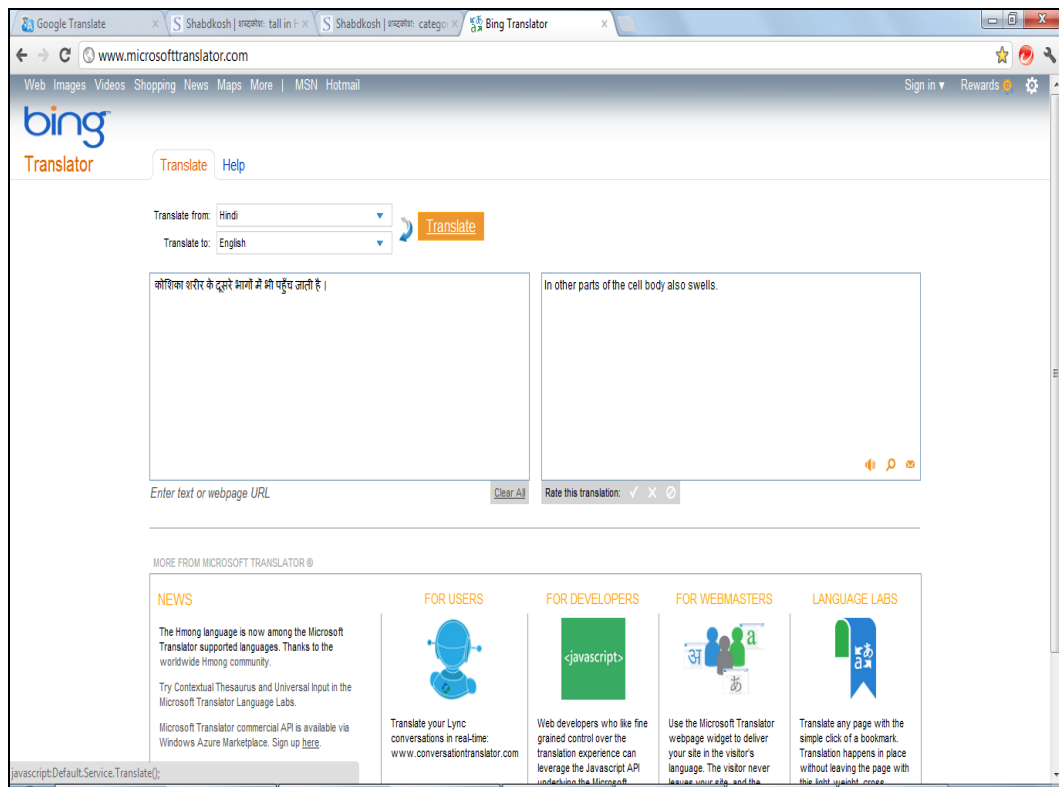
5. हिंदी वाक्य: आँखों को क्षति या रेटिनोपैथी ।

गूगल अनुवाद: Damage to eyes or retinopathy.

उपर्युक्त प्रथम वाक्य के अनुवाद में हिंदी के खिलाये शब्द को अंग्रेजी में वैसे ही रख दिया गया है। इससे इस वाक्य के अनुवाद का भाव अंग्रेजी पाठकों के लिए अर्थपूर्ण नहीं है, गूगल में जब किसी वाक्य के लिए शब्द नहीं मिलते तो वह लिप्यंतरण के रूप में अनुवाद करता है। द्वितीय वाक्य में Bad Smelling के अनुवाद से अंग्रेजी वाक्य का अनुवाद, अर्थ की दृष्टि से संपूर्ण गलत है। तृतीय वाक्य का अनुवाद संदिग्धता या अस्पष्टता के वर्ग में पाया जाता है। इस वाक्य में Mouth to mouth शब्द के अनुवाद से वाक्य में अस्पष्टता पायी गयी है। तो कुछ वाक्यों का अनुवाद अंशतः सही पाया गया है जैसे कि चतुर्थ वाक्य का अनुवाद। कुछ वाक्य का अनुवाद एकदम सटीक करता है। गूगल में यह विशेष रूप से पाया गया कि सामान्य और लघु वाक्य का अनुवाद पूर्ण रूप से सही करता है, लेकिन जब बृहद वाक्य अनुवाद के लिए दिए गए तो उसमें अस्पष्टता के साथ त्रुटिग्रस्त अनुवाद प्राप्त हुआ है।

गूगल द्वारा सही अनुवाद प्राप्त करने का प्रतिशत 70 है। इस प्रायोगिक अनुवाद से यह ज्ञात हुआ है कि अनुवाद में क्या त्रुटियाँ हो रही हैं, और त्रुटियाँ क्यों हो रही है। इसके लिए यह देख सकते हैं कि गूगल के अनूदित वाक्य की संरचना में हमेशा बदलाव होता रहता है। इसका कारण यह पाया गया कि वाक्य में कभी क्रिया (Verb) की तो कभी संज्ञा और विशेषक की भूमिका महत्वपूर्ण होती है। इस तंत्र से प्राप्त अनुवाद फिर भी अच्छा कहा जा सकता है।

त्रुटि विश्लेषण के दूसरी ओर माइक्रोसॉफ्ट बिंग ट्रांसलेटर के अनुवाद को भी देखा गया है कि वह कैसे अनुवाद करता है। नीचे इस तंत्र के अनुवाद की प्रक्रिया को दिखाया गया है।



Screen Shot 7: Microsoft Bing Translation Process

इस तंत्र की अनुवाद प्रक्रिया को फलक-चित्र (screen-shot) द्वारा देखा जा सकता है। इस तंत्र में अनुवाद के लिए हिंदी के वाक्य दिए गए, जिसका अंग्रेजी अनुवाद प्राप्त किया गया। सामान्यतः बिंग अनुवादक (Bing Translator) के अनुवाद में वाक्य के स्तर पर विविध त्रुटियाँ पायी गयीं। जिसमें अर्थ के साथ वाक्य में कुछ शब्द का अनुवाद लिप्यंतरण के रूप में प्रदान किया है, वह कॉर्पोरा (स्वास्थ्य और पर्यटन) के वाक्य के अनुवाद से लिया गया है। उदाहरण के लिए निम्नलिखित वाक्य देख सकते हैं।

1. हिंदी वाक्य : दंतचिकित्सक से दाँतों की जाँच नियमित रूप से कराएँ ।
बिंग अनुवाद : Dantachikitsak of teeth regularly check.
2. हिंदी वाक्य : सीढ़ी चढ़ना, बागवानी, घर के छोटे-मोटे काम या डांस करने जैसी हल्की-फुल्की शारीरिक गतिविधियाँ ज़रूर करते रहें ।
बिंग अनुवाद : Ladder climbing, gardening, home-work or dance like mushy coarse physical activities do, of course.
3. हिंदी वाक्य : एच.आई.वी. क्या है ?
बिंग अनुवाद : H. v.?

4. हिंदी वाक्य : नाश्ते में रोज दो या तीन अंडे जरूर खाएँ।
बिंग अनुवाद : Breakfast every day two or three eggs must avoid.
5. हिंदी वाक्य : चबाएँ शुगर रहित चुड़ंग गम ।
बिंग अनुवाद : Chabaen sugar free gum chuing.
6. हिंदी वाक्य : चिकुनगुनिया के रोगियों को लम्बे समय तक जोड़ों की पीड़ा हो सकती जो उनकी उम्र पर निर्भर करती है।
बिंग अनुवाद : Chunking long joints of patients who depend on their age might be.
7. हिंदी वाक्य : नियमित व्यायाम से आप हृदय रोग, कोलन कैंसर, ब्लड प्रेशर और डायबिटीज जैसी बीमारियों से बचाव कर सकते हैं ।
बिंग अनुवाद : Regular exercise, blood pressure heart disease, colon cancer, you and from diseases like diabetes.

उपर्युक्त वाक्यों के प्रथम वाक्य के अनुवाद में 'दंतचिकित्सक' शब्द का लिप्यंतरित अनुवाद हुआ है, तो दूसरे वाक्य में Domestic work के जगह Home work अनुवाद हुआ है। बिंग अनुवादक के लिए जब वाक्य बड़े हो तो अनुवाद में अक्सर त्रुटियाँ होती हैं, लेकिन यह लघु वाक्य में भी प्राप्त होती हैं। जैसा कि तृतीय वाक्य में देखा जा सकता है एच.आई.वी. का अनुवाद H.V किया है। चतुर्थ वाक्य में अनुवाद की सकारात्मकता को नकारात्मक रूप में बदल दिया, जिससे अनुवाद पूर्णतः गलत हुआ है। पाँचवें वाक्य का अनुवाद लिप्यंतरित किया गया है। जब बिंग अनुवाद की प्रक्रिया में कंप्यूटर द्वारा कुछ वाक्यों का आसान रूप में अनुवाद कर पाना असमर्थ होता है, तो ऐसे वाक्य को वह लिप्यंतरित कर देता है। छठे वाक्य के अनुवाद में यह देख सकते हैं कि 'चिकुनगुनिया' का अनुवाद Chunking कर दिया है। एक शब्द के अनुवाद से यह संपूर्ण वाक्य का अर्थ अस्पष्ट लगता है। जब वाक्य बड़ा होता है और वाक्य में एक से अधिक शब्द होते हैं तो अनुवाद में हमेशा त्रुटियाँ होती हैं, यह सातवें वाक्य से स्पष्ट होता है।

इस बिंग अनुवाद के त्रुटि विश्लेषण से यह स्पष्ट होता है कि कोई भी मशीनी अनुवाद का तंत्र त्रुटिरहित पूर्णतः अनुवाद प्रदान नहीं करता, उसमें त्रुटियाँ अवश्य होती हैं। इन त्रुटियों का विश्लेषण करना इस उद्देश्य से आवश्यक है कि इस अनुवाद प्रक्रिया से इन तंत्रों के अनुवाद की कार्यक्षमता और गुणवत्ता ज्ञात हो गयी। वास्तव में त्रुटि विश्लेषण का प्रयोग भाषा शिक्षण के संदर्भ में भाषा सीखने में होनेवाली कठिनाई का पता करने के लिए निदानात्मक विधि (Diagnostic Method) के रूप में किया गया था। लेकिन आज दूसरी ओर मशीनी अनुवाद में

प्राप्त अनुवाद की त्रुटियों का विश्लेषण के लिए इसका प्रयोग किया जा रहा है। उपर्युक्त वाक्य का त्रुटि विश्लेषण स्वास्थ्य कॉर्पोरा के अनुवाद के आधार पर किया गया है।

पर्यटन कॉर्पोरा के आधार पर गूगल और बिंग अनुवादक तंत्र से प्राप्त अनुवाद की तुलना और त्रुटि विश्लेषण टेबल 2 में दिखाई गई हैं।

आगत हिंदी वाक्य (Input Hindi Sentences)	निर्गत गूगल अनुवाद (Google Output)	निर्गत बिंग अनुवाद (Bing Translator Output)	सही अनुवाद (Actual Translation)
गढ़ मुक्तेश्वर का नाम शिव बल्लभपुर पड़ा ।	Garh Mukteshwar Shiva's name was Ballapur.	Citadel Mukteshwar name Shiva ballabhpur.	Garh Mukteshwar was named as Shiv Ballabhpur.
ट्रेन में प्रकृति का पूरा आनंद लेते हुए गोवा जाने का अलग ही मजा है।	Taking the train to Goa to enjoy nature's own fun.	Enjoy the nature train, Goa is the different fun.	Going to Goa having full joy of the nature in train has a different fun.
दमन का नजदीकी रेलवे स्टेशन गुजरात का 'वापी' है।	Suppression of the nearest railway station in 'Gujarat's <u>Valsad</u> is.	Proximity to the railway station of repression. ' Gujarat ' vapi	The nearby railway station of Daman is Vapi of Gujarat.
महेश मूर्ति की विशालता एवं विस्पर्यात्पादकता अपूर्व है।	Mahesh Murthy is the vastness and विस्पर्यात्पादकता unique.	Mahesh Murthy is a unique emerging and visparyatpadkata.	The largeness and wonderfulness of the Mahesha statue is unique.
आक्कुलम झील नौकारोहण की दृष्टि से महत्वपूर्ण है ।	Akkulm lake is important in <u>terms</u> of <u>haul</u> .	Akkulam is important from the point of view of the Lake <u>naukarohan</u> .	Akkulam lake is important in view of boating.

टेबल 2 : गूगल और बिंग मशीनी अनुवाद तंत्र का निर्गत अनुवाद

प्रस्तुत टेबल में गूगल और बिंग अनुवादक के द्वारा मशीनी अनुवाद की गुणवत्ता को देखा जा सकता है। मशीन में अनुवाद के लिए हिंदी के वाक्य दिए गए और अंग्रेजी में निर्गत अनुवाद प्राप्त किया गया। टेबल में प्राप्त इस अनुवाद की तुलना के आधार से त्रुटियों का विश्लेषण किया गया है।

प्रथम गूगल वाक्य में अनुवाद की वाक्य संरचना को बदल दिया गया है, ऐसा इसलिए हुआ है कि जब अनुवाद के लिए गूगल में किसी वस्तु या स्थान का प्रयोग किया जाता है तो अनुवाद में मशीन, सभी नामों (Name) को समान समझकर आउटपुट प्रदान करती है। इससे अनुवाद में मूल वाक्य के नाम और स्थान बदलने की संभावना बनी रहती है। यहाँ बिंग तंत्र का अनुवाद सही अवस्था में पाया गया है। संदर्भ के अनुसार वाक्य के कुछ शब्दों का अनुवाद और लिप्यंतरण मानव अपनी बुद्धि और पूर्व सांस्कृतिक ज्ञान के आधार पर करता है, लेकिन यह मशीन को समझ पाना मुश्किल है। बिंग अनुवाद के प्रथम वाक्य के शब्द 'Citadel' से यह ज्ञात होता है। द्वितीय वाक्य में दोनों तंत्रों गूगल और बिंग अनुवाद के द्वारा प्राप्त अनुवाद त्रुटिग्रस्त हैं, तो तृतीय वाक्य में 'दमन' और 'वापी' यह स्थान के नाम हैं, जिनका क्रमशः अनुवाद 'suppression' और 'Valsad' किया गया है और बिंग में भी यह अनुवाद त्रुटिग्रस्त पाया गया है। चतुर्थ और पंचम वाक्य में गूगल द्वारा अनूदित रूप में शब्द को वैसे ही (As it is) रख दिया गया है, तो बिंग अनुवाद में भी कुछ शब्दों का वैसे ही अनुवाद किया गया है।

उपर्युक्त टेबल में गूगल और बिंग के रेखांकित वाक्यों (शब्द) से त्रुटियों का विवरण मिलता है, कि दोनों मशीनी अनुवाद तंत्रों द्वारा अनुवाद कितना प्रतिशत सही प्राप्त हो सकता है। इन त्रुटियों के विवरण और विश्लेषण से देखा जा सकता है कि मशीनी अनुवाद में त्रुटियाँ हो सकती हैं और त्रुटियाँ क्यों हो रही हैं यह भी ज्ञात हुआ कि प्रत्येक भाषा की संरचना, साहित्य, संस्कृति से एक भाषा दूसरी भाषा से भिन्न है, इसलिए शब्द, वाक्य, और अर्थ के स्तर पर भिन्नता पायी जाती है।

मशीन के लिए जब अनुवाद दिया जाता है तो वह संदर्भ के अनुसार अनुवाद नहीं प्रदान करता, क्योंकि उसके लिए उस मशीन में वह सामग्री समावेशित (Feed) करना आवश्यक होता है। वह किसी सीमित क्षेत्र पर सही अनुवाद प्रदान कर सकता है। इस त्रुटि विश्लेषण से यह स्पष्ट होता है कि मशीनी अनुवाद के तंत्र में और सुधार करने की जरूरत है। फिर भी गूगल का अनुवाद अधिक मात्रा (70 प्रतिशत) में सही पाया जाता है और अर्थ को बनाए रखने के साथ कुछ अनूदित वाक्यों को वह देवनागरी में लिप्यंतरित करता है। दूसरी तरफ़ बिंग का अनुवाद 60 प्रतिशत ही बोधगम्य (comprehensible) हो पाता है। इन तंत्रों के अनुवाद में यह पाया गया कि अधिकतर अंग्रेजी के शब्दों को देवनागरी में अनूदित करता है तो कुछ शब्दों का अनुवाद रोमन में प्रदान करता है। इसलिए इसमें हिंदी और अंग्रेजी की शब्दावली में सुधार करने के साथ शब्द के अर्थ में स्पष्टता के लिए प्रत्येक शब्द के अर्थ को बद्ध (Bound) या निश्चित कर देना चाहिए ताकि अनुवाद अर्थपूर्ण और गुणवत्तापूर्ण प्राप्त हो।

1.5 संबंधित शोध कार्य (Related Research work)

मशीनी अनुवाद से संबंधित अनेक कार्य हुए हैं लेकिन संज्ञापद के सम्मिलन को लेकर यह भिन्न शोध कार्य है, उससे संबंधित संज्ञापद को लेकर निम्नलिखित शोध कार्य हुए हैं।

1. मर्णोद के. वर्मा ने मिशिगन विश्वविद्यालय (University of Michigan) से पीएच.डी में 'अंग्रेजी और हिंदी में संज्ञापद की संरचना' (The Structure of the Noun Phrase in English and Hindi, 1965-66) विषय पर शोध किया है। यह अभी पुस्तक के रूप में प्रकाशित है। इस शोध में अंग्रेजी और हिंदी संज्ञापद का विस्तृत अध्ययन कर दोनों भाषाओं के संज्ञापदों का तुलनात्मक अध्ययन किया गया है।
2. आई.आई.टी.(IIT) कानपुर में हिंदी से अंग्रेजी में अनुवाद हेतु प्रो.आर.एम.के. सिन्हा के निर्देशन में 'अणुभारती' तंत्र का निर्माण (1995) किया गया है। यह तंत्र उदाहरण आधारित प्रणाली (Example Base System) पर आधारित है जिसमें स्रोत भाषा से लक्ष्य भाषा में कॉर्पोरा का निर्माण किया है। यह स्रोत भाषा और लक्ष्य भाषा के वाक्य के अनुवाद युग्म (Translation pairs) को शब्द के स्तर पर सम्मिलन (Mapping) कर अनुवाद निर्गत (Output) करता है, यह कार्य तंत्र निर्माण आधारित है।
3. गणितशास्त्र विभाग, आई.आई.टी. (IIT) नई दिल्ली, में दीपा गुप्ता ने पीएच.डी. का शोध प्रबंध 'अंग्रेजी से हिंदी में मशीनी अनुवाद के लिए उदाहरण आधारित दृष्टिकोण का योगदान' (Contributions to English to Hindi Machine Translation Using Example – Based Approach, 2005) पर अध्ययन किया है, इस शोध को डॉ. नीलाद्रि चटर्जी ने निर्देशित किया है। इस शोध में अंग्रेजी से हिंदी में मशीनी अनुवाद के लिए उदाहरण आधारित मशीनी अनुवाद तंत्र के विकास पर अध्ययन किया गया है।
4. कंप्यूटर विज्ञान विभाग, पंजाबी विश्वविद्यालय, पटियाला में विशाल गोयल ने पीएच.डी. का शोध प्रबंध 'हिंदी से पंजाबी के लिए मशीनी अनुवाद तंत्र का विकास' (Development of a Hindi to Punjabi Machine Translation System, 2010) पर शोध किया है। इस शोध में हिंदी से पंजाबी भाषा में मशीनी अनुवाद तंत्र के विकास के लिए व्याकरण और वर्तनी का भाषा के स्तर पर तुलनात्मक अध्ययन किया है और तंत्र विकसित करने के लिए उपकरणों टूल्स (Tools) का निर्माण किया है।
5. विशिष्ट संस्कृत अध्ययन केंद्र, जवाहरलाल नेहरू विश्वविद्यालय, नई दिल्ली, में एम.फिल. का लघु-शोध प्रबंध कार्य कुमार नृपेन्द्र पाठक ने 'संस्कृत-हिंदी संज्ञापद सम्मिलन की

चुनौतियाँ' (Challenges in Sanskrit - Hindi Noun Phrase Mapping -2011) विषय पर शोध किया है। इस शोध को डॉ. गिरीश नाथ झा ने निर्देशित किया है। इस शोध में संस्कृत और हिंदी संज्ञापद का तुलनात्मक अध्ययन-विश्लेषण कर उसे स्थानांतरण व्याकरण के आधार पर सम्मिलित (Map) किया गया है, जिससे संज्ञापद को लेकर दोनों भाषाओं के अनुवाद में क्या बदलाव आया और कौन से पद उसमें जुड़ गए हैं, यह देखा गया है। साथ ही भारत के मशीनी अनुवाद तंत्रों (System) का वर्णन भी किया गया है।

6. भाषाविज्ञान केंद्र, जवाहरलाल नेहरू विश्वविद्यालय, नई दिल्ली, में एम.फिल. का लघु शोध-प्रबंध कार्य 2011, में ईशा बैनर्जी ने “*Intra-phrasal agreement in Hindi Noun Phrases: A grammar checker approach*” इस विषय पर किया है। इस शोध को प्रो. वैष्णा नारंग और डॉ. गिरीश नाथ झा ने निर्देशित किया है। इस शोध में हिंदी वाक्य और विशेषतः संज्ञा पदबंध (Noun Phrases) की व्याकरणिक जाँच करने के लिए व्याकरण जाँचक (grammar checker) का निर्माण किया गया है। हिंदी संज्ञापद के लिए POS Tag (पद वर्ग टैग) के आधार पर माइक्रोसॉफ्ट रिसर्च इंडिया (MSRI IL) के डाटा का प्रयोग किया गया है, साथ ही शीर्ष विशेषक (Modifier-head) और शीर्ष परसर्ग (Head-Postposition) के अनुसार संज्ञापद का सम्मिलन (Map) और तुलना की गयी है।

1.6. उपलब्ध अन्य शोध कार्यों से वर्तमान शोध की भिन्नता

अन्य शोध कार्यों से यह ज्ञात होता है, कि संज्ञा पदबंध से संबंधित कार्य विविध संस्थानों, विश्वविद्यालयों में हुए हैं। विशेषतः जवाहरलाल नेहरू विश्वविद्यालय, नई दिल्ली में संस्कृत भाषा के विकास को लेकर अनेक शोध कार्य हुए हैं, लेकिन हिंदी-अंग्रेजी भाषा के कॉर्पोरा को लेकर संज्ञा पदबंध पर शोध कार्य नहीं हुआ है। यह शोध कार्य हिंदी से अंग्रेजी संज्ञापद सम्मिलन पर आधारित है।

- इस शोध कार्य में भारतीय भाषा कॉर्पोरा उपक्रम (ILCI) के कॉर्पोरा का प्रयोग किया गया है। यह कॉर्पोरा स्वास्थ्य और पर्यटन के क्षेत्र के 50,000 वाक्यों का है, लेकिन नमूने के तौर पर 2000 वाक्यों का प्रयोग किया है। इस शोध में हिंदी से अंग्रेजी में अनूदित समानांतर कॉर्पोरा का वाक्यसंरचना के स्तर पर तुलनात्मक अध्ययन कर अनुवाद में पायी जानेवाली त्रुटियों का विश्लेषण भी किया गया है ।

- हिंदी और अंग्रेजी भाषा की प्रकृति समान न होने के कारण वाक्यसंरचना की तरह संज्ञापद में भी बदलाव पाया जाता है। जैसे अंग्रेजी में विशेषक (modifier) संज्ञा से पहले या बाद, दोनों स्थानों पर प्रयुक्त होते हैं, लेकिन हिंदी में विशेषक संज्ञा से पहले प्रयुक्त होते हैं, बाद में नहीं। विशेषतः संज्ञापद के स्तर पर कारक के प्रयोग का अध्ययन किया जाएगा।
- हिंदी-अंग्रेजी मशीनी अनुवाद (HAMT) के लिए दोनों भाषाओं के वाक्य के संज्ञापद को नियमबद्ध करने के लिए स्थानांतरण व्याकरण का प्रयोग किया गया है और हिंदी से अंग्रेजी में वाक्यसंरचना के स्तर पर संज्ञापदों का व्यतिरेकी विश्लेषण किया गया है।
- ILCI कॉर्पोरा के आधार पर मशीन द्वारा संज्ञापद को सम्मिलित (Mapping) किया गया है और मशीन द्वारा अनुवाद में पाए जानेवाली त्रुटियों का सुधार करने का प्रयास किया गया है। अंतः में संज्ञापद सम्मिलन के बाद दोनों भाषा के अनुवाद में पाए जानेवाले बदलाव के लिए स्थानांतरण व्याकरण नियम बनाए जाएंगे।

इस शोध कार्य में हिंदी-अंग्रेजी भाषा के अनुवाद में संज्ञापद के स्तर पर हमेशा क्या बदलाव होता है और इस बदलाव के वजह से वाक्यसंरचना में अर्थ का जो बदलाव होता है- उसे त्रुटि विश्लेषण अध्ययन से स्पष्ट किया जाएगा और उसे सुधार करने के लिए नियम बनाए जाएंगे। इसके लिए स्थानांतरण व्याकरण नियम का प्रयोग मशीनी अनुवाद में किया जाएगा, जिससे सटीक अनुवाद प्राप्त होने के लिए सहायता हो सकती है। यह शोध हिंदी से अंग्रेजी में म.अनु. तंत्र निर्माण करनेवाले शोधार्थी या संस्थान के लिए सहायक सिद्ध होगा, लेकिन इसके लिए और भी संसाधन (Resource) भाषा सामग्री (Corpora), शब्द संचयन (Lexical Data), रूपवैज्ञानिक विश्लेषक (Morphological Analyzer), विश्लेषक (Parser), वर्तनी जाँचक (Spell Checker), व्याकरण जाँचक (Grammar Checker), प्रजनक (Generator) आदि के साथ अल्गोरिदम (Algorithm) का प्रयोग भी तंत्र निर्माण के लिए आवश्यक होगा।

1.7 शोध प्रविधि (Research Methodology)

मशीनी अनुवाद का विकास 'अभिकलनात्मक भाषाविज्ञान' (Computational Linguistics) के अनुप्रयोगिक (Applied) क्षेत्र के रूप में हुआ है। इसलिए मशीनी अनुवाद के शोध में प्रयोगात्मक पद्धति का ज्यादा प्रयोग किया जाता है, लेकिन प्रयोगात्मक पद्धति की रचना सैद्धांतिक नियमों के आधार पर होती है।

मशीनी अनुवाद के आधार पर संज्ञापद सम्मिलन के लिए प्रस्तुत शोध कार्य में प्रयोगात्मक पद्धति का प्रयोग किया जा रहा है। यह शोध भारतीय भाषा कॉर्पोरा उपक्रम (ILCD) परियोजना के हिंदी से अंग्रेजी भाषा में टैग किए हुए कॉर्पोरा पर आधारित है। इसमें प्रथम हिंदी और अंग्रेजी भाषा की वाक्यसंरचना में भिन्नता देखी जाएगी और विशेष तौर पर, संज्ञापद में अनुवाद के स्तर पर आए बदलाव का त्रुटि विश्लेषण आधारित अध्ययन होगा और हिंदी से अंग्रेजी में मशीनी अनुवाद के लिए स्थानांतरण व्याकरण (Transfer Grammar) के कुछ नियम बनाए जाएंगे। इन नियमों का प्रयोग हिंदी और अंग्रेजी के संज्ञापद सम्मिलन (Mapping) के लिए किया जाएगा। इस सम्मिलन से अनुवाद का स्तर (गुणवत्तापूर्ण, गलत या असंदिग्ध है) ज्ञात हो जाएगा और आगे अनुवाद में सुधार हेतु प्रयास किए जाएंगे।

1.8 निष्कर्ष में यह देखा जा सकता है कि विश्व के साथ भारत में मशीनी अनुवाद के अनेक तंत्रों का निर्माण किया गया है, लेकिन अनुवाद में गुणवत्ता का अभाव रहा है। वास्तव में मशीनी अनुवाद सीमित क्षेत्र के विषय में अनुवाद प्रदान करने के लिए सक्षम है। जैसे सी-डैक पुणे ने मंत्र राजभाषा का निर्माण किया है। वह कार्यालयीन और प्रशासनिक दस्तावेजों का अनुवाद सक्षम रूप में करता है, लेकिन जब साहित्य के अनुवाद की बात आती है तो वहाँ अनुवाद में दुरुहता के साथ गलतियाँ होती हैं और प्राप्त अनुवाद में संदिग्धताएँ उत्पन्न होती हैं। इसके लिए मानव को मशीन की सहायता करनी पड़ेगी ताकि अनुवाद सटीक और गुणवत्तापूर्ण प्राप्त हो। वर्तमान में मशीनी अनुवाद तंत्र विकसित करने के लिए बृहत् स्तर पर प्रयास चल रहा है। इसके लिए भारत सरकार की तरफ से विविध परियोजनाएँ चलायी जा रही हैं। इससे यह संभावना दर्शायी जा सकती है कि भविष्य में प्रत्येक भाषा में मशीनी अनुवाद तंत्र विकसित होंगे जो अनुवाद की कार्यक्षमता के साथ गुणवत्ता भी प्रदान करेंगे ।

अध्याय – द्वितीय

हिंदी और अंग्रेजी की वाक्यसंरचना
और संज्ञापदों का व्यतिरेकी विश्लेषण

अध्याय – द्वितीय

हिंदी और अंग्रेजी की वाक्यसंरचना और संज्ञापदों का व्यतिरेकी विश्लेषण

(Hindi-English Sentence Structure and Contrastive Analysis of Noun Phrase)

प्रस्तुत अध्याय में हिंदी-अंग्रेजी वाक्यसंरचना के साथ वाक्य के प्रकारों को स्पष्ट किया गया है और हिंदी-अंग्रेजी संज्ञापद का अध्ययन कर उनके घटकों की तुलनात्मक चर्चा की गई है, साथ ही संज्ञा पदबंध के प्रकारों को स्पष्ट किया गया है। दूसरे भाग में दोनों भाषाओं के संज्ञापद (Noun Phrase) का कॉर्पोरा आधारित व्यतिरेकी विश्लेषण (Contrastive Analysis) किया गया है। यह कॉर्पोरा हिंदी से अंग्रेजी में भारतीय भाषा कॉर्पोरा उपक्रम (ILCI) के स्वास्थ्य (Health) और पर्यटन (Tourism) क्षेत्र का है। इस टैग किए हुए कॉर्पोरा के वाक्य का प्रयोग व्यतिरेकी विश्लेषण में किया गया है। इस व्यतिरेकी विश्लेषण में हिंदी-अंग्रेजी संज्ञापद में होनेवाले बदलाव पर विशेष महत्व दिया गया है।

भाषा एक संप्रेषण का माध्यम है जो मानव समूह में परस्पर संवाद का संप्रेषण कराती है। यह संप्रेषण का संबंध मानव और मानवेतर प्राणी में भी पाया जाता है। लेकिन मानव अपनी अभिव्यक्ति को मुखोच्चरित ध्वनियों द्वारा व्यक्त करता है। इसलिए कहा जाता है कि भाषा और समाज का संबंध एक दूसरे के साथ अभिन्न है।

प्रत्येक भाषा की संरचना अलग होती है और उस भाषा की संरचना में ध्वनि, व्याकरण, और अर्थ का स्थान महत्वपूर्ण होता है। “*ध्वनि भाषा का भौतिक एवं मूर्त पक्ष है जिसे हम ज्ञानेन्द्रियों के द्वारा ग्रहण करते हैं। व्याकरण भाषा की संरचना का व्यवस्थापरक पक्ष है, और अर्थ भाषा का मानसिक पक्ष है जो मूलतः वक्ता या श्रोता के मस्तिष्क में निहित होता है। ध्वनियाँ स्वयं में अर्थहीन होती हैं, लेकिन व्याकरण व्यवस्था के माध्यम से शब्द तथा वाक्यों का आकार प्राप्त कर, वे अर्थयुक्त हो जाती हैं। इसलिए कहा जाता है कि ध्वनि को अर्थ तक ले जाने वाली कड़ी व्याकरण है।*”¹ इन घटकों के साथ भाषा में वाक्य का भी स्थान महत्वपूर्ण है, क्योंकि वाक्य ही भाषा की संरचना को सार्थक बनाने का कार्य करता है। वाक्य एक या अधिक शब्दों से बनता है, जो संदर्भ विशेष में व्याकरणिक दृष्टि से पूर्ण होता है और वाक्य में विचार की अभिव्यक्ति के साथ अर्थ संरचना भी निहित होती है। हर वाक्य में एक क्रिया अवश्य होती है। इसलिए भाषा की संपूर्ण संरचना में वाक्य की भूमिका प्रमुख होती है।

¹ सिंह, सूरजभान (1985) : हिंदी का वाक्यात्मक व्याकरण, पृ. 20

2.1 हिंदी-अंग्रेजी भाषा की संरचना

हिंदी भाषा के विकास के बारे में कहा जाता है, कि भारत में आर्यों के आने के बाद से भारत में अनेक भाषाओं का विकास हुआ है। वैदिकी, संस्कृत, पालि, प्राकृत, अपभ्रंश, हिंदी, बांग्ला , ओड़िया, असमिया, मराठी, सिन्धी, पंजाबी, आदि भाषाओं का समावेश आर्यभाषा परिवार में किया जाता है।² आज संवैधानिक रूप से हिंदी भारत की राजभाषा है और लिपि देवनागरी है। साथ ही अंग्रेजी भाषा का प्रयोग सह-राजभाषा (द्वितीय भाषा) के रूप में किया जाता है। हिंदी भाषा के बोलने वालों की संख्या भारत में 42 करोड़ 20 लाख (41.3%) (2001 की जनगणना के अनुसार) है। हिंदी का प्रयोग प्रमुख रूप से उत्तर प्रदेश, बिहार, मध्यप्रदेश, हरियाणा, हिमाचल प्रदेश, झारखंड, राजस्थान, उत्तरांचल, छत्तीसगढ़ आदि राज्यों में किया जाता है। इस भाषा की विभिन्न बोलियाँ (Dialects) भी पायी जाती हैं, जिनका प्रयोग अलग-अलग प्रदेशों में किया जाता है।³ इस अनुरूप हिंदी और अंग्रेजी भाषा की संरचना (संज्ञापद के आधार पर) को नीचे टेबल द्वारा देख सकते हैं।

हिंदी	अंग्रेजी
कर्ता, कर्म, क्रिया शब्दक्रम	कर्ता, क्रिया, कर्म शब्दक्रम
परसर्ग का प्रयोग	पूर्वसर्ग का प्रयोग
दो लिंग : पुल्लिंग और स्त्रीलिंग	तीन लिंग : पुल्लिंग, स्त्रीलिंग और नपुंसक लिंग (Neuter)
वचन : एकवचन और बहुवचन [बहुवचन में कर्तृकारक (कमरा), तिर्यक कारक (Oblique case) (कमरे में) का प्रयोग]	वचन : एकवचन और बहुवचन [बहुवचन में संज्ञा के बाद -s, -es, -z, (boy-boys) आदि का प्रयोग]
हिंदी वचन का प्रभाव : संज्ञा (वह लड़का), सर्वनाम (वह-वे, यह-ये), विशेषण (छोटा कमरा-छोटे कमरे) और क्रिया पर	अंग्रेजी वचन का प्रभाव : संज्ञा, पुरुषवाचक सर्वनाम (he-they), निर्देशवाचक सर्वनाम (this-these) और क्रिया पर
संज्ञा के बाद परसर्ग का प्रयोग (सर्वनाम के साथ कारक प्रयोग) (राम ने, उसमें)	संज्ञा के पहले और बाद में पूर्वसर्ग का प्रयोग

² शर्मा, रामकिशोर (2007) : भाषाविज्ञान हिन्दी भाषा और लिपि, पृ. 216

³ Kachru Yamuna (1984) : Hindi, p.1

कर्तृकारक का प्रयोग (ergative)	कर्मकारक का प्रयोग (accusative)
संबंध कारक में (का, की, के) का प्रयोग	संबंध कारक के लिए सिर्फ of का प्रयोग
समुच्चयबोधक का प्रयोग दो या दो से अधिक शब्द या वाक्य को जोड़ने का काम करता है। और, कि, लेकिन, पर, या, तथा, एवं, इसलिए आदि का प्रयोग अधिक होता है।	अंग्रेजी में and, or, but, then, because, so, either or का प्रयोग अधिक किया जाता है।
वाच्य का प्रयोग : कर्तृवाच्य और कर्मवाच्य	वाच्य : कर्तृवाच्य और कर्मवाच्य, भाववाच्य
विशेषण का संज्ञा के रूप में प्रयोग (सुंदर लड़की) क्रिया के रूप में प्रयोग (अच्छा करना)	विशेषण के पहले the उपपद का प्रयोग होता है तो उसके बाद संज्ञा शब्द प्रयुक्त नहीं होता। (The poor)
संज्ञा का (लिंग से) क्रिया पर प्रभाव (सुधा खाना पकाती है/ सुधीर खाना पकाता है)	संज्ञा का क्रिया पर प्रभाव नहीं पड़ता (Sudha cookes food/ Sudhir cooks food)

टेबल 2.1 : हिंदी-अंग्रेजी भाषा की तुलनात्मक संरचना

2.1.1 हिंदी-अंग्रेजी भाषा के वाक्य का स्वरूप

अगर भारतीय भाषाओं की संरचना की बात करें तो इन भाषाओं की संरचना एक दूसरे के समान है। लेकिन हिंदी भाषा की तुलना अंग्रेजी भाषा से करें तो यह ज्ञात होगा कि दोनों भाषाओं की संरचना में भिन्नता पाई जाती है। इस भिन्नता को वाक्य के अंगों के आधार पर समझा जा सकता है। व्याकरणिक इकाई के रूप में वाक्य ऐसी रचना है जिसमें कम-से-कम दो अवयवों या अंगों (Constituents) का होना आवश्यक है। वे दो अंग-उद्देश्य (Subject) और विधेय (Predicate) हैं। उद्देश्य (Subject) वाक्य का वह अंश होता है, जो विषय-वस्तु के बारे में मुख्य जानकारी देता है, यह सूचित करने वाले शब्द को उद्देश्य कहते हैं और विधेय उद्देश्य के विषय में सूचना देनेवाला (गौण) अंश होता है। जैसे -

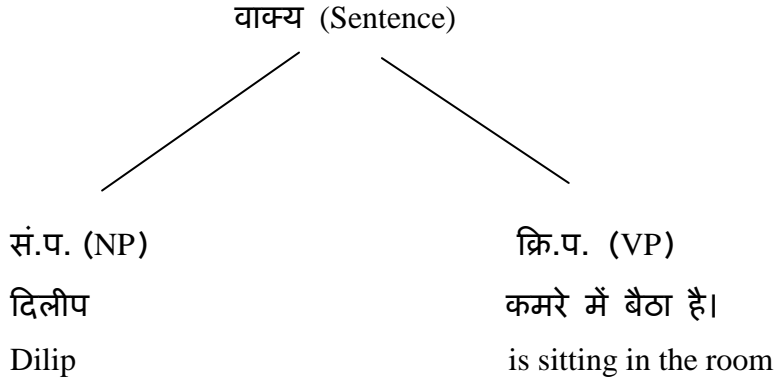
उद्देश्य (Subject)

सभी अतिथि
The bus
आपकी सभी लड़कियाँ

विधेय (Predicate)

भोजन कर रहे हैं।
will not stop here
सुंदर हैं।

उद्देश्य हमेशा संज्ञा पदबंध (Noun Phrase) के रूप में प्रयोग किया जाता है। जिसमें संज्ञा और संज्ञा का विस्तार दोनों शामिल हैं। अंग्रेजी में इसके लिए NP का प्रयोग होता है और विधेय हमेशा क्रिया के रूप में उद्देश्य के बाद प्रयुक्त होता है जिसके लिए VP का प्रयोग होता है। जिसमें मोटे रूप से क्रिया, कर्म, पूरक, क्रिया विशेषण आदि घटक शामिल होते हैं।



उपर्युक्त हिंदी वाक्य में 'दिलीप' उद्देश्य है और 'कमरे में बैठा है' जो 'कहाँ बैठा है' की जानकारी दे रहा है वह विधेय है। अंग्रेजी वाक्य में सिर्फ 'दिलीप' का ही समावेश उद्देश्य में होता है जिसे संज्ञापद कहते हैं और अन्य वाक्य का समावेश विधेय के रूप में किया जाता है जिसे क्रियापद कहते हैं।

2.1.2 वाक्य संरचना के आधार

वाक्य की संरचना में जिस तरह कर्ता, क्रिया, कर्म को वाक्य के प्रमुख अवयव माना जाता है, उसी तरह वाक्यों में कुछ आंतरिक घटक होते हैं, जिनका विस्तृत वर्णन विभिन्न विद्वानों यथा, राजमणि शर्मा और सूरजभान सिंह ने अपनी पुस्तकों⁴ में किया है। वाक्य में जिस तरह शब्दों और पदबंधों का समावेश होता है, उसी तरह वाक्य में विभिन्न घटकों के बीच एक प्रकार की क्रमबद्धता (Sequence) होती है, जिसे हम पदक्रम या शब्दक्रम (Word Order) कहते हैं।

2.1.2.1 शब्दक्रम या पदक्रम (Word order):

अनेक भाषाओं की वाक्यात्मक संरचना में पद का स्थान महत्वपूर्ण है। पदक्रम से तात्पर्य वाक्य में प्रयुक्त पदों या पदबंधों के उस क्रम से है, जिस क्रम में वे वाक्य में सहज रूप से प्रयुक्त होते हैं। इसके अंतर्गत मोटे रूप से कर्ता, कर्म, पूरक (complement), क्रियाविशेषण (adverbs) और क्रिया (verb) आदि पदबंधों के प्रयोग-स्थान शामिल हैं। पदक्रम के लिए कभी-

⁴ शर्मा, राजमणि (2010) : हिन्दी भाषा का इतिहास और स्वरूप, पृ. 292

सिंह, सूरजभान (2003) : अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ.197

कभी शब्दक्रम का प्रयोग होता है। हिंदी तथा अधिकांश भारतीय भाषाओं में वाक्यों का पदक्रम कर्ता-कर्म-क्रिया (SOV) है, लेकिन अंग्रेजी का पदक्रम कर्ता-क्रिया-कर्म (SVO) होने के साथ ही इस भाषा का पदक्रम स्थिर रहता है।

जैसे – सुशील घर जाता है —————> कर्ता, कर्म, क्रिया
Sushil is going to home —————> Subject, verb, object

अंग्रेजी भाषा की तुलना में हिंदी भाषा का पदक्रम अस्थिर है, लेकिन संस्कृत, रूसी आदि भाषाओं का पदक्रम स्वतंत्र है। संस्कृत विभक्ति प्रधान भाषा है, इसलिए इस भाषा में पदक्रम का कोई कठोर नियम नहीं है। (वर्मा, 2007, पृ.150) अंग्रेजी वाक्य संरचना में पदक्रम का व्याकरणिक महत्व होता है और उनमें परिवर्तन बहुत कम वाक्य संरचना में होता है। (*I purchased these boots from Gole market*), लेकिन हिंदी में पदक्रम का व्याकरणिक संरचना के साथ परिवर्तन होता है। जैसे- 1. मैंने ये जूते गोल मार्केट से खरीदे। 2. ये जूते मैंने गोल मार्केट से खरीदे। 3. गोल मार्केट से मैंने खरीदे ये जूते। अंग्रेजी में वाक्य के घटकों के स्थान परिवर्तन के लिए कुछ नियम होते हैं। उदाहरण के लिए, प्रश्नवाचक वाक्यों में सहायक क्रियाएँ अनिवार्य रूप से वाक्य के प्रारंभ में स्थानांतरित हो जाती हैं। जैसे – *Are you Mr. Jones?* लेकिन हिंदी में इस तरह वाक्य के पदक्रम में और भी रूप मिल सकते हैं।

2.1.2.2 अन्विति

अन्वय (Agreement) का अर्थ है- व्याकरणिक संरचना में एकरूपता। विभिन्न भाषाओं में विशेषण-विशेष्य, कर्ता-क्रिया, कर्म-क्रिया आदि विभिन्न व्याकरणिक कोटियों में लिंग, वचन, पुरुष तथा मूल और विकृत रूप की अनुरूपता होती है। हर भाषा के अन्वय के नियम अलग-अलग होते हैं। संस्कृत में कर्ता-क्रिया में लिंग का अन्वय नहीं है, लेकिन हिन्दी में होता है।

रामः पठति – राम पढ़ता है।

सीता पठति – सीता पढ़ती है।

ऐसे ही हिन्दी के विशेषण में भी वचन तथा लिंग का अन्वय है किन्तु अंग्रेजी में नहीं –

अच्छा छात्र – *Good Student*, अच्छी छात्रा – *Good Student*

अच्छे छात्र – *Good Students*, अच्छी छात्राएँ – *Good Students*

हिन्दी की क्रिया, कभी कर्ता के अनुरूप होती है – ‘राम गया’ ‘नीलिमा गई’ तो कभी कर्म के अनुरूप – ‘राम ने रोटी खाई’ ‘नीलिमा ने आम खाया’ आदि ।

हिंदी में अन्विति का क्षेत्र अंग्रेजी से अधिक व्यापक है। हिंदी में अन्विति – वचन, पुरुष और लिंग तीनों स्तरों पर मिलती है। इसके अलावा अंग्रेजी में अन्विति कर्ता और क्रिया के बीच ही होती है, लेकिन हिंदी में अन्विति का प्रभाव संज्ञा, विशेषण, और क्रिया पर होता है।

अन्वय की प्रमुख स्थितियाँ, कर्ता-क्रिया का अन्वय, कर्म और क्रिया का अन्वय, कर्ता कर्म से निरपेक्ष क्रिया, विशेषण और विशेष्य का अन्वय, संबंध और संबंधी का अन्वय, सर्वनाम और संज्ञा का अन्वय आदि के रूप में पायी जाती है। (तिवारी, 2005, पृ. 264)

2.1.2.3 कारक संबंध (Case relationship)

वाक्य के विभिन्न पदबंध आपस में कुछ व्याकरणिक संबंधों से (Grammatical relations) जुड़े होते हैं, जिससे उनका सही अर्थ बोध प्रतीत होता है। क्रिया और संज्ञा पदबंधों के बीच यह व्याकरणिक संबंध कारक कहलाता है। वाक्य में कारक संबंध का विशेष महत्त्व है। वाक्य में कारक की सूचना दो प्रकार से मिलती है – कारक विभक्तियों से (जैसे- *ने, को, में, से, के लिए* आदि) और पद के प्रयोग-स्थान से। हिंदी व्याकरण में कारक के 6 प्रकार हैं, *कर्ता (ने), कर्म(को), करण (से,के द्वारा), संप्रदान (को, के लिए), अपादान (से), अधिकरण (में,पर)*। उसी तरह अंग्रेजी की संरचना में भी कारक चिह्नों (*of, from, at, with, on, about, in, for*) का प्रयोग होता है। वाक्य विश्लेषण के संदर्भ में कारक संकल्पना (concept) में आज काफ़ी परिवर्तन आया है। फिल्मोर ने अलग से एक कारक व्याकरण (1968) की रचना की। उनके अनुसार, वाक्य एक क्रिया तथा एक से अधिक संज्ञा पदबंधों की ऐसी रचना है, जिसमें प्रत्येक संज्ञा, क्रिया के साथ, एक विशेष कारक संबंध के साथ जुड़ी रहती है।

2.1.2.4 अर्थ-संगति (Semantic concordance)

व्याकरण और अन्विति की दृष्टि से सही होते हुए भी वाक्य अर्थ-संगति (semantic concordance) की दृष्टि से दोषपूर्ण हो सकता है जैसे- **कल मैंने दूसरे के विचारों को खाया।*

उक्त वाक्य अर्थ की दृष्टि से असंगत है। इस वाक्य से शब्द जहाँ एक व्याकरणिक रचना है वहीं अर्थ के स्तर पर किसी वस्तु या व्यापार (action) का प्रतीक (symbol) भी है। जिस प्रकार वाक्य में व्याकरण की रचना संबंधी अपनी कुछ आकांक्षाएँ होती हैं उसी तरह वस्तु कार्य-व्यापार संबंधी अपनी कुछ आकांक्षाएँ हो सकती हैं। उपर्युक्त वाक्य में *खाना* कार्य व्यापार (action) का कर्म कोई खाद्य पदार्थ हो सकता है, विचार नहीं। इसी तरह वाक्य के हर घटक को अर्थ की दृष्टि से भी अन्य घटकों के अनुकूल होना चाहिए, अर्थात् उनके बीच अर्थ-संगति होनी चाहिए।

2.1.2.5 अध्याहार या लोप (Ellipsis)

सामान्यतः वाक्य के गठन या निर्मिति के समय में बहुत से शब्द अनावश्यक मानकर छोड़ दिए जाते हैं। वास्तविक जीवन में अनेक बार वाक्य के कुछ अंशों का लोप कर दिया जाता है और कभी-कभी केवल एक-दो शब्दों का ही प्रयोग किया जाता है। फिर भी अर्थ-बोध (comprehension) होता है। भाषा के इस लक्षण को अध्याहार कहते हैं। जैसे- *नमस्कार* (*good morning*), *कौन?* (*Who is there?*) आदि को अध्याहार या अल्पांग वाक्य कह सकते हैं। कर्ता, क्रिया वाक्यांश का अध्याहार होता है। हर भाषा में लोप के अलग-अलग नियम होते हैं एक ही भाषा में शब्दों का कहीं लोप होता है तो किसी वाक्य में लोप नहीं होता।

इस वाक्य संरचना के आधार पर वाक्य की अंतर्भूत संरचना को ज्ञात किया गया है। वाक्य केवल शब्द का जाल नहीं है, उसमें व्याकरणिक घटक के साथ अर्थ की संरचना का भी समावेश होना आवश्यक है, तभी वाक्य अर्थ की दृष्टि से सार्थक हो सकता है। इसलिए वाक्य संरचना के आधार को यहाँ स्पष्ट किया गया है।

2.1.3 वाक्य के प्रकार

सामान्यतः वाक्य के तीन प्रकारों को दो आधारों पर विभाजित किया गया है। अ) संरचना के आधार पर और ब) अर्थसंरचना के आधार पर। संरचनागत प्रकारों में से वाक्य के तीन प्रमुख प्रकार- सरल वाक्य, संयुक्त वाक्य, और मिश्र वाक्य हैं। इस प्रकारों को नीचे देखा जा सकता है-⁵

2.1.3.1 सरल वाक्य (Simple Sentence)

वाक्य रचना की मूल इकाई सरल वाक्य है। जिस वाक्य में एक संज्ञा उद्देश्य (कर्ता) और क्रिया एक विधेय होती है, उस वाक्य को सरल वाक्य (Simple Sentence) तथा साधारण वाक्य कहते हैं। जैसे- *'सुहास रोता है'* इसमें *'सुहास'* उद्देश्य है और *'रोता है'* विधेय है। कभी-कभी वाक्य में एक से अधिक उद्देश्य और विधेय का प्रयोग होता है पर उस वाक्य की संरचना संयुक्त या मिश्रित नहीं होती, जैसे- *'सुहास हँसता-खेलता है'*, *'रजत गाता-बजाता है'* ये वाक्य भी सरल वाक्य के अंतर्गत आते हैं। वाक्य का आंतरिक विश्लेषण सरल वाक्य के आधार पर ही किया जाता है। किसी भी भाषा के वाक्य सँचों (Sentence patterns) का निर्धारण भी सरल वाक्यों के आधार पर ही होता है। (सिंह, 2003, पृ.205) सरल वाक्यों के उद्देश्य और विधेय का कभी-कभी विस्तार भी हो सकता है, तथा इनके प्रश्नवाचक, निषेधात्मक तथा प्रश्नवाचक-निषेधात्मक रूपान्तर भी हो सकते हैं, जैसे- *'क्या श्याम रोता है?'*, *'श्याम नहीं रोता'*, *'क्या गौरव नहीं रोता'*

⁵ तिवारी, भोलानाथ (2005) : हिन्दी भाषा, पृ. 256-258

आदि वाक्यों में। (तिवारी, 2005; पृ.256) साथ ही इस सरल वाक्य में कर्ता के साथ उसके विस्तारक विशेषण और क्रिया के साथ विस्तारक कर्म एवं क्रिया-विशेषण का प्रयोग हो सकता है। जैसे- 'अच्छा बच्चा मीठा दूध अच्छी तरह पीता है।' इसका भी समावेश सरल वाक्य में किया जाता है।

2.1.3.2 उपवाक्य

उपवाक्य को भी प्रकार के रूप में भोलानाथ तिवारी और राजमणि शर्मा ने अपनी पुस्तक⁶ में विश्लेषित किया है और इस वाक्य को संयुक्त वाक्य और मिश्र वाक्य से अलग माना है। जब दो या दो से अधिक सरल वाक्यों को मिलाकर एक वाक्य बना देते हैं तो उस एक वाक्य में जो अन्य वाक्य मिले होते हैं, उन्हें उपवाक्य कहते हैं। वस्तुतः इस वाक्य के मिलान से बने बड़े वाक्य को जब 'वाक्य' कहते हैं तो उसके भीतर के छोटे वाक्यों को 'उपवाक्य' कहते हैं। उदाहरण के लिए 'श्याम आया और घनश्याम गया।' वाक्य में दो उपवाक्य 'श्याम आया' 'घनश्याम गया' हैं।

उपवाक्य के भी दो प्रकार हैं -

- क) आश्रित उपवाक्य- जो गौण तथा दूसरे पर आश्रित होता है उसे आश्रित उपवाक्य कहते हैं। जैसे- 'वह लड़का चला गया जो सबसे अच्छा था' इसमें 'जो सबसे अच्छा था' आश्रित उपवाक्य है।
- ख) प्रधान उपवाक्य- जो वाक्य किसी दूसरे पर आश्रित नहीं होता है, उसे प्रधान उपवाक्य कहते हैं। उपर्युक्त उदाहरण में 'वह लड़का चला गया' प्रधान वाक्य है।

2.1.3.3 संयुक्त वाक्य (Compound Sentence)

जब एक से अधिक सरल वाक्यों को मिलाकर एक वाक्य बनाया जाता है और इन दो या अधिक वाक्यों को समानाधिकरण समुच्चयबोधक (Conjunction) (जैसे पर, किन्तु, और, लेकिन या, and, but, or आदि) द्वारा जोड़ा जाता है, तो उस वाक्य को संयुक्त वाक्य कहते हैं। संयुक्त वाक्य में एक से अधिक प्रधान उपवाक्य रहते हैं और इन प्रधान उपवाक्यों के साथ इनके आश्रित उपवाक्य भी रहते हैं। लेकिन कोई भी वाक्य एक-दूसरे पर आश्रित नहीं रहते, इससे यह न समझ लेना चाहिए कि उनमें परस्पर आश्रय कुछ भी नहीं होता। इसका कारण है कि वाक्य में अर्थ संगति की दृष्टि से इस तरह का प्रयोग होता है।⁷

⁶ शर्मा, राजमणि (2010) : हिन्दी भाषा का इतिहास और स्वरूप, पृ. 298

⁷ गुरु कामताप्रसाद, (1920) : हिंदी व्याकरण पृ. 433

पिताजी दुकान में हैं और माताजी खाना बना रही हैं।

Father is in the shop and mother is cooking food.

उक्त वाक्य में 'पिताजी दुकान में हैं' तथा 'माताजी खाना बना रही हैं।' दोनों सरल वाक्य हैं जिन्हें समुच्चयबोधक द्वारा जोड़ा गया है। इस दोनों वाक्यों में कोई भी मुख्य या गौण वाक्य नहीं है, दोनों का संबंध समानस्तरीय है। संयुक्त वाक्य एक दूसरे पर आश्रित न होते हुए भी संयुक्त वाक्य के दो उपवाक्य आपस में किसी-न-किसी प्रकार के संबंध से जुड़े होते हैं। ये संबंध चार प्रकार के होते हैं। (सिंह, 2003, पृ. 206)

क) संयोजक संबंध (Coordinative) - जब एक साधारण वाक्य दूसरे साधारण या मिश्र वाक्य से संयोजक अव्यय द्वारा जुड़ा होता है। जैसे- *पेन हाथ से गिरा और टूट गया।* संयोजक संबंध सूचित करने वाले शब्द हैं : *और, तथा, एवं, फिर, बल्कि, and, then, not only..but also,* आदि। उदाहरण के लिए - *सचिन गया और मनीष आया।*

ख) विभाजक संबंध (Alternative) - जब संयुक्त वाक्य के दो उपवाक्य में दो घटनाओं, स्थितियों, या कार्य-व्यापारों में से किसी एक के ग्रहण या दोनों के त्याग की सूचना देते हैं तो उनके बीच विभाजक संबंध माना जाता है। इसके लिए *या, अथवा, या..या, न..न, नहीं तो (अन्यथा), चाहे..चाहे, न कि, or, either..or, neither..nor, or else, otherwise, and not* .आदि शब्दों का दोनों भाषाओं में प्रयोग किया जाता है। उदाहरण के लिए- *चाहे तुम यहाँ रहो चाहे होटल में।*

ग) विरोधवाची संबंध (Adversive) - जब संयुक्त वाक्य के दो उपवाक्य दो घटनाओं, स्थितियों, या कार्य-व्यापारों के बीच विरोध (Contrast) या विरोधाभास (contrastive) की सूचना दें तो उनके बीच विरोधवाची संबंध माना जाता है। उसमें प्रमुख विरोधवाची शब्द: *लेकिन (किंतु, परंतु, मगर, पर), बल्कि, प्रत्युत, but, whereas, even, though, much as.* आदि का प्रयोग होता है। उदाहरण के लिए- *वह मेहनत तो कुछ करता नहीं पर फल की आशा करता है।*

घ) परिणामवाची संबंध (Resultative)

जब एक साधारण वाक्य दूसरे साधारण या मिश्रित वाक्य का परिणाम होता है, तो उसे परिणामवाची संबंध कहते हैं। इसमें *इसलिए, अतः, और, so, therefore, and* आदि का प्रयोग समुच्चयबोधक शब्द के रूप में होता है। जैसे- *आज मुझे बहुत काम है इसलिए आज घर नहीं आऊँगा।*

2.1.3.4 मिश्र वाक्य (Complex Sentence)

जब एक से अधिक सरल वाक्यों को इस प्रकार जोड़ा जाता है कि उसमें एक प्रधान तथा शेष आश्रित उपवाक्य हों तो इस वाक्यात्मक संरचना को मिश्र वाक्य कहते हैं। 'जो सज्जन होता है, उसका सभी लोग आदर करते हैं।' इस वाक्य में 'उसका सभी लोग आदर करते हैं' प्रधान उपवाक्य है तथा 'जो सज्जन होता है' आश्रित उपवाक्य है। (शर्मा, 2010; पृ. 208) प्रधान उपवाक्य और आश्रित उपवाक्य के बीच परस्पर आश्रय-आश्रित का संबंध होता है। प्रधान उपवाक्य वाक्य का मुख्य कथन होता है और आश्रित उपवाक्य का आश्रय होता है। आश्रित उपवाक्य कुछ समुच्चयबोधक शब्दों (Conjunction) का (जो, जहाँ, कि, ताकि, who, that, where, which, so that आदि) द्वारा मुख्य उपवाक्य से जुड़े होते हैं, जो अक्सर इन उपवाक्यों के शुरू में जुड़ते हैं। जैसे—

मैंने कल एक लड़का देखा जो बिल्कुल आपकी तरह था।

I saw a boy yesterday who was like you.

मुख्य उपवाक्य बिना आश्रित उपवाक्य के भी प्रयुक्त होने की क्षमता रखता है, लेकिन आश्रित उपवाक्य अपने अर्थ की पूर्णता के लिए, मुख्य उपवाक्य पर आश्रित रहता है और स्वतंत्र रूप से प्रयुक्त नहीं हो सकता। इसलिए मिश्र वाक्य में प्रधान उपवाक्य एक ही रहता है, पर आश्रित उपवाक्य एक से अधिक रह सकते हैं। आश्रित उपवाक्य तीन प्रकार के होते हैं।⁸ संज्ञा उपवाक्य, विशेषण उपवाक्य, और क्रियाविशेषण उपवाक्य।

क) संज्ञा उपवाक्य (Noun Clause) - जब आश्रित उपवाक्य किसी संज्ञा अथवा सर्वनाम के स्थान पर आता है तब उसे संज्ञा उपवाक्य कहते हैं। जैसे—

रमेश ने बताया कि माँ बहुत बीमार हैं।

Ramesh told that mother is very sick

'माँ बहुत बीमार हैं' आश्रित उपवाक्य है और यह प्रधान उपवाक्य 'रमेश ने बताया' के प्रयोग में हुआ है।

ख) विशेषण उपवाक्य (Adjective Clause) - जो आश्रित उपवाक्य मुख्य उपवाक्य की संज्ञा शब्द की विशेषता बतलाता है वह विशेषण उपवाक्य कहलाता है। जैसे—

यहाँ जो घड़ी मेज पर रखी है, वह मुझे पुरस्कारस्वरूप मिली है।

The watch which is here kept on the table, I got it as a prize.

⁸ सिंह सूरजभान, (2003): अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ. 207

उपर्युक्त वाक्य में 'यहाँ जो घड़ी मेज पर रखी है' वह विशेषण उपवाक्य है।

ग) क्रिया विशेषण उपवाक्य- (Adverbial Clause) जब आश्रित उपवाक्य प्रधान उपवाक्य की क्रिया की विशेषता बताते हैं, तो उन्हें क्रिया विशेषण उपवाक्य कहते हैं। जैसे-

जब वह मेरे पास आया तब मैं सो रहा था।

When he came to me , I was sleeping.

'जब वह मेरे पास आया' यह क्रिया विशेषण उपवाक्य है।

2.2 हिंदी संज्ञा पदबंध (Hindi Noun Phrase)

जब एक से अधिक शब्द बँधकर एक व्याकरणिक इकाई (संज्ञा, क्रिया, विशेषण) का कार्य करे तो उस पद-समूह को पदबंध कहते हैं। और जो पदबंध किसी संज्ञा का काम करे, उसे संज्ञा पदबंध कहते हैं। जैसे- 'दिनभर काम कर रात में पढाई करनेवाला विद्यार्थी कैसे नहीं उत्तीर्ण होगा'। इस वाक्य में 'दिनभर काम कर रात में पढाई करनेवाला' यह पदों का समूह है तथा 'संज्ञा' का काम कर रहा है, अंतः यह 'संज्ञा पदबंध' है। हिंदी संज्ञा पदबंध में संज्ञा एक अनिवार्य घटक होता है जिसे पदबंध का शीर्ष (Head) कहा जाता है। संज्ञा के साथ वाक्य में सर्वनाम निर्धारक, और विशेषण का प्रयोग होता है। हिंदी अंग्रेजी दोनों भाषाओं में भिन्नता के अनुरूप अंग्रेजी में विशेषक (Modifier) संज्ञा से पहले और बाद में, तथा दोनों स्थानों पर प्रयुक्त होते हैं, लेकिन हिंदी में ये विशेषक संज्ञा से पहले प्रयुक्त होते हैं, बाद में नहीं। साथ ही विशेष रूप से हिंदी में संज्ञा पर लिंग, वचन, और कारक का प्रभाव हमेशा रहता है। इस दृष्टि से यहाँ लिंग वचन और कारक को विश्लेषित किया गया है।⁹

क) लिंग : हिंदी में पुल्लिंग और स्त्रीलिंग ये दो लिंग हैं। पुरुष जाति का जहाँ बोध होता है उसे पुल्लिंग कहते हैं, जैसे- लड़का, नर आदि। स्त्री जाति का जहाँ बोध होता है उसे स्त्रीलिंग कहते हैं। लड़की, नारी ये शब्द स्त्री जाति का बोध कराते हैं। संज्ञा में प्रत्ययों के प्रयोग से लिंग बदल जाता है। सामान्यतः पुल्लिंग में '-आ, -आव, -पा, -पन, -न' प्रत्ययों का प्रयोग शब्द के अंत में होता है। जैसे- मोटा, चढाव, बुढापा, लडकपन, लेन-देन आदि शब्दों में प्रत्ययों का समावेश अतंनिहित है। अन्य कई प्रकार के शब्दों का भी समावेश पुल्लिंग में किया जाता है। उसी तरह पुल्लिंग से स्त्रीलिंग बनाने के लिए स्त्री प्रत्यय '-ई, -इया, -इन, -नी, -आनी, -आइन, -आ, -नी' (घोड़ी, चिड़िया, धोबिन, मोरनी, नौकरानी, पंडिताइन, वाला आदि शब्दों में) का प्रयोग होता है।

⁹ Koul N Omkar, (2008) : Modern Hindi Grammar, p.33-35

ख) वचन : हिंदी में वचन के एकवचन और बहुवचन दो प्रकार हैं। शब्दों के जिस रूप से एक ही वस्तु का बोध हो, उसे एकवचन कहते हैं, जैसे- *बच्चा, पुस्तक* । और शब्दों के जिस रूप से अनेकता का बोध होता हो, उसे बहुवचन कहते हैं। जैसे- *कपड़े, टोपियाँ, गुरुजन* आदि। संज्ञाओं का एकवचन से बहुवचन में रूपांतर करने के लिए प्रत्ययों का प्रयोग होता है। हिंदी में पुल्लिंग रूप ‘-आ’ का रूपांतर बहुवचन ‘-ए’ में होता है, जैसे- *लड़का-लड़के* । वचन के दो प्रकारों के प्रत्ययों का प्रयोग टेबल में दिया गया है।

प्रत्यय	एकवचन	प्रत्यय	बहुवचन
-अ	पुस्तक	-एँ	पुस्तकें
-आ	घोड़ा	-ए	घोड़े
-आ	कला	-एँ	कलाएँ
-आ	गुड़िया	-आँ	गुड़ियाँ
-उ	धातु	-एँ	धातुएँ
-ऊ	वधू	-एँ	वधुएँ

टेबल 2.2 : एकवचन और बहुवचन के प्रत्यय

कुछ शब्दों के रूप ‘एकवचन’ और ‘बहुवचन’ में समान होते हैं, जैसे- *जल-जल, प्रेम-प्रेम* । जब संज्ञाओं के साथ परसर्ग (*ने, से, को* आदि) का प्रयोग होता है, तो संज्ञाओं का बहुवचन बनाने के लिए ‘-ओं’ का प्रयोग होता है। जैसे- *लड़के को बुलाओ/लड़कों को बुलाओ* । परसर्ग के साथ प्रयुक्त होनवाले रूप को तिर्यक् या विकारी रूप कहते हैं। संबोधन में भी ‘ओ’ जोड़कर बहुवचन बनाया जाता है। जैसे- *बच्चो ! ध्यान से सुनो* ।

ग) कारक : हिंदी में वाक्यात्मक और अर्थात्मक दृष्टि से तीन प्रकार के कारक पाए जाते हैं। क) विकारी रूप (Oblique case) - जिसके साथ कारक चिन्ह (परसर्ग) का प्रयोग होता है। जैसे- *लड़की ने फूल तोड़ा*। ख) अविकारी रूप (Direct case) - जिस शब्द के साथ कारक चिन्ह का प्रयोग नहीं होता, जैसे- *श्याम गया*। ग) संबोधन कारक (Vocative case) - जिसका प्रयोग किसी को संबोधित करने के लिए किया जाता है। जैसे- *‘ओ सोहन’* आदि।

रूप-रचना की दृष्टि से कारक चिन्ह को परसर्ग कहते हैं। इसमें भी कारक के प्रकार पाए जाते हैं- कर्ताकारक में (*ने*) का प्रयोग, कर्मकारक में (*को*), करण कारक में (*से, के साथ, के द्वारा*),

संप्रदान में (के लिए, को), अपादान में (से पृथक), संबंध कारक (का, के, की), अधिकरण कारक में (में, पर), संबोधन में (हे! अरे!) आदि कारक चिह्नों का प्रयोग होता है। ये लिंग, वचन, कारक संज्ञा को प्रभावित करनेवाले घटक के रूप में शामिल हैं। जिन्हें यहाँ विस्तृत रूप से स्पष्ट किया गया है। (तिवारी, 2005, पृ.139)

संज्ञा पदबंध में निश्चित ही कुछ घटकों का होना अनिवार्य होता है जिनके बिना संज्ञा की विशेषता ज्ञात नहीं होती। इन घटकों के बारे में सूरजभान सिंह (2003) ने अपनी पुस्तक में चर्चा की है।¹⁰ उनके अनुसार हिंदी संज्ञा पदबंध के प्रमुख घटकों में निर्धारक, पूर्व-विशेषक, शीर्ष, और पश्च विशेषक का समावेश होता है।

2.2.1 निर्धारक (Determiners)

हिंदी संज्ञा पदबंध में पहला स्थान निर्धारकों का होता है। लेकिन हिंदी में निर्धारक नाम नहीं दिया जाता तथा निश्चय या अनिश्चयवाचक आर्टिकल (*the, a/ an*) का प्रयोग नहीं होता है। हिंदी में अंग्रेजी के 'a' के लिए किसी शब्द का प्रयोग नहीं होता लेकिन कुछ स्थितियों में 'एक' का प्रयोग होता है। जैसे- *a boy is standing outside.* (बाहर एक लड़का खड़ा है।) यहाँ 'एक' अनिश्चयवाचक है। इसी प्रकार अंग्रेजी के 'the' के लिए भी किसी भी शब्द का प्रयोग नहीं होता, पर कुछ स्थितियों में सर्वनाम 'वह' का प्रयोग होता है। अगर हिंदी में संज्ञा के पहले किसी निर्धारक का प्रयोग नहीं होता हो तो, वह मूल रूप से निश्चित संज्ञाएँ मानी जाती है। जैसे- *लड़की बीमार है।* हिंदी में निर्धारक के रूप में 'सभी, दोनों, एक, दो, वे, कुछ, कई, पहला, थोड़ा, बहुत' आदि का प्रयोग होता है। अंग्रेजी में इनका प्रयोग निर्धारक की कोटि में रखा जाता है पर हिंदी में इन्हें सर्वनाम या विशेषण के कोटि में रखा जाता है। संज्ञा की अनिश्चित स्थिति को प्रकट करने के लिए हिंदी में प्रश्नवाचकों के अतिरिक्त संज्ञा के पूर्व अनिश्चित निर्धारक प्रयुक्त होते हैं। अनिश्चित निर्धारक 'एक' के अलावा अनिश्चित सर्वनाम और अनिश्चित परिमाणक आते हैं। अनिश्चित सर्वनाम- *कोई, कुछ* आदि का प्रयोग निर्धारक के रूप में अधिक होता है।

2.2.2 पूर्व विशेषक (Pre-modifiers)

हिंदी संज्ञा पदबंध में विशेषकों का प्रयोग संज्ञा से पहले होता है, जैसे - गरम चाय, पुराना मकान, टूटी कुरसी। हिंदी में संज्ञा के विशेषक कई तरह के हो सकते हैं, जैसे- गुणवाचक (सुंदर, कठिन, सबसे बड़ा, छोटे से छोटा), कृदंत (बहता (पानी), संबंधित, विचारणीय), -वाला (चायवाला, हरावाला, जानेवाली (लड़की)), संख्या या परिमाण (एक, दस, दोनों, पहला, तिगुना,

¹⁰ सिंह, सूरजभान (2003): अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ.88, 85

कुछ (पैसे), थोड़ा (आराम)), का-विशेषक (लड़के के, उसका, जाने का (मौका)), संज्ञा विशेषक (महिला (डॉक्टर)), सार्वनामिक, आदरार्थ उपाधि आदि पू. वि. शब्दों का समावेश हिंदी में किया जाता है। वाक्य में एक से अधिक विशेषकों वाले संज्ञा पदबंधों में शब्दों का क्रम नियम इस प्रकार निर्धारित किया जाता है।

[का-विशेषक+संकेतवाचक+संख्यावाचक+गुणवाचक+संज्ञा]

जैसे- *आपकी ये चारों सफ़ेद कमीजें।* इस वाक्य को 'ये आपकी चारों सफ़ेद कमीजें।' लिख सकते हैं, पर इस वाक्य में क्रम विपर्यय स्पष्ट होता है।

2.2.3 पश्च विशेषक (Post modifiers)

हिंदी में संज्ञा के बाद पश्च विशेषक का प्रयोग नहीं होता, केवल पूर्ण विशेषण उपवाक्य के रूप में ही इनका प्रयोग होता है। हिंदी के कुछ संज्ञा पदबंधों में संज्ञा (शीर्ष) के बाद विशेषणों का शैलीगत या मुहावरेदार प्रयोग मिलता है। जैसे- *चाय गरमा-गरम, चाट मसालेदार।* इसके विपरीत अंग्रेजी में पश्च विशेषक के प्रयोग की प्रवृत्ति अधिक प्रबल है। अनुवाद की दृष्टि में पश्च विशेषकों के अनेक भेद हिंदी से अधिक अंग्रेजी में पाए जाते हैं। जैसे- *विदेश में रह रहे भारतीय (Indians abroad)।*

उक्त घटकों के आधार पर हिंदी संज्ञापद संरचना के अनुरूप संज्ञापद नियम बनाए गए हैं।

हिंदी संज्ञा. संरचना नियम= सामान्य शब्द + निर्धा. + स. (विशे.पद.) + स. (संबंध. उप.)

उदाहरण - शहर का एक अच्छा आदमी जो विदेश में रहता है।

2.3 अंग्रेजी संज्ञा पदबंध (English Noun Phrase)

अंग्रेजी संज्ञा पदबंध में भी चार घटक पाए जाते हैं जो संज्ञा की संरचना में महत्वपूर्ण भूमिका निभाते हैं। अंग्रेजी संज्ञा पदबंध के घटक और हिंदी संज्ञापद के घटकों में भिन्नता है। वह निम्न रूप से स्पष्ट होगी।

2.3.1 निर्धारक (Determiners)

अंग्रेजी सं.प. में निर्धारक का काम एक संज्ञा के अर्थ को निश्चित (definite), या अनिश्चित (indefinite) आयाम प्रदान करने और उसके परिणाम (quantity) आदि का बोध कराने

के लिए होता है। अंग्रेजी संज्ञा पदबंध में सबसे पहला स्थान निर्धारक (determiner) का होता है। यदि पदबंध में एक से अधिक निर्धारक एक साथ प्रयुक्त होते हैं तो उनके प्रयोग-स्थानों का नियत क्रम अलग-अलग प्रकार का होता है, वह है- प्रथम स्थान- *all, both, half, double, twice, one third, such, what*; मध्य स्थान - Article : *the, a, an*; Dem : *these, this, that*; Poss.: *My, your, his, her, their*; अंतिम स्थान - Numerals: *one, two, first, second, little, several, Quantifiers: more, less* आदि। जैसे - *My first three choices are these.*
ये निर्धारक संज्ञापद को निश्चित क्रम प्रदान करते हैं।

2.3.2 पूर्व विशेषक (Pre-modifiers)

अंग्रेजी में पूर्व विशेषक का स्थान वाक्य में निर्धारक के बाद आता है। अंग्रेजी संज्ञा पदबंध में पूर्व विशेषक तीन रूपों में व्यक्त होता है। विशेषण के रूप में, जैसे- *a serious mistake, an early riser, criminal neglect*. दूसरा कृदंत के रूप में, (Participle) *a born orator, married life, trembling hands*. तीसरा संज्ञा के रूप में (Noun) *a lady doctor, city life, goods train*.

रचना की दृष्टि से पूर्व विशेषक एक विशेषक का हो सकता है और एक से अधिक विशेषक के समूह का भी हो सकता है। जब एक से अधिक विशेषणों को मिलाकर पूर्व विशेषकों का विस्तार किया जाता है तो सामान्य रूप दो विशेषणों या संज्ञाओं को हाइफन से जोड़कर एक मिश्र विशेषक का रूप दे दिया जाता है, जैसे- *good-looking girl*। इसी तरह अन्य विशेषणों का निर्माण किया जाता है। शीर्ष घटक का समावेश संज्ञा, सर्वनाम और विशेषण में निहित है।

2.3.3 पश्च विशेषक (Post modifiers)

अंग्रेजी में संज्ञा से पहले और बाद में भी विशेषकों का प्रयोग होता है। संज्ञा के बाद प्रयुक्त होने वाले विशेषकों को पश्च विशेषक (Post modifiers) कहा जाता है। हिंदी में सामान्यतः केवल संज्ञा के पहले ही विशेषणों/विशेषकों का प्रयोग होता है, संज्ञा के बाद नहीं। लेकिन अंग्रेजी में संज्ञा के बाद विशेषक कई रूपों में प्रयुक्त होते हैं। विशेषण के रूप में, क्रियाविशेषण के रूप में, परसर्गीय पदबंध, संबंधवाचक उपवाक्य, भावार्थक संज्ञा, और कृदंत, क्रिया विशेषण आदि के रूप में पश्च विशेषक का प्रयोग होता है।

उपर्युक्त घटकों के आधार पर अंग्रेजी संज्ञा पदबंध संरचना नियम को देख सकते हैं-

अंग्रेजी संज्ञा.संरचना नियम = निर्धा.+ पू.विशे. + स. (शीर्ष)+ पश्च विशे.

उदाहरण के लिए - *All the delegates coming from abroad*

[Determiner + Pre-modifier + Noun (Head) + Post-modifier] (Verma, 1971, p.12)

2.4 हिंदी और अंग्रेजी संज्ञा पदबंध घटकों की तुलना

जब दोनों भाषाओं की संरचना और प्रकृति में भिन्नता रहती है तो उनके घटकों में भी भिन्नता पाई जाती है। इस भिन्नता को हिंदी और अंग्रेजी के संज्ञा पदबंध घटकों के आधार पर टेबल में दिखाया गया है।

	हिंदी संज्ञा पदबंध (Hindi Noun Phrase)	अंग्रेजी संज्ञा पदबंध (English Noun Phrase)
निर्धारक (determiner)	सभी, दोनों, एक, दो, ये, वे, कुछ, कई, बहुत, हरेक, आपका, उसका, मेरा, अगला, पहला, बाकी, थोड़ा, काफी, आदि।	All, both, such, the, a, an, this, these, my, some, any, every, either, whose, one, first, next, many, few, several, more
वाक्य (sentence)	प्रिंसिपल मूलतः एक अध्यापक होता है। बाहर एक लड़का खड़ा है।	(whose book is this) (Half the hall was empty) (All the boys are guilty)
पूर्वविशेषक (pre-modifier)	छोटा, सुन्दर, तैयार, विवाहित, अर्जित, का बना,-वाला, का, ग्रामीण, स्पष्टवादी, जन्मजात आदि	Pretty, tall, bright, soft-spoken, odd, successful, mature, eloquent
वाक्य (sentence)	ये आपकी चारों सफ़ेद कमीजें।	<u>Good-looking</u> girl, <u>come what may</u> approach <u>Mumbai City Zimkhana Club</u> Members
संज्ञा Noun (Head)	लड़का, पेड़, जीवन, अध्यापक, बैठक, भय आदि	Boy, friend, idea, news, politics, job
पश्च विशेषक (Post-determiner)	विशेषण उपवाक्य -जो .. (लड़का जो बाहर बैठा है)	(The lady) with blue eyes, on board, in the corner working here, way out (Next flight) to Delhi, (The girl) Next door, (the question), (the President) elect, (prince) charming
वाक्य (sentence)	उन्होंने लड़की अच्छी चुनी है। मुहावरेदार प्रयोग में - चाट मसालेदार	All officers <u>concerned</u> , There is no <u>way out</u> , The property (that) he owns

टेबल 2.3 : हिंदी-अंग्रेजी संज्ञा पदबंध घटकों की तुलनात्मक संरचना

इन घटकों से दोनों भाषाओं में अनुवाद के स्तर पर संज्ञा, सर्वनाम, विशेषण, निर्धारक में बदलाव को रेखांकित किया गया है, जिससे दोनों भाषाओं में अनुवाद के लिए जानकारी मिलती है।

2.5 संज्ञा पदबंध के प्रकार

संज्ञा पदबंध वाक्य में कर्ता, कर्म, क्रिया के पूरक और परसर्ग के रूप में कार्य करता है। इसलिए इसकी संरचना वाक्य में सामान्य संज्ञा पदबंध, संयुक्त संज्ञा पदबंध, मिश्र संज्ञा पदबंध के रूप में हो सकती है। इन संज्ञा पदबंध प्रकारों को यमुना काचरू¹¹ की पुस्तक के आधार पर यहाँ विश्लेषित किया गया है।

2.5.1 सामान्य संज्ञा पदबंध (Simple Noun Phrase)

सामान्य सं.प. में सामान्य संज्ञा या विशेषक का प्रयोग होता है। हिंदी में विशेषक का प्रयोग हमेशा संज्ञा से पहले होता है तो अंग्रेजी में विशेषक का प्रयोग संज्ञा के पहले और अंत में भी होता है।

1. कुत्ता बाहर है (The dog is outside)
2. कोई लड़का जा रहा है। (Any boy is going)
3. सचिन ने एक अच्छी शर्ट पहनी है। (Sachin is wearing a good shirt)

उपर्युक्त प्रथम वाक्य सामान्य है। दूसरे वाक्य में 'कोई' (अनिश्चित निर्धारक) को अंग्रेजी के (निश्चित निर्धारक) 'A' से रेखांकित किया गया है। तृतीय वाक्य में संज्ञा के साथ विशेषक का प्रयोग हुआ है, इस वाक्य में 'एक' के लिए 'a' का प्रयोग हुआ है। अनिश्चित निर्धारक 'एक' के लिए अंग्रेजी में कभी (a) तो कभी (one) का प्रयोग होता है। हिंदी के संज्ञा पदबंध के उदाहरण में- कुछ लड़के (Some boys), कोई आदमी (a man) यह आम (this mango) आदि के विश्लेषण से यह ज्ञात होता है। 'कोई' या 'कुछ' के लिए अंग्रेजी में (Some, a) का प्रयोग होता है।

संज्ञापद में निर्धारक के रूप में परिमाणवाचक का समावेश होता है जिसमें संख्यावाचक को विभाजित किया जा सकता है। इसमें गणनावाची संख्या (एक, दो इ.) क्रमवाची संख्या (पहला, दूसरा) आंशिक संख्यावाची (Fractional) में (आधा, चतुर्थ) गुणा प्रयोग (गुना, दोगुना इ.) योगात्मक (दोनों), अन्य परिणामवाचक (कुछ, लगभग, करीब) आदि का प्रयोग होता है।

विशेषण और सीमाकारी प्रत्यय (Limiter participles) भी सामान्य संज्ञापद की जानकारी देते हैं। जैसे- सिर्फ़ ही, बस...ही के प्रयोग से संज्ञा का बोध होता है, इसके लिए 'ही' प्रत्यय का प्रयोग होता है। उदाहरण के लिए - सिर्फ़ ये दो दर्जन केले ही काफी हैं। (Only these two dozen bananas are enough.)

¹¹ Kachru, Yamuna (2006) : Hindi, p.132-134

2.5.2 संयुक्त संज्ञा पदबंध (Compound Noun Phrase)

मिश्रित और संयुक्त संज्ञा पदबंध में दो संज्ञाओं को एक साथ समुच्चयबोधक अव्यय (और, या, लेकिन, पर) द्वारा जोड़ा जाता है।

1. श्याम और घनश्याम क्रिकेट खेल रहे हैं।
Shyam and Ghanshyam are playing cricket.

2. उसने कुर्सियाँ और मेज़ एक साथ लगा दीं।
She/he put the chairs and table together.

उपर्युक्त दोनों हिंदी और अंग्रेजी के वाक्यों में 'श्याम और घनश्याम' (Shyam and Ghanshyam) और 'कुर्सियाँ और मेज़' (chairs and table) संयुक्त संज्ञा पदबंध हैं, जिनमें समुच्चयबोधक का प्रयोग हुआ है।

2.5.3 मिश्र संज्ञा पदबंध (Complex Noun Phrase)

मिश्र संज्ञा पदबंध के दो प्रकार हैं: क) पूरक उपवाक्य ख) संज्ञा उपवाक्य और अन्य संज्ञा पदबंध में संबंधवाचक वाक्य ।

क) पूरक उपवाक्य (Complement Clause): कि-उपवाक्य की तरह उस वाक्य का प्रयोग हिंदी और अंग्रेजी में भी समान रूप में देखा जाता है।

जैसे- 1. यह सूचना कि आयकर घटाया जा रहा है गलत है।

The information that the income tax is being reduced is incorrect.

2. रमेश ने ये कहा कि सुरेश सब से खुश है ।

Ramesh said that suresh was happy with everyone.

उपर्युक्त प्रथम वाक्य में, संज्ञा पदबंध 'यह सूचना' का पूरक 'कि आयकर घटाया जा रहा है' है तो दूसरे वाक्य में 'कि सुरेश सब से खुश है' संज्ञा (रमेश ने ये कहा) का पूरक है। अंग्रेजी में पूरक की शुरुआत that से होती है। (Singh, 2003,p.91)

ख) संज्ञा उपवाक्य (Noun clause)

जिस उपवाक्य को वाक्य में संज्ञा की जगह प्रयुक्त करते हैं, उसे संज्ञा उपवाक्य कहते हैं। अक्सर संज्ञा उपवाक्यों के शुरु में कि (that) का प्रयोग होता है, इसीलिए इसे कि-उपवाक्य या that-clause भी कहा जाता है, जैसे- सुरेश ने बताया कि पिताजी बहुत बीमार हैं, उसी तरह

अंग्रेजी में Ramesh said *that he was not well* . विकल्प या संदेह का भाव व्यक्त करने के लिए अंग्रेजी में संज्ञा उपवाक्य से पहले वाक्य *if* या *whether* का इस्तेमाल होता है जैसे- *I don't know if/whether I am right*. इसी वर्ग के कुछ संज्ञा उपवाक्य प्रश्नवाचक के रूप में प्रयुक्त होते हैं। संज्ञा उपवाक्य में संज्ञा और सर्वनाम शब्द के साथ संबंधात्मक रूप का प्रयोग होता है जो क्रियावाचक संज्ञा का कार्य करता है।

जैसे- सं.प. (कृते का) (भौंकना)

ग) संबंधवाचक उपवाक्य (Relative Clause)

अगर वाक्य की शुरुआत में संज्ञा या सर्वनाम शीर्ष को रखा जाता है, तो संबंधवाचक उपवाक्य का प्रयोग संज्ञा या सर्वनाम शीर्ष के बाद होता है। जैसे- वह *जो जा रहा* है। संबंधवाचक सर्वनाम 'जो' और प्रश्नवाचक सर्वनाम 'कौन' के रूप निश्चयवाचक सर्वनामों के अनुसार बनते हैं। 'जो' के विकृत रूप दोनों वचनों में क्रमशः 'जिस' और 'जिन' है तथा 'कौन' के 'किस' और 'किन' हैं, जिनका प्रयोग संबंधवाचक वाक्य और उपवाक्य में किया जाता है।

घ) कृदंत/वाला विशेषक (Participial Modifiers): हिंदी में *-वाला* बहुप्रयोजनी विशेषक है जो अंग्रेजी में अनेक तरह से व्यक्त होने वाले विशेषकों की जगह इस्तेमाल हो सकता है, जिनमें क्रियाओं के सहयोग से बनने वाले संज्ञा पदबंध भी और विशेषणों से बने संज्ञा पदबंध भी शामिल हैं। अंग्रेजी में तीनों कालों में इस विशेषक का प्रयोग होता है। जैसे - *लंबे काले बालों वाली महिला* (The lady with long black hair.) (Kachru, 2006; p.136.)

संज्ञा पदबंध के इन प्रकारों को हिंदी अंग्रेजी वाक्यों के उदाहरण के द्वारा स्पष्ट करने का प्रयास किया गया है। इस अध्ययन के बाद दोनों भाषाओं के संज्ञापदों का व्यतिरेकी विश्लेषण किया जाएगा, जिसके लिए कॉर्पोरा के कुछ वाक्य नमूने के तौर पर लिए गए हैं।

2.6 हिंदी अंग्रेजी संज्ञापदों का व्यतिरेकी विश्लेषण

यह व्यतिरेकी विश्लेषण कॉर्पोरा आधारित हिंदी-अंग्रेजी संज्ञापद का है। इस शोध के लिए उपयोग किया गया कॉर्पोरा 'भारतीय भाषा कॉर्पोरा उपक्रम'¹²(ILCI) की परियोजना से लिया गया है जो जवाहरलाल नेहरू विश्वविद्यालय में चल रही है। इस परियोजना के निर्देशक डॉ. गिरीश नाथ

¹² भारतीय भाषा कॉर्पोरा उपक्रम, यह परियोजना जवाहरलाल नेहरू विश्वविद्यालय, नई दिल्ली में विशिष्ट संस्कृत अध्ययन केंद्र में चलाई जा रही है, इस परियोजना का कॉर्पोरा हिंदी से अंग्रेजी में समांतर वाक्यों का है इसका स्वरूप पर्यटन और स्वास्थ्य के क्षेत्र का है। यह भारत सरकार के सूचना प्रौद्योगिकी विभाग द्वारा वित्तपोषित है।

झा हैं। यह कॉर्पोरा भारतीय भाषा कॉर्पोरा उपक्रम (ILCI) का हिन्दी से अंग्रेजी भाषा में अनूदित वाक्यों का है। उसका स्वरूप स्वास्थ्य (health) और पर्यटन (tourism) क्षेत्र का है। इसे टैग करने के लिए भारतीय मानक ब्यूरो (Bureau of Indian Standards- BIS) टैगसेट का प्रयोग किया है और अंग्रेजी कॉर्पोरा को टैग करने के लिए Penn (University of Pennsylvania) के टैगसेट का प्रयोग किया है। इस टैग किए हुए कॉर्पोरा के वाक्य के आधार पर इस अध्याय में संज्ञापद का व्यतिरेकी विश्लेषण किया गया है। व्यतिरेकी विश्लेषण के लिए प्रयोग किए गए कॉर्पोरा का स्वरूप 2000 टैग वाक्य का है। इस कॉर्पोरा के कुछ प्रमुख नमूने के तौर पर के वाक्यों का यहाँ व्यतिरेकी विश्लेषण कर दोनों भाषाओं के संज्ञापद की संरचना में होनेवाले बदलाव और जुड़नेवाले घटकों को स्पष्ट कर संज्ञापद को रेखांकित किया गया है।

सामान्यतः व्यतिरेकी विश्लेषण में दोनों भाषाओं की भाषिक विशिष्टताओं का विवेचन किया जाता है, जिसमें दोनों भाषाओं के विभिन्न स्तर के समान और असमान तत्वों की जानकारी प्राप्त होती है, जिसका प्रयोग अनुवाद के लिए किया जा सकता है। प्रारंभ में व्यतिरेकी विश्लेषण की संकल्पना भाषा शिक्षण के लिए प्रयोग में लायी गई। फिर इसका प्रयोग भाषा और अनुवाद से जुड़े सभी विषयों के लिए किया गया। अनुवाद में स्रोत भाषा के कथन को लक्ष्य भाषा में रूपांतरित किया जाता है, उसी तरह व्यतिरेकी विश्लेषण में स्रोत भाषा की संरचना का लक्ष्य भाषा की संरचना के साथ तुलनात्मक अध्ययन किया जाता है। यहाँ हिंदी-अंग्रेजी वाक्य में संज्ञापद का व्यतिरेकी विश्लेषण वाक्य के द्वारा स्पष्ट किया गया है।

1) [दाँतों\N_NN से\PSP] [आपका\PR_PRP आत्मविश्वास\N_NN] भी\RP_RPD बढ़ता\V_VM है\V_VAUX ।\RD_PUNC

अंग्रेजी- [Your\PRP\$ self-confidence\NN] also\RB increases\VBZ [with\IN teeth\NNS] .\.

दोनों वाक्यों के व्यतिरेकी विश्लेषण में 'दाँतो से' के लिए 'With teeth' संज्ञा के साथ हिंदी में परसर्ग और अंग्रेजी में पूर्वसर्ग का प्रयोग हुआ है। 'आपका आत्मविश्वास' के लिए 'Yours self-confidence' का प्रयोग हुआ है, जो सर्वनाम है।

2) चबाएँ\V_VM [शुगर\N_NN रहित\N_NN चुड़ंग\N_NN गम\N_NN] ।\RD_PUNC

अंग्रेजी - Chew\V B [the\DT sugar-free\NN chewing\NN gum\NN] .\.

दोनों वाक्यों में कोष्ठक में रेखांकित किया हुआ संज्ञा पदबंध है। जिसमें हिंदी वाक्य के 'शुगर रहित' के लिए अंग्रेजी में 'The sugar-free' शब्द के साथ निर्धारक का प्रयोग हुआ है।

3) [नियमित\JJ व्यायाम\N_NN से\PSP]आप\PR_PRP हृदय\N_NN रोग\N_NN ,\RD_PUNC कोलन\N_NNP कैंसर\N_NN ,\RD_PUNC ब्लड\N_NN प्रेशर\N_NN और\CC_CCD डायबिटीज\N_NN जैसी\CC_CCS [बीमारियों\N_NN से\PSP बचाव\N_NN] कर\V_VM सकते\V_VAUX हैं\V_VAUX |\RD_PUNC

अंग्रेजी - You\PRP can\MD prevent\VB heart\NN diseases\NNS ,\, colon\NN cancer\NN ,\, blood\NN pressure\NN and\CC diabetes\NN like\IN [diseases\NNS by\IN regular\JJ exercise\NN] .\.

दोनों वाक्यों को तुलनात्मक स्वरूप में देखें तो यह ज्ञात होता है कि हिंदी में परसर्ग का प्रयोग संज्ञा के बाद होता है तो अंग्रेजी में पूर्वसर्ग का प्रयोग संज्ञा के पहले होता है। जैसे 'नियमित व्यायाम से' के लिए 'by regular exercise' में प्रयोग हुआ है।

4) [बच्चों\N_NN को\PSP फास्ट\JJ फूड\N_NN ,\RD_PUNC कुरकुरे\N_NN ,\RD_PUNC आइसक्रीम\N_NN] नहीं\RP_NEG खिलायें\V_VM |\RD_PUNC

अंग्रेजी- Do\VBP not\RB feed\VB [children\NNS fast\JJ food\NN ,\, kurkure\NN ,\, ice-cream\NN] .\.

हिंदी वाक्य में एक से अधिक संज्ञाओं का प्रयोग हुआ है साथ ही परसर्ग के बाद विशेषण का भी प्रयोग हुआ है। 'बच्चों को फास्ट..' संज्ञापद के लिए 'children fast food' से रेखांकित किया है, लेकिन इस हिंदी वाक्य के परसर्ग का अंग्रेजी में लोप (Drop) हुआ है।

5) [लैपरोस्कोपिक\N_NNP विधि\N_NN में\PSP] [सामान्य\JJ चिकित्सा\N_NN के\PSP मुकाबले\N_NN दोगुना\QT_QTF समय\N_NN] लगता\V_VM है\V_VAUX |\RD_PUNC

अंग्रेजी- [Laparoscopic\NN method\NN] takes\VBZ [double\JJ time\NN in\IN comparison\NN with\IN the\DT simple\JJ treatment\NN] .\.

दोनों वाक्य में संज्ञापदों की दो रचनाएँ हैं। हिंदी वाक्य में प्रथम रेखांकित कोष्ठक के संज्ञापद को अंग्रेजी के व्यतिरेकी स्वरूप की तुलना में देखा जाए तो अंग्रेजी में प्रथम कोष्ठक में पूर्वसर्ग का प्रयोग नहीं हुआ है। दूसरे कोष्ठक के संज्ञापद 'सामान्य चिकित्सा के मुकाबले' के लिए 'In comparison with the simple treatment' में हिंदी परसर्ग के लिए अंग्रेजी पूर्वसर्ग का प्रयोग रेखांकित किया गया है।

6) [इनमें\PR_PRP विटामिन\N_NN ई\N_NNP प्रचुर\QT_QTF मात्रा\N_NN में\PSP] पाया\V_VM जाता\V_VAUX है\V_VAUX ,\RD_PUNC जो\PR_PRL न\RP_NEG केवल\QT_QTF मोतियाबिंद\N_NNP ,\RD_PUNC [कई\QT_QTF प्रकार\N_NN के\PSP अंधेपन\N_NN] और\CC_CCS अन्य\JJ नेत्र\N_NN रोगों\N_NN से\PSP बचाने\V_VM के\PSP अलावा\CC_CCD [पराबैंगनी\N_NNP किरणों\N_NN से\PSP] पहुँच\V_VM रहे\V_VAUX नुकसान\N_NN से\PSP भी\RP_RPD आँखों\N_NN की\PSP रक्षा\N_NN करता\V_VM है\V_VAUX ।\RD_PUNC

अंग्रेजी- [Plenty\NN of\IN vitamin-E\NN] is\VBZ found\VBN in\IN these\DT ,\, which\WDT not\RB only\RB saves\VBZ from\IN cataract\NN ,\, [many\JJ kind\NN of\IN blindness\NN] and\CC other\JJ eye\NN diseases\NNS but\CC also\RB [protect\NN eyes\NNS from\IN the\DT harms\NNS of\IN ultraviolet\NN rays\NNS] .\.

इस बृहद् वाक्य में एक से अधिक संज्ञापदों के साथ विशेषण और क्रियापद का प्रयोग हुआ है। हिंदी वाक्य में प्रथम कोष्ठक में संज्ञा के पहले परिणामवाचक (प्रचुर) का अंग्रेजी में (*plenty of vitamin-e*) संज्ञा के रूप में प्रयोग हुआ है। ठीक उसके बाद अधिकांश वाक्य में परसर्ग का प्रयोग हिंदी में हुआ है तो अंग्रेजी में पूर्वसर्ग का प्रयोग संज्ञा के पहले और संज्ञा के बाद में हुआ है। इस तरह के वाक्य में संज्ञापद का व्यतिरेकी विश्लेषण करना कठिन हो जाता है। कोष्ठक में दोनों भाषाओं के टैग वाक्य को रेखांकित करने से यह ज्ञात होता है कि परसर्ग के लिए अधिकांश जगह अंग्रेजी वाक्य में पूर्वसर्ग का लोप हुआ है और कुछ संज्ञाओं के साथ निर्धारक का प्रयोग हुआ है।

7) [नई\JJ जानकारी\N_NN को\PSP किसी\DM_DMQ पुरानी\JJ जानकारी\N_NN से\PSP] जोड़कर\V_VM याद\N_NN रखने\V_VM का\PSP प्रयास\N_NN करें\V_VM ।\RD_PUNC

अंग्रेजी- Try\VB to\TO remember\VB [the\DT new\JJ information\NN by\IN relating\VBG it\PRP with\IN any\DT old\JJ information\NN] .\.

वाक्य के कोष्ठक में संज्ञापद 'नई जानकारी' '*The new information*' में अंग्रेजी में (the) निर्धारक का प्रयोग हुआ है। हिंदी और अंग्रेजी के अधिकांश वाक्य में संज्ञा के पहले विशेषण का प्रयोग हुआ है। उसी तरह 'पुरानी जानकारी' के लिए '*old Information*' का प्रयोग हुआ है।

8) [पीतज्वर\N_NNP का\PSP] [अफ्रीका\N_NNP ,\RD_PUNC अमेरिका\N_NNP ,\RD_PUNC यूरोप\N_NNP तथा\N_NNP कैरिबियन\N_NNP द्वीप\N_NN समूह\N_NN के\PSP इतिहास\N_NN में\PSP] बड़ा\RP_INTF महत्वपूर्ण\JJ स्थान\N_NN है\V_VAUX ।\RD_PUNC

अंग्रेजी- [Yellow\JJ fever\NN] has\VBZ an\DT very\RB important\JJ place\NN [in\IN the\DT history\NN of\IN Afric\NNP ,\, America\NNP ,\, Europe\NNP and\CC the\DT Caribbean\NNP Islands\NNP] .\.

वाक्य में एक से अधिक संज्ञापद हैं। हिंदी वाक्य के दूसरे कोष्ठक में संज्ञापद 'के इतिहास में' के लिए अंग्रेजी में 'In the history of' को रेखांकित किया जाता है।

9) [एक\QT_QTC माह\N_NN से\PSP] लगातार\N_NN बुखार\N_NN आना\N_NN \V_VM \RD_PUNC

अंग्रेजी- Frequent\JJ fever\NN [for\IN a\DT month\NN] .\.

कोष्ठक में रेखांकित वाक्य से गणवाची संख्या (एक) संज्ञा (माह) के पहले आयी है, साथ में परसर्ग (से) का प्रयोग हुआ है जिसे अंग्रेजी के वाक्य में कोष्ठक में संज्ञापद से रेखांकित किया है।

10) [शरीर\N_NN के\PSP किसी\PR_PRI भाग\N_NN या\CC_CCD अंग\N_NN से\PSP] असामान्य\JJ रूप\N_NN से\PSP रक्त\N_NN या\CC_CCD मवाद\N_NN आना\N_NN \V_VM \RD_PUNC

अंग्रेजी- Abnormal\JJ passing\NN of\IN blood\NN or\CC pus\NN [from\IN any\DT part\NN or\CC organ\NN of\IN body\NN] .\.

दोनों वाक्यों को कोष्ठक से रेखांकित किया गया है। 'शरीर के' को 'of body' और 'किसी भाग या अंग से' को 'From any part or organ' से रेखांकित किया जाता है।

11) [इंटेंसिव\N_NN फेस\N_NN जो\PR_PRL] [2-3\QT_QTC मास\N_NN तक\PSP] चलता\N_NN \V_VM है\N_NN \RD_PUNC

अंग्रेजी- [Intensive\JJ phase\NN which\WDT] goes\VBZ on\IN [for\IN 2-3\CD months\NNS] .\.

यहाँ हिंदी वाक्य में संज्ञा के साथ संबंधवाचक 'जो' के लिए अंग्रेजी 'Which' का प्रयोग हुआ है, और दूसरे कोष्ठक के में गणनावाची संख्या के बाद संज्ञा और परसर्ग (तक) के लिए (for) का प्रयोग हुआ है।

12) [पूर्वोत्तर\N_NNP रेलवे\N_NN की\PSP एक\QT_QTC लाइन\N_NN भटनी\N_NNP से\PSP] तथा\CC_CCD दूसरी\QT_QTO छपरा\N_NNP से\PSP इलाहाबाद\N_NNP सिटी\N_NN तक\PSP] जाती\N_NN \V_VM है\N_NN \RD_PUNC

अंग्रेजी- [One\CD line\NN from\IN Bhatni\NNP and\CC another\DT from\IN Chhapra\NNP of\IN the\DT Northeastern\NNP Railway\NNP] goes\VBZ up\RP to\TO Allahabad\NNP City\NNP .\.

इस वाक्य में भी कई संज्ञापदों के साथ गणनावाची संख्या के साथ परसर्ग और पूर्वसर्ग का प्रयोग हुआ है। हिंदी वाक्य में 'पूर्वोत्तर रेलवे की' के लिए अंग्रेजी में '*of the Northeastern Railway*' और 'एक लाइन भटनी से' के लिए '*one line from Bhatni*' को रेखांकित किया गया है।

13) [चित्रकोट\N_NNP में\PSP इन्द्रावती\N_NNP नदी\N_NN] विस्तारित\N_NN होकर\N_VVM [मनमोहक\JJ जलप्रपात\N_NN का\PSP] निर्माण\N_NN करती\N_VVM है\N_VVAUX ।\RD_PUNC

अंग्रेजी- Having\VBG expanded\VCB [in\IN Chitrakot\NNP the\DT Indrāvati\NNP river\NN] forms\VBZ [the\DT charming\JJ waterfall\NN] .\.

हिंदी वाक्य के कोष्ठक में संज्ञापद 'चित्रकोट में' के लिए अंग्रेजी में '*in chitrakot*' संज्ञा के साथ पूर्वसर्ग प्रयुक्त हुआ है। दूसरे कोष्ठक में निर्धारक '*the*' का प्रयोग अंग्रेजी में हुआ है, लेकिन इस तरह हिंदी में निर्धारक की संरचना नहीं है।

14) [पुर्तगालियों\N_NN और\CC_DCC डचों\N_NN ने\PSP तलशेशरी\N_NNP पर\PSP] अधिकार\N_NN जमाने\N_VVM की\PSP चेष्टा\N_NN की\PSP थी\N_VVAUX ।\RD_PUNC

अंग्रेजी- [The\DT Portuguese\NNP and\CC the\DT Dutch\NNP] had\VCB tried\VCB to\TO capture\VB Thalassery\NNP .\.

वाक्य में दो संज्ञाओं को संयोजक 'और' से तथा अंग्रेजी में '*and*' से जोड़ा गया है। अंग्रेजी वाक्य में संज्ञा के साथ निर्धारक का प्रयोग हुआ है तो हिंदी में कारक चिह्न 'ने' का लोप हुआ है।

15) [भारतीय\N_NN चिकित्सा\N_NN पद्धति\N_NN में\PSP] आयुर्वेद\N_NN का\PSP भी\RP_RPD [पर्यटन\N_NN के\PSP क्षेत्र\N_NN में\PSP] महत्वपूर्ण\JJ योगदान\N_NN है\N_VVAUX ।\RD_PUNC

अंग्रेजी- [Among\IN Indian\JJ medical\JJ systems\NNS] Ayurveda\NNP also\RB has\VBZ an\DT important\JJ contribution\NN [in\IN the\DT area\NN of\IN tourism\NN] .\.

हिंदी वाक्य में परसर्ग 'में' अंग्रेजी वाक्य के कोष्ठक में समावेशित है। दूसरे कोष्ठक में 'पर्यटन के क्षेत्र में' में दो परसर्गों का प्रयोग हुआ है, जिसे अंग्रेजी में '*in the area of tourism*' से रेखांकित किया गया है।

16) [1684\QT_QTC में\PSP ब्रिटिश\N_NNP ईस्ट\N_NNP इंडिया\N_NNP कंपनी\N_NN ने\PSP] [केरल\N_NNP में\PSP अपना\PR_PRF पहला\QT_QTO व्यापार\N_NN केन्द्र\N_NN अंचुतेंगु\N_NNP में\PSP] ही\RP_RPD स्थापित\V_VM किया\N_NN था\N_NN I\RD_PUNC

अंग्रेजी- [In\IN 1684\CD British\N_NNP East\N_NNP India\N_NNP company\NN] founded\VBD [its\PRP\$ first\JJ business\NN center\NN in\IN Kerala\N_NNP at\IN Anchuthengu\N_NNP].\.

दोनों वाक्यों में संज्ञा का बृहद् पदबंध है, इस पदबंध में संज्ञाएँ एक से अधिक हैं लेकिन संज्ञा के साथ हमेशा प्रयोग में आनेवाले परसर्ग और पूर्वसर्ग को रेखांकित किया गया है, जो वाक्य में संज्ञा के साथ अहम भूमिका निभाते हैं। सामान्यतः वाक्य में संज्ञा एक से अधिक बार आती है। लेकिन वाक्य में क्रिया का प्रयोग एक या दो बार होता है। जब वाक्य संयुक्त होता है तब वाक्य में एक से अधिक बार संज्ञा के साथ क्रिया का भी प्रयोग होता है और निर्धारक, परसर्ग, विशेषक का भी।

इस संज्ञापद के व्यतिरेकी विश्लेषण में मुख्य रूप से यह पाया गया है कि दोनों भाषाओं की वाक्य संरचना में संज्ञा के साथ या संज्ञा के पहले विभिन्न घटकों का प्रयोग में आना अनिवार्य है, उसके बिना संज्ञा की विशेषता और उसका महत्व वाक्य में पता नहीं चलेगा। इस उद्देश्य से संज्ञा के साथ हिंदी-अंग्रेजी में निर्धारक का प्रयोग होता है। हिंदी में निर्धारक का प्रयोग नहीं होता लेकिन विशिष्ट संदर्भ के साथ कुछ अनिश्चित निर्धारक प्रयुक्त होते हैं। अंग्रेजी में निर्धारक (the, a, an) का प्रयोग होता है। हिंदी में परसर्ग तो अंग्रेजी में पूर्वसर्ग का प्रयोग होता है। जैसे- (का, के, की से) के लिए अंग्रेजी में (of), (पर,को) के लिए (on), (के लिए, के) को अंग्रेजी में (for), (में पर) के लिए (in, at, on), (में) के लिए (in), (से) के लिए (from), (के साथ) के लिए (with), (के बारे में, के विषय में, के संबंध में) के लिए (about) आदि परसर्ग के लिए अंग्रेजी में पूर्वसर्ग का प्रयोग होता है। गणवाचक संख्याएँ भी संज्ञा के बारे में जानकारी देती हैं। इसलिए वाक्य में संज्ञा के साथ ये घटक महत्वपूर्ण रूप से उसकी विशेषता बताते हैं। इस व्यतिरेकी विश्लेषण में संज्ञापद के संदर्भ में विशेष रूप से अध्ययन किया गया है।

निष्कर्षतः इस अध्याय में हिंदी अंग्रेजी वाक्यसंरचना के साथ उनके घटक और उनके प्रकारों को दोनों भाषाओं के स्तर पर बताएं गए हैं। फिर संज्ञा पदबंध को स्पष्ट करते हुए उनके घटकों की तुलना टेबल में दिखायी गयी है। साथ ही संज्ञापदबंध के प्रकार को स्पष्ट करते हुए संज्ञापदों का कॉर्पोरा के वाक्य के आधार पर व्यतिरेकी विश्लेषण किया गया है। इस व्यतिरेकी विश्लेषण में संज्ञा के साथ प्रयुक्त होनेवाले घटकों को दोनों भाषाओं की क्रमसंरचना के अनुरूप कोष्ठक में रेखांकित किया गया है।

अध्याय – तृतीय

मशीनी अनुवाद में स्थानांतरण
व्याकरण की भूमिका

अध्याय - तृतीय

मशीनी अनुवाद में स्थानांतरण व्याकरण की भूमिका

(Role of Transfer Grammar in Machine Translation)

प्रस्तुत अध्याय में, मशीन अनुवाद में स्थानांतरण व्याकरण (Transfer Grammar) के प्रयोग पर चर्चा की जा रही है। प्रथमतः स्थानांतरण व्याकरण के विभिन्न पहलुओं के अध्ययन हेतु स्थानांतरण व्याकरण क्या है, मशीनी अनुवाद में स्थानांतरण व्याकरण का प्रयोग किस तरह से किया जाता है और स्थानांतरण व्याकरण के नियम कौन से हैं, इन प्रश्नों पर विस्तृत चर्चा की गयी है। साथ ही अंतरराष्ट्रीय सूचना प्रौद्योगिकी संस्थान, हैदराबाद (IIT Hyderabad) में स्थानांतरण व्याकरण (TG) पर किए गए कार्य का अध्ययन किया गया है, जो संपर्क मशीनी अनुवाद तंत्र पर आधारित है।

“संपर्क मशीनी अनुवाद तंत्र (Sampark Machine Translation System) में स्थानांतरण व्याकरण के तीन उप घटकों का समावेश किया गया है, वे हैं- संरचनात्मक अंतरण (Structural Transfer), शाब्दिक अंतरण (Lexical Transfer) और लिप्यंतरण (Transliteration)। संरचनात्मक अंतरण एक भारतीय भाषा से (IL-1) अन्य भारतीय भाषा (IL-2) में वाक्य स्तर पर होता है। इस तंत्र में नियम आधारित (Rule Based) घटकों का समावेश किया गया है। जो नियमों के समुच्चय (Set of rules) का प्रयोग एक भाषा से दूसरी भाषा में अंतरित करने के लिए करता है।”¹ इन व्याकरणिक नियमों का प्रयोग विश्व स्तर में नियम आधारित तंत्र (rule-based system) के साथ अन्य तंत्रों (सांख्यिकीय आधारित तंत्र) में भी किया गया है। लेकिन भारत में मशीनी अनुवाद के क्षेत्र में स्थानांतरण व्याकरण के नियमों का निर्माण एवं प्रयोग नए रूप में हुआ है, जिस पर अभी शोध हो रहे हैं।

3.1 स्थानांतरण व्याकरण का परिचय

“मानवीय भाषा की दो प्रमुख विशेषताएँ हैं- उसकी रचनात्मक क्षमता और उसकी सूक्ष्म नियमबद्धता। भाषाशास्त्र का लक्ष्य केवल यह नहीं कि विशिष्ट भाषाओं के गठन का अध्ययन हो, बल्कि यह भी है कि उन सूक्ष्म नियमों का अविष्कार किया जाए जो मानव भाषा की परिभाषा

¹ Avinesh P.V.S, (2010) : *Transfer Grammar Engine and Automatic Learning of Reorder Rules in Machine Translation* (M.Sc- Thesis), p.18

बाँधते हैं।² मानव भाषा के मुक्त रूप से संचार करने के लिए और दो भाषाओं के बीच संप्रेषण माध्यम के लिए अनुवाद का क्षेत्र आज बहुत महत्वपूर्ण बन चुका है। अनुवाद जब एक भाषा से दूसरी भाषा में होता है, तो अनुवाद के लिए दोनों भाषाओं की संरचना को मूल बनाए रखने का प्रयास मानव अनुवादक करता है। उसी तरह मशीन के माध्यम से जब एक भाषा से दूसरी भाषा में अनुवाद किया जाता है, तो (मशीन के लिए बनाए गए नियम) स्थानांतरण व्याकरण नियम के माध्यम से दोनों भाषाओं की वाक्यात्मक संरचना की विलक्षणताओं और असमानताओं को स्पष्ट करते हुए समान अभिव्यक्तियों में अनुवाद किया जाता है। इस अनुवाद में स्थानांतरण व्याकरण के द्वारा स्रोत भाषा (Source Language) के शब्द और अभिव्यक्तियों में उपलब्ध एकाधिक पर्यायी शब्दों में से लक्ष्य भाषा की वाक्यसंरचना के अनुरूप उसके समानार्थी शब्दों और अभिव्यक्तियों का एक क्रम में अनुवाद किया जाता है।³

वाक्य के शब्द क्रम बदलाव से स्रोत भाषा के वाक्य के कौन-कौन से शब्दों या अभिव्यक्तियों को लक्ष्य भाषा वाक्य में किन-किन प्रयोग स्थानों पर अंतरित किया गया है, यह ज्ञात होता है। जब स्रोत भाषा का वाक्य, अनुवाद के लिए कंप्यूटर में दिया जाता है, तब अनुवाद करने के लिए कंप्यूटर में सामान्य जानकारी, शब्द संरचना, अर्थ संबंधी जानकारी और वाक्यविन्यास संबंधी जानकारी के साथ 'स्थानांतरण शब्दकोश' (Transfer Lexicon) का समावेश कंप्यूटर में होता है। यह स्थानांतरण शब्दकोश द्विभाषी कोश होता है, जिसमें स्रोत भाषा के सभी संभावित शब्दों या वाक्यांशों के लिए संदर्भ के अनुकूल समानार्थी शब्दों का समावेश होता है। इस शब्दकोश में उसके शब्दवर्ग (Part of Speech), अनुवाद पर्याय (Translation equivalent), मुख्य रूपावली (Paradigm), व्याकरणिक लक्षण और पदक्रम (Word order) संबंधी नियमों का समावेश होता है।

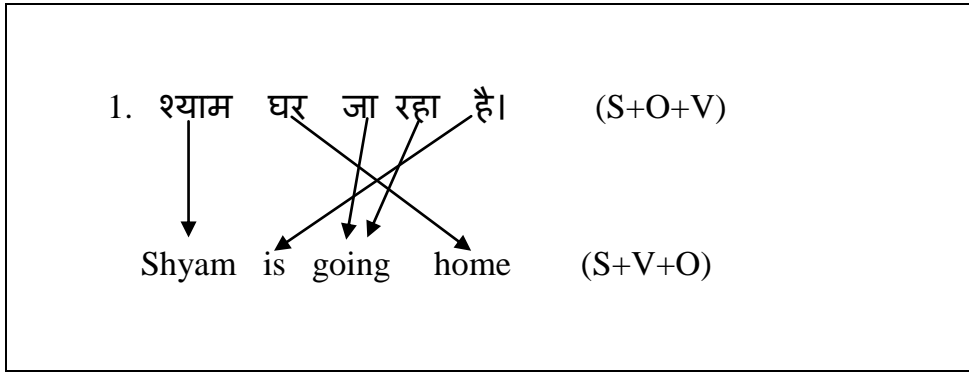
अगर स्रोत और लक्ष्य भाषाएँ सजातीय (Cognate) हों, जैसे (हिन्दी-बंगला, हिन्दी-मराठी, अंग्रेजी-फ्रांसीसी) तो भाषाओं के बीच स्थानांतरण अधिक सहज या सकारात्मक हो सकता है, लेकिन यदि भाषाएँ विजातीय (Non-cognate) या भिन्न संरचना प्रणाली की हुईं तो स्थानांतरण की प्रक्रिया अधिक जटिल हो सकती है, जिसके लिए सामान्य नियमों से अलग नए नियमों का निर्माण करना पड़ता है।⁴ विशेषतः जब हिंदी से अंग्रेजी में मशीन अनुवाद तंत्र के माध्यम से अनुवाद करना है, तो इसके लिए नए नियमों का निर्माण किया जाता है।

² काचरू, यमुना (1997) : हिंदी व्याकरण के कुछ प्रकरण, पृ. 2

³ सिंह, सूरजभान (2003) : अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ. 265

⁴ वही, पृ. 266

प्रत्येक भाषा की वाक्यसंरचना भिन्न है, उसी तरह हिंदी-अंग्रेजी भाषाओं की वाक्यसंरचनाओं में भिन्नता पायी जाती है। इस भिन्नता की वजह से अनुवाद में वाक्यसंरचना, शब्दक्रम और अर्थसंरचना की समस्याएँ आती हैं। इस समस्या को सुलझाने के लिए स्थानांतरण नियम बहुत महत्वपूर्ण होते हैं। मशीन को सभी भाषाओं का ज्ञान न होने की वजह से उसे प्रशिक्षित (Trained) करने के लिए दो भाषाओं की संपूर्ण जानकारी के साथ, नियमों का समावेश करना होता है ताकि इन नियमों के माध्यम से अच्छा अनुवाद प्राप्त हो सके। हिंदी और अंग्रेजी भाषा में वाक्य स्तर पर भिन्नता निम्नलिखित आकृति से स्पष्ट होती है। सामान्यतः हिंदी भाषा की वाक्यसंरचना कर्ता+कर्म+क्रिया (SOV) पर आधारित है और अंग्रेजी भाषा की वाक्यसंरचना कर्ता+क्रिया+कर्म (SVO) पर आधारित है। जैसे-



आकृति 3.1: हिंदी-अंग्रेजी वाक्य की व्यतिरेकी संरचना और शब्दक्रम

उपर्युक्त आरेख के वाक्य से स्पष्ट होता है कि यहां हिंदी भाषा के वाक्य का अंग्रेजी के वाक्य में व्यतिरेकी शब्दक्रम में बदलाव हुआ है। यह बदलाव हमेशा वाक्यसंरचना और शब्द संरचना के स्तर पर मशीन अनुवाद में पाया जा सकता है। इस आधार पर हिंदी से अंग्रेजी अनुवाद में वाक्य संरचना के लिए उदाहरण के तौर पर निम्नलिखित नियम का निर्माण किया जा सकता है-

नियम : कर्ता + कर्म + क्रिया (SOV) → कर्ता + क्रिया + कर्म (SVO)

इस नियम का प्रयोग कर एक भाषा की संरचना का दूसरी भाषा में रूपांतरण किया जा सकता है। इस आधार पर मशीन के लिए कुछ नियम बनाए जाते हैं, इन नियमों को स्थानांतरण व्याकरण नियम (TG rules) कहते हैं। इसी अनुरूप अगर देखा जाए तो हिंदी-अंग्रेजी मशीनी अनुवाद तंत्र में यह प्रमुख नियम होगा, जो शब्दक्रम को भी स्पष्ट करता है और दोनों भाषा का संरचना के साथ उनसे प्राप्त अनुवाद इस नियम के द्वारा प्राप्त होगा।

3.2 स्थानांतरण व्याकरण (TG) का मशीनी अनुवाद में प्रयोग

मशीनी अनुवाद में स्थानांतरण व्याकरण की भूमिका महत्वपूर्ण है। इस व्याकरण नियम निर्माण के लिए दोनों भाषा की संपूर्ण संरचना की जानकारी होना आवश्यक है। इस नियम के माध्यम से मशीन द्वारा प्राप्त अनुवाद में गुणवत्ता को और सुधारा जा सकता है। मशीन में नियमों का समावेश (Installation) करने के साथ उसे प्रयुक्त करने के लिए अल्गोरिथम (Algorithm) बनाने की आवश्यकता होती है, जो नियमों को व्यवस्थित रूप में समावेशित और संचालित करता है। इस अल्गोरिथम के माध्यम से मशीन द्वारा संपूर्ण क्रियाकलाप (Process) किया जाता है।

स्रोत भाषा (SL) के वाक्य को शाब्दिक और संरचनात्मक एकरूपता के साथ लक्ष्य भाषा (TL) में अंतरित करने के लिए प्रथमतः मशीन में समावेशित पद विश्लेषक (Parser) वाक्य के पदों की व्याख्या करता है।⁵ वह वाक्य के विभिन्न पदों को उनके शब्द वर्गों (संज्ञा, क्रिया, विशेषण, क्रियाविशेषण, आदि) में विभक्त करता है और वाक्य के विभिन्न घटकों के बीच मौजूद प्रकार्यात्मक संबंधों (जैसे- कर्ता, कर्म, क्रिया, पूरक आदि) की पहचान करता है। यह कार्य पद विश्लेषक उन व्याकरणिक नियमों और लक्षणों के आधार पर करता है जो उपयुक्त अल्गोरिथम के जरिए उसके स्मृतिकोश (Memory) में पहले से संचित होते हैं। इसलिए नियमों की सूक्ष्मता और शुद्धता जितनी अधिक होगी, उतना ही पद विश्लेषण (Parsing) अच्छा प्राप्त होगा और तंत्र सही अनुवाद करने में अधिक कार्यक्षम होगा।⁶

द्वितीय प्रक्रिया में स्थानांतरण (Transfer) के द्वारा स्रोत भाषा से लक्ष्य भाषा की वाक्य संरचना के स्वरूप, शाब्दिक और संरचनात्मक अभिव्यक्ति के लिए द्विभाषी शब्दकोश के आधार पर उपयुक्त समान शब्दों का चयन किया जाता है। इस प्रक्रिया में स्थानांतरण व्याकरण के नियम महत्वपूर्ण भूमिका निभाते हैं।⁷

तृतीय प्रक्रिया में लक्ष्य भाषा में अनुवाद प्रदान करने के लिए उत्पादक (Generator) का प्रयोग किया जाता है। यह स्रोत भाषा के शब्दों, पदबंधों, और वाक्यों के लिए लक्ष्य भाषा में उपयुक्त समानार्थी शब्दों, पदबंधों और वाक्यों का चयन करता है। यह अनुवाद की प्रक्रिया तीन माध्यमों से मशीन के द्वारा होती है। इस प्रक्रिया में अनुवाद केवल शब्द की बुनियाद पर होने के बजाय उसकी अर्थ संरचना (Semantic Knowledge) का ज्ञान मशीन को देना होता है।⁸

⁵ रस्तोगी, कविता (2010) : समसामायिक अनुप्रयुक्त भाषाविज्ञान, पृ. 101-102

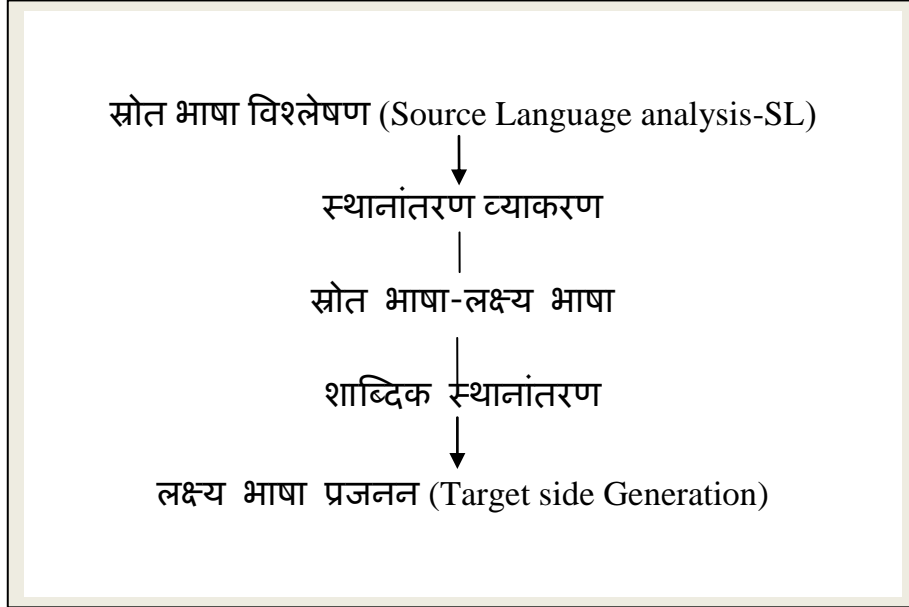
⁶ सिंह, सूरजभान (2003) : अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ. 270

⁷ Mohamed Azza Abd El- Moniem (2000) : *Machine Translation of Noun Phrases: From English to Arabic* (Master of Science thesis) p.24

⁸ सिंह, सूरजभान (2003) : अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ. 271

इस अर्थ संरचना के अनुरूप अगर सही अनुवाद होता है तो दोनों भाषाओं में समान अभिव्यक्तियों का सही संप्रेषण होगा जिससे अनुवाद की कार्यक्षमता और गुणवत्ता बढ़ेगी।

मशीनी अनुवाद में स्थानांतरण व्याकरण की प्रक्रिया किस तरह से होती है यह आरेख द्वारा देखा जा सकता है -



आकृति 3.2 : मशीनी अनुवाद की संरचना और स्थानांतरण व्याकरण की प्रक्रिया ⁹

उपर्युक्त आरेख से स्पष्ट होता है कि स्थानांतरण व्याकरण, मशीनी अनुवाद में किस तरह संपूर्ण अनुवाद की प्रक्रिया से गुजरकर अनुवाद प्रदान करने में सबसे महत्वपूर्ण घटक है। प्रथमतः वाक्य को विश्लेषण की प्रक्रिया से गुजरकर स्रोत भाषा वाक्य का लक्ष्य भाषा में स्थानांतरण किया जाता है। जिसमें लक्ष्य भाषा के अनुकूल शब्दों का चयन किया जाता है। लेकिन इसके लिए हिंदी प्रजनक (Hindi generator) और अंग्रेजी प्रजनक (English generator) का भी विकास करना होता है। मशीनी अनुवाद में हिंदी प्रजनक (Hindi generator) हिंदी वाक्य को अंग्रेजी में पर्यायी शब्द संग्रह के द्वारा आउटपुट (Generate) प्रदान करने में सहायक होता है।

3.3 संपर्क मशीनी अनुवाद तंत्र में स्थानांतरण व्याकरण नियम का प्रयोग

संपर्क मशीन अनुवाद तंत्र का निर्माण भारत में 11 संस्थाओं के संयुक्त प्रयासों से “Indian language to Indian language machine translation” संघ परियोजना के तहत किया जा रहा है, जो

⁹ Mala, Christopher (2009) : “Implementation of Transfer Grammar in Telugu-Hindi Machine Translation System”

Proceedings of SCONLI-2009: 3rd Students Conference of Linguistics in India (research paper)

टी. डी. आई. एल. के आई. टी. विभाग (भारत सरकार) द्वारा वित्तपोषित है। इस ILMT परियोजना ने 9 भारतीय भाषाओं के लिए भाषा प्रौद्योगिकी का विकास किया है और 18 भाषाओं का विकास भाषा-युग्म (Language Pairs) के साथ करने के लिए प्रयास कर रहा है। संपर्क तंत्र में भाषा का विश्लेषण करने के लिए और मशीन को सीखने (भाषा संसाधन) की प्रक्रिया के साथ जुड़ने के लिए कम्प्यूटेशनल पाणिनियन व्याकरण (CPG) का प्रयोग किया गया है। इसमें पारंपरिक नियम और शब्दकोश आधारित अल्गोरिथम के साथ सांख्यिकीय मशीन अधिगम (Statistical Machine learning) का प्रयोग किया गया है। वर्तमान में संपर्क तंत्र के अंतर्गत 6 भाषा युग्म तंत्रों पंजाबी-हिंदी, हिंदी-पंजाबी, तमिल-तेलुगु, हिंदी-उर्दू, हिंदी-तमिल, तेलुगु-हिंदी का विकास किया गया है।

संपर्क तंत्र विश्लेषण- अंतरण- उत्पादक प्रणाली पर आधारित है। इसमें प्रथमतः स्रोत भाषा का विश्लेषण किया जाता है, उसके बाद शब्दावली और स्रोत भाषा की संरचना का अंतरण (Transfer) लक्ष्य भाषा में किया जाता है और अंत में लक्ष्य भाषा की संरचना उत्पादित (Generate) होती है। प्रत्येक चरण में कई "मॉड्यूल" में से 13 प्रमुख मॉड्यूल हैं। इस दृष्टिकोण से विशेष रूप से किसी एक 'भाषा विश्लेषक' (Language analyzer) को विकसित किया जा सकता है। फिर अन्य भाषाओं के लिए जनरेटर के प्रयोग से विभिन्न भाषाओं को भाषा-युग्म के साथ विकसित किया जा सकता है। उदाहरण के लिए –

पंजाबी से हिंदी मशीनी अनुवाद तंत्र का निर्माण करने के लिए, पंजाबी विश्लेषक को हिंदी जनरेटर में शामिल किया जा सकता है। क्योंकि भारतीय भाषाओं की व्याकरणिक संरचना समान हैं। केवल इसके लिए सतही पद-विच्छेदन (Shallow Parsing) किया जाता है। इसलिए संपर्क तंत्र में ट्रान्सफर ग्रामर इंजिन प्रमुख घटक है।¹⁰ संपर्क की तरह संस्कृत-हिंदी मशीनी अनुवाद तंत्र के लिए गहरे स्तर के पार्सर का विकास किया जा रहा है।¹¹ इसमें स्थानांतरण व्याकरण नियम का भी समावेश किया गया है। इन नियमों का निर्माण विभिन्न भाषाओं के तंत्र के लिए आई.आई.टी. हैदराबाद में किया जा रहा है। कुछ भाषाओं के लिए बनाए गए नियमों की चर्चा यहां उदाहरण के साथ की जा रही है। संपर्क तंत्र के नियम निर्माण के लिए आई. आई. टी. हैदराबाद के संगणक विभाग के शोधार्थी अविनेश (P.V.S. Avinesh) के शोध प्रबंध में, संपर्क तंत्र के लिए बनाए गए नियमों को स्पष्ट किया गया है। साथ ही तेलुगु-हिंदी मशीनी अनुवाद तंत्र के लिए स्थानांतरण व्याकरण पर क्रिस्टोफर माला (Christopher Mala) का शोध पत्र महत्वपूर्ण है।

¹⁰ http://tdil-dc.in/index.php?option=com_vertical&parentid=74 retrieved on (05.06.2012)

¹¹ Pathak, Kumar Nripendra (2011) : *Challenges in Sanskrit-Hindi Noun Phrase Mapping*, M.Phil (dissertation), JNU, New Delhi, p.74

इस कार्य से संबंधित तेलुगु-हिंदी भाषा के नियमों को यहाँ आधार रूप में लिया गया है।¹² यहाँ नियमों के क्रम में प्रथमतः तेलुगु के नियम दिए गए हैं फिर उसे हिंदी में स्थानांतरण के अनुरूप दिखाया गया है।

3.4 तेलुगु-हिंदी भाषा में स्थानांतरण से होनेवाला बदलाव और स्थानांतरण व्याकरण नियम¹³

अगर स्रोत भाषा (तेलुगु) की संरचना में दो संज्ञापद एक क्रम में हैं और वे अनुवांशिक संबंध (Genitive relation) से एक दूसरे से संबंधित हैं। तो लक्ष्य भाषा (हिंदी) में संबंधकारक (Genitive case) का प्रयोग होता है। उदाहरण के लिए-

(तेलुगु) रामुडी दुकानमु (rAmudi dukaanamu)

(हिंदी) राम की दुकान (Ram's shop)

उपर्युक्त तेलुगु वाक्य के हिंदी वाक्य के साथ सम्मिलन से यह ज्ञात होता है हिंदी में “की” संबंधकारक का प्रयोग हुआ है। इस आधार पर नियम का निर्माण किया गया है।

$NP\sim 1(\{\langle case="gen", cm="UNDEF">\})NP\sim 2 \Rightarrow NP\sim 1(\{\langle case="gen", cm="kI">\})NP\sim 2$

नियम: संप.1 (कारक “अनिश्चित”) संप.2 \rightarrow संप.1 (कारक “की”) संप.2

1. तेलुगु में जातिवाचक संज्ञा (Common noun) के साथ कारक चिह्न “*ti*” (ती) “*iMiti*” (इमीटी) का प्रयोग होता है और व्यक्तिवाचक संज्ञा (Proper Noun) में “*du*” (दू) “*rAmudu*” (रामुदू) का प्रयोग होता है। लेकिन हिंदी में इसके लिए ‘KA’ (का) एक ही कारक चिह्न का प्रयोग होता है।

Rule: $NP\sim 1(\{\langle case="o", tam="ti">\})NP\sim 2 \Rightarrow NP\sim 1(\{\langle case="o", cm="kA">\})NP\sim 2$

2. “है” का प्रयोग

स्रोत भाषा (तेलुगु) में संज्ञा पद (NP~1) के साथ विशेषण (NP~2) का प्रयोग किया जाता है, तो लक्ष्य भाषा (हिंदी) वाक्य के अंत में काप्यूल योजक (Copula) का प्रयोग किया जाता है।

¹² Avinesh P.V.S (2010) : *Transfer Grammar Engine and Automatic Learning of Reorder Rules in Machine*

Translation (M.Sc- Thesis), p.04

¹³ Mala Christopher (2009) : “*Implementation of Transfer Grammar in Telugu-Hindi Machine Translation System*”

Proceedings of SCONLI-2009: 3rd Students Conference of Linguistics in India (research paper)

उदा.1. (तेलुगु) *रामुडु मंचीवाडु* (*rAmudu maNchivAdu*)

(हिंदी) *राम अच्छा लडका है।*

(Ram is a good boy.)

उपर्युक्त दोनों वाक्यों की संरचना पर ध्यान देने से यह ज्ञात होता है कि हिंदी वाक्य में “है” का प्रयोग योजक के रूप में हुआ है, लेकिन तेलुगु में इस तरह के किसी योजक का प्रयोग नहीं होता। उपर्युक्त वाक्य के आधार पर इस नियम का निर्माण किया गया है।

Rule: NP~1 NP~2({{<lcat= “adj”>}}) => NP~1 NP~2 +VGF ({{hai%VM<root= “hai”, lcat= “v”, gen= “m”, num= “sg”, per= “3”, tam= “hai”>}})

इस नियम के आधार पर इसे हिंदी में लिखा जा सकता है।

नियम : संप.1 संप.2 (विशे.) → संप.1 संप.2 + क्रिया (मूल शब्द “है”)

उदा.2. अगर स्रोत भाषा (तेलुगु) में क्रिया (verb) का प्रयोग नहीं है तो लक्ष्य भाषा वाक्य के अंत में “है” का प्रयोग किया जाता है।

उदा. (तेलुगु) *राम लक्ष्मनुलु अन्नादम्मलु* (*rAma laxmaNulu annadammulu*)

(हिंदी) *राम और लक्ष्मण भाई हैं।*

(Ram and laxman are brothers.)

Rule: VGF ({{. %SYM}}) => VGF ({{+ {<root= “hai”, lcat= “v”, tam= “hai”>}} {.%SYM}})

नियम : क्रिया → क्रिया (मूल शब्द “है”)

इस नियम में, VGF (क्रिया) के लिए प्रयोग हुआ है, उदाहरण के तेलुगु वाक्य से हिंदी में स.क्रिया के रूप में “है” का प्रयोग हुआ है।

3. ‘और’ का प्रयोग

अगर स्रोत भाषा (तेलुगु) में दो संज्ञापद (NP~1, NP~2) लंबे स्वर के साथ हैं, तो वाक्य में कारक का प्रयोग निहित है। लेकिन लक्ष्य भाषा (हिंदी) में दो संज्ञापदों के बीच में समुच्चयबोधक (Conjunction) का प्रयोग होता है।

उदा. (तेलुगु) *ई तरगतिलो कुर्चीलु बल्लेलु उन्नाई.* (*ee taragatilo kurchilu ballalu unnayi*)

(हिंदी) *इस कक्षा में कुर्सी और मेज हैं।*

(There are chairs and tables in this class)

इस तेलुगु वाक्य के हिंदी में सम्मिलन से 'और' इस समुच्चयबोधक शब्द का प्रयोग हुआ है, लेकिन तेलुगु में इसका प्रयोग नहीं हुआ है। इस आधार पर यह नियम देख सकते हैं।

Rule: NP~1((({<cm="A">}))) NP~2((({<cm="A">}))) => NP~1((({<cm="O">}))) + CCP ((({Aur %CC<root="Aur", lcat="conj">}))) NP~2((({<cm="A">})))

नियम: सप.1 सप 2 (कारक “लो”) → सप 1 (कारक “में”) + (समु. “और”)

इस नियम से तेलुगु में 'तरगतिल्लो' शब्द में कारक का प्रयोग निहित है, साथ ही वाक्य में समुच्चयबोधक शब्द का प्रयोग नहीं हुआ है। लेकिन हिंदी में कारक का प्रयोग और समुच्चयबोधक शब्द “और” का प्रयोग भी हुआ है।

4. 'ने' का प्रयोग

स्रोत भाषा (तेलुगु) में, संज्ञापद 1 (NP~1) का संज्ञापद 2 (NP~2) और क्रिया पदबंध (VP) के साथ प्रत्यक्ष प्रयोग होता है। तो लक्ष्य भाषा में संज्ञापद 1 (NP~1) के साथ “ने” का प्रयोग होता है, और बाकी संरचना तेलुगु वाक्य की तरह ही होती है।

उदा. (तेलुगु) रामु पुस्तकम छदीवाडु (rAmu pustakaM chadivAdu)

(हिंदी) राम ने पुस्तक पढ़ी।

(Ram read a book)

Rule: NP~1 ((({<case="d",cm="">}))) NP~2 ((({<case="o">}))) VGF((({<tam="A">}))) => NP~1 ((({<case="o",cm="ne">}))) NP~2((({<case="o">}))) VGF((({<tam="A">})))

नियम: सप. 1 सप. 2 (कारक “0”) क्रिप. (क्रिया “छ.”) → सप. 1 (कारक “ने”) सप. 2 (कारक “0”) क्रिप. (क्रिया “प.”)

तेलुगु के नियम में संज्ञापद का क्रिया पदबंध के साथ प्रत्यक्ष प्रयोग हुआ है, तो हिंदी में संज्ञापद के साथ कारक चिह्न “ने” का प्रयोग हुआ है।

5. तेलुगु में “samayaAnni” (समयानी) शब्द का मूल शब्द “samayam” (समय) है और “ni” (नि) कारक चिह्न है। लेकिन हिंदी में “samaya” (समय) के साथ कोई कारक चिह्न का प्रयोग

नहीं हुआ है। इसका मतलब तेलुगु के कुछ वाक्यों में कारक चिह्न का प्रयोग किया जाता है, तो हिंदी के कुछ वाक्य में कारक चिह्न का लोप (dropped) होता है।

उदा. (तेलुगु) समयानी

(हिंदी) समय

Rule: NP (((<root="samayaM",cm="ni">))) => NP (((<root="samayaM",cm="O">)))

नियम : संप. (मूल शब्द “समय, कारक “नि”) → संप. (मूल शब्द “समय”, कारक “0”)

6. तेलुगु भाषा के वाक्य में, अपरिमित पुनरुक्त क्रिया (non-finite reduplicated verb) का प्रयोग काल पुनरुक्त (tense reduplication) में भी होता है, जैसे- चेप्पी चेप्पी शब्द का प्रयोग। लेकिन हिंदी में काल पुनरुक्ति (tense reduplication) का प्रयोग नहीं होता। हिंदी में पुनरुक्ति क्रिया के लिए उपयुक्त शब्द “बता बता कर” का प्रयोग हो सकता है लेकिन इस तरह का शब्द प्रयोग हिंदी में किसी विशेष संदर्भ में ही किया जाता है।

उदा. (तेलुगु) चुस्सी चुस्सी / विनी विनी

(हिंदी) देख देख कर / सुन सुन कर

Rule: VGNF~1(((<lcat="v", tam="i">))) VGNF~2(((<lcat="v",tam="i">))) => VGNF~1(((<tam="0">))) VGNF~2(((<tam="i">)))

7. तेलुगु में विधेय क्रिया के अंत में “-ani” (-अनि) का प्रयोग होता है। उदाहरण के लिए क्रिया “cheppamani” (चेप्पामनि) का अर्थ “कह रहा” है। जिसका “चेप्पु” यह आधार रूप है। हिंदी में कृदंत विशेषण (The participle) “ani” (-अनि) का अर्थ “रहा” है। तेलुगु में “-आनि” का किसी शब्द के साथ प्रयोग होता है जैसे (चेप्पामनि)। लेकिन हिंदी में “रहा” यह शब्द सहायक क्रिया है। इसे मुख्य क्रिया के साथ परसर्ग के रूप में निर्दिष्ट किया गया है।

Rule: VGF (((<tam="\$x.ani">))) => VGF (((<tam="\$x">))) +NP (((ani %PSP<root="ani", lcat="psp">)))

उदा. (तेलुगु) मोहन वाल्लु पणी चेस्तुन्नारानी चेपेडु (mohan vaallu panii chestunnarani cheppedu)

(हिंदी) मोहन ने ऐसा कहा वे लोग काम कर रहे हैं।

(Mohan said that they are doing the work)

उपर्युक्त वाक्य के सम्मिलन में तेलुगु में क्रिया के अंत में “-ani” (-अनि) शब्द का प्रयोग कारक के रूप में हुआ है, जिसके लिए हिंदी में “ने” कारक चिह्न का प्रयोग हुआ है। इस उदाहरण के आधार पर नियम का निर्माण किया गया है -

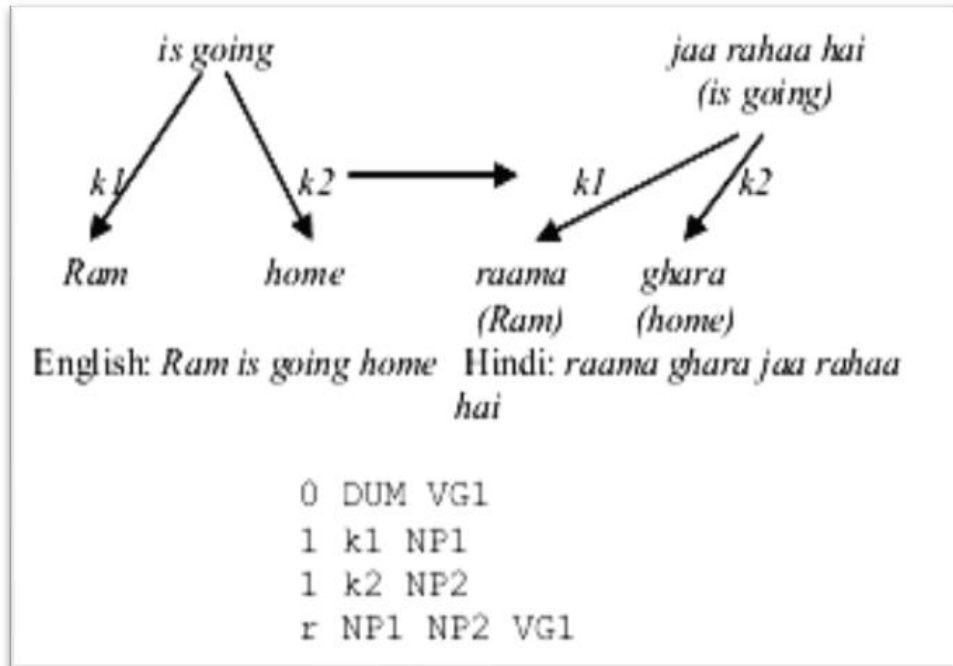
Rule: $VG (\{< tam = “\$x ani” > \}) \Rightarrow VG (\{< tam = “\$x” > \}) + NP (\{ @ aisaa \% PRP \{ root = “@ aisaa” \} \})$

नियम: क्रिया (“-अनि”) \longrightarrow क्रिया (“0”) + संप. (PRP “ऐसा”)

उपर्युक्त सभी स्थानान्तरण व्याकरण के नियमों का निर्माण तेलुगु-हिंदी भाषा की संरचना के सम्मिलन से किया गया है। इन नियमों का निर्माण संपर्क तंत्र को विकसित करने के लिए हैदराबाद विश्वविद्यालय में किया गया है। इस संपर्क मॉड्यूल के आधार पर शक्ति तंत्र का भी विकास नियम आधारित प्रणाली पर किया गया है। इस शक्ति तंत्र कि चर्चा नीचे प्रस्तुत है।

3.5 शक्ति मशीनी अनुवाद तंत्र में स्थानान्तरण व्याकरण का प्रयोग

शक्ति (2004) तंत्र नियम आधारित प्रणाली पर आधारित है। यह अंग्रेजी से हिंदी में अनुवाद प्रदान करता है। इस तंत्र में स्थानान्तरण व्याकरण का प्रयोग पुनःक्रम संरचना (Reordering) के लिए किया गया है।



आकृती 3.3 : शक्ति मशीनी अनुवाद तंत्र में प्रयुक्त स्थानान्तरण व्याकरण नियम¹⁵

इस आरेख के अंतर्गत दिए गए उदाहरण - अंग्रेजी में 'Ram' (NP1), 'is going' (VG1) और 'home' (NP2) है। हिंदी में 'राम' संज्ञा (NP1) है, और 'जा रहा है' क्रिया (VG1) है और 'घर' संज्ञा² (NP 2) है। वाक्य को तीन क्रमों में विभाजित किया गया है, जिसमें अंग्रेजी के स्रोत वाक्य को हिंदी में विश्लेषित किया गया है। वाक्य में दो संज्ञाएं हैं, जिसका संबंध k1 (राम) और k2 (घर) के साथ दर्शाया गया है। r (वाक्यसंरचना) यह स्रोत भाषा के वाक्य को लक्ष्य भाषा के वाक्य का क्रम परिवर्तन प्रदर्शित कर रहा है। इस तंत्र में अंग्रेजी हिंदी शब्द आरेखित समांतर कार्पस (word aligned parallel corpus) का प्रयोग किया गया है। जिसमें अंग्रेजी वाक्य को आश्रयता पार्सर (Dependency Parser) के द्वारा विश्लेषित किया जाता है और हिंदी वाक्य के समांतर रखा जाता है। फिर शब्द के आधार पर लक्ष्य भाषा में अनुवाद किया जाता है।

इसमें अनुवाद की सुविधा के लिए कारक-संबंध का समावेश किया गया है। भारतीय पारंपरिक भाषा में कारक का अधिक महत्व है। कारक संबंध वाक्यात्मक और अर्थात्मक संबंध की जानकारी सूचित करता है। इन कारकों में कर्ता, कर्म, करण, संप्रदान, अपादान, अधिकरण आते हैं। इन कारकों के लिए इस तंत्र में नियम बनाए गए हैं। इन नियमों का निर्माण प्रशिक्षित भाषावैज्ञानिकों के द्वारा किया गया है। वर्तमान में इस तंत्र में मानव निर्मित नियमों का प्रयोग किया जा रहा है। लेकिन यह नियम सभी कारकों के लिए नियमबद्ध करने में पूर्णतः सक्षम नहीं है।¹⁴

स्थानांतरण व्याकरण का प्रयोग वाक्य स्तर पर होता है। इस व्याकरण में कारक के नियमों का समावेश करने के लिए फिल्मोर के कारक व्याकरण का समावेश करना आवश्यक है। वाक्य की संरचना पर भाषावैज्ञानिक चाम्स्की और फिल्मोर का कार्य वाक्यात्मक संबंध पर है। चाम्स्की ने रूपांतरण प्रजनक व्याकरण का निर्माण किया और फिल्मोर ने वाक्य का अर्थ संबंध स्पष्ट करते हुए हिंदी भाषा के लिए कारक व्याकरण के नियम बताएँ हैं। इन नियमों को उदाहरण के द्वारा देख जा सकता है -

Rule: C → NP + K

कारक → संज्ञापद + परसर्ग

इस नियम में कारक का विस्तार संज्ञापद और परसर्ग के रूप में किया गया है। परसर्ग के लिए नियम में k का प्रयोग हुआ है। फिल्मोर ने स्पष्ट किया है कि वाक्य में कारक कोटि का चयन

¹⁴ Yadav, Vineet, (ed.) (2009) : "Inducing Transfer Grammar from Word Aligned Corpus", Proceedings of ICON-2009: 7th International Conference on Natural Language Processing, Macmillan Publishers, India.

आवश्यक है, क्योंकि कारक का संबंध भाषिक तत्त्व से है जो भाषा के आधारभूत घटक से संबंध रखते हैं।

Rule: संज्ञा पदबंध → { सर्वनाम
निर्धारक संज्ञा }

संज्ञा पदबंध का विकास सर्वनाम के रूप में या निर्धारक के रूप में हो सकता है। संज्ञापद वाक्य में महत्वपूर्ण भूमिका निभाते हैं। कारक के इस नियम को नीचे वाक्यात्मक उदाहरण से देख सकते हैं।

Rule: अभि + (कारक) – अभि. + स. + (कारक)/ – (क्रिया + प्रेरणार्थक)

मैंने लड़के को सुलाया।

“I put the boy to sleep”

उपर्युक्त नियम के आधार पर इस वाक्य में यदि किसी क्रिया के मूल कारक साँचों में अभिकर्ता कारक हो (लेकिन अभिकर्ता के साथ और कारक हो या न हो) और वह क्रिया [+ प्रेरणार्थक] गुण से चिह्नित हो तो उस क्रिया के मूल कारक साँचों में संप्रदान कारक जुड़ जाएगा। उक्त वाक्य में मूल क्रिया ‘सोना’ का कारक साँचा अभिकर्ता है, लेकिन जब क्रिया[+प्रेरणार्थक] गुण से युक्त हो जाती है तब इसका कारक साँचा अभि + सं. हो जाता है। यह वाक्य के आधार से स्पष्ट होता है।¹⁵

इस तरह फिल्मोर ने कारक का संबंध अर्थ तत्त्व से जोड़कर इन नियमों का निर्माण किया है। इन कारकों का प्रयोग अधिकांश भारतीय भाषाओं में होता है। जब अनुवाद के स्तर पर दोनों भाषाओं की भिन्नताएँ स्पष्ट होती हैं, तब कारक का संबंध वाक्य में अर्थतत्त्व की दृष्टि से महत्वपूर्ण हो जाता है।

हिंदी अंग्रेजी संरचना के स्तर पर स्थानान्तरण व्याकरण नियम नमूने के तौर पर यहाँ स्पष्ट किए गए हैं। हिंदी वाक्य में ‘अच्छा लड़का’ और ‘अच्छी लड़की’ के लिए अंग्रेजी में ‘Good boy’ और ‘Good girl’ अनुवाद होता है। अंग्रेजी वाक्य में विशेषण पर लिंग, वचन का कोई प्रभाव नहीं पड़ता, लेकिन हिंदी में विशेषण पर लिंग, वचन का प्रभाव पड़ता है। उसी तरह दूसरा वाक्य हिंदी में ‘बूढ़ा आदमी’ का अंग्रेजी में ‘The old man’ अनुवाद हुआ है। अंग्रेजी वाक्य

¹⁵ बालचंद्रन, लक्ष्मीबाई (1988): हिंदी का कारक व्याकरण, पृ. 78-79, 81.

के सम्मिलन में 'The' निर्धारक का प्रयोग हुआ है, लेकिन हिंदी में इस तरह कोई शब्द प्रयुक्त नहीं होता। इसलिए विशेषण के तौर पर यहां इस नियम का निर्माण किया जा सकता है।

नियम: स.प.1 : स.प.2 [नि.+विशे.+ स.] → [विशे. + स.]

इस नियम के आधार पर हिंदी-अंग्रेजी की संरचनात्मक भिन्नता के कॉर्पोरा आधारित संज्ञापद सम्मिलन (Mapping) के बाद चतुर्थ अध्याय में स्थानान्तरण व्याकरण के नियम का निर्माण किया जाएगा। यह नियमों का निर्माण निम्नलिखित घटकों का सम्मिलन करने के बाद किया जाएगा -

- हिंदी के परसर्ग और अंग्रेजी के पूर्वसर्ग के लिए स्थानान्तरण व्याकरण नियम
- हिंदी और अंग्रेजी के कारकों के लिए स्थानान्तरण व्याकरण नियम
- संयुक्त वाक्यसंरचना और समुच्चयबोधक के लिए स्थानान्तरण व्याकरण नियम
- हिंदी - अंग्रेजी विशेषण के लिए स्थानान्तरण व्याकरण नियम
- हिंदी - अंग्रेजी संज्ञापद के साथ प्रयुक्त संख्यावाचक के लिए स्थानान्तरण व्याकरण नियम
- व्याकरणिक कोटियों के लिए स्थानान्तरण व्याकरण नियम
- हिंदी - अंग्रेजी शब्दक्रम के लिए स्थानान्तरण व्याकरण नियम

इन नियम का निर्माण सैद्धांतिक रूप से किया जाएगा। वास्तव में इन नियमों का निर्माण व्यावहारिक रूप से होना महत्वपूर्ण है। व्यावहारिक तौर पर इन नियमों का समावेश अल्गोरिदम के माध्यम से कंप्यूटर में किया जाता है। इन समावेशित नियमों से म. अनु. तंत्र बहुत अधिक कार्यक्षम रूप से अनुवाद कर सकता है। पाश्चात्य देशों में स्थानान्तरण व्याकरण पर किए गए कार्य से मशीनी अनुवाद के अनेक तंत्रों का निर्माण किया गया है। आज भारत में सभी भाषाओं को प्राकृतिक भाषा संसाधन (NLP) के साथ जोड़कर इन नियमों का निर्माण किया जा रहा है। अतः यह कह सकते हैं कि आज कोई भी भाषा बिना संसाधन के नहीं रह सकती, इसलिए सभी भाषाओं का विकास होना आवश्यक है।

निष्कर्ष:

स्थानान्तरण व्याकरण नियम का प्रयोग किसी भी मशीनी अनुवाद तंत्र में किया जा सकता है। जब दो भाषाओं में संरचनात्मक भिन्नता पायी जाती है, तो दोनों भाषाओं की संरचनात्मक भिन्नता को सम्मिलन के द्वारा प्राप्त किया जा सकता है। इस प्राप्त सम्मिलन के आधार पर किसी भी भाषा के लिए नियमों का निर्माण किया जा सकता है। हिंदी और अंग्रेजी भाषा में उसी अनुरूप संरचनात्मक भिन्नता है। यह भिन्नता शब्दक्रम से लेकर परसर्ग-पूर्वसर्ग, संयुक्त और

समुच्चयबोधक वाक्य, विशेषण, उपपद, लिंग, वचन, कारक आदि तक में पायी जाती है। इस भिन्नता को दूर करने के लिए सशक्त नियमों का निर्माण करना आवश्यक है। इसीलिए इस अध्याय में कुछ स्थानान्तरण व्याकरण नियम आधारित संपर्क और शक्ति तंत्र के नियमों के उदाहरण के साथ चर्चा की गयी है। चतुर्थ अध्याय में हिंदी अंग्रेजी संज्ञापद सम्मिलन के आधार पर इस तरह नियमों का निर्माण किया जाएगा।

अध्याय – चतुर्थ

कॉर्पोरा आधारित हिंदी - अंग्रेजी
संज्ञापद का सम्मिलन

अध्याय - चतुर्थ

कॉर्पोरा आधारित हिंदी - अंग्रेजी संज्ञापद का सम्मिलन

(Corpora based Hindi-English Noun Phrase Mapping)

4.1 सम्मिलन का परिचय (Introduction of Mapping)

तृतीय अध्याय में स्थानांतरण व्याकरण (Transfer Grammar) का परिचय और स्थानांतरण व्याकरण का मशीनी अनुवाद (MT) में प्रयोग किस तरह होता है, यह देखा गया है। साथ ही, संपर्क मशीनी अनुवाद तंत्र (SMTS) के लिए स्थानांतरण व्याकरण के बनाए गए नियमों (Rules) की चर्चा की गई है।

चतुर्थ अध्याय में, मशीनी अनुवाद हेतु-कॉर्पोरा (Corpora) के आधार पर हिंदी-अंग्रेजी संज्ञापद (Noun Phrase) का सम्मिलन (Mapping) किया गया है। इस सम्मिलन के द्वारा हिंदी से अंग्रेजी में प्राप्त अनुवाद के संज्ञापद स्तर पर होने वाले बदलाव को सम्मिलित किया गया है और अनुवाद का सुधार करने के लिए मशीन के लिए सैद्धांतिक नियम बनाए गए हैं। यह शोध मुख्य रूप से कॉर्पोरा के अध्ययन पर आधारित है जिसके लिए 'भारतीय भाषा कॉर्पोरा उपक्रम' (ILCI) का हिन्दी भाषा में 50,000 वाक्य का कॉर्पोरा है, इस शोध में 2000 वाक्यों को नमूने के तौर पर लिया गया है। इसका स्वरूप स्वास्थ्य (health) और पर्यटन (tourism) क्षेत्र का है, जिसका मूलतः हिन्दी से अंग्रेजी में अनुवाद कर दो भाषाओं का समानान्तर कॉर्पोरा तैयार किया है। उसे टैग करने के लिए 'भारतीय मानक ब्यूरो' (Bureau of Indian Standards- BIS) टैगसेट का प्रयोग किया है और अंग्रेजी कॉर्पोरा को टैग करने के लिए Penn. (University of Pennsylvania) टैगसेट का प्रयोग किया गया है।

मशीनी अनुवाद के लिए स्रोत भाषा हिंदी और लक्ष्य भाषा अंग्रेजी में अनुवाद के लिए पाठ दिया जाए, तो मशीन द्वारा प्राप्त अनुवाद में इन दो भाषाओं की संरचना में बदलाव पाया जाता है। यह बदलाव इस अध्याय में संज्ञापद स्तर पर देखना है, जिसके लिए संज्ञापद का सम्मिलन (Mapping) किया जाएगा। इस सम्मिलन के द्वारा यह ज्ञात होगा कि हिंदी भाषा के कौनसे घटकों का अंग्रेजी भाषा के अनुवाद में बदलाव हुआ है और अंग्रेजी में कौन से घटकों का समावेश किया गया है। इस अनुवाद के बदलाव में अंग्रेजी प्रजनक (English Generator) की भूमिका प्रमुख होती है। वह हिंदी भाषा संरचना के वाक्य को अंतरण शब्दकोश, द्विभाषी व्याकरण, पार्सर, आदि के द्वारा अंग्रेजी की संरचना में अनुवाद करता है। यह अनुवाद

वाक्यात्मक, शाब्दिक, तथा अर्थात्मक हो सकता है लेकिन, मशीन से अर्थात्मक दृष्टि से अनुवाद प्राप्त करने के लिए उपर्युक्त घटकों का समावेश किया जाना आवश्यक है। वाक्य में संज्ञापद का विशेष महत्व होता है। इस शोध में मशीनी अनुवाद के लिए संज्ञापद के सम्मिलन हेतु हिंदी से अंग्रेजी भाषा के कारक के साथ प्रयुक्त होने वाले परसर्ग-पूर्वसर्ग, सर्वनाम, विशेषण, संख्यावाचक विशेषण, समुच्चयबोधक, लिंग, वचन, शब्दक्रम आदि का अध्ययन किया गया है।

4.2 हिंदी - अंग्रेजी संज्ञापद सम्मिलन के प्रमुख घटक

एक भाषा का दूसरी भाषा में सम्मिलन करने का तात्पर्य यह है कि अनुवाद करते समय दोनों भाषाओं की संरचना की तुलना करना। इसमें भाषा में होने वाले बदलाव और जुड़ने वाले घटकों को सम्मिलित किया जा सकता है। यहाँ स्रोत भाषा हिंदी है और लक्ष्य भाषा अंग्रेजी है। इस भाषा के अनुवाद स्वरूप दोनों में भिन्नता दिखाई देगी। इस भिन्नता को निम्नलिखित बिंदुओं के आधार पर देखा जा सकता है। हिंदी में संज्ञापद के साथ परसर्ग का प्रयोग संज्ञा के बाद होता है। जैसे- (राम ने) लेकिन अंग्रेजी में पूर्वसर्गों का प्रयोग संज्ञा के पहले होता है। जैसे- (on the table) इस आधार पर हिंदी में परसर्ग 'ने' का प्रयोग तो अंग्रेजी में 'on' का प्रयोग परसर्ग और पूर्वसर्ग के क्रम को दर्शाता है। दूसरी ओर हिंदी वाक्य के परसर्ग का अंग्रेजी पूर्वसर्ग में संदर्भ और अर्थ की दृष्टि से लोप हो जाता है। जैसे- 'प्रेमशंकर ने पुस्तक पढ़ी'। (Premshankar read a book) इस वाक्य के सम्मिलन में हिंदी परसर्ग 'ने' का अंग्रेजी में लोप हुआ है। इस आधार पर हिंदी परसर्ग और अंग्रेजी पूर्वसर्गों का सम्मिलन अन्य बिन्दुओं या घटकों के साथ किया जा सकता है।

- हिंदी से अंग्रेजी भाषा में मशीनी अनुवाद की महत्वपूर्ण समस्या **शब्दक्रम (Word order) या पदक्रम** की है। इसे मशीन के द्वारा अनुवाद प्राप्त करने के लिए नियम बनाने की आवश्यकता सर्वप्रथम होती है। जिससे उन भाषाओं की संरचना को देखते हुए मशीन के लिए नियम बनाकर उससे अच्छा अनुवाद प्राप्त किया जा सकता है। हिंदी भाषा में शब्दक्रम का स्वरूप कर्ता-कर्म-क्रिया (SOV) है, इस क्रम संरचना में क्रिया का स्थान वाक्य के अंत में होता है और अंग्रेजी संरचना का शब्दक्रम कर्ता-क्रिया-कर्म (SVO) है। यह क्रम संरचना क्रिया, कर्ता और कर्म के बीच प्रयुक्त होती है। हिंदी के वाक्य में पदक्रम अधिक लचीला होता है। जैसे- 'मैंने ये दवाइयाँ सरोजिनी मार्केट से खरीदीं' 'ये दवाइयाँ मैंने खरीदीं सरोजिनी मार्केट से' (I purchased these medicines from Sarojini market.) इस तरह से हिंदी में और भी वाक्यों का निर्माण हो सकता है। इससे इस

पदक्रम को सम्मिलित करना अनुवाद के लिए महत्वपूर्ण होता है, जिसमें भाषा की संरचना के अनुसार अनुवाद प्राप्त करने के लिए नियम बनाए जा सकते हैं।

- हिंदी-अंग्रेजी की संरचना में हिंदी के परसर्ग (Postposition) और अंग्रेजी के पूर्वसर्ग (Preposition) में अंतर पाया जाता है। संस्कृत में परसर्गों का समावेश कारक की विभक्तियों में किया जाता है, लेकिन हिंदी में 'परसर्ग' का प्रयोग नए रूप में किया जा रहा है। (सिंह, 2003, पृ. 35) जब अनुवाद में स्रोत भाषा (हिंदी) और लक्ष्य भाषा (अंग्रेजी) की वाक्यसंरचना के शब्दों के क्रम में स्थानांतरण होगा, तो हिंदी में विशेषक (Modifier) संज्ञा से पहले प्रयुक्त होगा। लेकिन अंग्रेजी में विशेषक (Modifier) संज्ञा से पहले और बाद, दोनों स्थानों पर प्रयुक्त होते हैं। जैसे- 'मंडल के सभी सदस्य' (All the members on the board), इन भिन्नता को सम्मिलित (Mapping) करने की आवश्यकता है। कारक के लिए नियम- हिन्दी में कारकों- कर्ता, कर्म, करण, संप्रदान, अपादान, अधिकरण, संबंधकारक, संबोधन कारक में परसर्गों का प्रयोग होता है। हिंदी परसर्गों में *ने, से को, पर, में, को, का, की, के लिए* आदि का समावेश होता है तो अंग्रेजी में *by, for, from, to, of, on, at, in, with* आदि पूर्वसर्गों का समावेश तथा प्रयोग संज्ञा के साथ होता है।
- हिंदी में संज्ञा या संज्ञा पदबंध के बदले सर्वनाम का प्रयोग होता है। हिंदी के पुरुषवाचक सर्वनाम के लिए अंग्रेजी में विभिन्न निम्न सर्वनामों का प्रयोग होता है जैसे- उत्तम पुरुष में (*मैं-I*), (*हम-we*) मध्यम पुरुष में (*आप, तुम, तू- you*) अन्य पुरुष में (*वह-he, she, that, यह, ये, वे- It, this, these, they, those*) आदि। हिंदी में सर्वनाम का प्रयोग कर्ताकारक के रूप में होता है जिसके साथ कारक प्रयोग समावेशित होता है। जैसे- (*मुझको- to me, हमको- to us, इसको/उसको-to him, तुझको/तुमको- To you*) आदि सर्वनाम। उसी प्रकार अपादान कारक में सर्वनाम के साथ 'से' का प्रयोग होता है लेकिन अंग्रेजी में इसके लिए पूर्वसर्ग शब्द के साथ जुड़कर प्रयोग नहीं होता। एकवचन और बहुवचन में (*मुझसे-me, तुझसे-you, आपसे you, इससे-him, her, it, them*) का प्रयोग होता है। कर्ताकारक 'ने' का भी सर्वनाम के साथ जुड़कर प्रयोग होता है। जैसे- (*मैंने, तुमने, आपने*) आदि।

हिंदी-अंग्रेजी में सर्वनाम सम्मिलन की दृष्टि से समस्या आदरार्थ शब्द के प्रयोग में आती है जिसके लिए हिंदी के समान अंग्रेजी में कोई आदरसूचक शब्दों का प्रयोग नहीं होता। सामान्यत आदरार्थ में 'आप' शब्द का प्रयोग होता है, फिर भी आदरार्थ के लिए संदर्भ के अनुसार विभिन्न शब्दों का प्रयोग होता है। उदाहरण के लिए- 'उनके पिता आए हैं' (*His father has come.*) इस वाक्य में 'उनके' शब्द आदरार्थ प्रयुक्त हुआ है जिसके लिए अंग्रेजी

में 'his' का प्रयोग हुआ है। इस वाक्य से अन्य रूप में भी आदरार्थ शब्द का प्रयोग होता है, जिसका अंग्रेजी अनुवाद के लिए सम्मिलन करना आवश्यक है।

- **हिंदी-अंग्रेजी में विशेषण** को लेकर भी भिन्नताएँ हैं। हिंदी में विशेषण पर लिंग-वचन-कारक का परिणाम या परिवर्तन होता है लेकिन अंग्रेजी में लिंग-वचन के कारण इस प्रकार का कोई परिवर्तन नहीं होता। जैसे- *सुंदर कमरा, सुंदर कमरे, सुंदर लड़की, (Large room/ rooms), अच्छे लड़के (Good boys) अच्छी लड़कियाँ (good girls)* दूसरी ओर हिंदी में द्विरुक्ति विशेषण का प्रयोग होता है जैसे- 'अच्छे-अच्छे पकवान' के लिए अंग्रेजी में (*good-good food*) का प्रयोग नहीं होता बल्कि (*good food*) का ही प्रयोग होगा। विशेषण के अन्य प्रकारों में हिंदी के संख्यावाचक विशेषणों का प्रयोग संज्ञा के रूप में किया जाता है। यह अंग्रेजी में अनुवाद सम्मिलन के लिए महत्वपूर्ण हो जाता है। जैसे- 'एक चम्मच नमक' (a spoon of salt) में 'एक' लिए उपपद (निर्धारक) 'a' का प्रयोग हुआ है। दूसरा वाक्य 'चार-चार घंटे पर फ़ीवर को मापें' (*Measure the fever every four hours.*) के सम्मिलन से ज्ञात होता है कि हिंदी और अंग्रेजी में संख्यावाचक विशेषण का प्रयोग भी अर्थ की दृष्टि से वाक्य के विभिन्न संदर्भों में संज्ञा के साथ होता है। हिंदी में -वाला एक बहुप्रयोजनी विशेषक का प्रयोग होता है जो अंग्रेजी में अनेक तरह से व्यक्त होने वाले विशेषकों की जगह प्रयुक्त होता है, जिनमें क्रियाओं के सहयोग से बनने वाले संज्ञा पदबंध और विशेषणों से बने संज्ञा पदबंध शामिल हैं। जैसे- *लंबे काले बालों वाली महिला (The lady with long black hair)* (सिंह, 2003:91) आदि वाक्य के सम्मिलन से यह ज्ञात होता है कि हिंदी में -वाला विशेषक के लिए अंग्रेजी में किसी शब्द का प्रयोग नहीं होता इसलिए भाषिक संरचना का सम्मिलन करना आवश्यक है।
- कारक के परसर्ग और पूर्वसर्गों का जिस तरह से संज्ञा या सर्वनाम के साथ संबंध होता है, उसी तरह **संबंधसूचक अव्यय (Indeclinable)** का संबंध संज्ञा के साथ होता है। जैसे- '*लड़का नहाने को गया है*' (*The boy has gone to take a bath*) अथवा '*नहाने के लिए गया/गई है*' (*he/she has gone to take a bath*) । इसके साथ संबंधसूचक से जितने संबंध आते हैं उनमें हिंदी कारक परसर्ग का प्रयोग नहीं होता। जैसे- *लड़का नदी तक गया (the boy went up to the river) मुसाफिर पेड़ तले बैठा है (Passenger has sat under the tree), नौकर साँप के पास पहुँचा (The servant reached near the snack)* आदि वाक्यों में अंग्रेजी के अधिकरण कारक पूर्वसर्ग के अंतर्गत लिंग, वचन, के अनुसार शब्दवर्ग में बदलाव नहीं होता। जिन भाषाओं में विभक्तियों की संख्या अधिक रहती है (जैसे संस्कृत में है) उनमें संबंधसूचकों का प्रयोग अधिक नहीं होता। लेकिन भिन्न-भिन्न भाषाओं में रूप के जो भेद

दिखाई देते हैं, उनका विशेष कारण यही है कि संबंधसूचकों का उपयोग किसी में स्वतंत्र रूप से तथा किसी में प्रत्यय के रूप में होता है। (गुरु, 2011:196) इसलिए इन वाक्यों के प्रयोग से संबंधसूचक का हिंदी से अंग्रेजी में सम्मिलन करना आवश्यक हो जाता है।

- संयुक्त और मिश्र वाक्य में **समुच्चयबोधक (Conjunction)** का प्रयोग होता है। इस समुच्चयबोधक का प्रयोग वाक्य में संज्ञा के साथ भी होता है। जैसे- *राम और श्याम जा रहे हैं। (Ram and Shyam are going) रमेश आएगा या उसके पिता फोन करेंगे। (Ramesh will come or his father will call.)* समुच्चयबोधक के दो प्रकार हैं- एक है समानाधिकरण (Coordinating) और दूसरा है अधीनतासूचक (Subordinating)। जिनमें *और, एवम, तथा, व* के लिए *and* और *या, वा, अथवा* के लिए संदर्भ के अनुसार अंग्रेजी में (*or*) का प्रयोग होता है। संज्ञा के साथ प्रमुख रूप से (और-and, या/वा-or, पर-but, नहीं तो-otherwise, यदि-if) आदि संयोजकों का प्रयोग होता है। हिंदी में संयोजक ‘-कि’ का प्रयोग भी होता है जैसे-*जया मुझसे मिली न ताकि उसको (Jaya met me not him)। दूसरा प्रश्नवाचक वाक्य ‘श्याम पढ़ता है कि सोता है?’ (Does Shyam study or sleep?)* इन वाक्यों का सम्मिलन समुच्चयबोधक के प्रयोग से हिंदी अंग्रेजी के लिए महत्वपूर्ण है।
- अंग्रेजी में **उपपदों का प्रयोग** होता है, लेकिन हिंदी में उपपदों का प्रयोग नहीं होता लेकिन उसके सजात दूसरे रूपों का प्रयोग होता है। अंग्रेजी में उपपद (a, an, the) का प्रयोग होता है। जैसे- ‘*एक लड़की आपसे यहाँ मिलने आयी थी। (A girl had been here to see you.)* इस वाक्य में ‘एक’ के लिए अंग्रेजी में ‘a’ का प्रयोग हुआ है। इस तरह हिंदी से अंग्रेजी में अनुवाद करते समय संदर्भ के अनुरूप ‘एक’ का अनुवाद अंग्रेजी में ‘a’ होता है। मगर दूसरे वाक्य में ‘*चार लड़के आपसे यहाँ मिलने आए थे*’ का अनुवाद (*four boys had been here to see you*) होता है। यह मशीन के अनुवाद के लिए चुनौतियाँ पैदा करता है कि संदर्भ के अनुसार किस तरह अनुवाद करें।
- **हिंदी में दो लिंग** हैं पुल्लिंग और स्त्रीलिंग। हिंदी में लिंग के अंतर्गत प्राणीवाचक (animate) संज्ञाओं का समावेश जैसे लड़का (पु.) और लड़की (स्त्री.) में किया है, उसी तरह अप्राणीवाचक या अचेतन (inanimate) संज्ञाओं को भी यादृच्छिक (arbitrary) रूप से पुल्लिंग और स्त्रीलिंग वर्गों में रखा गया है। दूसरी ओर अंग्रेजी में तीन लिंग हैं पुल्लिंग, स्त्रीलिंग और नपुंसक लिंग। जब हिंदी से अंग्रेजी के लिए अनुवाद दिया जाएगा तो लिंग के अनुसार अनुवाद में भी बदलाव होगा। हिंदी में लिंग का प्रभाव विशेषण और क्रिया रूपों पर पड़ता है, लेकिन सर्वनाम पर नहीं, जबकि अंग्रेजी में लिंग का प्रभाव केवल अन्य पुरुष (3rd person) सर्वनामों पर पड़ता है। जैसे- *मैं (m/f) कल आपसे (m/f)*

मिलूँगा/मिलूँगी (*I will meet you.*) हिंदी में प्राणीवाचक संज्ञा के लिंग पता न होने पर उसे पुल्लिंग में रखा जाता है, तो अंग्रेजी में इसका समावेश (he) में किया जाता है। (Man is mortal. He must die one day) (सिंह, 2003:111)

- हिंदी और अंग्रेजी में दो वचन होते हैं एकवचन (Singular) और बहुवचन (Plural)। (सिंह, 2003:113) हिंदी में वचन का प्रभाव संज्ञा (*एक लड़का-दो लड़के*), सर्वनाम (*वह-वे, उसे-उन्हें*), क्रिया (*लड़का हँस रहा है - लड़के हँस रहे हैं*), विशेषण (*छोटा कमरा-छोटे कमरे*) आदि पर पड़ता है। अंग्रेजी में वचन लिंग का प्रभाव संज्ञा (boy-boys), पुरुषवाचक सर्वनाम (he-they), निर्देशवाचक सर्वनाम (demonstrative pronoun- this-these) और क्रियाओं (is-are, lives-live) पर पड़ता है, लेकिन विशेषण पर नहीं पड़ता। जैसे- लंबा लड़का, लंबे लड़के (the tall boy/boys) आदि वाक्यों में सम्मिलन के लिए यह भिन्नता जान लेना महत्वपूर्ण है।
- हिंदी-अंग्रेजी संरचना में भिन्नता के स्तर पर आर.एम.के. सिन्हा और अनिल ठाकुर¹ ने अपने शोध पत्र में समय का सम्मिलन के बारे में चर्चा की है। दुनिया में विभिन्न सामाजिक और सांस्कृतिक मान्यताएँ हैं। उसी तरह भारत में भी समय को दर्शाने के लिए विभिन्न शब्दों का प्रयोग होता है। इसे अनुवाद के सम्मिलन के तौर पर देखना जरूरी हो जाता है। अंग्रेजी में सुबह और श्याम के लिए 'am. और pm.' शब्द का प्रयोग होता है लेकिन हिंदी में इसके लिए और भी शब्द के प्रयोग होते हैं। समय की निश्चितता के लिए उपर्युक्त शब्द का प्रयोग हिंदी में नहीं होता। जैसे- हिंदी के वाक्य का अंग्रेजी में 'वह पाँच बजे सुबह/पूर्वाह्न को आया' (*He arrived at 5 a.m.*) वह ग्यारह बजे (सुबह/पूर्वाह्न) को आया (*He arrived at 11 a.m.*), वह तीन बजे अपराह्न को आया (*He arrived at 3 p.m.*), वह पाँच बजे श्याम को/अपराह्न को आया (*He arrived at 5 p.m.*), वह सात बजे रात को/अपराह्न को आया (*He arrived at 7 p.m.*) आदि वाक्यों के सम्मिलन से यह ज्ञात होता है कि अंग्रेजी के शब्द के लिए हिंदी में अनेक शब्दों का प्रयोग होता है लेकिन निश्चित शब्द का प्रयोग नहीं होता जिस तरह अंग्रेजी में (am. pm.) शब्द का प्रयोग होता है। इसलिए समय का सम्मिलन करना भारतीय भाषाओं के लिए महत्वपूर्ण है।

इस हिंदी-अंग्रेजी की संरचनात्मक भिन्नता को देखते हुए यह स्पष्ट होता है कि अगर मशीन से अनुवाद करना है तो विभिन्न समस्याएँ आ सकती हैं जिसके लिए दो भाषाओं का सम्मिलन कर

¹ Sinha, R.M.K. et.al. (2005) : 'Divergence Patterns in Machine Translation between Hindi and English' Proceedings of MT Summit X, Thailand, 12-16 September.

उन भाषाओं की भिन्नता को ज्ञात किया जा सकता है और उस आधार पर मशीन के लिए स्थानांतरण व्याकरणिक नियम बनाए जा सकते हैं। इस नियम बनाने के साथ मशीन में संचालित करने के लिए प्रमुख घटक 'हिंदी-अंग्रेजी विश्लेषक (Hindi- English Analyzer), ट्रान्सफर लेक्सिकॉन, हिंदी से अंग्रेजी प्रजनक (Hindi-English Generator) की भूमिका महत्वपूर्ण रहती है।

4.3 अंग्रेजी प्रजनक (English Generator)

स्रोत भाषा (हिंदी) से लक्ष्य भाषा (अंग्रेजी) में अनुवाद करने की प्रक्रिया विश्लेषण-अंतरण-प्रजनक पर आधारित है। स्रोत भाषा के किसी भी वाक्य या पाठ के अनुवाद के लिए उसका विश्लेषण होता है और विश्लेषण के बाद लक्ष्य भाषा संरचना के लिए शब्दावली का चयन किया जाता है और प्रजनक की प्रक्रिया में लक्ष्य भाषा की प्रकृति और संरचना के अनुरूप तथा व्याकरणिक नियमों का प्रयोग कर अनुवाद मशीन द्वारा प्रदान किया जाता है। (Dave et al, 2001) इसी तरह हिंदी की संरचना का अंग्रेजी में अनुवाद करने के लिए प्रजनक के लिए कुछ व्याकरणिक नियमों का निर्माण आगे इस अध्याय में किया जाएगा। यहाँ हिंदी-अंग्रेजी की संरचना को संज्ञापद स्तर का कॉर्पोरा आधारित सम्मिलन कर इसके लिए नियम बनाए जा रहे हैं। सम्मिलन के घटक निम्नप्रकार हैं।

- हिंदी के कारक परसर्ग और अंग्रेजी के पूर्वसर्ग का सम्मिलन
- सर्वनाम का सम्मिलन
- विशेषण (संख्यावाचक विशेषण) का सम्मिलन
- संबंधबोधक अव्यय का सम्मिलन
- संयुक्त और मिश्र वाक्यसंरचना में समुच्च्यबोधक का सम्मिलन
- हिंदी और अंग्रेजी वाक्यसंरचना में उपपदों का सम्मिलन
- व्याकरणिक कोटियों- लिंग, वचन और पुरुष का सम्मिलन
- हिंदी अंग्रेजी के शब्दक्रम का सम्मिलन

ये घटक मशीन अनुवाद में अंग्रेजी प्रजनक के लिए चुनौतियाँ उत्पन्न करते हैं। हिंदी के कारक परसर्ग और पूर्वसर्गों में अधिकतर भिन्नताएँ पाई जाती हैं। जैसे- कर्ताकारक परसर्ग में 'ने' का प्रयोग होता है लेकिन अंग्रेजी में पूर्वसर्ग का लोप (\emptyset) हो जाता है, तो कभी-कभी अंग्रेजी के कर्मवाच्य वाक्य में 'ने' के लिए 'by' का प्रयोग होता है। दूसरी तरफ़ यही भिन्नता संबंधकारक में मिलती है, जिसमें परसर्ग 'का, की, के' का प्रयोग होता है जिसके लिए अंग्रेजी में 'of' पूर्वसर्ग का प्रयोग होता है। अनुवाद के लिए समस्या इस तरह के कारक में आती है। इसी तरह विशेषण

में 'अच्छा लड़का, अच्छे लड़के' (good boy, good boys) होता है। यहाँ हिंदी में लिंग व वचन के अनुसार विशेषण में बदलाव हो रहा है। शब्दक्रम में भी दोनों भाषाओं में भिन्नता (हिंदी में कर्ता-कर्म-क्रिया तथा अंग्रेजी में कर्ता-क्रिया-कर्म) है। इससे अनुवाद में भी क्रम बदलेगा। इस भिन्नता के लिए नियम बनाकर अंग्रेजी प्रजनक द्वारा निर्गत अनुवाद में सुधार किया जा सकता है।

4.4 स्थानांतरण व्याकरण (Transfer Grammar)

मशीन अनुवाद में स्थानांतरण व्याकरण (Transfer Grammar) के दो प्रयोजन हैं। पहला, यह अनुवाद प्रक्रिया के दौरान स्रोत भाषा (Source Language) के शब्दों और अभिव्यक्तियों के लिए लक्ष्य भाषा (Target Language) के समानार्थी शब्द और अभिव्यक्तियों को उपलब्ध कराता है। दूसरे, इन समानार्थी शब्दों और अभिव्यक्तियों को लक्ष्य भाषा की वाक्य संरचना में ठीक जगह पर और ठीक क्रम में स्थापित होने के निर्देश देता है। भारत में आई.आई.टी. हैदराबाद ने संपर्क तंत्र के लिए स्थानांतरण व्याकरण नियमों का निर्माण किया है जिसमें एक भाषा से अनेक भाषाओं के लिए नियमों का निर्माण किया जा रहा है। यह नियम एक भारतीय भाषा (IL-1) से दूसरी भारतीय भाषा (IL-2) में अनुवाद के लिए महत्वपूर्ण है। इस पर अविनेश² (IIT HYDERABAD) के शोध-प्रबंध कार्य में संपर्क तंत्र (अंग्रेजी-हिंदी मशीनी अनुवाद तंत्र) निर्माण में स्थानांतरण व्या. नियमों के प्रयोग को स्पष्ट किया गया है, साथ ही कुमार नृपेन्द्र³ (JNU) के द्वारा इस नियमों के आधार पर हिंदी-अंग्रेजी संज्ञापद के सम्मिलन के लिए इन नियमों का निर्माण किया जा रहा है। यह नियमों का निर्माण सैध्दांतिक रूप से है।

4.5 कारक (परसर्ग-पूर्वसर्गों) का सम्मिलन और स्थानांतरण व्याकरणिक नियम

हिंदी के कारक परसर्ग (Postposition) में *ने, को, से, पर, में, का, के, की, द्वारा, के लिए, हे, अरे* आदि का समावेश किया जाता है। संस्कृत में इसे विभक्ति कहते हैं तो अंग्रेजी में इसे पूर्वसर्ग (Preposition) कहते हैं। हिंदी में संज्ञा के बाद परसर्ग प्रयुक्त होते हैं, पहले नहीं। लेकिन अंग्रेजी में पूर्वसर्गों का प्रयोग संज्ञा से पहले और बाद, दोनों स्थानों पर होता है। हरिवंश सिंह⁴ ने कारक परसर्ग पर विस्तृत अध्ययन किया है। वे अपनी पुस्तक में लिखते हैं कि "हिंदी में परसर्ग का प्रयोग आधुनिक रूप में किया जा रहा है। यह संस्कृत के 'उपसर्ग' के आधार पर बनाया

² Avinesh, P.V.S (2010) : *Transfer Grammar Engine and Automatic Learning of Reorder Rules in Machine Translation* (M.Sc- Thesis)

³ Pathak, Kumar Nripendra (2011) : *challenges in Sanskrit-Hindi Noun Phrase Mapping*, M.Phil (dissertation), JNU, New Delhi

⁴ सिंह, हरिवंश (2003) : *कारक व्याकरण की दृष्टि से हिन्दी परसर्ग विचार*, पृ. 35

गया शब्द है, जो एक रूपिम है जिसका संज्ञा शब्द के साथ प्रयोग होता है, जैसे- घर में, मेज पर। यह परसर्ग व्याकरणिक अंश है और रूपिम भी है।” इसी तरह कारक का विभाजन कर्ता, कर्म, करण, संप्रदान, अपादान, संबंधकारक, अधिकरण, संबोधन आदि में किया गया है। नीचे टेबल में परसर्गों और पूर्वसर्गों की भूमिकाओं का सह संबंध दिखाया गया है।

भूमिका	परसर्ग (postposition)	पूर्वसर्ग (preposition)
कर्ताकारक (agentive)	ने, Ø	Ø, by (only in passive constructions)
कर्मकारक (Objective/Patient or Experiencer)	को	to
करणकारक (Instrumental)	से, के द्वारा	With, by, through
संप्रदान (Dative, Beneficiary)	को, के लिए	to, for
अपादान (ablative)	से	from
अधिकरण (Locative)	में, पर	In, at, on
संबंधकारक (Genitive)	का, की, के	Of, 's'
संबोधन (vocative)	हे, अरे, ऐ	O!

टेबल 4.1 : हिंदी - अंग्रेजी कारक परसर्ग और पूर्वसर्गों की भिन्नता

4.5.1 कर्ता कारक (agentive) का विभाजन और स्थानांतरण व्या. नियम⁵

जिसे रूप से क्रिया करने वाले का बोध होता है वह 'कर्ता' कारक कहलाता है। इसमें 'ने' परसर्ग का प्रयोग होता है। तो कुछ संदर्भ के अनुसार परसर्ग और पूर्वसर्ग का लोप होता है।

1. अध्ययनों\N_NN ने\PSP यह\DM_DMD प्रमाणित\N_NN कर\V_VM दिया\V_VAUX है\V_VAUX
ENG - Studies\NNS have\VBP proved\VBN that\IN

2. सी.डी.सी.\N_NN ने\PSP बताया\V_VM है\V_VAUX कि\CC_CCS मृत्यु\N_NN दर\N_NN 15\QT_QTC
से\PSP 50\QT_QTC प्रतिशत\N_NN तक\PSP मानी\V_VM है\V_VAUX ।\RD_PUNC

⁵ इस अध्याय में हिंदी-अंग्रेजी के उदाहरणों में प्रयोग किए गए टैग के लेबल की सूची पृ. viii में दी गई है।

ENG - CDC\NNP has\VBZ said\VBN that\IN the\DT death\NN rate\NN is\VBZ considered\VBN to\TO be\VB 15\NNS.\.

उपर्युक्त दोनो हिंदी के वाक्यों में कर्ता के साथ 'ने' परसर्ग का प्रयोग हुआ है परंतु अंग्रेजी में पूर्वसर्ग का लोप (Ø) हुआ है।

3. महाराज\N_NN शिव\N_NNP ने\PSP भगवान\N_NNP परशुराम\N_NNP से\PSP यहाँ\DM_DMD शिव\N_NNP मंदिर\N_NN की\PSP स्थापना\N_NN कराई\V_VM थी\V_VAUX ।\RD_PUNC

ENG - Maharaj\NNP Shiv\NNP had\VBD Lord\NNP Parashuram\NNP establish\VB the\DT Shiva\NNP Temple\NNP here\RB.\.

4. सभी\QT_QTF ने\PSP गंगा\N_NNP स्नान\N_NN कर\V_VM पूजा-अर्चना\N_NN ,\RD_PUNC कथा\N_NN आदि\N_NN की\PSP ।\RD_PUNC

ENG- having\VBG bathed\VBN in\IN the\DT Ganges\NNP ,\ , all\DT offered\VBN prayers\NNS and\CC religious\JJ stories\NNS.\.

कर्ताकारक में 'ने' परिमाणवाचक के साथ भी प्रयोग हो सकता है।

5. बिल्ली\N_NN दूध\N_NN पी\V_VM गई\V_VAUX ।\RD_PUNC

ENG - The\DT cat\NN drank \VB the\DT milk\NNS.\.

इस कर्ताकारक वाक्य में परसर्ग का प्रयोग नहीं हुआ है। तथा अंग्रेजी में भी पूर्वसर्ग का लोप (Ø) हुआ है।

6. चन्द्रादित्य\N_NNP मंदिर\N_NN का\PSP निर्माण\N_NN नाग\N_NNP राजा\N_NN चन्द्रादित्य\N_NNP ने\PSP करवाया\V_VM था\V_VAUX

ENG - Chandraditya\NNP Temple\NNP was\VBD built\VBN by\IN the\DT Naga\NNP king\NN Chandradita\NNP

हिंदी में जब 'ने' का प्रयोग कर्ता के साथ होता है तो अंग्रेजी में कर्मवाच्य अनुवाद में उसके लिए 'by' का प्रयोग होता है। 'ने' का प्रयोग सामान्यतः सकर्मक धातुओं के साथ भूतकाल में होता है।

स्थानांतरण व्याकरणिक नियम : (TG Rules)

- हिंदी कर्ताकारक में 'ने' परसर्ग का अंग्रेजी पूर्वसर्ग में अधिकांश वाक्यों में लोप (Ø) होता है।
- अर्थ के अनुसार कभी-कभी हिंदी एवं अंग्रेजी वाक्य में कर्ताकारक परसर्ग का प्रयोग नहीं होता।
- हिंदी वाक्यों में जब 'ने' का प्रयोग कर्ता के साथ होता है तो अंग्रेजी में कर्मवाच्य अनुवाद के लिए 'by' का प्रयोग होता है।

स्थानान्तरण व्याकरणिक नियम :

- हिंदी कर्म कारक परसर्ग 'को' का अंग्रेजी में लोप तब होता है जब 'को' का प्रयोग कर्ता के साथ होता है। (उदाहरण के लिए 7, 8 देखें।)
- हिंदी में कर्म कारक 'को' का प्रयोग जब कर्म के साथ होता है तो अंग्रेजी में 'to' का प्रयोग कर्म के साथ होता है। (उदाहरण के लिए 9, 10 देखें।)
- प्रयोज्य कर्म में 'को' के लिए 'to' का प्रयोग प्रायः होता है।

4.5.3 करण कारक (Instrumental) का विभाजन और स्थानान्तरण व्या. नियम

करण कारक संज्ञा के उस रूप को कहते हैं जिससे क्रिया के साधन का बोध होता है। करण कारक में परसर्ग 'से', 'के द्वारा' का प्रयोग होता है।

12. योग\N_NN से\PSP काबू\N_NN हो\V_VM सकता\V_VAUX है\V_VAUX लकवा\N_NN I\RD_PUNC

ENG- Paralysis\NN may\MD be\VB controlled\VBN with\IN yoga\NN .\.

13. सामान्यतः\RB सुसाध्य\JJ ट्यूमरों\N_NN को\PSP शल्य-चिकित्सा\N_NN द्वारा\PSP निकाला\V_VM जा\V_VAUX सकता\V_VAUX है\V_VAUX

ENG- Generally\RB ,\, easy\JJ tumors\NNS can\MD be\VB weeded\VBN out\RP through\IN surgery\NN

14. गर्भाशय-ग्रीवा\N_NN का\PSP कैंसर\N_NN रक्त\N_NN प्रवाह\N_NN के\PSP द्वारा\PSP भी\RP_RPD फैल\V_VM सकता\V_VAUX है\V_VAUX I\RD_PUNC

ENG- Cervical\JJ cancer\NN can\MD spread\VB through\IN blood\NN circulation\NN .\.

15. विशेष\JJ लैन्स\N_NN या\CC_CCD यंत्र\N_NN के\PSP द्वारा\PSP नेत्र\N_NN तंत्रिका\N_NN की\PSP जाँच\N_NN करेंगे\V_VM कि\PSP कितनी\PR_PRQ क्षति\N_NN पहुँची\V_VM है\V_VAUX I\RD_PUNC

ENG- We\PRP will\MD check\VB the\DT eye\NN nerves\NNS with\IN special\JJ lens\NN or\CC equipment\NN to\TO find\VB how\WRB much\JJ of\IN damage\NN has\VBZ occurred\VBN .\.

16. नारायणपाल\N_NNP का\PSP यह\DM_DMD विष्णु\N_NNP मंदिर\N_NN पुरातत्व\N_NN विभाग\N_NN द्वारा\PSP संरक्षित\V_VM स्मारक\N_NN है\V_VAUX I\RD_PUNC

ENG- This\DT Vishnu\NNP Temple\NNP of\IN Narayanpal\NNP is\VBZ a\DT protected\JJ monument\NN by\IN the\DT Archaeological\NNP Department\NNP .\.

स्थानान्तरण व्याकरण नियम (TG rules)

- हिंदी के वाक्य में जब अ-प्राणिवाचक संज्ञाओं के साथ करण कारक परसर्ग, 'से' और 'के द्वारा' का प्रयोग होता है तो अंग्रेजी में इसके लिए 'with' का प्रयोग होता है। (उदाहरण के लिए 12,15 देखें)
- हिंदी परसर्ग, 'के द्वारा' का प्रयोग 'प्रक्रिया या कार्य-प्रणाली' के संदर्भ में किया जाता है तो वहाँ अंग्रेजी पूर्वसर्ग, 'through' का प्रयोग होता है। (उदाहरण के लिए 13,14 देखें)
- हिंदी परसर्ग 'द्वारा' का प्रयोग यदि कर्मवाच्य (Passive) वाक्यों में प्राणिवाचक और अ-प्राणिवाचक संज्ञा के साथ होता है तो अंग्रेजी पूर्वसर्ग 'by' का प्रयोग होता है। (उदाहरण के लिए 16 देखें)

4.5.4 संप्रदान कारक (Beneficiary) का विभाजन और स्थानान्तरण व्या. नियम

जिसके लिए कुछ कार्य किया जाता है या जिसे कुछ दिया जाता है उसे संप्रदान कारक कहते हैं। इसमें परसर्ग 'को, के लिए' का प्रयोग होता है।

17. इंग्वाईनल\N_NNP हर्निया\N_NN के\PSP उपचार\N_NN के\PSP लिए\PSP I\RD_PUNC

ENG- For\IN the\DT treatment\NN of\IN inguinal\JJ hernia\NN .\.

18. दीव\N_NNP से\PSP मुंबई\N_NNP के\PSP लिए\PSP रोजाना\N_NN कई\QT_QTF उड़ानें\N_NN भरी\N_NN जाती\N_NN हैं\N_NN I\RD_PUNC

ENG- Several\JJ flights\NNS take\NNS off\RP for\IN Mumbai\N_NNP from\IN Diu\N_NNP .\.

19. संयुक्त\N_NNP टीका\N_NN बच्चों\N_NN और\CC_CCD वयस्कों\N_NN के\PSP लिए\PSP बना\N_NN है\N_NN I\RD_PUNC

ENG - The\DT combined\N_NN vaccine\NN is\VBZ made\N_NN for\IN children\NNS and\CC adults\NNS .\.

20. कणिमंगलम\N_NNP पीतल\N_NN से\PSP बने\N_NN शिल्पों\N_NN और\CC_CCD दीपों\N_NN के\PSP लिए\PSP प्रसिद्ध\N_NN है\N_NN I\RD_PUNC

ENG - Kanimangalam\N_NNP is\VBZ famous\JJ for\IN sculptures\NNS and\CC the\DT lamps\NNS made\N_NN of\IN brass\NN .\.

उपर्युक्त सभी हिंदी वाक्यों में परसर्ग 'के लिए' का अंग्रेजी में 'for' अनुवाद हुआ है।

21. राजा *प्रजा को* अन्नदान देता है।

ENG – The\DT king\NN gives\VBZ grains\NNS to\TO the\DT state\NN .\.

22. रवि ने श्याम को एक किताब दी।

ENG – Ravi\NNP gave\VBD a\DT book\NN to\TO Shyam\NNP .\.

इन (21, 22) वाक्यों में संप्रदान में 'को' परसर्ग का प्रयोग हुआ है जिसके लिए अंग्रेजी में 'to' परसर्ग का प्रयोग हुआ है। वाक्य 22 में 'श्याम को' संप्रदान कारक हैं, जिसमें 'श्याम' के लिए क्रिया की गई है। वाक्य में 'लेना और देना' जैसी क्रियाओं के संदर्भ में अप्रत्यक्ष कर्मों (Indirect object) की आवश्यकता होती है, जिनके लिए क्रिया की जाती है। वाक्य नं. 22 को अन्य तरह से भी प्रयोग किया जा सकता है। उदाहरण के तौर पर 22a

22a. रवि ने श्याम के लिए एक किताब दी।

Ravi\NNP gave\VBD a\DT book\NN for\IN Shyam\NNP .\.

स्थानान्तरण व्याकरण नियम (TG rules)

- संप्रदान कारक में परसर्ग 'के लिए' का प्रयोग हुआ है, जिसके लिए अंग्रेजी में पूर्वसर्ग 'for' का प्रयोग होता है।

परसर्ग 'को' का प्रयोग जब संप्रदान कारक के संदर्भ में होता है तो अंग्रेजी पूर्वसर्ग 'to' का ही प्रयोग किया जाता है। इस स्थानान्तरण व्याकरण नियम का प्रयोग कर्म कारक में भी किया जाता है।

4.5.5 अपादान कारक (Ablative) का विभाजन और स्थानान्तरण व्याकरण नियम

संज्ञा के जिस रूप से क्रिया के अलगाव या पृथक्ता का बोध होता है, उसे अपादान कारक कहते हैं। अपादान में परसर्ग 'से' का प्रयोग होता है। करण और अपादान दोनों में परसर्ग 'से' का प्रयोग होता है लेकिन दोनों में अंतर है। करण में परसर्ग 'से' साधन का बोध करवाता है और अपादान के परसर्ग 'से' से अलगाव का पता चलता है।

23. असामान्य\JJ यौन\JJ संबंधों\N_NN से\PSP दूर\N_NN रहें\N_VVM I\RD_PUNC

ENG - Keep\VB away\RB from\IN uncommon\JJ sexual\JJ relations\NNS .\.

24. पथरी\N_NNP से\PSP बचाव\N_NN के\PSP कुछ\QT_QTF उपाय\N_NN -\RD_PUNC

ENG - Some\DT solutions\NNS of\IN safety\NN from\IN stone\NN .\

उपर्युक्त वाक्य में संज्ञा के साथ 'से' परसर्ग के लिए अंग्रेजी में संज्ञा के पहले 'from' पूर्वसर्ग का प्रयोग हुआ है।

25. लक्सर\N_NNP से\PSP एक\QT_QTC लाइन\N_NN हरिद्वार\N_NNP तक\PSP गई\N_VVM है\N_VAUX
I\RD_PUNC

ENG - One\CD line\NN has\VBZ gone\VBN from\IN Luxar\NNP to\TO Haridwar\NNP .\.

काल या स्थान के आरम्भ के बारे में जानकारी देने का कार्य भी अपादान में होता है ।

26. डायबिटीज\N_NNP से\PSP होने\V_VM वाली\PSP समस्याएँ\N_NN ।\RD_PUNC

ENG - Problems\NNS occurring\VBG from\IN diabetes\NN .\.

स्थानांतरण व्याकरण नियम (TG rules)

- हिंदी में अपादान परसर्ग 'से' के लिए अंग्रेजी में 'from' का प्रयोग होता है। अपादान कारक में एक ही परसर्ग का प्रयोग होता है।

4.5.6 अधिकरण कारक (Locative) का विभाजन और स्थानांतरण व्याकरण नियम

शब्द के जिस रूप से क्रिया के आधार का बोध होता है उसे अधिकरण कारक कहते हैं। अधिकरण में परसर्ग 'में', 'पर' का प्रयोग होता है।

27. अगर\CC_CCS कैंसर\N_NN सीमित\N_NN क्षेत्र\N_NN में\PSP और\CC_CCD प्रारम्भिक\JJ चरण\N_NN में\PSP है\V_VAUX

ENG - If\IN cancer\NN is\VBZ in\IN a\DT limited\JJ area\NN and\CC in\IN initial\JJ stage\NN

28. वागामण\N_NNP में\PSP ढलान\N_NN पर\PSP चीड़\N_NN का\PSP वन\N_NN है\V_VAUX ।\RD_PUNC

ENG - There\EX is\VBZ pine\VB forest\NN at\IN the\DT slopes\NNS in\IN Vagaman\NNP .\.

29. धरमपुरा\N_NNP रोड\N_NN पर\PSP भगवान\N_NN बालाजी\N_NNP का\PSP भव्य\JJ मंदिर\N_NN है\V_VAUX ।\RD_PUNC

ENG - The\DT grand\JJ temple\NN of\IN Lord\NNP Balaji\NNP is\VBZ on\IN the\DT Dharampura\NNP road\NN .\.

30. यही\DM_DMD नहीं\RP_NEG जीभ\N_NN पर\PSP सफेद\JJ परत\N_NN जम\V_VM जाती\V_VAUX है\V_VAUX ।\RD_PUNC

ENG - Not\RB only\RB this\DT .\T white\JJ coat\NN also\RB forms\VBZ on\IN tongue\NN .\.

31. घर\N_NN में\PSP कीटनाशक\N_NN दवाओं\N_NN का\PSP छिड़काव\V_VM करवाये\V_VAUX ।\RD_PUNC

ENG - Get\VB insecticides\NNS sprayed\VBN at\IN home\NN .\.

स्थानांतरण व्याकरण नियम (TG rules)

- यदि हिंदी वाक्यों में परसर्ग 'में' का प्रयोग स्थिति और स्थान को दर्शाने के लिए होता है तो अंग्रेजी पूर्वसर्ग, 'in' का प्रयोग होता है। (उदाहरण के लिए वाक्य 27,28 देखें।)

- हिंदी वाक्यों में स्थान के बारे में जानकारी देने वाले संज्ञा शब्द के साथ परसर्ग 'पर' का प्रयोग होता है जिसके लिए अंग्रेजी में अधिकरण पूर्वसर्ग 'on' का प्रयोग होता है। (उदाहरण के लिए वाक्य 29, 30 देखे।)
- कुछ स्थितियों में, अपवाद के तौर पर, अधिकरण परसर्ग 'में' का प्रयोग स्थानवाचक संज्ञाओं के साथ होता है, जैसे- घर में, कमरे में, जिनके लिए अंग्रेजी पूर्वसर्ग 'at' का प्रयोग होता है और ऐसा अंग्रेजी वाक्यों की व्याकरणिक संरचना में होता है।

4.5.7 संबंध कारक (Genitive) का विभाजन और स्थानांतरण व्याकरण नियम

संज्ञा के जिस रूप से उसकी वाच्य वस्तु का संबंध किसी दूसरी वस्तु के साथ सूचित होता है उस रूप को संबंध कारक कहते हैं। संबंध कारक में परसर्ग 'का, के, की' का प्रयोग होता है। संबंधकारक का रूप संबंधी शब्द के लिंग, वचन के अनुरूप बदलता है।

32. पहले\QT_QTO ओ.आर.एस.\N_NNP का\PSP घोल\N_NN पिलायें\V_VM ,\RD_PUNC

ENG - First\JJ feed\NN the\DT mixture\NN of\IN ORS\NNP .\.

33. डायरिया\N_NNP का\PSP घरेलू\JJ उपचार\N_NN क्या\PR_PRQ है\V_VAUX ?\RD_PUNC

ENG - What\WP is\VBZ the\DT home\NN treatment\NN of\IN diarrhoea\NN .\.

34. राम\N_NNP का\PSP भाई\N_NN I\RD_PUNC

ENG - Ram's\NNP brother\NN

35. कालाजार\N_NNP के\PSP प्रारंभिक\JJ लक्षण\N_NN क्या\RP_RPQ हैं\V_VM ?\RD_PUNC

ENG - What\WP are\VBZ the\DT primary\JJ symptoms\NNS of\IN kalaazar\NN .\.

36. घास\N_NN के\PSP मैदान\N_NN ,\RD_PUNC जलप्रपात\N_NN ,\RD_PUNC चीड़\N_NN वृक्ष\N_NN आदि\N_NN पीरुमेडू\N_NNP की\PSP खूबसूरती\N_NN में\PSP चार\QT_QTC चाँद\N_NN लगा\V_VM देते\V_VAUX हैं\V_VAUX I\RD_PUNC

ENG - The\DT field\NN of\IN grass\NN ,\ , waterfall\NN ,\ , pine\VB tree\NN etcetera\NN increase\NN the\DT beauty\NN of\IN Peerumedu\NNP fourfold\RB .\

37. स्तन\N_NN की\PSP लगभग\N_NN 80\QT_QTC प्रतिशत\N_NN गाँठें\N_NN कैंसर\N_NN नहीं\RP_NEG होतीं\V_VM

ENG - About\IN 80\CD percent\NN of\IN the\DT cysts\NNS of\IN the\DT breasts\NNS are\VBZ not\RB cancer\NN

38. अस्पताल\N_NN की\PSP सूची\N_NN जहाँ\PR_PRL कैंसर\N_NN के\PSP इलाज\N_NN की\PSP सुविधा\N_NN है\|V_VAUX |\RD_PUNC

ENG - List\NN of\IN hospitals\NNS where\WRB the\DT treatment\NN of\IN cancer\NN is\VBZ available\JJ .\.

39. बच्चों\N_NN को\PSP लम्बे\JJ समय\N_NN तक\PSP अपना\PR_PRF दूध\N_NN पिलाना\|V_VM |\RD_PUNC

ENG - Feeding\NNP children\NNS one\NN 's\POS own\JJ milk\NN for\IN a\DT long\JJ time\NN .\.

40. जॉन के पास दो किताबें हैं।

ENG – John\NNP has\VBZ two\CD books\NNS .\.

स्थानान्तरण व्याकरण नियम (TG rules)

- जब हिंदी वाक्यों में संज्ञा के साथ संबंधकारक परसर्ग 'का, के, की' का प्रयोग होता है, वहाँ अंग्रेजी पूर्वसर्ग 'of' का प्रयोग होता है, (उदाहरण 32, 35,37), अंग्रेजी वाक्यों में प्रत्यय "s' का भी प्रयोग होता है (उदाहरण 34)। यहाँ परसर्ग 'का' कर्ता-कर्म भाव की भूमिका दर्शाता है।
- हिंदी में कुछ शब्द जैसे- 'स्वयं, अपना, अपने-आप, आपका' का प्रयोग संज्ञा के साथ संबंध दर्शाते हैं तो इन्हें सुस्पष्ट संबंधसूचक चिन्हक (Explicit genitive marker) भी कहा जाता है। जिनका अंग्रेजी में रूपांतर 'own' और 'self' होता है। ये शब्द परसर्ग और पूर्वसर्ग की श्रेणी में नहीं आते हैं। हिंदी में इनको सर्वनाम में वर्गीकृत किया जाता है और अंग्रेजी में विशेषण की श्रेणी में रखा जाता है। (उदाहरण 39)।
- जब हिंदी वाक्य में परसर्ग 'के पास' का प्रयोग संज्ञा के साथ संबंधसूचकता दर्शाने के लिए किया जाता है, तो अंग्रेजी वाक्यों में इसके लिए 'has' को प्रयोग होता है, जो सहायक क्रिया (Helping verb) के वर्ग में आते हैं। (उदाहरण 40)

4.5.8 संबोधन कारक (Vocative) का विभाजन और स्थानान्तरण व्याकरण नियम

जिससे किसी से बुलाने अथवा सचेत करने का भाव प्रकट हो उसे संबोधन कारक कहते हैं और उसमें संबोधन चिन्ह (!) लगाया जाता है। जिसमें 'अरे, हे, ऐ' शब्दों का प्रयोग होता है।

41. ए\RP_INJ लड़के\N_NN !\RD_PUNC

ENG - Hey\UH young\JJ boy\N_NN!\.

42. शंकर\N_NNP जी\N_NN ने\PSP कहा\V_VM कि\CC_CCS हे\RP_INJ देवी\N_NN सतयुग\N_NN में\PSP सभी\QT_QTF तीर्थ\N_NN पुण्य\N_NN का\PSP फल\N_NN देते\V_VM हैं\V_VAUX ,\RD_PUNC

ENG - Shankarji\NNP said\VBD that\IN O\NNP Devi\NNP ,\, all\PDT the\DT pilgrimages\NNS give\VBP the\DT fruits\NNS of\IN meritorious\JJ actions\NNS in\IN the\DT Satayuga\NNP ,\,

43. 'भाइयों\N_NN और\CC_CCD बहनों\N_NN !\RD_PUNC आपका\PR_PRP स्वागत\N_NN करते\V_VM हुए\V_VAUX मुझे\PR_PRP बड़ी\JJ खुशी\N_NN हो\V_VM रही\V_VAUX है\V_VAUX !\RD_PUNC

ENG - ' Ladies\NNS and\CC_Gentlemen\NNS !\UH I\PRP am\VBP very\JJS pleased\VBD to\TO welcome\VB you\PRP .\.' (Kacharu: 2006, 51)

स्थानान्तरण व्याकरण नियम (TG rules)

- हिंदी वाक्यों में संज्ञा के साथ 'संबोधन' करने वाले शब्दों जैसे 'ए, हे' का अंग्रेजी में रूपांतरण 'Hey' और 'o' आदि संबोधनसूचक शब्दों में किया जाता है। (उदाहरण के लिए 41, 42 देखें)
- संबोधन के लिए हिंदी और अंग्रेजी में संज्ञा के साथ आगे या पीछे एक विस्मयबोधक चिन्ह (!) लगता है (उदाहरण के लिए 43 देखें)

यहाँ इन उपर्युक्त कारक परसर्ग के लिए नियम बनाएँ गए जिनका संज्ञा के साथ वाक्य में प्रयोग प्रमुख रूप से होता है। हिंदी में परसर्ग का प्रयोग बाद में होता है और अंग्रेजी में पूर्वसर्ग का प्रयोग संज्ञा के पहले या बाद में होता है। हिंदी में अन्य परसर्ग- 'पर, ऊपर, अन्दर, आगे, नीचे, के पास, पीछे, बाहर, बिना, बीच, भीतर' आदि का प्रयोग भी संदर्भ विशेष के अनुसार अलग-अलग रूप में किया जाता है। संज्ञा के स्थान पर सर्वनाम के प्रयोग में कारक का प्रयोग होता है परंतु लिंग के कारण इसका रूप नहीं बदलता। इसके के बारे में (4.2) में चर्चा की गयी है। इन नियमों को मशीन में संचालित करना होगा तो कारक-विश्लेषक का होना आवश्यक है जिसके माध्यम से कारक का विभाजन कर नियमों को मशीनी अनुवाद के लिए उपयोग किया जा सकता है।

4.6 विशेषण का सम्मिलन (Adjective Mapping) और स्थानान्तरण व्याकरणिक नियम

संज्ञा या सर्वनाम की विशेषता बताने वाले शब्द को 'विशेषण' कहते हैं तथा जिसकी विशेषता प्रकट की जाती है उसे 'विशेष्य' कहते हैं। हिंदी में कुछ ऐसे विशेषण हैं, जो संज्ञा के लिंग और वचन के अनुरूप अपना रूप बदलते हैं। जैसे-

44. नियमित\JJ रूप\N_NN से\PSP कराएँ\V_VM दाँतों\N_NN की\PSP जाँच\N_NN !\RD_PUNC

ENG -Get\VB the\DT teeth\NNS checked-up\JJ regularly\RB .\.

45. बासी\JJ भोजन\N_NN व\CC_CCD सड़क\N_NN किनारे\N_NN बिकने\N_NN वाले\N_NN सामान\N_NN का\N_NN सेवन\N_NN नहीं\RP_NEG करें\N_NN ।\RD_PUNC

ENG -Do\N_NN not\RB eat\VB stale\JJ food\NN and\CC the\DT things\NNS being\VBG sold\N_NN at\N_NN road\N_NN sides\NNS .\.

46. सुन्दर\JJ लॉन\N_NN एवं\N_NN पुष्प\N_NN उद्यान\N_NN पर्यटकों\N_NN का\N_NN मन\N_NN मोह\N_NN लेते\N_NN हैं\N_NN ।\RD_PUNC

ENG -Beautiful\JJ lawns\NNS and\CC flower\N_NN parks\NNS hypnotise\N_NN the\DT tourists\NNS .\.=

हिंदी के अधिकांश वाक्यों में संज्ञा के पहले विशेषण प्रयुक्त होते हैं और अंग्रेजी में संज्ञा के पहले और बाद में विशेषण का प्रयोग होता है।

47. प्राचीन\JJ काल\N_NN में\N_NN पितामह\N_NN ब्रह्मा\N_NN जी\N_NN ने\N_NN कांची\N_NN में\N_NN माँ\N_NN भगवती\N_NN के\N_NN दर्शन\N_NN के\N_NN लिए\N_NN दुष्कर\JJ तपस्या\N_NN की\N_NN थी\N_NN ।\RD_PUNC

ENG -In\N_NN ancient\JJ times\NNS the\DT Pitamah\N_NN Brahma\N_NN had\N_NN done\N_NN difficult\JJ penance\N_NN to\N_NN see\N_NN Mother\N_NN Bhagawati\N_NN .\.

कॉर्पोरा के वाक्य में कभी – कभी एक से अधिक बार विशेषण आए हैं।

48. हेयर\JJ पिन\N_NN मोड़\N_NN लिये\N_NN हुए\N_NN यह\N_NN मनोरम\JJ घाटी\N_NN पर्यटकों\N_NN का\N_NN मन\N_NN मोह\N_NN लेती\N_NN है\N_NN ।\RD_PUNC

ENG - With\N_NN a\DT hairpin\N_NN curve\N_NN this\DT charming\JJ valley\N_NN hypnotises\N_NN tourists\NNS .\.

49. बचेली\N_NN का\N_NN अयस्क\N_NN अच्छे\JJ ग्रेड\N_NN का\N_NN माना\N_NN जाता\N_NN है\N_NN ।\RD_PUNC

ENG -The\DT ores\NNS of\N_NN Bachel\N_NN are\N_NN considered\N_NN to\N_NN be\N_NN good\JJ grade\N_NN .\.

50. नानी\N_NN दमन\N_NN का\N_NN मुख्य\JJ आकर्षण\N_NN संत\N_NN जैरोम\N_NN का\N_NN किला\N_NN है\N_NN ।\RD_PUNC

ENG -The\DT main\JJ attraction\N_NN of\N_NN Nanidaman\N_NN is\N_NN the\DT Saint\N_NN Jerome\N_NN Fort\N_NN .\.

51. केरल\N_NN के\N_NN प्रमुख\JJ पड़ोसी\JJ राज्य\N_NN तमिलनाडु\N_NN और\N_NN कर्नाटक\N_NN हैं\N_NN ।\RD_PUNC

ENG -Major\JJ neighboring\JJ states\NNS of\N_NN Kerala\N_NN are\N_NN Tamil\N_NN Nadu\N_NN and\N_NN Karnataka\N_NN .\.

52. कांजार\N_NN जलप्रपात\N_NN पूरे\JJ वर्ष\N_NN जलयुक्त\N_NN रहता\N_NN है\N_NN ।\RD_PUNC

ENG -Kanjar\NNP waterfall\NN remains\VBZ filled\VBN with\IN water\NN whole\JJ year\NN .\.

विशेषण सम्मिलन के लिए और स्थानांतरण व्याकरण नियम (TG rules) के लिए संभावित बिंदु

- हिंदी में आकारांत (*आ* में अंत होने वाले) विशेषण शब्द विशेष्य शब्द के लिंग तथा वचन से प्रभावित होते हैं। जैसे - *एक छोटा कमरा, दो बड़े कमरे, मेरी बड़ी लड़की* इत्यादि।
- अन्य विशेषण शब्द, जो आकारांत नहीं होते, अविकारी होते हैं, अर्थात् उनके रूप में कोई परिवर्तन नहीं होता। जैसे- '*सुंदर कमरा, सुंदर कमरे, सुंदर लड़की*' में विशेष्य के लिंग तथा वचन में परिवर्तन के बावजूद भी विशेषण '*सुंदर*' का रूप नहीं बदलता। (उदा. के लिए 46 देखें)
- अंग्रेजी विशेषण पर विशेष्य के लिंग तथा वचन का कोई प्रभाव नहीं पड़ता। जैसे- *Large room/ rooms, young boys, young girls.* अंग्रेजी में सभी विशेषण अविकारी होते हैं।
- हिंदी और अंग्रेजी दोनों में कुछ विशेषणों का प्रयोग संज्ञा के रूप में भी होता है और अंग्रेजी में ऐसे शब्दों से पहले अधिकांशतः आर्टिकल्स '*the*' का प्रयोग होता है। जैसे- *The English and the Scotch, the poor!* दूसरी ओर हिंदी में '*the*' के समान किसी शब्द का प्रयोग नहीं होता परंतु संकेतवाचक शब्दों 'उस' और 'इस' का प्रयोग किया जाता है उदाहरण के तौर पर

अ. *उस अमीर को देखो।*

ब. *इस गरीब ने यह काम किया।*

- हिंदी में बहुलता का भाव व्यक्त करने के लिए द्विरुक्ति (Reduplication) का प्रयोग विशेषण के साथ भी किया जाता है, जैसे- '*अच्छे-अच्छे पकवान, बड़े-बड़े लोग, गरम-गरम चाय*' आदि शब्द। जबकि अंग्रेजी में द्विरुक्ति जैसे (**good-good, big-big*) शब्दों का प्रयोग नहीं होता। लेकिन कभी-कभी द्विरुक्ति विशेषणों का प्रयोग संदर्भ की महत्ता बताने के लिए भी होता है, जिसके लिए अंग्रेजी में समकक्ष शब्द न होने के कारण 'adverbial modifiers' का प्रयोग होता है। जैसे- '*very big people*'

संख्यावाचक (CARDINALS)

57. ढाई\QT_QTC -\RD_PUNC तीन\QT_QTC वर्षों\N_NN से\PSP हर\QT_QTF दो\QT_QTC -\RD_PUNC
तीन\QT_QTC माह\N_NN पर\PSP फीवर\N_NN हो\V_VM जाता\V_VAUX है\V_VAUX I\RD_PUNC

ENG - Fever\NN occurs\VBZ at\IN an\DT interval\NN of\IN two\CD to\TO three\CD months\NNS since\IN two\CD and\CC half\NN to\TO three\CD years\NNS .\.

58. एक\QT_QTC गिलास\N_NN पानी\N_NN में\PSP दो\QT_QTC चम्मच\N_NN चीनी\N_NN एवं\CC_CCD
एक\QT_QTC चुटकी\JJ नमक\N_NN डालकर\V_VM उबालकर\V_VM दें\V_VM I\RD_PUNC

ENG - Put\VB two\CD spoons\NNS of\IN sugar\NN and\CC a\DT pinch\NN of\IN salt\NN in\IN a\DT glass\NN of\IN water\NN and\CC boil\NN it\PRP .\.

59. दो\QT_QTC सप्ताह\N_NN तक\PSP इन\PR_PRP लक्षणों\N_NN के\PSP बने\V_VM रहने\V_VAUX पर\PSP
इन्हें\PR_PRP चिकित्सक\N_NN को\PSP दिखाएँ\V_VM I\RD_PUNC

ENG - Show\NNP these\DT to\TO the\DT doctor\NN if\IN these\DT symptoms\NNS stay\VBP for\IN two\CD weeks\NNS .\.

स्थानांतरण व्याकरण नियम (TG rules)

- हिंदी संख्यावाचक विशेषण में जैसे- 'एक, दो, तीन' शब्दों का प्रयोग संज्ञा के साथ मात्रा दर्शाने के लिए किया जाता है, और कुछ ऐसे शब्द 'ढाई, तिहाई, चौथाई' आदि शब्द भी इसी श्रेणी में आते हैं। अंग्रेजी में 'एक' का अनुवाद संदर्भ के अनुसार 'a' or 'one' भी हो सकता है, परंतु शब्द जैसे 'ढाई, तिहाई, चौथाई' आदि के लिए अंग्रेजी में कोई एक समकक्ष शब्द न होने के कारण यह मशीनी अनुवाद के लिए चुनौती उत्पन्न हो सकती है। जैसे-

ढाई\QT_QTC → two\CD and\CC half\NN (उदा. 57 देखें।)

एक\QT_QTC → a\DT (उदा. 58 देखें।)

दो\QT_QTC → two\CD (उदा. 59 देखें।)

क्रमसंख्यावाचक (ORDINALS)

60. एड्स\N_NNP का\PSP पहला\QT_QTO मरीज\N_NN भारत\N_NNP में\PSP 1986\QT_QTC में\PSP
पाया\V_VM गया\V_VAUX I\RD_PUNC

ENG - The\DT first\JJ patient\NN of\IN AIDS\NNP in\IN India\NNP was\VBD found\VBN in\IN 1986\CD .\.

61. केरल\N_NNP में\PSP आवास\N_NN केन्द्रों\N_NN के\PSP विकास\N_NN का\PSP दूसरा\QT_QTO
चरण\N_NN संगमकाल\N_NN माना\N_NN जाता\N_NN है\N_NN ।\RD_PUNC

ENG - The\DT second\JJ stage\NN of\IN the\DT development\NN of\IN the\DT living\NN centers\NNS in\IN Kerala\NNP is\VBZ supposed\VBN to\TO be\VB the\DT Sangam\NNP Age\NNP .\.

* हिंदी क्रम संख्यावाचक विशेषण में 'पहला, दूसरा' संज्ञा के साथ प्रयोग होकर उनके क्रम को प्रदर्शित करते हैं। इनका अंग्रेजी में रूपांतर 'first' and 'second' होता है, जिनको हमेशा विशेषण (JJ) टैग से अंकित (Tagged) किया जाता है।

4.7 संयुक्त और मिश्र वाक्यसंरचना में समुच्चयबोधक (Conjunctive) का सम्मिलन

हिंदी में संयोजक शब्द दो प्रकार के होते हैं जो एक वाक्य में दो शब्दों को जोड़ते हैं जैसे- और, तथा, व, या,... आदि। दूसरे प्रकार के संयोजक जो एक वाक्य से दूसरे वाक्यों को जोड़ते हैं उनको वाक्य संयोजक भी कह सकते हैं, जैसे- हालांकि, बल्कि, यदि..तो, जब..तो, क्योंकि, इसलिए आदि शब्दों का प्रयोग संयोजक में होता है।

62. हमारे\PR_PRP मसूढ़ों\N_NN और\CC_CCD दाँतों\N_NN के\PSP बीच\PSP बैक्टीरिया\N_NN मौजूद\N_NN होते\N_NN हैं\N_NN ।\RD_PUNC

ENG - Bacteria\NNS stay\VB between\IN our\PRP\$ gums\NNS and\CC teeth\NNS .\.

63. ये\DM_DMD शरीर\N_NN और\CC_CCD दिमाग\N_NN में\PSP आक्सीजन\N_NN का\PSP प्रवाह\N_NN बढ़ाते\N_NN हैं\N_NN ।\RD_PUNC

ENG - They\PRP increase\VB the\DT flow\NN of\IN oxygen\NN in\IN body\NN and\CC brain\NN

64. कालाजार\N_NNP ,\RD_PUNC मलेरिया\N_NNP या\CC_CCD फिर\RB हाई\JJ फीवर\N_NN हो\N_NN सकता\N_NN है\N_NN ।\RD_PUNC

ENG - Kalajar\NNP ,\ , Malaria\NNP or\CC high\JJ fever\NN may\MD have\VB happened\VBN .\.

65. चश्मे\N_NN या\CC_CCD काँटैक्ट\N_NN लेंस\N_NN का\PSP रखरखाव\N_NN करें\N_NN ।\RD_PUNC

ENG - Maintain\VB the\DT glass\NN or\CC contact\NN lens\NN .\.

66. मुँह\N_NN में\PSP लगातार\N_NN घाव\N_NN व\CC_CCD छाले\N_NN होना\N_NN ।\RD_PUNC

ENG - Wounds\NNS or\CC blisters\NNS in\IN mouth\NN .\.

67. एड्स\N_NNP से\PSP बचाव\N_NN एवं\CC_CCD प्रचार\N_NN प्रसार\N_NN ।\RD_PUNC

ENG - Prevention\NNP of\IN AIDS\NNP and\CC propagation\NN .\.

68. बल्कि\CC_CCD साँसों\N_NN की\PSP दुर्गंध\JJ से\PSP भी\RP_RPD छुटकारा\N_NN मिलता\N_NN है\N_NN ।\RD_PUNC

ENG - But\CC one\CD also\RB gets\VBZ relief\NN from\IN stinking\JJ breath\NN .\.

69. हालांकि\CC_CCS योग\N_NN पक्षाघात\N_NN की\PSP समस्या\N_NN का\PSP स्थायी\JJ समाधान\N_NN करता\N_NN है\N_NN ,\RD_PUNC लेकिन\CC_CCS यह\DM_DMD समय\N_NN लेता\N_NN है\N_NN एवं\CC_CCD श्रमसाध्य\N_NN है\N_NN ।\RD_PUNC

ENG - Although\IN yoga\NN solves\VBZ the\DT problem\NN of\IN paralysis\NN permanently\RB but\CC it\PRP takes\VBZ time\NN and\CC is\VBZ tough\JJ .\.

स्थानान्तरण व्याकरण नियम (TG rules)

- हिंदी में (BIS Tagset) के अनुसार दो शब्दों को जोड़ने के लिए (CC_CCD) और दो वाक्यों को जोड़ने के लिए (CC_CCS) का प्रयोग किया जाता है।
- अंग्रेजी में (Penn Tagset) के अनुसार संयोजक शब्दों के लिए (CC) टैग है।
- यदि मशीनी अनुवाद के लिए हिंदी के शब्दों को संयोजक 'CC_CCD' से अंकित किया जाता है तो अंग्रेजी संयोजक शब्दों के लिए 'CC' से अंकित कर सम्मिलन (Mapping) किया जाए तो हम नीचे दिए गए शब्दों से सही अनुवाद प्राप्त कर सकते हैं।
- अ. और\CC_CCD → and\CC (उदा. 62, 63 देखें।)
- ब. या\CC_CCD → or\CC (उदा. 64, 65 देखें।)
- स. व\CC_CCD → or\CC (उदा. 66 देखें।)
- द. एवं\CC_CCD → and\CC (उदा. 67 देखें।)
- इ. बल्कि\CC_CCD → But\CC (उदा. 67 देखें।)
- * हालांकि\CC_CCS → Although\IN उदा के लिए (69)

परंतु वाक्य संयोजक शब्दों के लिए अंग्रेजी में कोई विशेष टैग नहीं है, जो मशीन अनुवाद के लिए एक महत्वपूर्ण चुनौती है।

4.8 व्याकरणिक कोटियों का सम्मिलन (Mapping of grammatical Categories)

व्याकरणिक कोटियों में विभिन्न घटकों का समावेश है, लेकिन यहाँ सिर्फ हिंदी-अंग्रेजी अनुवाद में होने वाले बदलाव को देखते हुए कुछ घटकों का अध्ययन किया गया है और उसे संरचना के आधार पर सम्मिलित (Mapped) किया गया है।

4.8.1 वचन (Number)

हिंदी और अंग्रेजी में एकवचन (Singular) और बहुवचन (Plural) दो वचन होते हैं। सामान्यतः संज्ञा के जिस रूप से एक ही वस्तु का बोध होता है, उसे एकवचन कहते हैं, जैसे- लड़का, कपड़ा, रंग, रूप। और संज्ञा के जिस रूप से एक से अधिक वस्तुओं का बोध होता है, उसे बहुवचन कहते हैं; जैसे- लड़के, कपड़े, रंगों में, रूपों से इत्यादि।

1) हिंदी में वचन का आदरार्थ प्रयोग (Honorific use of number) - हिंदी में बहुवचन रूप एक से अधिक की सूचना देता है। किसी व्यक्ति के लिए बहुवचन रूप का प्रयोग कर हिंदी में आदर का बोध व्यक्त किया जाता है, जो अंग्रेजी में संभव नहीं है। उदा. मेरे पिताजी कल आएँगे। (My father will come tomorrow), आप लोग (pl.)\ अगले माह आप कहाँ जा रहे हैं? (Where are you people going next month?)

2) हिंदी और अंग्रेजी दोनों में ही शब्दों के बाद कुछ प्रत्ययों के योग से बहुवचन रूप बनते हैं (लड़की-लड़कियाँ, girl-girls) और कुछ शब्द एकवचन और बहुवचन दोनों में समान रहते हैं। बहुवचन में प्रत्यय का प्रयोग इस बात पर निर्भर करता है कि संज्ञा शब्द आकारांत (ending in आ) है, इकारांत (ending in इ) है, ईकारांत (ending in ई) है, उकारांत (ending in उ) है, या व्यंजनांत (ending in consonant) है अथवा याकारांत है।

3) अंग्रेजी में बहुवचन रूपों का प्रभाव संज्ञा (boy-boys), (बहुवचन में संज्ञा के साथ -s, -es, -en प्रत्यय प्रयोग होता है।) पुरुषवाचक सर्वनाम (he-they), निर्देशकवाचक सर्वनाम (this-these) और क्रियाओं (is-are, lives-live) पर पड़ता है; लेकिन हिंदी में बहुवचन रूपों का प्रभाव इनके साथ-साथ विशेषण पर भी पड़ता है, जैसे- छोटा कमरा- छोटा कमरे, जो अंग्रेजी में नहीं होता है।

4) हिंदी आकारांत पुल्लिंग शब्दों का बहुवचन बनाने के लिए अंत्य 'आ' के स्थान में 'ए' लगाते हैं; जैसे लड़का-लड़के, घोड़ा-घोड़े, दूधवाला- दूधवाले

5) अकारांत स्त्रीलिंग शब्दों का बहुवचन अंत्य स्वर में बदलकर 'एँ' करने से बनता है; जैसे बहिन- बहिनें, आँख- आँखें।

4.8.2 लिंग (Gender) : हिंदी में दो लिंग होते हैं- पुल्लिंग (Masculine) और स्त्रीलिंग (Feminine)। अंग्रेजी में तीन लिंग हैं- पुल्लिंग (Masculine), स्त्रीलिंग (Feminine) और नपुंसकलिंग (Neuter)।

1) हिंदी में लिंग का प्रभाव विशेषण और क्रिया रूपों पर पड़ता है, लेकिन सर्वनाम पर नहीं होता, जबकि अंग्रेजी में लिंग का प्रभाव केवल अन्य पुरुष (3rd person) सर्वनामों पर पड़ता है। जैसे- मैं

(m/f) कल आपसे (m/f) मिलूँगा/मिलूँगी।, मेरा बड़ा लड़का (m), मेरे बड़े लड़के (m/pl), मेरी बड़ी लड़की (f), मेरी बड़ी लड़कियाँ (f/pl)

2) हिंदी में पशुओं का मानवीकरण (Personification of animals) के लिए 'वह' या 'यह' का प्रयोग होता है, तो अंग्रेजी में he और she का प्रयोग किया जाता है। जैसे 'I have not yet given my dog his morning food.' इस वाक्य से (उसका) का अर्थ अभिप्रेरित होता है।

3) अगर हिंदी में किसी प्राणिवाचक संज्ञा के लिंग का उल्लेख या पता न हो तो उसे हिंदी के पुल्लिंग में समावेश किया जाता है, उसके लिए अंग्रेजी में 'he' का प्रयोग किया जाता है। जैसे- 'वह एक न एक दिन अवश्य मृत्यु को प्राप्त होगा।' (He must die one day.)

4) हिंदी में विशेषतः 'पशुओं, शिशुओं, और छोटे बच्चों' के लिए अलग से कोई पुरुषवाचक सर्वनाम (Personal pronoun) का प्रयोग नहीं होता बल्कि 'वह, यह, उसका, अपना' का प्रयोग होता है तो अंग्रेजी में इनके लिए 'it' का प्रयोग किया जाता है। 'The cat raised its tail in anger'⁶

4.9 शब्दक्रम सम्मिलन (Word order Mapping)

हिंदी और अन्य भारतीय भाषाओं में शब्दों का क्रम परिवर्तनशील है, लेकिन अंग्रेजी में शब्दों का क्रम निर्धारित है। सामान्य शब्द क्रम में शब्दों के परिवर्तन से अधिकांश रूप से अर्थ में बदलाव होता है। हिंदी भाषा में शब्दक्रम का स्वरूप कर्ता+कर्म+क्रिया (SOV) है, इस क्रम संरचना में क्रिया का स्थान वाक्य के अंत में होता है, तो अंग्रेजी का शब्दक्रम कर्ता+क्रिया+कर्म (SVO) पर आधारित है। इस क्रम संरचना में क्रिया कर्ता और कर्म के बीच प्रयुक्त होती है। हिंदी के वाक्य में पदक्रम अधिक लचीला होता है। जैसे-

हिंदी : मैंने ये दवाइयाँ सरोजिनी मार्केट से खरीदीं।

अंग्रेजी : I purchased these medicines from Sarojini market.

उपर्युक्त हिंदी के वाक्य को और भी क्रमों में देखा जा सकता है

ये दवाइयाँ मैंने सरोजिनी मार्केट से खरीदीं ।

मैंने सरोजिनी मार्केट से खरीदीं ये दवाइयाँ।

सरोजिनी मार्केट से मैंने खरीदीं ये दवाइयाँ।

ये दवाइयाँ सरोजिनी मार्केट से खरीदीं मैंने।

⁶ सिंह, सूरजभान (2003) : अंग्रेजी-हिंदी अनुवाद व्याकरण, पृ. 111

लेकिन उसी तरह अंग्रेजी वाक्य में शब्दक्रम परिवर्तन संभव नहीं।

*From Sarojini market I purchased these medicines, * These medicines from Sarojini market I purchased, * Purchased I these medicines from Sarojini market.

उपर्युक्त अंग्रेजी वाक्य से ज्ञात होता है की अंग्रेजी का शब्दक्रम स्थिर है, जिसमें बदलाव विशेष रूप से, प्रश्नवाचक वाक्य में सहायक क्रियाओं का वाक्य के प्रारंभ में स्थांनातंत्रित हो जाने से होता है। जैसे - *Are you Mr. Shahrukh?* इसके विपरीत हिंदी के प्रश्नवाचक वाक्यों में पदक्रम अधिक लचीला है, जैसे- *क्या आप शाहरुख हैं? आप क्या शाहरुख हैं?* आदि। अंग्रेजी में यह प्रश्नवाचक वाक्य का शब्दक्रम अपवाद है। इस तरह का शब्दक्रम बदलाव हिंदी से अंग्रेजी अनुवाद में भी होता है। संभावित नियम है :

नियम : कर्ता + कर्म + क्रिया (SOV) → कर्ता + क्रिया + कर्म (SVO)

यह नियम, हिंदी अंग्रेजी संरचना के लिए प्रयुक्त किया जाता है।

इस अध्याय में संज्ञापद का सम्मिलन कर उनके साथ प्रयोग किए जाने वाले घटकों के लिए नियम बनाए गए हैं। जिनको मशीनी अनुवाद के लिए प्रयुक्त किया जा सकता है। यह नियम सैद्धांतिक होने से इसके लिए अल्गोरिदम बनाकर इसका प्रयोग किया जा सकता है। तब यह नियम हिंदी से अंग्रेजी अनुवाद में सहायक हो पाएँगे।



उपसंहार

उपसंहार (Conclusion)

प्रस्तुत लघु शोध-प्रबंध के माध्यम से यह दर्शाने का प्रयास किया गया है कि मशीनी अनुवाद के क्षेत्र में किए जा रहे प्रयासों से प्रत्येक भारतीय भाषा में मशीनी अनुवाद तंत्र निर्माण के लिए आज देश के प्राइवेट संस्थान, उच्च शिक्षा संस्थान, आई.आई.आई.टी., सी-डैक, भारतीय भाषा के लिए प्रौद्योगिकी विकास उपक्रम आदि संस्थान इस क्षेत्र में कार्यरत हैं। मशीनी अनुवाद एक व्यापक क्षेत्र है और इस विषय पर आज अनेक शोध हो रहे हैं। यह शोध भी हिंदी से अंग्रेजी मशीनी अनुवाद तंत्रों का विकास करने के लिए से सहायक हो सकता है। यह शोध अनुवाद के संबंध में होने के कारण मशीनी अनुवाद के लिए चुनैतियों पर भी विचार करता है। जिस तरह साहित्य के क्षेत्र में मानव द्वारा अनुवाद किया जाता है उसी तरह क्या मशीन अनुवाद कर पाएगी ? इस प्रश्न पर भी विचार होता है। अनुवाद जब एक भाषा से दूसरी भाषा में किया जाता है, तो उसके भावार्थ के साथ सांस्कृतिक यथार्थ का पूर्ण रूप से अनुवाद होना आवश्यक है। तभी इस अनुवाद के माध्यम से दोनों भाषाओं में संप्रेषण साध्य हो सकता है। इस दृष्टिकोण और प्रश्नों की दिशा में अभी शोध होना बाकी है। फिर भी विभिन्न विषय क्षेत्रों में मशीन के माध्यम से सटीक और शुद्ध रूप से अनुवाद किया जा रहा है। अगर प्राकृतिक भाषा संसाधन (Natural Language Processing) के उद्देश्य से देखा जाए तो किसी भी भाषा को विकसित करना प्रा.भा.स. का उद्देश्य है। इसलिए यह कह सकते हैं कि मशीन अनुवाद के क्षेत्र में आगे शोध होते रहेंगे, तो भारत की सभी भाषाओं के लिए मशीनी अनुवाद तंत्र का निर्माण करना आसान होगा और भारत में सूचना प्रौद्योगिकी का विकास होगा।

यह शोध हिंदी-अंग्रेजी मशीनी अनुवाद तंत्र निर्माण के लिए सहायक सिद्ध हो, इस उद्देश्य से इस शोध में हिंदी-अंग्रेजी वाक्यों का कॉर्पोरा आधारित संज्ञापद का सम्मिलन (Mapping) किया है। इसके लिए हिंदी और अंग्रेजी संज्ञापद का अध्ययन किया है और विभिन्न उदाहरणों के द्वारा संज्ञापद के साथ प्रयोग किए जाने वाले कारक, परसर्ग, पूर्वसर्ग, निर्धारक, विशेषक, सर्वनाम, विशेषण, संयोजक, संख्यावाचक आदि का कॉर्पोरा के उदाहरणों के आधार पर विश्लेषण किया गया है। इस शोध की सार्थकता भविष्य में यह हो सकती है कि, अगर हिंदी-अंग्रेजी नियम आधारित मशीनी अनुवाद तंत्र का निर्माण किया जा रहा है तो इस तंत्र के लिए मानव निर्मित इन नियमों का प्रयोग किया जा सकता है। इसके प्रयोग से इस तंत्र के निर्माण में कार्यक्षमता और शुद्धता में सुधार किया जा सकता है।

इस शोध के प्रथम अध्याय में भारतीय मशीनी अनुवाद तंत्र के बारे में चर्चा की गयी है। वर्तमान में भारतीय भाषा के लिए कार्य कर रहे विश्व स्तर पर मशीनी अनुवाद तंत्रों का भी अध्ययन किया गया है। इन तंत्रों के अनुवाद की शुद्धता की जाँच के लिए इन तंत्रों में अनुवाद के लिए वाक्य दिए गए हैं। इसके लिए वैश्विक मशीनी अनुवाद तंत्रों (गूगल, माइक्रोसाफ्ट, बेबेल फिश आदि) में कार्पोरा के हिंदी के 2000 वाक्य अनुवाद के लिए दिए गए और अनुवाद प्राप्त करने के उपरांत उसका अनुवाद कितना प्रतिशत सही है- इसकी जाँच त्रुटि विश्लेषण के माध्यम से की गई है। इस त्रुटि विश्लेषण से यह ज्ञात हुआ कि ये तंत्र कितना प्रतिशत अनुवाद भारतीय भाषाओं के लिए सही करते हैं और इन्हें पुनः विकसित करने के लिए किस सामग्री की आवश्यकता है इसका परिक्षण भी किया गया है। इन तंत्रों को अधिक विकसित करने के लिए बृहत कार्पोरा, वर्तनी जाँचक, रूपवैज्ञानिक विश्लेषक, हिंदी से अंग्रेजी शब्दकोश संकलन, हिंदी प्रजनक, अंग्रेजी प्रजनक, स्थानांतरण व्याकरणिक नियम (Transfer Grammar rule) आदि का मशीन में समावेश कर संचालित करने की आवश्यकता है।

द्वितीय अध्याय में हिंदी और अंग्रेजी की वाक्य संरचना के स्तर पर दोनों भाषाओं की संरचना की भिन्नता को अनुवाद के दृष्टि से स्पष्ट किया गया है। इस भिन्नता को वाक्य संरचना और दोनों भाषाओं के शब्दक्रम के तौर पर देखा गया है। साथ ही दोनों भाषाओं के विभिन्न वाक्य के प्रकारों को प्रमुख रूप से स्पष्ट किया गया है। दूसरी ओर हिंदी-अंग्रेजी संज्ञापद का व्यतिरेकी विश्लेषण कार्पोरा के आधार पर किया गया है। इस अध्ययन में व्यतिरेकी संरचना में आनेवाले संज्ञापद को उदाहरण के माध्यम से स्पष्ट किया है और एक भाषा से दूसरी भाषा के अनुवाद में संज्ञापद के साथ होनेवाले बदलाव को दोनों भाषाओं के संज्ञापद से रेखांकित किया गया है। साथ ही हिंदी और अंग्रेजी के संज्ञापद में कारक, निर्धारक, विशेषक, आदि में होनेवाले बदलावों को भी रेखांकित किया गया है, क्योंकि इन घटकों का प्रयोग संज्ञा के साथ हमेशा होता है। इसलिए हिंदी और अंग्रेजी की संरचना में इन घटकों का होना महत्वपूर्ण है जिनके माध्यम से इन भाषाओं में अनुवाद के स्तर पर अर्थ अभिप्रेत होता है। इस व्यतिरेकी विश्लेषण में इन प्रमुख घटकों का अध्ययन किया गया है।

तृतीय अध्याय में स्थानांतरण व्याकरण की जानकारी दी है। इस व्याकरण का प्रयोग मशीनी अनुवाद के लिए अधिक महत्वपूर्ण है। इस व्याकरण के माध्यम से मशीन के लिए एक भाषा से दूसरी भाषा में अनुवाद के लिए नियम बनाए जाते हैं। इस तरह नियमों का प्रयोग कर नियम आधारित तंत्र निर्माण के लिए सामग्री विकसित करने का कार्य हैदराबाद विश्वविद्यालय कर रहा है, जिसने संपर्क तंत्र का निर्माण किया है। यह तंत्र इस नियम के आधार पर वर्तमान में 6 भाषाओं में भाषा युग्मों (Language pairs) के साथ अनुवाद प्रदान कर रहा है। यह अनुवाद

तंत्र अन्य भाषाओं के लिए भी विकसित किया जा रहा है। संपर्क तंत्र में जिस तरह से स्थानान्तरण व्याकरण के नियमों का निर्माण किया है, उसी तरह चतुर्थ अध्याय में स्थानान्तरण व्याकरण नियम संज्ञापद के लिए बनाए गए हैं। इस नियमों का निर्माण हिंदी-अंग्रेजी नियम आधारित तंत्रों के लिए सहायक हो सकता है।

चतुर्थ अध्याय में हिंदी अंग्रेजी दोनों भाषाओं के कार्पोरा आधारित 2000 वाक्यों का संरचना के साथ संज्ञापद का सम्मिलन (Mapping) किया गया है और पाया गया कि इस सम्मिलन में दोनों भाषाओं में क्या-क्या भिन्नता आ रही है। इस सम्मिलन के बाद हिंदी से अंग्रेजी में अनुवाद के लिए अंग्रेजी प्रजनक की भूमिका महत्वपूर्ण मानी जाती है। इस भिन्नता को सम्मिलित करते हुए हिंदी से अंग्रेजी के लिए सैद्धांतिक रूप से स्थानान्तरण व्याकरणिक नियमों का निर्माण किया गया है। ये नियम दोनों भाषाओं के अनुवाद में आनेवाली समस्याओं को सहज बना सकते हैं। इस तरह के नियम किसी भी भाषा के तंत्र निर्माण के लिए निर्मित किए जा सकते हैं। इस शोध के प्रमुख अध्ययन के रूप में कार्पोरा के आधार पर हिंदी अंग्रेजी वाक्य में आनेवाले संज्ञापद का विस्तृत अध्ययन करते समय पाया गया कि संज्ञापद के साथ प्रयोग होनेवाले कारक, विशेषण, संख्यावाचक, अव्यय, संयोजक, शब्दक्रम, लिंग, वचन आदि घटकों संबंधी अनेक चुनौतियाँ अनुवाद में होती हैं। इन घटकों को सम्मिलन के माध्यम से नियमबद्ध किया जा सकता है। फिर भी इन घटकों का प्रयोग दोनों भाषाओं में अलग-अलग संदर्भों में होता है, वह अनुवाद की दृष्टि से मशीन को सिखाना (Machine learning) आवश्यक है।

प्रत्येक शोध की समय सीमा होती है उस समय के अंतर्गत शोध को नया रूप देना होता है। भविष्य में इस शोध को बढ़ाने के लिए यह प्रयास हो सकता है कि संज्ञापद के बाद क्रियापद पर शोध किया जा सकता है, जिससे हिंदी से अंग्रेजी म. अनुवाद तंत्र के निर्माण के लिए पूर्ण नियमों का निर्माण होगा। अगर इस तंत्र का निर्माण किया जाएगा तो इसके लिए विभिन्न विषय सामग्री संकलन, रूपवैज्ञानिक विश्लेषक, वर्तनी जाँचक, टैगर, द्विभाषी शब्दकोश, आदि का संकलन और निर्माण करना होगा। अतः इस शोध प्रबंध के निष्कर्ष के रूप में मशीनी अनुवाद को प्रगत बनाने के लिए कुछ सुझाव दिए जा सकते हैं।

1. म.अनु. की सफलता के लिए या निर्माण के लिए दोनों भाषाओं की भाषाई सामग्री (Corpus) का विकास करना आवश्यक होता है।
2. मशीनी अनुवाद के शब्दकोश (Lexicon) को अधिक विस्तृत बनाया जाना चाहिए साथ ही, इसमें मुहावरों (Idioms) और लोकोक्तियों के लिए अलग से कोश बनाया जाना चाहिए।

3. यदि नियम आधारित पद्धति (Rule Based Method) से मशीनी अनुवाद प्रणाली का निर्माण करना है तो इसमें इसके व्याकरण संबंधी नियमों को अधिक स्पष्ट बनाना चाहिए। भाषा में हो रहे प्रयोगों को अधिक से अधिक नियम के रूप में बनाकर व्याकरण में देना चाहिए। यदि सांख्यिकी आधारित अनुवाद प्रणाली (Statistical Based Translation System) का निर्माण हो तो इस अध्ययन के लिए अधिक से अधिक सभी क्षेत्रों के कॉर्पस कंप्यूटर में समावेशित करना आवश्यक है।
4. अनुवाद में बहुअर्थकता की समस्या के लिए संदर्भ को समझने योग्य प्रणाली विकसित करने का प्रयत्न करना चाहिए, और एक से अधिक अर्थ चयन की उपयुक्तता होनी चाहिए।
5. भाषाओं की व्याकरण संबंधी सभी विशेषताओं को ध्यान में रखते हुए दृढ़ और सुस्पष्ट नियम तैयार किए जाएंगे, तो अनुवाद सहज और सरल रूप में प्राप्त किया जा सकता है।

अतः ये सुझाव मशीनी अनुवाद के लिए महत्वपूर्ण रूप से सहायक हो सकते हैं।



1. संदर्भ ग्रंथ-सूची

2. परिशिष्ट

1. संदर्भ ग्रंथ-सूची

(Bibliography)

प्राथमिक स्रोत : आधार सामग्री

- जवाहरलाल नेहरू विश्वविद्यालय, के विशिष्ट संस्कृत अध्ययन केन्द्र की परियोजना “भारतीय भाषा कॉर्पोरा उपक्रम” का कॉर्पोरा (कॉर्पोरा का स्वरूप स्वास्थ्य और पर्यटन क्षेत्र : हिंदी-अंग्रेजी के टैग वाक्य)

द्वितीय स्रोत : हिंदी पुस्तकें

1. अय्यर, एन.ई. विश्वनाथ (2007) : *अनुवाद कला*, प्रभात प्रकाशन, दिल्ली ।
2. बाहरी, हरदेव (2007) : *व्यावहारिक हिन्दी व्याकरण तथा रचना*, लोकभारती प्रकाशन, इलाहाबाद ।
3. बालचंद्रन, लक्ष्मीबाई (1988) : *हिंदी का कारक व्याकरण*, केन्द्रीय हिन्दी संस्थान, आगरा ।
4. भारतभूषण (1984) : *हिंदी पंजाबी क्रिया पदबंध*, केन्द्रीय हिन्दी संस्थान, आगरा ।
5. भारतभूषण (1984) : *हिंदी-पंजाबी क्रिया पदबंध : व्यतिरेकी विश्लेषण*, केन्द्रीय हिंदी संस्थान, आगरा।
6. भाटिया, कैलाश चन्द्र (1987) : *अंग्रेजी-हिंदी शब्दों का ठीक प्रयोग*, प्रभात प्रकाशन, नई दिल्ली।
7. गुरु, कामता प्रसाद (1920) : *हिंदी व्याकरण*, काशी नागरी प्रचारिणी सभा, काशी ।
8. गोस्वामी, कृष्ण कुमार (2012) : *अनुवाद विज्ञान की भूमिका*, राजकमल प्रकाशन, दिल्ली ।
9. जैन, वृषभ प्रसाद (1995) : *अनुवाद और मशीनी अनुवाद*, सारांश प्रकाशन, नई दिल्ली।
10. काचरू, यमुना (1997) : *हिंदी रूपान्तरणात्मक व्याकरण के कुछ प्रकरण*, केन्द्रीय हिंदी संस्थान, आगरा ।
11. किशोरीदास बाजपेयी (1959) : *हिंदी शब्दानुशासन*, नागरी प्रचारिणी सभा, वाराणसी ।
12. मल्होत्रा, विजय कुमार(2007) : *कंप्यूटर के भाषिक अनुप्रयोग*, वाणी प्रकाशन, नई दिल्ली।
13. नगेन्द्र, (संपा.), (1993) : *अनुवाद विज्ञान:सिद्धांत और प्रयोग*, हिंदी माध्यम कार्यान्वय निदेशालय, दिल्ली विश्वविद्यालय।
14. नजीर, आरिफ (2005) : *हिन्दी में अनुवाद की भूमिका और द्विभाषी कम्प्यूटरीकरण*, अनंत प्रकाशन, नई दिल्ली ।
15. पालीवाल, रीतारानी (2004) : *अनुवाद प्रक्रिया और परिदृश्य*, वाणी प्रकाशन, दिल्ली ।

16. पाण्डेय, राम कमल (1985) : *त्रुटि विश्लेषण: सिध्दांत और व्यवहार*, केन्द्रीय हिन्दी संस्थान, आगरा।
17. राजगोपालन (1973) : *हिंदी का भाषावैज्ञानिक व्याकरण*, केन्द्रीय हिंदी संस्थान, आगरा ।
18. रस्तोगी, कविता (2010) : *समसामायिक अनुप्रयुक्त भाषाविज्ञान*, अविराम प्रकाशन, दिल्ली ।
19. रेड्डी, विजयराघव (1998): *हिंदी-शिक्षण और भाषा-विश्लेषण*, साहित्य सहकार, नई दिल्ली।
20. रेड्डी, विजयराघव (1987) : *हिंदी तेलुगु संज्ञा पदबंध*, केन्द्रीय हिन्दी संस्थान, आगरा ।
21. सहगल मनमोहन (2008) : *हिन्दी शोध तंत्र की रूपरेखा*, पंचशील प्रकाशन, जयपूर।
22. सहाय, चतुर्भुज (2007) : *हिंदी पदविज्ञान*, कुमार प्रकाशन, आगरा ।
23. सहाय, चतुर्भुज; पिल्ले टी.के.नारायण (संपा.), (1992) : *हिंदी क्रिया, काल, पक्ष एवं वृत्ति*, केन्द्रीय हिन्दी संस्थान, आगरा ।
24. शर्मा, राजमणि (1996) : *आधुनिक भाषा विज्ञान*, वाणी प्रकाशन, नई दिल्ली ।
25. शर्मा, राजमणि (2004) : *अनुवाद विज्ञान*, हरियाणा साहित्य अकादमी, पंचकूला ।
26. शर्मा, राजमणि (2010): *हिन्दी भाषा इतिहास और स्वरूप*, वाणी प्रकाशन, नई दिल्ली ।
27. शर्मा, रमेश चन्द्र (1999) : *हिंदी-अंग्रेजी व्याकरणिक संरचना*, अक्षरिणी प्रकाशन, नई दिल्ली।
28. शर्मा, रामकिशोर (2010) : *भाषा चिन्तन के नए आयाम*, लोकभारती प्रकाशन, इलाहाबाद।
29. शर्मा, रामकिशोर (2007) : *भाषाविज्ञान हिन्दी और भाषा और लिपि*, लोकभारती प्रकाशन, इलाहाबाद ।
30. श्रीवास्तव, रवीन्द्रनाथ; गोस्वामी, कृष्णकुमार (2010) : *अनुवाद सिध्दांत और समस्याएँ*, आलेख प्रकाशन, दिल्ली ।
31. सिंह, हरिवंश (2003) : *कारक व्याकरण की दृष्टि से हिन्दी परसर्ग विचार*, केन्द्रीय हिन्दी संस्थान, आगरा।
32. सिंह, सूरजभान (2003) : *अंग्रेजी-हिंदी अनुवाद व्याकरण*, प्रभात प्रकाशन, नई दिल्ली ।
33. सिंह, सूरजभान (2000) : *हिंदी का वाक्यात्मक व्याकरण*, साहित्य सहकार, नई दिल्ली ।
34. तिवारी, बालेन्दु शेखर (संपा.), (2010) : *अनुवाद विज्ञान*, प्रकाशन संस्थान, नई दिल्ली ।
35. तिवारी, भोलानाथ (2011) : *अनुवाद विज्ञान सिध्दांत एवं प्रविधि*, किताबघर प्रकाशन, दिल्ली ।
36. तिवारी, भोलानाथ (2005) : *हिन्दी भाषा*, किताब महल, इलाहाबाद ।
37. तिवारी, भोलानाथ (1999) : *हिंदी भाषा की संरचना*, वाणी प्रकाशन, नई दिल्ली ।
38. तिवारी, भोलानाथ (1986) : *हिंदी की रूप संरचना*, साहित्य सहकार, नई दिल्ली ।
39. त्रिपाठी, रामदेव (1986) : *हिन्दी भाषानुशासन*, बिहार हिन्दी ग्रन्थ अकादमी, पटना ।
40. वर्मा, शिवेन्द्र किशोर (2005) : *भाषा अध्ययन के विविध पक्ष*, आलेख प्रकाशन, नई दिल्ली ।

41. वर्मा, शिवेन्द्र किशोर (1990) : *भाषाविज्ञान की आधुनिक प्रवृत्तियाँ और हिन्दी भाषा शिक्षण*, केन्द्रीय हिन्दी संस्थान, आगरा।

English Books:

42. Abbi, Anvita (1980) : *Semantic Grammar of Hindi: A Study in Reduplication*, Bahri Publications, New Delhi.
43. Arnold Douglas, Balkan Lorna, Meijer Siety, Humphreys R. Lee, Sadler Louisa (1994) : *Machine Translation: an Introductory Guide*, Blackwell Publishers, London UK.
44. Bharti Akshar, Chaitanya Vineet, Sangal Rajeev. (2000) : *Natural Language Processing : a Paninian Perspective*, Prentice –Hall of India, New Delhi.
45. Chomsky, N. (1957) : *Syntactic Structure*, The Hagun Mouton.
46. Chomsky, N. (1965) : *Aspect of the theory of Syntax*, Cambridge Mass, MIT Press.
47. Dorr, Bonnie (1993) : *Machine Translation a view from the lexicon*, The MIT Press, USA.
48. Hook, Peter Edwin (1974) : *The Compound Verb in Hindi*. Ann Arbor: The University of Michigan.
49. Jacobs, R.A. & Rosenbaum, P.S. (1968): *English Transformational Grammar*, Blaisdell, London.
50. Jacobsen, Bent (1977) : *Transformational Generative Grammar: an introductory survey its genesis and development*, North–Holland publishing company, Amsterdam. New York. Oxford
51. Kachru, Yamuna (1966) : *An Introduction to Hindi Syntax*, The University of Illinois, Urbana.
52. Kachru, Yamuna (1980) : *Aspects of Hindi Grammar*, Manohar Publications, New Delhi
53. Kachru, Yamuna (1984) : *Hindi*, John Benjamin's Publishing Company, Amsterdam/Philadelphia.
54. Karunkaran K. & Jaykumar M. (ed.), (1988) : *Translation As Synthesis*, Bahari Publications, New Delhi.
55. Kellog, S. H. (3rd edi.1963) : *A Grammar of the Hindi Language*, Routledge and Kegan Paul, London.
56. Koul ,Omkar N. (2008) : *Modern Hindi Grammar*, Dunwoody Press, USA.

57. Koul, Omkar N. (ed.), (1982): *Topics in Hindi Linguistics, Vol 2*, Bahri Publications, New Delhi.
58. Koutsoudas, Andreas (1966): *Writing Transformational Grammars: an introduction*, McGraw- Hill Book Company, New York.
59. Lutze Luthar & Bahadur Singh (1970) : *Hindi As A Second Language*, Radhakrishna Prakashan, New Delhi.
60. McGregor, R. S. (1995): *Outline of Hindi Grammar*, Oxford University Press, Oxford. (3rd edition).
61. Roderick A. Jacobs, Peter S. Rosenbaum, Paul M. Postal (Edi.), (1968): *English transformational grammar*, Blaisdell Publishing Company, Waltham, Massachusetts /Toronto/ London.
62. Sharma, Aryendra (1958) : *A Basic Grammar of Modern Hindi*, Central Hindi Directorate, New Delhi. (5th Edi.1994)
63. Sinha, K.Binod (1986) : *Contrastive Analysis of English and Hindi Nominal Phrase*, Bahari Publications, New Delhi.
64. Slocum, Jonathan (ed.), (1988) : *Machine Translation Systems*, Cambridge University Press, New York.
65. Thakur K P. (2011) : *A Practical Guide to English Grammar*, Bharati Bhawan Publisher & distributors, New Delhi.
66. Thomson A.J., Martinet A.V. (edi.), (1987): *A Practical English Grammar*, Oxford University Press, New Delhi.
67. Verma, M.K. (1971): *The Structure of Noun Phrase in English and Hindi*, Motilal Banarsidass, Delhi.
68. Whitelock Peter & Kilby Kieran (1995): *Linguistic and Computational Techniques in Machine Translation System Design*, UCL Press limited, University Press Limited, London.
69. Wren & Martin, (2012): *High School, English Grammar & Composition*, S.Chand & Company LTD. New Delhi.
70. Yule, George (1996) : *The study of Language*, Cambridge University Press, United Kingdom.

Research Papers (शोध पत्र):

71. D. Gupta and N. Chatterjee (2003) : *Identification of Divergence for English to Hindi EBMT*, In Proceeding of MT Summit-IX: 141-148.

72. Dave Shachi, Parikh Jignashu, Bhattacharyya Pushpak, (2001): *Interlingua-Based English-Hindi Machine Translation and Language Divergence*, Source: Machine Translation, Vol. 16, No. 4, 251-304 Published by: Springer
73. Durgesh, Rao (2001) : *Machine Translation in India: Brief survey*, In Proceeding of SCALLA 2001 Conference, Bangalore, India,
74. Dwivedi, Sanjay Kumar & Sukhadeve Pramod Premdas (2010) : *Machine Translation System in Indian Prospectives*, IJCA Journal of Computer Science 6 (10): 1111-1116, ISSN 1549-3636 © 2010 Science Publications.
75. Mala Christopher, (2009): *Implementation of Transfer Grammar in Telugu-Hindi Machine Translation System*, Proceedings of SCONLI-2009: 3rd Students Conference of Linguistics in India.
76. Naskar Sudip & Bandyopadhyay Sivaji, (2004) : *Use of Machine Translation in India: Current Status*, AAMT Journal NO.36, Oct.2004.
77. R Ananthakrishnan, M Kavitha, J Hegde Jayprasad, Chandra Shekhar, Ritesh Shah, Sawani Bade, M Sasikumar, (2006) : *MaTra: A Practical Approach to Fully-Automatic Indicative EnglishHindi Machine Translation*, In the proceedings of MSPIL-06, 2006.
78. R.M.K. Sinha and Anil Thakur (2005) : *Translation Divergence in English-Hindi MT*, In the Proceeding of EAMT Xth Annual Conference, Budapest, Hungary, 30-31 May.
79. Sinha, R. M. K., Jain, R. and Jain A. (2001): *Translation from English to Indian languages: ANGLABHARTI Approach*, In proceedings of Symposium on Translation Support System STRANS 2001. February 15-17, IIT Kanpur, India
80. Sitender, Bawa Seema (2012) : *Survey of Indian Machine Translation Systems*, JCST Vol.3, Issue 1, Jan-March 2012, ISSN:0976-8491(online)

Thesis and Dissertation (शोध प्रबंध और लघु-शोध प्रबंध) :

81. कांबले प्रकाश अभिमन्यु (2009) : *हिंदी-मराठी मशीनी अनुवाद में शाब्दिक अस्पष्टता (समान उच्चारण वाले शब्दों के संदर्भ में) एम.फिल.* (लघु-शोध प्रबंध), ज.ने.वि; नई दिल्ली.
82. Asamidinova Aainura (2007): *Knowledge Base for Russian-English Machine Translation Divergences*, (Ph.D Thesis) JNU, Delhi.
83. Avinesh, P.V.S. (2010) : *Transfer Grammar Engine and Automatic Learning of Reorder Rules in Machine Translation*, M.Sc.(Thesis), LTRC-IIIT, Hyderabad.
84. Banerjee, Esha (2011): *Intra-phrasal agreement within the Hindi Noun Phrase : Grammar Checker Approach*, M.Phil (dissertation), JNU, New Delhi.

85. Choudhary, Narayan Kumar (2006) : *Developing a Computational Framework for the verb morphology of Great Andamanese*, M.Phil (dissertation), JNU, New Delhi
86. Goyal, Vishal (2010) : *Development of a Hindi to Punjabi Machine Translation System* PhD (Thesis), Department of Computer Science, Punjabi University, Patiala.
87. Gupta, Deepa (2005) : *Contributions to English to Hindi Machine Translation Using Example Based Approach*, PhD (Thesis), Department of Mathematics, IIT, Delhi.
88. Mathur, Prashant (2011) : *Automatic Translation of Noun Compounds from English to Hindi*, M.Sc.(Thesis), LTRC-IIIT, Hyderabad.
89. Mohamed Azza Abd Ei- Moniem, (2000) : *Machine Translation of Noun Phrases: From English to Arabic*, M.Sc.(Computer Science Thesis), Faculty of Engineering, Cairo University, Giza, Egypt.
90. Pathak, Kumar Nripendra (2011) : *challenges in Sanskrit-Hindi Noun Phrase Mapping*, M.Phil (dissertation), JNU, New Delhi

पत्रिकाएँ (Magazines) :

91. गवेषणा, अंक-93, जनवरी-मार्च (2009) : *भारत में कोशविज्ञान पर विशेष*, केंद्रीय हिन्दी संस्थान, आगरा ।
92. गवेषणा, अंक-99, अप्रैल-जून (2011) : *शिक्षा पर विशेष*, केंद्रीय हिन्दी संस्थान, आगरा ।
93. गवेषणा, अंक-90, अप्रैल-जून (2008) : *भाषाविज्ञान आज*, केंद्रीय हिन्दी संस्थान, आगरा।
94. गवेषणा, अंक-89, जनवरी-मार्च, (2008) : *भाषा शिक्षण पर विशेष*, केंद्रीय हिन्दी संस्थान, आगरा।
95. गवेषणा, अंक-88, अक्टूबर-दिसंबर (2007) : *केन्द्रीय हिन्दी संस्थान, आगरा, संपादक- शंभूनाथ*
96. अनुवाद शतक विशेषांक, अंक:100-101, जुलाई-दिसंबर (1999) : *भारतीय अनुवाद परिषद, नई दिल्ली।*
97. Vishwbharat, A Ministry of Information Technology Publication on Language Technology in India, Issues, May,2001., January, 2003; October, 2003; Jan-April-2004,

शब्दकोश (Dictionary) :

98. भाषाविज्ञान शब्दावली (1992) : *वैज्ञानिक तथा तकनीकी शब्दावली आयोग, नई दिल्ली।*
99. Bulcke, Father K. (2002) : *An English- Hindi Dictionary*, S. Chand & Company Ltd., New Delhi.

Encyclopaedia (विश्वकोश)

100. David, Crystal (1987) : The Cambridge Encyclopaedia of Language, Cambridge University Press, New Work.

INTERNET WEBSITE and links

http://en.wikipedia.org/wiki/Google_Translate

<http://translate.google.com/>

<http://www.microsofttranslator.com/>

<http://www.shabdkosh.com/>

http://en.wikipedia.org/wiki/Yahoo!_Babel_Fish

http://babelfish.yahoo.com/translate_txt

http://www.tdil.mit.gov.in/Research_Effort.aspx

<http://www.cse.iitk.ac.in/users/rmk/>

<http://www.languageinindia.com>

<http://tdil-dc.in>

<http://anusaaraka.iiit.ac.in/>

<http://ltrc.iiit.ac.in>

<http://www.mantra-rajbhasha.cdac.in/mantrarajbhasha/>

<http://www.cdacindia.com/html/about/success/mantra.aspx>

<http://pune.cdac.in>

<http://www.ncst.ernet.in/matra/>

<http://shakti.iiit.ac.in/>

<http://www.languagetechnologies.ac.in/lerc/mat/mat.htm>

<http://www.cfilt.iitb.ac.in/>

http://www.au-kbc.org/research_areas/nlp/projects/anusaaraka.html

<http://sampark.iiit.ac.in>

<http://www.languageinindia.com>

2. परिशिष्ट (Appendix)

भारतीय भाषा कॉर्पोरा उपक्रम (ILCI) के टैग वाक्य

1. अध्ययनों\N_NN ने\PSP यह\DM_DMD प्रमाणित\N_NN कर\V_VM दिया\V_VAUX है\V_VAUX
ENG - Studies\NNS have\VBP proved\VBN that\IN
2. सी.डी.सी.\N_NN ने\PSP बताया\V_VM है\V_VAUX कि\CC_CCS मृत्यु\N_NN दर\N_NN 15\QT_QTC से\PSP 50\QT_QTC प्रतिशत\N_NN तक\PSP मानी\V_VM है\V_VAUX I\RD_PUNC
ENG- CDC\NNP has\VBZ said\VBN that\IN the\DT death\NN rate\NN is\VBZ considered\VBN to\TO be\VB 15\NNS .\.
3. डब्ल्यू.एच.ओ\N_NNP ने\PSP 2001\QT_QTC में\PSP जारी\JJ फैक्ट\N_NN शीट\N_NN में\PSP कहा\V_VM था\V_VAUX
ENG- The\DT WHO\WP has\VBZ said\VBN in\IN a\DT fact\NN sheet\NN issued\VBN in\IN 2001\CD
4. सभी\QT_QTF ने\PSP गंगा\N_NNP स्नान\N_NN कर\V_VM पूजा-अर्चना\N_NN ,\RD_PUNC कथा\N_NN आदि\N_NN की\PSP I\RD_PUNC
ENG- having\VBG bathed\VBN in\IN the\DT Ganges\NNP ,\ , all\DT offered\VBN prayers\NNS and\CC religious\JJ stories\NNS .\.
5. भगवान\N_NN महावीर\N_NNP स्वामी\N_NN ने\PSP अपनी\PR_PRF दिव्य\JJ वाणी\N_NN में\PSP प्रवचन\N_NN दिए\V_VM I\RD_PUNC
ENG- Bhagwan\NNP Mahavir\NNP Swami\NNP preached\VBD in\IN their\PRP\$ divine\JJ voice\NN .\.
6. चन्द्रादित्य\N_NNP मंदिर\N_NN का\PSP निर्माण\N_NN नाग\N_NNP राजा\N_NN चन्द्रादित्य\N_NNP ने\PSP करवाया\V_VM था\V_VAUX
ENG- Chandraditya\NNP Temple\NNP was\VBD built\VBN by\IN the\DT Naga\NNP king\NN Chandradita\NNP
7. पुर्तगालियों\N_NN और\CC_CCD डचों\N_NN ने\PSP तलशेरी\N_NNP पर\PSP अधिकार\N_NN जमाने\V_VM की\PSP चेष्टा\N_NN की\PSP थी\V_VAUX I\RD_PUNC
ENG- The\DT Portuguese\NNP and\CC the\DT Dutch\NNP had\VBD tried\VBN to\TO capture\VB Thalassery\NNP .\.
8. ताजा\JJ साँसें\N_NN और\CC_CCD चमचमाते\JJ दाँत\N_NN आपके\PR_PRP व्यक्तित्व\N_NN को\PSP निखारते\V_VM हैं\V_VAUX I\RD_PUNC
ENG- Fresh\JJ breath\NN and\CC shining\JJ teeth\NNS enhance\VB your\PRP\$ personality\NN .\.

9. ये\DM_DMD दाँतों\N_NN को\PSP गंदा\JJ और\CC_CCD साँसों\N_NN को\PSP बदबूदार\JJ बना\V_VM देते\V_VAUX हैं\V_VAUX ।\RD_PUNC

ENG- They\PRP make\VB teeth\NNS dirty\JJ and\CC breath\NN stinky\NN .\.

10. दाँतों\N_NN को\PSP ठीक\JJ से\PSP साफ\N_NN करें\V_VM ।\RD_PUNC

ENG- Clean\NNP your\PRP\$ teeth\NNS properly\RB .\.

11. बच्चों\N_NN को\PSP फास्ट\JJ फूड\N_NN ,\RD_PUNC कुरकुरे\N_NN ,\RD_PUNC आइसक्रीम\N_NN नहीं\RP_NEG खिलायें\V_VM ।\RD_PUNC

ENG- Do\VB not\RB feed\VB children\NNS fast\RB food\NN ,\, kurkure\NN ,\, ice-cream\NN .\.

12. नियमित\JJ व्यायाम\N_NN से\PSP मस्तिष्क\N_NN को\PSP ज्यादा\QT_QTF आक्सीजन\N_NN मिलती\V_VM है\V_VAUX

ENG- Brain\NN gets\VBZ more\JJR oxygen\NN with\IN regular\JJ exercise\NN ,\

13. माँ\N_NN को\PSP दूध\N_NN के\PSP लिए\PSP प्रोत्साहित\N_NN करना\V_VM ।\RD_PUNC

ENG- Encouraging\VBG mother\NN for\IN milk\NN .\.

14. सिलिंडरिकल\JJ ग्लासेस\N_NN को\PSP हमेशा\N_NST पहनना\N_NN है\V_VAUX ।\RD_PUNC

ENG- Cylindrical\JJ glasses\NNS have\VB to\TO be\VB worn\JJ always\RB .\.

15. वर्ष\N_NN 2004-2005\QT_QTC में\PSP 560\QT_QTC अप्रशिक्षित\JJ दाईयों\N_NN को\PSP प्रशिक्षित\JJ किया\V_VM गया\V_VAUX ।\RD_PUNC

ENG- 560\CD untrained\JJ midwives\NNS have\VB been\VBN trained\VBN in\IN the\DT year\NN of\IN 2004-2005\CD .\.

16. मुख्य\JJ स्नान\N_NN 13\QT_QTC नवंबर\N_NNP को\PSP है\V_VAUX ।\RD_PUNC

ENG- The\DT holy\JJ dip\NN is\VBZ on\IN 13\CD November\NNP .\.

17. राजकुमार\N_NN चन्द्रदेव\N_NNP को\PSP विरासत\N_NN में\PSP चन्द्रगिरी\N_NNP का\PSP क्षेत्र\N_NN मिला\V_VM ।\RD_PUNC

ENG- Prince\NNP Chandradeva\NNP got\VBD the\DT area\NN of\IN Chandragiri\NNP in\IN inheritance\NN .\.

18. ताड़ी\N_NNP के\PSP रसभरे\JJ फलों\N_NN को\PSP खाना\N_NN अच्छा\JJ लगता\V_VM है\V_VAUX ।\RD_PUNC

ENG- It\PRP feels\VBZ good\JJ to\TO eat\VB juicy\NN fruits\NNS of\IN palm\NN .\.

19. दमन\N_NNP के\PSP दोनों\QT_QTO भागों\N_NN को\PSP विभक्त\N_NN करने\V_VM वाली\PSP नदी\N_NN दमनगंगा\N_NNP नदी\N_NN है\V_VAUX ।\RD_PUNC

ENG- The\DT river\NN dividing\VBG Daman\NNP into\IN two\CD parts\NNS is\VBZ Damanganga\NNP river\NN .\.

20. प्राचीन\JJ केरल\N_NNP को\PSP इतिहासकार\N_NN तमिल\N_NNP भूभाग\N_NN का\PSP अंग\N_NN समझते\V_VM थे\V_VAUX ।\RD_PUNC

ENG- The\DT historians\NNS considered\VBD the\DT ancient\JJ Kerala\NNP as\IN a\DT part\NN of\IN the\DT Tamil\NNP region\NN .\.

21. मछली\N_NN खाने\V_VM से\PSP न\RP_NEG केवल\QT_QTF आपकी\PR_PRP त्वचा\N_NN ,\RD_PUNC बाल\N_NN ,\RD_PUNC दिमाग\N_NN स्वस्थ\JJ रहता\V_VAUX बल्कि\CC_CCD नजर\N_NN भी\RP_RPD तेज\JJ होती\V_VM है\V_VAUX ।\RD_PUNC

ENG- By\IN eating\VBG fish\NN not\RB only\RB your\PRP\$ skin\NN ,\, hair\NN ,\, brain\NN remain\VBP healthy\JJ but\CC eyes\NNS also\RB sharpen\VBP .\.

22. सामान्यतः\RB सुसाध्य\JJ ट्यूमरों\N_NN को\PSP शल्य-चिकित्सा\N_NN द्वारा\PSP निकाला\V_VM जा\V_VAUX सकता\V_VAUX है\V_VAUX

ENG- Generally\RB ,\, easy\JJ tumors\NNS can\MD be\VB weeded\VBN out\RP through\IN surgery\NN

23. दाँतों\N_NN से\PSP आपका\PR_PRP आत्मविश्वास\N_NN भी\RP_RPD बढ़ता\V_VM है\V_VAUX ।\RD_PUNC

ENG- Your\PRP\$ self-confidence\NN also\RB increases\VBZ with\IN teeth\NNS .\.

24. दंतचिकित्सक\N_NN से\PSP दाँतों\N_NN की\PSP जाँच\N_NN नियमित\JJ रूप\N_NN से\PSP कराएँ\V_VM ।\RD_PUNC

ENG- Get\VB the\DT teeth\NNS checked-up\JJ with\IN the\DT dentists\NNS regularly\RB .\.

25. हर\QT_QTF हफ्ते\N_NN इस\DM_DMD मिश्रण\N_NN से\PSP दाँतों\N_NN को\PSP साफ\JJ करें\V_VM ।\RD_PUNC

ENG- Clean\NNP the\DT teeth\NNS every\DT week\NN with\IN this\DT mixture\NN .\.

26. धूम्रपान\N_NN करने\V_VM से\PSP कैंसर\N_NN समेत\N_NN कई\QT_QTF बीमारियों\N_NN का\PSP खतरा\N_NN होता\V_VM है\V_VAUX ।\RD_PUNC

ENG- There\EX is\VBZ a\DT danger\NN of\IN several\JJ diseases\NNS including\NN cancer\NN with\IN smoking\NN .\.

27. लगातार\N_NN बुखार\N_NN से\PSP पीड़ित\N_NN हो\V_VM ?\RD_PUNC

ENG- I\PRP suffer\VBP from\IN fever\NN continuously\RB .\.

28. विशेष\JJ लैन्स\N_NN या\CC_CCD यंत्र\N_NN के\PSP द्वारा\PSP नेत्र\N_NN तंत्रिका\N_NN की\PSP जाँच\N_NN करेंगे\V_VM की\PSP कितनी\PR_PRQ क्षति\N_NN पहुँची\V_VM है\V_VAUX ।\RD_PUNC

ENG- We\PRP will\MD check\VB the\DT eye\NN nerves\NNS with\IN special\JJ lens\NN or\CC equipment\NN to\TO find\VB how\WRB much\JJ of\IN damage\NN has\VBZ occurred\VBN .\.

29. गर्भाशय-ग्रीवा\N_NN का\PSP कैंसर\N_NN रक्त\N_NN प्रवाह\N_NN के\PSP द्वारा\PSP भी\RP_RPD फैल\V_VM सकता\V_VAUX है\V_VAUX ।\RD_PUNC

ENG- Cervical\JJ cancer\NN can\MD spread\VB through\IN blood\NN circulation\NN .\.

30. कैंसर\N_NN की\PSP जाँच\N_NN पैथोलोजी\N_NN परीक्षण\N_NN के\PSP द्वारा\PSP की\PSP जाती\N_NN है\N_NN
 है\N_NN \RD_PUNC **ENG-**
 Cancer\NNP is\VBZ tested\VBN through\IN pathology\NN .\.
31. विकिरण\N_NN द्वारा\PSP कैंसर\N_NN कोशिकाओं\N_NN को\PSP जलाया\N_NN जाता\N_NN है\N_NN
 है\N_NN \RD_PUNC **ENG-**
 Cancer\NN cells\NNS are\VBP burnt\VBN through\IN laser\NN .\.
32. नारायणपाल\N_NNP का\PSP यह\DM_DMD विष्णु\N_NNP मंदिर\N_NN पुरातत्व\N_NN विभाग\N_NN
 द्वारा\PSP संरक्षित\N_NN स्मारक\N_NN है\N_NN \RD_PUNC **ENG-** This\DT
 Vishnu\NNP Temple\NNP of\IN Narayanpal\NNP is\VBZ a\DT protected\JJ monument\NN by\IN
 the\DT Archaeological\NNP Department\NNP .\.
33. पहले\QT_QTO ओ.आर.एस.\N_NNP का\PSP घोल\N_NN पिलायें\N_NN \RD_PUNC
ENG- First\JJ feed\NN the\DT mixture\NN of\IN O\NNP .\NP .\NP .\.
34. डायरिया\N_NNP का\PSP घरेलू\N_NN उपचार\N_NN क्या\PR_PRQ है\N_NN \RD_PUNC
ENG- What\WP is\VBZ the\DT home\NN treatment\NN of\IN diarrhoea\NN .\.
35. इस\PR_PRP समस्या\N_NN का\PSP प्रमुख\N_NN कारण\N_NN उच्च\N_NN रक्तचाप\N_NN का\PSP
 अधिक\N_NN बढ़ना\N_NN \RD_PUNC
ENG- The\DT main\JJ reason\NN of\IN this\DT problem\NN is\VBZ increasing\VBG high\JJ blood\NN
 pressure\NN
36. एच.आइ.वी.\N_NNP एड्स\N_NNP का\PSP अभी\N_NN तक\N_NN कोई\DM_DMI इलाज\N_NN
 नहीं\N_NN है\N_NN
ENG- The\DT treatment\NN of\IN H\NNP .\RP .\NP .\: is\VBZ not\RB available\JJ as\IN yet\RB
37. मोतियाबिंद\N_NNP का\PSP उपचार\N_NN केवल\QT_QTF शल्य-चिकित्सा\N_NNP द्वारा\PSP ही\N_NN
 सम्भव\N_NN है\N_NN \RD_PUNC
ENG- The\DT treatment\NN of\IN cataract\NN is\VBZ possible\JJ through\IN surgery\NN only\RB .\.
38. मुख\N_NN के\PSP ऊतकों\N_NN के\PSP परिवर्तन\N_NN को\PSP आसानी\N_NN से\N_NN देखा\N_NN
 और\CC_CCD अनुभव\N_NN किया\N_NN जा\N_NN सकता\N_NN है\N_NN \RD_PUNC
ENG- Change\NNP of\IN tissues\NNS of\IN mouth\NN can\MD be\VB easily\RB seen\VBN and\CC
 experienced\VBN .\.
39. कालाजार\N_NNP के\PSP प्रारंभिक\N_NN लक्षण\N_NN क्या\RP_RPQ हैं\N_NN \RD_PUNC
ENG- What\WP are\VBP the\DT primary\JJ symptoms\NNS of\IN kala-azar\NN .\.
40. साइटिका\N_NN एवं\CC_CCD स्लिप\N_NN डिस्क\N_NN के\PSP रोगी\N_NN इसका\PR_PRP
 अभ्यास\N_NN न\N_NN करें\N_NN \RD_PUNC
ENG- The\DT patients\NNS of\IN sciatica\NN and\CC slip\NN disk\NN should\MD avoid\VB its\PRP\$
 practice\NN .\.

41. घास\N_NN के\PSP मैदान\N_NN ,\RD_PUNC जलप्रपात\N_NN ,\RD_PUNC चीड़\N_NN वृक्ष\N_NN आदि\N_NN पीरुमेडू\N_NNP की\PSP खूबसूरती\N_NN में\PSP चार\QT_QTC चाँद\N_NN लगा\N_NN देते\N_NN हैं\N_NN ।\RD_PUNC
ENG- The\DT field\NN of\IN grass\NN ,\, waterfall\NN ,\, pine\VB tree\NN etcetera\NN increase\NN the\DT beauty\NN of\IN Peerumedu\NNP fourfold\RB .\.
42. स्तन\N_NN की\PSP लगभग\N_NN 80\QT_QTC प्रतिशत\N_NN गाँठें\N_NN कैंसर\N_NN नहीं\RP_NEG होती\N_NN ।\RD_PUNC
ENG- About\IN 80\CD percent\NN of\IN the\DT cysts\NNS of\IN the\DT breasts\NNS are\VB not\RB cancer\NN
43. अस्पताल\N_NN की\PSP सूची\N_NN जहाँ\PR_PRL कैंसर\N_NN के\PSP इलाज\N_NN की\PSP सुविधा\N_NN है\N_NN ।\RD_PUNC
ENG- List\NN of\IN hospitals\NNS where\WRB the\DT treatment\NN of\IN cancer\NN is\VBZ available\JJ .\.
44. अगर\CC_CCS कैंसर\N_NN सीमित\N_NN क्षेत्र\N_NN में\PSP और\CC_CCD प्रारम्भिक\JJ चरण\N_NN में\PSP है\N_NN ।\RD_PUNC
ENG- If\IN cancer\NN is\VBZ in\IN a\DT limited\JJ area\NN and\CC in\IN initial\JJ stage\NN
45. घर\N_NN में\PSP कीटनाशक\N_NN दवाओं\N_NN का\PSP छिड़काव\N_NN करवायें\N_NN ।\RD_PUNC
ENG- Get\VB insecticides\NNS sprayed\VB at\IN home\NN .\.
46. डेंगू\N_NNP बुखार\N_NN में\PSP कास्टायड\JJ दवाइयाँ\N_NN न\RP_NEG ली\N_NN जायें\N_NN ।\RD_PUNC
ENG- Non-steroidal\JJ drugs\NNS should\MD not\RB be\VB taken\VB in\IN dengue\NN fever\NN .\.
47. रक्त\N_NN में\PSP शर्करा\N_NN की\PSP अधिकता\N_NN दीर्घकालीन\JJ जटिलताएँ\N_NN ।\RD_PUNC
ENG- Excess\NN of\IN sugar\NN in\IN blood\NN ,\, long\JJ term\NN complexities\NNS .\.
48. शरीर\N_NN में\PSP अकड़न\N_NN ।\RD_PUNC
ENG- Stiffness\NN in\IN body\NN .\.
49. शहरी\JJ क्षेत्रों\N_NN में\PSP भी\RP_RPD आधे\QT_QTF से\N_NN अधिक\QT_QTF (\RD_SYM 56\QT_QTC प्रतिशत\N_NN)\RD_SYM प्रसव\N_NN घरों\N_NN में\PSP होते\N_NN हैं\N_NN ।\RD_PUNC
ENG- Even\RB in\IN urban\JJ areas\NNS more\JJR than\IN half\NN -LRB-\-LRB- 56\CD percent\NN -RRB-\-RRB- of\IN childbirths\NNS takes\VBZ place\NN at\IN home\NN .\.
50. इडुक्कि\N_NNP वन्यजीव\N_NN अभयारण्य\N_NN समुद्र\N_NN तल\N_NN से\N_NN 450\QT_QTC -\RD_PUNC 748\QT_QTC मीटर\N_NN ऊँचाई\N_NN पर\N_NN स्थित\N_NN है\N_NN ।\RD_PUNC
ENG- Idukki\NNP Wildlife\NNP Sanctuary\NNP is\VBZ situated\VB at\IN a\DT height\NN of\IN 450\CD -\: 748\CD metres\NNS above\IN the\DT sea\NN level\NN .\.
51. वागामण\N_NNP में\PSP ढलान\N_NN पर\N_NN चीड़\N_NN का\N_NN वन\N_NN है\N_NN ।\RD_PUNC

ENG- There\EX is\VBZ pine\VB forest\NN at\IN the\DT slopes\NNS in\IN Vagaman\NNP .\.

52. एच.आई.वी.\N_NNP जाँच\N_NN आसान\JJ ,\RD_PUNC सुरक्षित\JJ तथा\CC_CCD जन\N_NN स्वास्थ्य\N_NN केन्द्रों\N_NN पर\PSP उपलब्ध\N_NN है\V_VAUX I\RD_PUNC

ENG- H\NNP .\RP .\NP .\.: test\NN is\VBZ easy\JJ ,\ safe\JJ and\CC available\JJ at\IN public\JJ health\NN centres\NNS .\.

53. लसिका\N_NNP ग्रन्थियों\N_NN का\PSP एक\QT_QTC से\PSP अधिक\QT_QTF स्थान\N_NN पर\PSP आकार\N_NN बढ़ना\N_NN I\RD_PUNC

ENG- Increment\NN of\IN lymph\JJ nodes\NNS in\IN size\NN at\IN more\JJR than\IN one\CD place\NN .\.

54. शिप्रा\N_NNP अथवा\CC_CCD क्षिप्रा\N_NNP नदी\N_NN के\PSP तट\N_NN पर\PSP बसी\N_NN महाकाल\N_NNP की\PSP यह\DM_DMR उज्जैन\N_NNP नगरी\N_NN पतित\JJ पावनी\JJ है\N_NN I\RD_PUNC

ENG- Inhabited\VB on\IN the\DT banks\NNS of\IN Shipra\NNP or\CC Kshipra\NNP river\NN ,\ Ujjain\NNP City\NNP is\VBZ purgator\NN of\IN the\DT fallen\VB .\.

55. धरमपुरा\N_NNP रोड\N_NN पर\PSP भगवान\N_NN बालाजी\N_NNP का\PSP भव्य\JJ मंदिर\N_NN है\N_NN I\RD_PUNC

ENG- The\DT grand\JJ temple\NN of\IN Lord\NNP Balaji\NNP is\VBZ on\IN the\DT Dharampura\NNP road\NN .\.

56. यही\DM_DMD नहीं\RP_NEG जीभ\N_NN पर\PSP सफेद\JJ परत\N_NN जम\N_NN जाती\N_NN है\N_NN I\RD_PUNC

ENG- Not\RB only\RB this\DT .\T white\JJ coat\NN also\RB forms\VBZ on\IN tongue\NN .\.

57. दीव\N_NNP से\PSP मुंबई\N_NNP के\PSP लिए\PSP रोजाना\N_NN कई\QT_QTF उड़ानें\N_NN भरी\N_NN जाती\N_NN है\N_NN I\RD_PUNC

ENG- Several\JJ flights\NNS take\VB off\RP for\IN Mumbai\NNP from\IN Diu\NNP .\.

58. हायाटस\N_NNP हर्निया\N_NN को\PSP दुरुस्त\N_NN करने\N_NN के\PSP लिए\PSP I\RD_PUNC

ENG- To\TO repair\VB hiatus\NN hernia\NN .\.

59. कणिमंगलम\N_NNP पीतल\N_NN से\PSP बने\N_NN शिल्पों\N_NN और\CC_CCD दीपों\N_NN के\PSP लिए\PSP प्रसिद्ध\JJ है\N_NN I\RD_PUNC

ENG- Kanimangalam\NNP is\VBZ famous\JJ for\IN sculptures\NNS and\CC the\DT lamps\NNS made\VB of\IN brass\NN .\.

60. संयुक्त\JJ टीका\N_NN बच्चों\N_NN और\CC_CCD वयस्कों\N_NN के\PSP लिए\PSP बना\N_NN है\N_NN I\RD_PUNC

ENG- The\DT combined\VB vaccine\NN is\VBZ made\VB for\IN children\NNS and\CC adults\NNS .\.

61. लगातार\N_NN बुखार\N_NN से\PSP पीड़ित\N_NN हो\N_NN ?\RD_PUNC

ENG- I\PRP suffer\VB from\IN fever\NN continuously\RB .\.

62. शरीर\N_NN के\PSP किसी\PR_PRI भाग\N_NN या\CC_CCD अंग\N_NN से\PSP असामान्य\JJ रूप\N_NN से\PSP रक्त\N_NN या\CC_CCD मवाद\N_NN आना\N_NN \RD_PUNC

ENG- Abnormal\JJ passing\NN of\IN blood\NN or\CC pus\NN from\IN any\DT part\NN or\CC organ\NN of\IN body\NN .\.

63. डायबिटीज\N_NNP से\PSP होने\N_NN वाली\N_NN समस्याएँ\N_NN \RD_PUNC

ENG- Problems\NNS occurring\VBG from\IN diabetes\NN .\.

64. असामान्य\JJ यौन\JJ संबंधों\N_NN से\PSP दूर\N_NN रहें\N_NN \RD_PUNC

ENG- Keep\VB away\RB from\IN uncommon\JJ sexual\JJ relations\NNS .\.

65. सामान्यतः\N_NN ये\DM_DMD पथरियाँ\N_NN मूत्र\N_NN के\PSP रास्ते\N_NN शरीर\N_NN से\PSP बाहर\N_NN निकाल\N_NN दी\N_NN जाती\N_NN हैं\N_NN \RD_PUNC

ENG- Generally\RB these\DT stones\NNS are\VBP flushed\VBN out\RP from\IN the\DT body\NN through\IN the\DT urinary\JJ passage\NN .\.

66. पथरी\N_NNP से\PSP बचाव\N_NN के\PSP कुछ\QT_QTF उपाय\N_NN \RD_PUNC

ENG- Some\DT solutions\NNS of\IN safety\NN from\IN stone\NN .\.

67. दण्डक\N_NNP गुफा\N_NN कोटमसर\N_NNP से\PSP 8\QT_QTC किलोमीटर\N_NN की\N_NN दूरी\N_NN पर\N_NN स्थित\N_NN है\N_NN \RD_PUNC

ENG- Dandak\NNP cave\NN is\VBZ at\IN a\DT distance\NN of\IN 8\CD kilometres\NNS from\IN Kotamsar\NNP .\.

68. लक्सर\N_NNP से\PSP एक\QT_QTC लाइन\N_NN हरिद्वार\N_NNP तक\N_NN गई\N_NN है\N_NN \RD_PUNC

ENG- One\CD line\NN has\VBZ gone\VBN from\IN Luxar\NNP to\TO Haridwar\NNP .\.

69. ए\RP_INJ लड़के\N_NN!\RD_PUNC

ENG- Hey\UH young\JJ boy\N_NN!\.

70. शंकर\N_NNP जी\N_NN कहा\N_NN कि\CC_CCS हे\RP_INJ देवी\N_NN सतयुग\N_NN में\N_NN सभी\QT_QTF तीर्थ\N_NN पुण्य\N_NN का\N_NN फल\N_NN देते\N_NN हैं\N_NN \RD_PUNC

ENG- Shankarji\NNP said\VBD that\IN O\NNP Devi\NNP ,\, all\PDT the\DT pilgrimages\NNS give\VBP the\DT fruits\NNS of\IN meritorious\JJ actions\NNS in\IN the\DT Satayuga\NNP ,\,

71. 'भाइयों और बहनों' आपका स्वागत करते हुए मुझे बड़ी खुशी हो रही है।

ENG- 'Ladies and Gentlemen! I am very pleased to welcome you.'

72. नियमित\JJ रूप\N_NN से\PSP कराएँ\N_NN दाँतों\N_NN की\N_NN जाँच\N_NN \RD_PUNC

ENG- Get\VB the\DT teeth\NNS checked-up\JJ regularly\RB .\.

73. बासी\JJ भोजन\N_NN व\CC_CCD सड़क\N_NN किनारे\N_NN बिकने\N_NN वाले\N_NN सामान\N_NN का\N_NN सेवन\N_NN नहीं\RP_NEG करें\N_NN \RD_PUNC

85. बुखार\N_NN ज्यादा\QT_QTF लंबे\JJ समय\N_NN तक\PSP रहता\V_VM है\V_VAUX I\RD_PUNC
ENG- Fever\NN stays\NNS for\IN a\DT longer\JJR time\NN .\.
86. अधिक\QT_QTF प्यास\N_NN लगना\V_VM ,\RD_PUNC घुंधला\JJ दिखाई\V_VM देना\V_VAUX
 I\RD_PUNC
ENG- Feeling\VBG more\RBR thirsty\JJ and\CC dim\JJ vision\NN .\.
87. ढाई\QT_QTC -\RD_PUNC तीन\QT_QTC वर्षों\N_NN से\PSP हर\QT_QTF दो\QT_QTC -\RD_PUNC
 तीन\QT_QTC माह\N_NN पर\PSP फीवर\N_NN हो\V_VM जाता\V_VAUX है\V_VAUX I\RD_PUNC
ENG- Fever\NN occurs\VBZ at\IN an\DT interval\NN of\IN two\CD to\TO three\CD months\NNS
 since\IN two\CD and\CC half\NN to\TO three\CD years\NNS .\.
88. एक\QT_QTC गिलास\N_NN पानी\N_NN में\PSP दो\QT_QTC चम्मच\N_NN चीनी\N_NN एवं\CC_CCD
 एक\QT_QTC चुटकी\JJ नमक\N_NN डालकर\V_VM उबालकर\V_VM दें\V_VM I\RD_PUNC
ENG- Put\VB two\CD spoons\NNS of\IN sugar\NN and\CC a\DT pinch\NN of\IN salt\NN in\IN a\DT
 glass\NN of\IN water\NN and\CC boil\NN it\PRP .\.
89. कैंसर\N_NN 1000\QT_QTC से\PSP अधिक\QT_QTF बीमारियों\N_NN का\PSP एक\QT_QTC समूह\N_NN
 है\V_VAUX I\RD_PUNC
ENG- Cancer\NNP is\VBZ a\DT group\NN of\IN more\JJR than\IN 1000\CD diseases\NNS .\.
90. दो\QT_QTC सप्ताह\N_NN तक\PSP इन\PR_PRP लक्षणों\N_NN के\PSP बने\V_VM रहने\V_VAUX पर\PSP
 इन्हें\PR_PRP चिकित्सक\N_NN को\PSP दिखलाएँ\V_VM I\RD_PUNC
ENG- Show\NNP these\DT to\TO the\DT doctor\NN if\IN these\DT symptoms\NNS stay\VBP for\IN
 two\CD weeks\NNS .\.
91. वर्ष\N_NN 1952\QT_QTC में\PSP राष्ट्रीय\JJ स्तर\N_NN पर\PSP परिवार\N_NN कल्याण\N_NN
 कार्यक्रम\N_NN चलाया\V_VM गया\V_VAUX I\RD_PUNC
ENG- The\DT family\NN welfare\NN program\NN was\VBD run\VBN at\IN the\DT national\JJ
 level\NN in\IN the\DT year\NN of\IN 1952\CD .\.
92. एड्स\N_NNP का\PSP पहला\QT_QTO मरीज\N_NN भारत\N_NNP में\PSP 1986\QT_QTC में\PSP
 पाया\V_VM गया\V_VAUX I\RD_PUNC
ENG- The\DT first\JJ patient\NN of\IN AIDS\NNP in\IN India\NNP was\VBD found\VBN in\IN 1986\CD .\.
93. बोटिंग\N_NN और\CC_CCD बच्चों\N_NN का\PSP पार्क\N_NN दोनों\QT_QTO बच्चों\N_NN को\PSP
 भी\RP_RPD बेहद\RP_INTF पसन्द\N_NN आते\V_VM हैं\V_VAUX I\RD_PUNC
ENG- Children\NNS also\RB like\IN both\DT boating\VBG and\CC children\NNS 's\POS park\NN .\.
94. 8वीं\QT_QTO शताब्दी\N_NN की\PSP शिखर\N_NN शैली\N_NN के\PSP स्थान\N_NN पर\PSP गुंबदाकार\JJ
 नया\JJ शिखर\N_NN बनाया\V_VM गया\V_VAUX है\V_VAUX I\RD_PUNC
ENG- In\IN place\NN of\IN 8th\JJ century\NN arch\NN style\NN ,\, dome\NN like\IN new\JJ top\NN
 has\VBZ been\VBN built\VBN .\.
95. केरल\N_NNP में\PSP आवास\N_NN केन्द्रों\N_NN के\PSP विकास\N_NN का\PSP दूसरा\QT_QTO
 चरण\N_NN संघमकाल\N_NN माना\V_VM जाता\V_VAUX है\V_VAUX I\RD_PUNC

ENG- The\DT second\JJ stage\NN of\IN the\DT development\NN of\IN the\DT living\NN centers\NNS in\IN Kerala\NNP is\VBZ supposed\VBN to\TO be\VB the\DT Sangam\NNP Age\NNP .\.

96. हमारे\PRP मसूढ़ों\N_NN और\CC_CCD दाँतों\N_NN के\PSP बीच\PSP बैक्टीरिया\N_NN मौजूद\N_NN होते\N_NN हैं\N_NN ।\RD_PUNC

ENG- Bacteria\NNS stay\VB between\IN our\PRP\$ gums\NNS and\CC teeth\NNS .\.

97. बल्कि\CC_CCD साँसों\N_NN की\PSP दुर्गंध\JJ से\PSP भी\RP_RPD छुटकारा\N_NN मिलता\N_NN है\N_NN ।\RD_PUNC

ENG- But\CC one\CD also\RB gets\VBZ relief\NN from\IN stinking\JJ breath\NN .\.

98. बासी\JJ भोजन\N_NN व\CC_CCD सड़क\N_NN किनारे\N_NN बिकने\N_NN वाले\N_NN सामान\N_NN का\N_NN सेवन\N_NN नहीं\RP_NEG करें\N_NN ।\RD_PUNC

ENG- Do\VB not\RB eat\VB stale\JJ food\NN and\CC the\DT things\NNS being\VBG sold\VBN at\IN road\NN sides\NNS .\.

99. कालाजार\N_NNP ,\RD_PUNC मलेरिया\N_NNP या\CC_CCD फिर\RB हाई\JJ फीवर\N_NN हो\N_NN सकता\N_NN है\N_NN ।\RD_PUNC

ENG- Kalajar\NNP ,\ , Malaria\NNP or\CC high\JJ fever\NN may\MD have\VB happened\VBN .\.

100. एक\QT_QTC गिलास\N_NN पानी\N_NN में\N_NN दो\QT_QTC चम्मच\N_NN चीनी\N_NN एवं\CC_CCD एक\QT_QTC चुटकी\N_NN नमक\N_NN डालकर\N_NN उबालकर\N_NN दें\N_NN ।\RD_PUNC

ENG- Put\VB two\CD spoons\NNS of\IN sugar\NN and\CC a\DT pinch\NN of\IN salt\NN in\IN a\DT glass\NN of\IN water\NN and\CC boil\NN it\PRP .\.

101. चश्मे\N_NN या\CC_CCD कांटेक्ट\N_NN लेंस\N_NN का\N_NN रखरखाव\N_NN करें\N_NN ।\RD_PUNC

ENG- Maintain\VB the\DT glass\NN or\CC contact\NN lens\NN .\.

102. ये\DM_DMD शरीर\N_NN और\CC_CCD दिमाग\N_NN में\N_NN आक्सीजन\N_NN का\N_NN प्रवाह\N_NN बढ़ाते\N_NN हैं\N_NN ।\RD_PUNC

ENG- They\PRP increase\VB the\DT flow\NN of\IN oxygen\NN in\IN body\NN and\CC brain\NN

103. मुँह\N_NN में\N_NN लगातार\N_NN घाव\N_NN व\CC_CCD छाले\N_NN होना\N_NN ।\RD_PUNC

ENG- Wounds\NNS or\CC blisters\NNS in\IN mouth\NN .\.

104. एड्स\N_NNP से\N_NN बचाव\N_NN एवं\CC_CCD प्रचार\N_NN प्रसार\N_NN ।\RD_PUNC

ENG- Prevention\NNP of\IN AIDS\NNP and\CC propagation\NN .\.

105. हालांकि\CC_CCS योग\N_NN पक्षाघात\N_NN की\N_NN समस्या\N_NN का\N_NN स्थायी\JJ समाधान\N_NN करता\N_NN है\N_NN ,\RD_PUNC लेकिन\CC_CCS यह\DM_DMD समय\N_NN लेता\N_NN है\N_NN एवं\CC_CCD श्रमसाध्य\N_NN है\N_NN ।\RD_PUNC

ENG- Although\IN yoga\NN solves\VBZ the\DT problem\NN of\IN paralysis\NN permanently\RB but\CC it\PRP takes\VBZ time\NN and\CC is\VBZ tough\JJ .\.

