ISSN 0970-0277

OSMANIA PAPERS IN LINGUISTICS

Volume 44 & 45 2018 - 2019

Editor

Dr. Kalloji Susheel Kumar

Department of Linguistics, (CAS - IV Phase)
University College of Arts and Social Sciences
Osmania University
Hyderabad, Telangana, India – 500007

Osmania Papers in Linguistics

Volume – 44& 45 (2018-2019)

Editor:

Dr. Kalloji Susheel Kumar

Osmania Papers in Linguistics (OPiL) is an annual refereed research journal devoted to the publication of articles in the field of general and applied linguistics. It also includes book reviews and review articles. The journal publishes work primarily by the staff, students, and visiting faculty of the Department of Linguistics, Osmania University. However, articles may occasionally be invited from scholars outside Osmania on special topics. All contributions must be in English. Views expressed in OPiL are only those of the authors. Articles for publication, review copies, and communications relating to editorial matters should be sent to The Editor, OPiL, Department of Linguistics, Osmania University, Hyderabad 500007, India. Matters related to orders, payment, change in address or exchange may be sent to The Head, Department of Linguistics, Osmania University, and Hyderabad 500007. The price for OPiL (one Volume) is Rs. 500/- or US$ 50 (excluding postage). This publication is also supplied to institutions, societies and Departments of Linguistics in exchange for similar publications.

Contents **Pg.No.**

Mohd Hamid Raza1 and Dedy Suhery

A Phonological Study of Syllable Structure and Economy in Bahasa Indonesia: An OT Account 1

Mohd. H Masoud, Bedoor Sharaf Al- Deen, Mohd Hajeb and Nabil Asemi

Interaction Of Arabic Diacritics And Theta Roles 16

Jaya Verma

An Investigation on Syntactic Complexity in Indian Acts 29

N. Nirmeen and Dr. N. Vijayan

A Genderlect Study on Function Words Used in Suicide Notes 59

Srijani Mondal

The Language, Gender and Power in The Print Media and The Electronic Media in India– The Brief Case Study 69

Ms. Sanaz Montazer and Dr. Srikanth Muddam

Sentiment Analysis of Persian children's bedtime stories 79

Dr. Kalloji Susheel Kumar

Rhetorical Devices in 'Antarani Vasantam' 92

Vidyamrutha Imandi and C.S. Swathi

Sonority Profiles of Initial Consonant Clusters of Telugu Words 97

Dr.S.Santhi Prakash and Mrs.V.J.Vinila

Syntactic performance of Telugu speaking school going children with Hearing Impairment 112

Dr. Kalloji Susheel Kumar

Aesthetic Importance of Onomatopoeia in 'Antarani Vasantam' 128

K. Balu Naik

Morphophonemic processes of Banjara Language 132

Karishma Khatoon

Urdu Verb Agreements with Personal Pronouns Within Distinct Cases 140

Krishna Banothu and C.S. Swathi
Plural Formation in Lambadi Language 150
Gabriel Vinod Kumar Dolla
Language and Power: The Role of Language in Creating Political Demagogues 158
Khalid Mulla
Phenomenon of Code-Mixing in News18 Urdu TV Channel 168
Bipasha Patgiri
Book review of Aspects of Modern Assamese 177
Renuka Devi G
Implementation of UNL Based Machine Translation system for Telugu 182

A Phonological Study of Syllable Structure and Economy in Bahasa Indonesia: An OT Account

Mohd Hamid Raza[1] and Dedy Suhery[2]

(1) Department of Linguistics, Aligarh Muslim University, Aligarh (U.P.), India.
(2), Department of English Education, IAIN Langsa Aceh Indonesia

Abstract

This paper contains the phonological and economic properties of the syllable structures of the words in Bahasa Indonesia. It determines the behaviour of certain speech segments attached to the root words and elaborates the economy of the syllable structure in tokens. In Bahasa Indonesia, there are various types of segmental processes related to the addition or deletion of the phonemes that affected the root and altered the entire physical mechanism of the words. This paper aims to know the exact economic conditions of syllable structures in words after adding or deleting segments in Bahasa Indonesia. All the types of conflicts between the candidates will be manipulated with the help of constraint rankings in Optimality Theory (Prince and Smolensky, 1993). The general purpose of this paper is to reveal the whole criteria of OT principles and explored the actual framework of syllables within their marginal and obligatory components. The researcher governed the phonological property of consonant clusters with the help of faithfulness constraints and markedness constraints. The architecture of root words wholly varied from the artificial formulation of other words, but after the imposition of constraints, we revealed the concrete fact of linguistic items in Bahasa Indonesia. The groundwork of this paper led to the systematic phenomena of epenthesis and syllable structures of vowels or consonants within OT tenets. The researcher considered the typology of the syllable structure of words and the phonological observations of linguistic features. The generalization of each step of the syllable structure and economy of the words should be related to the positional variation of input and output candidates. The conflict between output candidates is solved base on the ranking features of an input candidate to find out an optimal form in Optimality Theory.

Keywords: Optimality Theory; Syllable Structure; Economy; Syllable Typology

1. Introduction

Bahasa Indonesia is an official language and language of unity in the Indonesian nation. It was inaugurated as the national

language after the Proclamation of Indonesian Independence, August 18, 1945. Its stipulation is contained in Article 36 of the 1945 Constitution, which states that "The State Language is Indonesian."Bahasa Indonesia has played an essential role in the country because it is spoken widely among Indonesian people and widespread throughout the nation (Alwi et al. 1993, pp. 1-2). It is used as the medium for science, literature and the arts, and cultural expression. Historically, the initial declaration of the use of Bahasa Indonesia as a national language was made in 1928. This declaration, commonly known as Sumpah Pemuda 'Youth Pledge,' says: "Kami poetera dan poeteri Indonesia mendjoenjoeng bahasa persatoean, bahasa Indonesia" (Alwi et al., 1993, p. 1). The oath means 'We, the young generation of Indonesia, respect the language of unity, Bahasa Indonesia.' According to Yusuf (2013, p. 3), Bahasa Indonesia has become the official language in Aceh since the province became a part of Indonesia in 1950 (Reid, 2005).

The Syllable Structure and Economy are the two critical factors of this paper regarding Bahasa Indonesia. It is notified that every language has its pattern of segment sequencing and formed a phonological building block of words. The syllable structure is a type of phonological process in which the different segments with different shapes and sizes come together and are fused to the formation of a syllable. The syllable structure may be monosyllabic (V, CV, VC, CVC, VCC, CVCC, etc.), disyllabic (CVCV, CVCCV, VCCV, etc.), trisyllabic (CVCCVC, VCCVCVC, CVCVCVC, etc.), and so on. In the term of Economy, it is generalized that Bahasa Indonesia is involved in adding certain types of elements and generating a new physical mechanism of the syllable structure of words. The word economy determined the addition or deletion of a single segment to the root structure of the words that affected the whole framework of the syllable structures. It examined that how many extraneous elements are possible to add or delete in the initial and non-initial position of words and how many syllables are likely to form in a specific environment of Bahasa Indonesia.

OT was first introduced by Alan Prince and Paul Smolensky (1993) and further extended by John McCarthy (1994) to organize the well-formed syllable structure of the words, but soon spread in other areas of Linguistics. According to Gussenhoven and Haikes (1998), Optimality Theory is a pertinent portion of the Phonology related to the thought of a universal set of constraints represented in a hierarchically ranked of language-specific facts. According to McCarthy (2002), "Gen is universal," which means that all produced candidates by Gen for a given input are the same in all languages? These candidates vary from language to language, and Gen's property is called "inclusive or freedom of analysis."Alan Prince and Paul Smolensky (1993) introduced that CON tells us what the substantive constraints are, from which grammars are built. The third significant key component of Universal Grammar is a precise definition of constraints, which is also referred to as EVAL that, spells out what it means to be optimal with respect to a ranking of CON.

2. Research Highlights

This study reveals the phonological properties of syllable structures and their economic observation regarding the addition and deletion of segments from the root words in Bahasa Indonesia. There is a straightforward way to formalize the various types of syllable structures and find out the numeral categories that how much a root word can bear the load of external segments. The groundwork of this study is that after the addition or deletion of segments from the root words, we will apply the principles of Optimality Theory to find out the exact dominating form of the candidate. This study will examine the conflicts between input and output candidates and will be solved based on constraints of OT. After the implications of constraint rankings will determine one of the best candidates for all output candidates as an optimal candidate. In the account of the economy, how many segments are possible to the addition or

deletion, and what will be happening with the physical mechanism of the syllable structure of words.

3. Graphical Abstracts

Correspondence diagram for the interaction between input and output candidate.

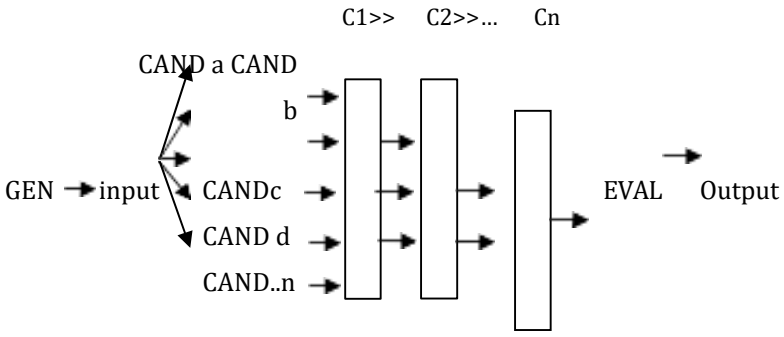


Figure 1: the interaction between constraints and their candidates

OT analysis of input candidate with the help of constraint rankings in the tableau

/input/	Constraint 1	Constraint 2	Constraint 3
a. Candidate 'a'	*!	*	*
b.Candidate 'b'		*!	
c.☞Candidate 'c'			*

Table: 3.1 The Consideration of constraints and candidates with the reference of input

4. Research Objectives

The research objectives of this study are relevant to the whole phenomena of epenthesis and elision at the level of syllable structures and their economic behavior of segments in words. There are various types of objectives, but this paper reveals some specific procedures of elements with the utilization of constraint rankings such as:

To know the typological structure and economy of the syllables based on the aspects of the addition or insertion of the speech segmentsin Bahasa Indonesia.

To explore the status of the output candidates based on the ranking features of the input to find out the best candidate with the help of the hierarchy of constraints.

5. Research Methods

In conducting research, we need research design. The research design refers to the strategy to integrate the different components of research projects cohesively and coherently. Some experts have different opinions about what is mean by research design. According to Creswell (2009:3), the research design is plans and procedures for research to detailed data collection and analysis methods.

Meanwhile, Bogdan and Biklen (1992) state that qualitative research is descriptive in which the data is collected in the form of words or pictures rather than numbers. Data in the form of quotes from documents, field notes, and interviews or excerpts from videotapes, audiotapes, or electronic communications are used to present the study's findings.

The researcher used data from the Big Indonesian Dictionary (KBBI). The author use documentation as a means of data collection (Arikunto, 1993:202). We select numerous samples from the lexicon that illustrate the problem's phenomena and then use Optimality Theory to evaluate the samples further. The analysis' findings should be able to explain the wonders of the situation in this study.

6. Data Analysis

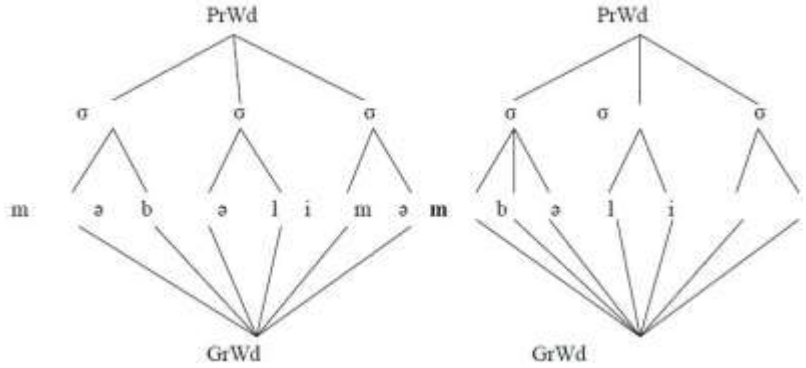
6.1 Phonological conditioning of syllable structure and economy

The syllable structure is a property of phonological process in which the speech segments are organized in a particular way to construct a building block of words in a language or across

the languages. In terms of the phonological conditioning, each segment is congregated with different manner to show the specific grammatical category of words. In the account of phonological conditioning, the syllable structure and economy revealed the factors of marginal and obligatory positions of the speech segments. The term economy stands for the determination of the frequency of syllable structures after the addition or deletion of segments in the words. For example,Bahasa Indonesia speakers have the process of addition or deletion of certain types of segments in initial, medial or final position of words:

/mə- + bəli/	⟶	[məmbəli]	'to buy'
/mə- + fitnah/	⟶	[məmfitnah]	'to slander'
/mə- + pəŋaruhi/	⟶	[məmpəŋaruhi]	'to influence'
/mə- + karaŋ/	⟶	[məŋgaraŋ]	'to compose'
/mə- + hilaŋ/	⟶	[məŋhilaŋ]	'to lose'
/mə- + goreŋ/	⟶	[məŋgoreŋ]	'to fry'
/mə- + dukuŋ/	⟶	[məndukuŋ]	'to support'
/mə- + cuci/	⟶	[məntʃutʃi]	'to wash'
/mə- + jəmur/	⟶	[mənjəmur]	'to dry'

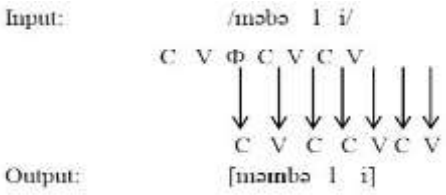
The essential representation of the prosodic and morphological structure of the syllables is concerned with the generalization of the addition or insertion of an external speech segment that broke the framework of the consonant clusters in Bahasa Indonesia:



In the above data, it is represented that there is the insertion of the nasal consonants after the prefix mə- to cover the features of the place and voicing assimilation. It is notified that the inserted segments break the process of the consonant clusters in the syllable structure of the words. To explain the phonological properties of the inserted segments, we have to use OT principles for the determination of an optimal candidate:

i. [mə.bə.li] satisfied with DEP-IO and *CODA, but violated to AGREE and AGREE (VOI)
ii. [məm.bə.li] satisfied with AGREE (VOI), but violated to DEP-IO, AGREE and *CODA
iii. [məb.bə.li] satisfied with AGREE and AGREE (VOI), but violated to DEP-IO and *CODA

In the account of OT principles, the correspondence diagram represented the familiar framework of the phonological process to generalize the ranking features of the input and output candidates:



In the correspondence diagram, the unbroken lines remained the position of the speech segments,while the broken line denoted the process of insertion of a voiced nasal consonant. Now, we have to focus on the level of the constraint rankings for the classification of the best candidate in Bahasa Indonesia:

Tableau for the analysis of an input candidate /mə.bə.li/ in Bahasa Indonesia

/mə.bə.li/	*CODA	DEP-IO	AGREE	AGREE (VOI)
a. [məm.bə.li]	*!	*	*	
b. ☞ [mə.bə.li]			*!	*
c. [məb.bə.li]	*!	*		

Table 5.1 a: phonological conditioning of the syllable structure in the words.

The strengthening of the constraints is based on the presence and absence factors of the dominance and precedence in a particular language or across the languages. In this framework, the power and potential of the output candidates are generalized through the significant aspects of the input candidate:

*CODA>> DEP-IO >> AGREE>> AGREE (VOI)

i. [məŋ.ka.raŋ] satisfied with AGREE (VOI), but violated to DEP-IO, AGREE, and *CODA
ii. [mə.ka.raŋ] satisfied with DEP-IO and *CODA, but violated to AGREE and AGREE (VOI)
iii. [məŋ.ga.raŋ] satisfied with AGREE and AGREE (VOI), but violated to DEP-IO and *CODA

Tableau for the analysis of an input candidate /mə.ka.raŋ/ in Bahasa Indonesia

/mə.ka.raŋ/	*CODA	DEP-IO	AGREE	AGREE (VOI)
a. [məŋ.ka.raŋ]	*!	*	*	
b. ☞ [mə.ka.raŋ]			*!	*
c. [məŋ.ga.raŋ]	*!	*		

Table 5.1 b: the process of anaptyxis in the syllable structure of the words.

In the account of the phonological conditioning of the words, it is determined that the anaptyxis is a prominent aspect of the phonological process in which an outer segment is inserted between a prefix and lexical item that did not maintain the property of the syllable structures.

i. [mə.du.kuŋ] satisfied with *COM-CODA and DEP-IO, but violated to AGREEand *CODA
ii. [mənd.uk.uŋ] satisfied withAGREE, but violated to *COM-CODA, *CODA,and DEP-IO
iii. [mən.duk.uŋ] satisfied withAGREE and *COM-CODA,but violated to DEP-IO and *CODA
iv. [mənd.u.kuŋ] satisfied withAGREE, but violated to DEP- IO, *CODA,and *COM-CODA

Tableau for the analysis of an input candidate /mə.du.kuŋ/ in Bahasa Indonesia

/mə.du.kuŋ/	*COM-CODA	DEP-IO	AGREE	*CODA
a.[mənd.uk.uŋ]	*!	*		**
b.[mə.du.kuŋ]			*!	*
c. ☞ [mən.duk.uŋ]		*!		***

d.[mənd.u.kuŋ]	*!	*		*

Table 5.1 c: economy in the syllable structure of the words

In the account of the economy, the frequency of the syllable structures is determined based on the factors of the constraint rankings. The constraint rankings revealed the phonological properties of the output and input candidates to find out the best candidate in a particular language or across the languages. There is a hierarchy of the constraints to denote the proper solution of the conflicts that happened among the output candidates:

COM-CODA>> DEP-IO >> AGREE>> *CODA

7. Results and Discussions

In Bahasa Indonesia, the structure of the syllables is not maintained due to the addition or insertion of a particular speech segment in words. It is determined that there is a common prefix mə- (to become active) added to the beginning of the words to know the vision of the inserted segment in words. It is revealed that the voiced bilabial nasal /m/ is inserted between prefix mə- and the root words beginning from voiceless and voiced bilabial stops /p/ and /b/ and voiceless labio-dental fricative /f/. The voiced alveolar nasal /n/ is inserted between prefix mə-and the root words beginning from voiced alveolar stop /d/, voiceless postalveolar affricate /tʃ/, and voiced palatal semivowel /j/. The voiced velar nasal /ŋ/ is inserted between prefix mə-and the root words beginning from voiceless and voiced velar stops /k and /g/ and voiceless glottal fricative /h/. The inserted segments agreed with the voicing and pattern (place and manner of the articulation) of the following speech segments of the root words to maintain the properties of the assimilations. The inserted segments increased the frequency of the syllables but not maintained the process of the consonant clusters in words. It is notified that these segments are inserted only in the condition of adding a common prefix mə-at the beginning of the words. It is

evaluated that Bahasa Indonesia native speakers need a prefix to insert a speech segment in words to maintain the phonological properties of the whole structure. They did not preserve the account of the consonant clusters due to the insertion of the speech segments in the syllable structure of the words.

In the OT account, it is discussed that there are different types of constraints applied on the structure of the output candidates that matched the ranking features to the input for the description of an optimal form in Bahasa Indonesia. The constraints are universally used to solve the conflicts between the output candidates to become optimal candidates in a particular language or across the languages. Each constraint has its unique level of features and powers that appeared in a hierarchy to know the factors of dominance and precedence in a specific environment of the languages. In tables 5.1a and b, there are four kinds of constraints applied on the structures of the output candidates similar to the ranking features of the input to mark the best candidate.

The candidate 'a' is satisfied with the lowest rank of the constraint AGREE (VOI) (voicing of the output segments must be correspondence to the input), while violated to the highest, a higher, and lower rank of the constraints *CODA(final marginal consonant must be in the syllables), DEP-IO(output segments must be input correspondence) and AGREE(similarity between preceding and following speech segments must be identical). It has the highest rank of the constraints, and the least number of the ranking features similar to the input candidate to become an optimal form. The candidate 'b' is satisfied with the highest and higher rank of the constraints *CODAand DEP-IO, while violated to thelower and lowest AGREEand AGREE (VOI). It has the minor rank of the constraints,and a maximum number of the ranking features similar to the framework of the input candidate. The candidate 'c' is satisfied with the lower and lowest rank of the constraints AGREEand AGREE (VOI), while violated to the highest and higher *CODAand DEP-IO. It has the

highest and higher rank of the constraints violated to the ranking features of the input candidate. It is finalized that the candidate 'b' has the minimum account of the violations (*) and fatal violations (*!) and maximum rank of the features similar to the structure of the input candidate compared to the candidate 'a' and 'c' in Bahasa Indonesia. Based on the factors of OT principles, it is evaluated that the candidate 'b' is declared the best candidate marked as an optimal candidate indicated by
☞ in Bahasa Indonesia.

In table 5.1c, four types of constraints are applied on the different kinds of output candidates matched to the ranking features of the input to find out an optimal candidate in Bahasa Indonesia. The candidate 'a' is satisfied with the lower rank of the constraint AGREE, while violated to the highest, higher, and lowest rank of the constraints *COM-CODA(coda must be simple in the syllables), DEP-IO, and *CODA. It has the procedure of the violations and fatal violations affected to the ranking features of the input candidate to become an optimal form. The candidate 'b' is satisfied with the highest and higher rank of the constraints *COM-CODAand DEP-IO while violating the lower and lowestAGREEand *CODA. It has the minor rank of the constraints within the maximum values of the ranking features similar to the input candidate to become an optimal candidate in Bahasa Indonesia.

The candidate 'c' is satisfied with the highest and higher rank of the constraints *COM-CODAand AGREE, while violated to the higher and lowest DEP-IO, and *CODA. The candidate 'd' is satisfied with the lower rank of the constraintAGREE, while violated to the highest, higher, and lowest rank of the constraints *COM-CODA,DEP-IO, and
*CODA. It is evaluated that the candidate 'b' has more identical ranking features similar to the input candidate compared to others in Bahasa Indonesia. So, the candidate 'b' became the best candidate marked an optimal form indicated by . In the hierarchy of the constraints, the dominant form is represented by

>> to know the power implemented on the output candidates. The left-most constraint is more powerful, while the right-most is least in the hierarchy used to solve the conflicts between the output candidates.

8. Conclusion

Phonological conditioning of the syllable structure and economy affected the sound systems of a particular word to cover assimilation and anaptyxis in Bahasa Indonesia. It is determined that Bahasa Indonesia native speakers account for the assimilation and anaptyxis within the conventional properties of the prefix and root word. They did not maintain the phonological properties of the consonant clusters due to the insertion of the speech segments in the syllable structure of the words. It is considered that the nasal sounds got the position between a common prefix and the roots. These nasal sounds are agreed with the voicing features of the beginning elements of the root words. It is notified that the economy of the syllable structures is generalized and evaluated base on OT principles to find out the best candidate. It is concluded that Bahasa Indonesia native speakers increased the economy of the syllable structures based on factors of the constraint rankings in a particular hierarchy.

Acknowledgments

We would like to express my special thanks to Professor Syed Imtiaz Hasnain, Dr. Hemanga Dutta (posthumously), and Professor Mohammed Jahangir Warsi for their patient guidance and enthusiastic encouragement in writing this research paper.

References

Al-Harbi, A. A. A. (2003). Acehnese Coda Condition: An Optimality-Theoritic Account. *Umm Al-Qura University Journal of Educational and Sciences and Humanities*,Vol. 15-No. 1 Thu-1-Qida 1423H, January 2003.

Aljumah, A. H., (2008). The Syllable Shape of Al-Ahsa Dialect: An OT Perspective. *Poznan Studies in Contemporary Linguistics* 44(2), pp. 155-177.

Alwi, H. (2007). *KamusBesar Bahasa Indonesia*. Jakarta: BalaiPustaka.

Alwi, H., Dardjowidjojo, S., Lapoliwa, H., &Moeliono, A. M. (1993). *Tata Bahasa Baku Bahasa Indonesia [Standard Grammar of Indonesian] (Second ed.)*. Jakarta: Department of Education and Culture of the Republic of Indonesia.

Akbar, O. M., Abdullah, W., Latif, S. N., &Ahmaddin, S. (1985). *Pemetaan Bahasa Aceh*, *Gayo*, *dan Alas [The Mapping of Acehnese*, *Gayo Language*, *and Alas Language]*. Jakarta: PusatPembinaan dan Pengembangan Bahasa

Arikunto, S. (2010). *ProsedurPenelitian: SuatuPendekatanPraktik*. Jakarta: PT. RinekaCipta.

Arka, I. W. (2008). Local Autonomy, Local Capacity Building, and Support for Minority Languages: Field Experiences from Indonesia. *Language Documentation & Conservation Special Publication* 1, 66-92.

Ayatrohaedi. (1999). KebijaksanaanPemerintahMengenai Bahasa Daerah [The Policy of Indonesian Government on Vernacular Languages].In A. Rosidi (Ed.), Bahasa Nusantara SuatuPemetaanAwal [The Initial Mapping of Languages in Indonesia]. Jakarta: DuniaPustaka Jaya.

Bogdan, R. C., & Biklen, S. K. (1992). Qualitative Research for Education: An Introduction to Theory and Methods. Boston: Allyn and Bacon.

Creswell, J. (2009). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. USA: Sage Publication Inc.

Clark, J. &Yallop, C. (1995). *An Introduction to Phonetics and Phonology*.2. ed. Oxford: Blackwell.

Gussenhoven, C. & Jacobs, H. (2017). *Understanding Phonology* (7 ed).Oxon: Routledge.

Gouskova, M. (2002). Falling Sonority Onsets, Loanwords, and Syllabic Contact. In M. Andronis, C. Ball, H. Elston and S. Neuval (eds.) CLS 37: the main session. Papersfrom the 37th meeting of the Chicago Linguistic Society (pp. 175-186) Chicago: The Chicago Linguistic Society.

Jarrah, A. S. I. (2013). Syllables and Syllable Structure in Arabic in the Light of the Optimality Theory.*Open Science Repository Language and Linguistics*, Online (open access), e70081958.

Kager, R. (1999). *Optimality Theory*. United Kingdom: Cambridge University Press.

Keating, P., Linker, W. & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops.*Journal of Phonetics*, 11,277-290.

Yusuf, Y. Q. (2013). A comparative study of vowels in the Acehnese language spoken in Kedah, Malaysia, and Aceh, Indonesia. (Doctoral), University of Malaya: Kuala Lumpur.

Prince, A. and P. Smolensky. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*.Ms. Rutgers University, New Brunswick and University of Colorado, Boulder. Later published in 2004 from Blackwell Publishing, Malden, MA.

Widodo, H. P., &Fardhani, A. E. (2011). The Language Rights of Indigenous Languages: An Approach to Maintaining Indonesia's Linguistic and Cultural Diversity. *In Q. Lê& T. Lê (Eds.)*, *Linguistic Diversity and Cultural Identity: A Global Perspective* (pp. 129-140). New York: Nova Science Publishers, Inc.

Interaction Of Arabic Diacritics And Theta Roles

Mohd. H Masoud[1], Bedoor Sharaf Al- Deen[2] Mohd Hajeb[3] Nabil Asemi[4]

1- Ph. D. Research Scholar, Department of English, Aligarh Muslim University, Aligarh, UP, India
2 & 4 Ph. D. Research Scholars, Department of Linguistics, Aligarh Muslim University, Aligarh, UP, India
3 Ph.D. student, English Department, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India

ABSTRACT

This paper aims to analyze diacritics in Arabic verbal sentence structures. Specifically, it aims to show to what extent Arabic verbal sentence structures having Diacritics (Altashkiil and Tanwīn) interact with Theta Roles. Diacritics play a significant role in exposing the semantic relations of words as in Arabic verbal syntactic structure using Altashkiil three letters, ambiguity may occur, and to some extent, it is unavoidable. Therefore, this paper also aims to show how ambiguity occurs in Arabic due to Diacritics and highlights the causes of those diacritics in Arabic verbal sentence structures through some examples taken as data analysis.

This study revealed that diacritics (Altashkiil and Tanwīn) interacted with Theta Roles differently. The absence of Altashkiil, though it causes an ambiguity for native speakers, the theory was appropriately applied. It also was found that though there is a variation in verbal sentence structures having Diacritic (Tanwīn), the presence of a case maker helped in identifying theta roles of arguments regardless of their positions.

Keywords: Theta theory, Diacritics, Arguments, Verbal Structures

1. Background of the Study

Arabic is a Semitic language that uses 28 basic letters. Besides, it uses different signs or symbols known as diacritics, diacritical signs, or marks known as characters. Aabed et al. (2007) explained that Arabic uses different signs or symbols as diacritics known as Harakāt. There are eight Arabic diacritics as “Fatha, Kasrah, Dhammah, Sukkon, Shaddah, TanweenFath, TanweenKasr, TanweenDham”. Arabic is written and read from right to left, and its letters mainly denote consonant sounds. Meanwhile, its short vowels are presented by diacritical marks.

These diacritics are the hallmarks of the Arabic language. The words in Arabic can be of both letters and diacritics, Jarrar et al., (2018). In this regard, Diab et al. (2007) stated that diacritization is at the last letter of the stem used to identify the syntactic role of words in a sentence. In Arabic writing, there are several types of diacritics, diacritics, short vowels, verbalization of diacritics, and various decorative signs Hssini and Lazrek, (2012).

Diacritics clarify linguistic features, morphologically about the structure of words, grammatically about the functions, semantically about the meaning of words, and other linguistic and voice features, Debili and Achour, (1998). Arabs tend to neglect the use of diacritics in writing. Attia (2008) states that when Arabs write, they commonly write without diacritics, and this makes an ambiguity in Arabic text, Tarrar et al., (2018). Attia (2008) stated that what usually differentiates between the subject (the nominative case) and the object (accusative case) are the diacritics used. They do not appear on the surface because they are usually omitted in modern writing. According to Zitouni, Sorensen, Sarikaya (2006), as cited in Chennoufi and Mazroui(2017), there are three groups in which Arabic diacritical marks are categorized. In representing diacritics, this paper is limited only to the two following diacritic types: (Altashkiil and Tanwīn)

1. Altashkiil, in this group, there are three diacritics, which are known as single short diacritics fatha, damma, and kasra.
   a. [fthh], it is a small line that represents a short /a/. Such a diagonal line is placed above a letter.
   b. [dmmh], it is a small cur-like diacritic which represents a short /u/ or /o/. Such a cur-like diacritic is placed above a letter.
   c. [ksrh]. The word kasrah means literally breaking. It represents a short /i/. Such a diagonal line is placed below a letter.

2. Tanwīn, this is the second group. It is a duplicate case ending diacritic could be Fatah, Damma, and Kashra, Zitouni et al. (2006). "Nunation is the suffixation of the so-called nunation suffix -n to nouns, nominal modifiers and adverbs" Jarrah (2016).

The paper's significance lies in showing how Arabic diacritics in Arabic verbal sentence structures interact with theta roles. Therefore, this paper aims to show how Arabic verbal sentence structures having Diacritics (Altashkiil and Tanwīn) interact with Theta roles. This paper also aims to show how am ambiguity occurs in Arabic due to Diacritics. It has been recognized that diacritics is a problematic issue in Arabic and may cause a misunderstanding and a breakdown of the relationship among communicators either for Arabic native language speakers or others. Therefore, these diacritics in Arabic verbal sentence structures are intended to investigate the interaction of such diacritics with Theta roles and discuss the causes of diacritics which is another significant point of this study.

Arabic is structured under the combinations of Subject, Verb, and Object (SVO), and Verb, Object and Subject, Verb, Subject, and Object (VSO), and Object, Verb and Subject (O.V.S). (SVO) is known as nominal sentence structure and (V.O.S.) known as verbal sentence structure. The scope of this paper is limited to only Arabic verbal sentence structures that have diacritics. Theta theory is applied to Arabic verbal structure. Therefore, this research paper is intended to show the interaction of two types of diacritics (Altashkiil and Tanwīn) with two thematic roles of Agent and Patient through examples that have been taken from Arabic as data analysis.

This study is divided into four sections. The first section is a brief background of the study. It includes the rationale of the study, this paper's objectives, and its limitation. The second section deals with the notion of theta roles. The third section

applies verbal sentence structures of Arabic. The last section is the conclusion.

2. The Notion of Theta Theory

Theta-Theory is the module used for determining an argument structure and theta or thematic role assignments of the sentence, for example, Agent, Patient, etc., Kuhns and Little (1981). They added that through these theta relations and general principles, arguments and their possible positions could be predicted and explained. Marrone (2013) traditionally labeled thematic relations into nine types: the Agent, Instrument, Experiencer, Theme (Patient), Cause, Sentient, Subject Matter/Target of Emotion, and Goal/Benefactor. The fundamental principle of Theta Theory is the Theta-Criterion Chomsky (1981). In Theta Theory, there are many concepts such as the Projection Principle and Theta Grids, which mean that theta roles are placed in a verb's theta grid in Government and Binding Theory (G.B.). GB. is an approach to Universal Grammar, including rules and principles which apply to all languages Moubaiddin, (2013). Though there are several linguistic principles and universal rules, there are many variations among languages, such as different sentence structures, nominal and verbal sentence structures in English and Arabic. It is agreed that every language has a basic word order, and all other word orders result from the movement of sentence constituents, and some rules and principles restrict this movement. In the typology of theta roles, there are many types of Theta roles. However, in this study, only two types are considered the Agent and the Patient, through which the interaction of Arabic diacritics is analyzed.

Type	Definition	Examples
Agent	Agent refers to the person, thing, the doer, or the initiator that carries out the action, and it is mostly the subject of a clause.	1- Ahmed studied English

Patient	Patient refers to the person/thing that undergoes the action expressed by the predicate.	2- Yousif hit his brother

In G.B., a set of theta roles (semantic roles) are assigned by a verb. In (1), the verb '*study'* is a two-place predicate and requires two arguments to which these roles are assigned. The thematic structure of the verb '*study'* assigns two thematic roles (agent and theme): Amhed [1(external): Agent; English 2 (internal): theme]. Similarly, the verb '*hit'* in (2) is a two-place predicate and requires two arguments. The thematic structure of the verb '*hit'* assigns two thematic roles (agent and patient). Yousif [1(external):Agent; his bother2(internal): Patient].

The theta-criterion (θ-criterion) states that each argument is assigned one and only one θ-role, and each θ-role is assigned to one and only one Argument Chomsky,(1981).

To make it clearer, every argument, e.g., N.P., in a sentence must receive one and only one thematic role (e.g., agent, patient, etc.); each thematic role that is associated with a theta-assigner (typically a verb) must be assigned to one and only one Argument. Each Argument is assigned one, and only one theta role, or each theta role is assigned to one and only one Argument (Chomsky, 1981). To illustrate, when applying theta criterion to the sentence [s [$_{NP}$ Yousif]$_{I}$ *hit* [$_{N.P.}$his bother], *Yousif* is assigned only the external theta role of Agent, and *his bother* is assigned only the internal theta role of the Patient. Since the scope of this research paper is only about two types of Arabic Diacritics: the short vowel and nunation in Arabic verbal sentence structure. The following section points out the connection between diacritics in verbal sentence structures in Arabic and Theta roles, particularly the agent and the patient.

3. Diacritics in Verbal Sentence Structure in Arabic

In this section, we discuss diacritics in Arabic verbal sentence structure. We will start with the short vowel and move

to nunation to point out the interaction between diacritics and theta theory.

3.1. Verbal Sentence Structure and *Altashkiil* (short vowel)

In Arabic, if the arguments in VSO word order end with one of the letters -a,-w, or –y (ا و ى), we face a challenge in distinguishing the subject from the object. Consider the example in (3).

(3) Dharaba Eysa musa
hit-past Eissa Mussa
(4) Qatala Akhi Sadiqy
kill-past my brother my friend

The example in (3) consists of a predicate 'Dharaba, with two arguments: Eissa and Mussa. However, we cannot tell which one is the subject and which one is the object. In Arabic, for the nouns or words to end with one of these vowel marks (-a -w -y), it is ambiguous to determine the subject from the object. Eissa or Mussa is either considered as a subject or as an object. Similarly, in (4), Akhi 'my bother' or Sadiqy 'my friend' is either considered a subject or an object in this sentence. This happens because the word ends with the short vowel mark –y. Moreover, it is also challenging to determine the subject from the object. That is, it is challenging to decide who beat whom? And who killed whom? Did Eissa beat Mussa?

Similarly, Did Akhi kill Sadiqi? Or Akhi was killed by Sadiqy? This ambiguity arises due to the lack of the respective nominative and accusative case markers. In Arabic, the subject is nominative and the basic nominative marker is-u. However, the object is accusative, and the accusative marker is –a. The examples in (3-4) lack such markers, which triggers ambiguity in the sentence structure. Had they been marked by–a or –u, no ambiguity would have been there.

Now we will consider applying theta theory to the data in (3,4), the verb Dharaba 'beat' is a two-place predicate, and the

thematic grid or theta grid of the verb Dharaba 'beat' is given the representation in (5).

(5) Dharaba 'beat.'

The representation in (5) indicates that the verb Dharaba 'beat' assigns two thematic roles (agent and patient). However, it is difficult to identify the above thematic roles with the corresponding argument. The predicate theta marks the N.P.s directly. Given this, the Theta criterion argues that i) each argument is assigned one and only one theta role and ii) each theta role is assigned to one and only one argument Haegeman, (1994). It is plausible to assume that Eissa is assigned only the theta role and Mussa also is assigned only theta role. Similarly, only one theta role is assigned to Akhi, and only theta role also is assigned to Sadiqy. Hence, the theta theory rightly predicts the following theta grid in (6) and (7).

(6) Dharaba 'beat.'

Agen t NP	Patien t N.P.
Eissa	Muss a

(7) Qatala (killed)

Agen t NP	Patien t N.P.
Akhi	Sadiqy

As shown in (6), the NP Eisaa bears one theta role (agent), and the NP Musaa receives another theta role (patient). This means that the two arguments [Eisaa] and [Mussa] stand in different semantic relations with the verb darab (hit-past). In the same manner, As shown in (7), the NP Akhi bears one theta role (agent), and the NP Sadiqy receives another theta role (patient). This means that the two arguments Akhi and Sadiqy, stand in different semantic relations with the verb qatala (killed- past).

3.2. Verbal Sentence Structure and Tanwīn (nunation)

In Arabic, there is another type of diacritics called nunation. "Nunation is the suffixation of the so-called nunation suffix -n to nouns, nominal modifiers and adverbs" Jarrah (2016). Al-Jahaawi (1982) has argued that the Arab phoneticians defined nunation as "a short vowel plus the nasal consonant *n,''* as cited in Jalabneh (2009). This type can be of three types TanweenDham, Tanweenfath, and TanweenKaser. Zitouni et al. (2006). let us take an example in (8).

(8) a. *Saad Mohammadun Salehan*

Help -past Mohammadun- 3sm-Nom Salehan- 3sm-Acc

'Mohammed helped Saleh.'

b. *Saad Salehan Mohammadun*

Help -past Salehan-3sm-Acc Mohammadun- 3sm-Nom

'Saleh helped Mohammed.'

c. *Saad Salehun Mohammadan*

Help -past Salehun-3sm- Nom Mohammadan- 3sm-Acc 'Saleh helped Mohammed.'

As mentioned earlier, the Theta criterion argues that i) each argument is assigned one and only one theta role and ii) each theta role is assigned to one and only one argument Haegeman, (2020). Therefore, it is clear that nouns having the vowel mark /un / as in the NP Mohammadun are assigned one and only theta role known as an agent, and the vowel mark /an/as in the NP Salehan is assigned one and only theta role known as a patient. Hence, the theta theory rightly predicts the following theta grid in (8a and b).

In (8a) (8b) and (8c), notice that the diacritic mark (un) is added to the nouns to indicate the subject, and the diacritic mark (an) is added to the nouns to indicate the object. Without these markers, it is too difficult to determine the subject or the object

of the sentence. In (8a) has the word order VSO while in (8b), has the word order V.O.S. In (8c), has the word order VSO where (un) is added to the noun, Saleh to indicate the subject, and the diacritic mark (an) is added to the noun, Mohammad to indicate the object.

In (8a), the NP Mohammadun is the subject because it has the nominative case/ un /, and the NP Salehan is the object because it has the accusative case/ an /. However, in (8c), the NP Salehun in the same position is the subject because it also has the nominative case/ un /, and the NP Mohammadan is the object because it has the accusative case/ an /.

In (8b), you can recognize that the NP Salehan in the subject position indicating an accusative case. The NP Salehan here is the object because it has the accusative case/ an /. And the NP Mohammadun in the object position, indicating a nominative case. The NP Mohammadun, here is the subject because it also has the nominative case/ un/.

Here, it would be too difficult to determine the object's subject. The misuse or absence of such marker [Tanwīn diacretics] would make it too difficult to determine the thematic roles, Agent and Patient appropriately.

Given this, we will apply theta theory to examples (8a) (8b) and (8c). The verb Saad 'help-past' is a two-place predicate, and it assigns two thematic roles (agent and patient). Consequently, the thematic grid of (8a) is given (9a), the thematic grid of (8b) is given (9b), and the thematic grid of (8c) is given (9c)

(9a) *Saad* 'Help –past.'

Agent NP	Patien t NP
Mohammad***un***	Saleh***an***

9b. *Saad* 'Help –past.'

Patien t N.P.	Agen t NP
Saleha***n***	Mohammad***un***

9c. *Saad* 'Help –past.'

Agen t N.P.	Patien t NP
Salehu***n***	Mohammad***an***

As shown in (9a), the NP Mohammad*un* bears one theta role (agent), and the NP Saleh*an* receives another theta role (patient). This means that the two arguments Mohammad*un* *and* Saleh*an*, stand in different semantic relations with the verb Saad (Help –past). Similarly, but not identically in (9b), the NP Saleh*an* bears one theta role (patient). And the NP Mohammad*un* receives another theta role (agent). This means that the two arguments Saleh*an* and Mohammad*un* stand in different semantic relations with the verb Saad (Help –past). In (9c), the NP Saleh*un* bears one theta role (agent), and the NP Mohammad*an* receives another theta role (patient). This means that the two arguments Saleh*un* and Mohammad*an* stand in different semantic relations with the verb Saad (Help –past).

4. Conclusions

The study utilized Theta Theory to analyze Arabic verbal sentence structures having Diacritics (Altashkiil and Tanwīn). In this study, it was found a) that the verbal sentence structures having diacritics (altashkiil/ causes an ambiguity due to the absence of case makers and the ambiguity is resolved by applying theta theory in that the nearest NP (Eissa) is the agent and the NP (Mussa) is the patient. And b) in nunation, theta theory is easily applied due to the presence of case makers regardless of the position of arguments.

It could be concluded that diacritics play a significant role in exposing the semantic relations of words as in Arabic verbal syntactic structure using Altashkiil three letters, ambiguity may occur, and to some extent, it is unavoidable. However, theta roles were appropriately applied regardless of the absence or presence of case makers in both Arabic Diacritics.

Abbreviations:

Subject, Verb, and Object (SVO),

Verb, Object, and Subject (V.O.S.),

Verb, Subject, and Object (V.S.O),

Object, Verb, and Subject (O.V.S).,

Government and Binding Theory (G.B.),

Noun Phrase (NP)

References

Aabed, M., Awaida, S., & Gutub, A. (2007). *Arabic Diacritics Based Steganography*. *February 2014*. https://doi.org/10.1109/ICSPC.2007.4728429

Attia.M. (2008). Handling Arabic Morphological and Syntactic Ambiguity within the LFG Framework with a View to Machine Translation. Ph.D. Dissertation. University of Manchester.

Chennoufi, A., &Mazroui, A. (2017). Morphological, syntactic and diacritics rules for automatic diacritization of Arabic sentences. Journal of King Saud University - Computer and Information Sciences, 29(2), 156–163. https://doi.org/10.1016/j.jksuci.2016.06.004

Chomsky, N. (1981). Lectures on Government and Binding, Foris, Dordrecht.

Debili, F., Achour, H., (1998).Voyellationautomatique de l'arabe. In: Proc. Work. Comput.Approaches to Semit.Lang. pp. 42– 49.

Diab, M., Hacioglu, K., Jurafsky, D. (2007).Automatic Processing of Modern Standard Arabic Text. In: Arabic Computational Morphology. Springer, Netherlands Dordrecht, pp. 159–179.

Haegeman, L. (1994). Introduction to government and binding theory. Oxford Uk and Cambridge USA: Blackwell.

Harbi, H. A.-. (2016). An Exploration Into The Various Types Of Ambiguity In the Arabic Language : A Pedagogical Study, 3(7), 2359–2366. https://doi.org/10.18535/ijsshi/v3i7

Hssini, M., & Lazrek, A. (2012). Problem of Multiple Diacritics Design for Arabic Script. 2(12), 48–53.

Jalabneh, A. (2009). Multiple functions of the nunations in Arabic syntax: A minimalist perspective. Dirasat: Human and Social Sciences, 36(3), 687–707.

Jarrah, M. (2016). Syntactic investigation of nunation in Haili Arabic Standard Arabic. Consider the following examples from Haili Arabic. 29(2016), 39–62.

Jarrar, M., Zaraket, F., Asia, R., & Amayreh, H. (2018). *Diacritic-Based Matching of Arabic Words*. 18(2). A.C.M. Trans. Asian Low-Resour. Lang. Inf. Process. 18, 2, Article 10 (December 2018), 21 pages.https://doi.org/10.1145/3242177

Kuhns, R. J., & Little, A. D. (1981). Government-Binding Theory and implemented in constraints as proposed by this theory , the requires few grammar rules and no explicit. 546–550.

Marrone, P. (2013). Chambers, RT. *Etica e Politica*, *15*(1), 583–605. https://doi.org/10.1093/acprof

Moubaiddin, A. (2013). *Sentences Investigating the Syntactic Structure of Arabic Sentences*. *January 2014*. https://doi.org/10.1109/ICCSPA.2013.6487275

Zitouni, I., Sorensen, J.S., Sarikaya, R. (2006). Maximum entropy based restoration of Arabic diacritics. In:

Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the A.C.L. – A.C.L. '06. Association for Computational Linguistics, Morristown, NJ, U.S.A., pp. 577–584. http://dx.doi.org/10.3115/1220175.1220248

An Investigation on Syntactic Complexity in Indian Acts

Jaya Verma

Assistant Professor, CVR College of Engineering, Hyderabad

ABSTRACT

There are number of countries that have English as a language of apex court, India is one of them. English, in India, crept into Indian education system as a subject with great effort of Lord Macaulay but dynamically it changed its status. It started its journey as foreign language, then it acquired the status of second language and now it is official language of Supreme Court of India. It would be interesting and at the same time quite important from pedagogical point of view that how far English used in India in legal documents still shares the same linguistic features established by the British.

The present study investigates the syntactic features of Indian acts. Studies conducted so far on legal English consensually attribute an adjective to legal English: complex. The present study aims to investigate the area where this complexity lie and would also justify the reason for its complexity. It would highlight those syntactical areas which contribute to its complexity. The present paper is divided into three sections, the first deals with definition of major concepts used in the paper such as stylistics, register, genre and legal English. The second section presents analyse of sentences from acts. The third and final section discusses the result of findings and gives substantiating conclusion to generalise those findings on acts. The research will employ triangulation method for analysis and interpretation of data. The genres chosen for the study is acts therefore, the present study does not generalize the findings of the research on all the genres of law but only on acts.

1. Introduction to Stylistics, Register and Genre

In the late sixties and early seventies, with the introduction of the concept of language description used in different texts, stylistics analysis came into existence. Stylistics is an essential branch of Applied Linguistics but it has always caused confusion and controversy due to its elusive nature and lack of framework. There are many difficulties related to its definition, approaches, and theories of interpretation. As far as matter of defining stylistics is concerned, it is difficult to single out one definition. Traditionally, stylistics was associated to linguistic analysis of fictional texts. It was introduced in literature as an alternative of

practical criticism as literary criticism rests solely on the subjective interpretation of texts. New approaches of interpretation of texts are expanding its periphery. Psychological and statistical analysis of texts have introduced cognitive and corpus linguistics respectively.

In recent years, stylistic studies have not only expanded its width but also its depth. Linguists have a started finding out reasons underneath the surface of linguistic features. Stylistic analysis of non-literary texts is popularly known as register analysis. Register analysis is study of language variety according to uses in different occupations. Crystal (1991: 295) describes register as "a variety of language defined according to its use in social situations, e.g. a register of scientific, religious, formal English." Bhatia (1991) views generalizing register features on all the genres of the particular register. He points out that labelling occupational variety such as scientific, medical, legal or even newspaper English is quite feasible but in reality such terms can be misleading as they represent the features of overall occupational variety while there is variation in genre conventions addresser-addressee relationships and communicative purpose.

Linguists feel that register analysis gives much generalized features of the text of a particular field and every text has many sub-texts. Every sub-text has its own linguistic features which may be different from overall registral features for example, general registral features of newspaper may vary when analysed on the basis of genre. This fact must also be taken into consideration that report on accident would be different from editorial article in terms of linguistic features. In a way, every register is comprised of many genres and to have deep analysis of text it should not be studied on the basis of genres. There are many theories of genre propounded by Martin (1984), Devitt (2004), Eggins (1994), Bhatia (2004), Freedman and Medway (1994), Swales (1990) which have given different stress to text.

The present paper focuses on syntactic features of legal English used in Indian Acts. Following the new trend of stylistics, this study does not confine legal register to a few genralized linguistic features rather it examines Indian acts intensely to come up with some representative syntactic features of genre of acts. It is necessary here to make the distinction between legal English and legalese clear for the better understanding of legal register. Legal English is formal academic English with some technical law jargon. It is used in Legal literature, legal textbooks, and correspondence: office memoranda, judicial opinions, and client letters. Although legal English is the style of English used by lawyers and other legal professionals in the course of their work, yet every text of legal genre is not written in the same manner. The correspondence on legal matters between companies is not done using legalese. These correspondences are done in formal Standard English with some technical legal terms. Undoubtedly, legal written English is formal but as far as its nearness to Standard English is concerned it deviates from norms at several syntactical spots.

Legalese is characterized by long complex sentences full of nominalizations, passives, oddly inserted adverbials, prepositional phrases and archaic vocabulary. Legalese is a term used, often derogatorily, to refer to the unique characteristics of legal English. Legalese is typically criticized for being overly complicated, dense, repetitive, and outdated (Candlin et al. 2002). The term legalese is used to describe the language used by the lawyers and court systems around the world, and it is a separate language that is not meant to be used and or understood by the layman or the people. To a large extent, legalese can be considered to be an extreme version of formal written language or can be said 'deviation' from standard formal English.

2. Studies on Legal English

Studies on legal studies commenced in late sixties of twentieth century. Evolution and historical development of legal

English have extensively covered by Mellinkoff (1963) and Tiersma (1999). Goutal (1976) conducted his study to trace back connection between judicial decisions and legal civilization of the country. Legal English has been analysed on different linguistic levels by many linguists. It was scrutinized at syntactic level by Gustafasson (1975). Gustafsson (1975) gave detailed description of syntactic properties of British and American legislative language in her doctoral thesis. Charrow and Charrow (1979) carried out their research on jury instruction and concluded that 'self-embedding' or 'center-embedding' create heaviness and make legal English less transparent. Hiltuman (1984) listed an account of structure and type of clausal embedding in legal English. He focuses on embedding, a syntactic property, which makes legal English complicated. Gotti (2001) presented insights into modality, Facchinetti (2001) conducted research on if-conditionals. Bhatia (2004) gave his views on legal written discourse in academic, professional and institutional contexts. Mackinlay (2004) analysed six hundred sentences from UK and EU legislative material to examine syntactic discontinuity as a major reason for incomprehensibility of legal text. They gave detail description of syntactic discontinuity. Gocheco (2011) investigates verb group in legal provision and explain how verb phrase could be continent which can make reading and understanding of legal text easy. Lehto (2012) studied diachronic changes in subordinate clauses. Subordinate clauses used in the genre of Early Modern English legal texts reveals that subordination in statutes vary during the sixteenth century but became more constant/ stable in the seventeenth century.

Social scientists, anthropologists, political scientists, many linguists as Carlen (1976), Charrow and Charrow (1979), O' Barr (1981) criticized archaic, convoluted and heavy use of language.

3. Methodology of the Present Study

In stylistic study, investigation is executed on the following levels: phonological, morphological, lexical, semantic, syntactic and graphological but the present research concentrates mainly on syntactic levels. The detail investigation of all the levels, no doubt, would have made research exhaustive and extensive but would not have been able to make the research deep. Therefore, the present research is attempt to present intensive study of syntactic features of legal English used in Indian statues.

The present research is a mix method research involving both qualitative and quantitative methodologies of research. Previous studies have made it evident that 'complexity' is a remarkable stylistic feature of legal English. Some studies justify its complexity and some oppose and demand for simplification. To investigate syntactical features of legal English, sentences from genre of acts were selected through systematic random sampling technique from official website of Supreme Court namely: http://indiacode.nic.in/ and were analysed on syntax theory of Phrase Structure grammar demonstrated in Bakshi (2000) and Quirk et al .(1985) by using parsing tree diagram. The research has attempted to answer the following questions:

- Does the English used in acts in India, too, replicate the linguistic features found out in earlier studies conducted by European countries?
- To what extent, these features are similar to or different to formal Standard English?
- Does the research prove that legal English used in acts is complex and if yes, what sort of complexity is it?
- What linguistic patterns are most commonly used, and what is their distribution?

For sentence analysis, the corpus of 300 sentences and for noun phrase analysis is 60 sentences taken from eleven acts ranging from year 2000 to 2007 namely:

- ☐ The Indian Christian Marriage Act 1872
- ☐ Parsi Marriage Act, 1936
- ☐ The Hindu Marriage Act, 1955
- ☐ Companies Act, 1956
- ☐ Copyright Rule 1957
- ☐ The Dowry Prohibition Act 1961
- ☐ Maternity Benefit Act, 1961
- ☐ The Cine Workers and Cinema Theatre Workers,1984
- ☐ Cable Television Networks 1995
- ☐ The Information Technology Act, 2000.
- ☐ Right to Information Act, 2005

As the present study is concerned with syntactic features of legal English and does not include discourse features, so selection of topic of acts has not affected result of the findings.

Crystal and Davy (1969), Mellinkoff (1963), Trosberg (1997), Gustafsson (1975) highlight complexity of legal language in their works and declare it 'complex'. Complexity in the language used in acts has been marked on the following syntactic parameters:

- ☐ Length sentences
- ☐ Number of complex sentences
- ☐ Horizontal relation of subordinate clauses
- ☐ High frequency of complex sentences
- ☐ Left branching of subordinate clauses
- ☐ Coordinated embedded clauses
- ☐ Syntactic discontinuity
- ☐ Self-embedding
- ☐ Complex noun phrases
- ☐ High frequency of adverbial clauses and phrases and their unusual positions

☐ Complex preposition

4. Analysis and Interpretation of Acts

4.1 Syntactic Features

4.1.1. Complexity

The language of law is language of arguments and justice; it cannot bear the tag of being complex without justifying the accusation levied on it. In terms of legal language it is important to make two issues clear: what we mean by complexity here and what elements create complexity.

There could be two types of complexity in legal texts: linguistic and non-linguistic. Non-linguistic complexity or legal (procedure) complexity is equally important to understand and make texts readable but here matter of discussion is linguistic complexity. Linguistic complexity is directly related to readability and comprehension. To understand the text without any syntactical and lexical difficulty is sign of comprehensibility but when reader faces problems in this regard it is due to complexity at these two levels. Complexity and comprehensibility are reversely proportionate to each other. It is quite challenging to trade off complexity and comprehensibility. Some linguists attached complexity to sentence structure (Masson and Waldron 1994**)** and other to vocabulary (Tanner 2010). Legal English is widely known for complexity at both the levels: syntactic and lexical. Gustafsson (n.d.) relates lexical complexity to semantic difficulties and refers it 'conceptual' complexity where in spite of knowing meaning of the particular word, word cannot be understood in lack of knowledge of legal principles.

In the present study elements of complexity has been measured at sentence level on the following parameters:

sentence length, structure of sentence and organization of sentence components:

4.1.1.1 Sentence length

Total no. of Sentences	Minimum	Maximum	Mean	Std. Deviation	Mode
300	7.00	404.00	71.3333	54.31706	53.00

Table: 1: Sentence length of acts

The above Table 1 exhibits that an average sentence lenght in acts is 71.33 words long and the longest sentence of Acts is 404 words and shortest 7 words. A sentence with 53 words has occurred most of the time. The average sentence length (71.33) is higher than any other register. Barber (1962 in Bhatia 1993) notified average sentence length in scientific English is 27.6 words and Gustafsson (n.d.) informed in journalism 20 to 21 words long. Davies and Vlachopolous (in Gibbons et al.2004) reported sentence length 74.3, 187 words per sentence respectively in different studies on legal language. Praksham (2004) conducted his study on Indian Evidence Act and reported average sentence length of 93 words. Tessuto's (2008) study on statutes also confirmed lengthy sentences, a prominent feature of legal English. Tessuto (2008) reported average sentence length ranging from 15 to 90 words. Very few simple sentences were reported in his studies, most of the sentences were complex.

The above statistical data reveals that average sentence length in acts is much more than sentence length of general English sentences. Cutts (2009) observes the average sentence length in general formal English is 15-20 words. Sentences are longer in legal English than an average sentence in general English.

Researches of Psycholinguistic have disclosed the fact that long sentences put adverse effect on readability and comprehensibility. Miller (in Tanner 2000) noted that the short- term memory can hold about seven unrelated units of information at any one time before it fails. It follows that the short-memory has to process long syntactically complex sentences which likely to contain too much information. Reformers and plain English supporters demand for improvement in this area of legal language. At the same time, many experts support its traditional form. Traditionalists held that in legal English sentences are long because it is required to make texts clear and precise. If simple and short sentences do not cover all the necessary guidelines of law, it is not appreciated in law. Law expert generally do not compromise on complexity to bring linguistic clarity. Gustafsson (n.d.) finds the use of lengthy sentence just a habit of adhering to old traditional system of drafting documents.

> The long sentence has tradition in English legal language. In the early days of legislation and jurisdiction the drafting of all legal documents was in the hands of a small number of professional lawyers and scribes, as the majority of people were illiterate. The handling of the case in court was crucially dependent on the form of plea. For the benefit of those drafting the pleas and other documents special form- book were published, and these books were then copied by generation after generation of lawyers. This practice gave birth to the tradition which among other things, demands that all relevant information pertaining to the same case has to be presented in one paragraph consisting of only one sentence. The more detailed and accurate description became, the longer the sentence grew.
>
> **(Gustafsson n.d.:26)**

Legal English sentences become lengthy due to establishing unambiguity, all-inclusiveness, and precision in documents (Bhatia 1993; Crystal and Davy 1969). They become long and odd in comparison to general English.

Tiersma (2006) lists out number of reasons for retaining old age characteristics. The first reason, he thinks is that the lawyers are more comfortable and safer to adhere to the well-known forms. Another reason is influenced by professional tactics of lawyers. Lawyers prefer their clients to remain totally dependable on their services. Originally lawyers used to charge for drafting documents on the basis of length of document. Lawyers sought out many ways to lengthen the documents. There were many ways to lengthen. One easy way for a lawyer to increase the length of a document was by employing 'recitals'.

Therefore, the discussion on sentence length concludes that longer sentences are of course the reason of incomprehensibility in legal English as no other genre uses such long sentences so readers do not encounter lengthy sentences elsewhere.

4.1.1.2. Horizontal relation of subordinate clauses

Sentences in acts are long because instead of constituting a new sentence from several same types of subordinate clauses attached to a single main clause, all the subordinate clauses occur in a sentence in horizontal relation. For example:

Any marriage solemnized, whether before or after the commencement of this Act, shall be voidable and may be annulled by a decree of nullity on any of the following grounds, namely:-

(a) that the marriage has not been consummated owing to the impotency of the respondent; or

(b) that the marriage is in contravention of the condition specified in clause (ii) of Section 5; or

(c) that the consent of the petitioner, or where the consent of the guardian in marriage of the petitioner was required under Section 5 as it stood immediately before the commencement of the Child Marriage Restraint (Amendment) Act, 1978, the consent of such guardian was obtained by force or by fraud as to the nature of the ceremony or as to any material fact or circumstance concerning the respondent; or

(d) that the respondent was at the time of the marriage pregnant by some person other than the petitioner.

The Hindu Marriage Act

In above sentence all *that*-clauses are directly subordinate clauses of single main clauses. The syntactic analysis of these types of sentences is difficult task. *That-clauses* in above example have neither occurred in hierarchical relation with one another nor embedded in one another rather they are at parallel or horizontal relation with one another. A reader has to pay equal attention on all the clauses equally because no one is superordinate of next clause rather share equal relation with main clause. Their representation can be in the following manner:

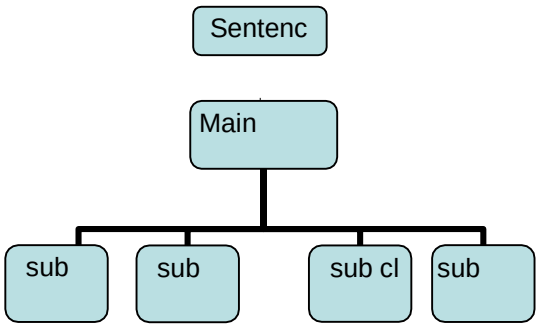


Figure1: visual representation of subordinating clauses sharing horizontal relation

4.1.1.3. Frequency of complex sentences

The sentences have been found longer in acts as most of the sentences have more than two clauses and majority of the sentences in acts are complex. The below figure 2 and table 2 substantiate this fact:

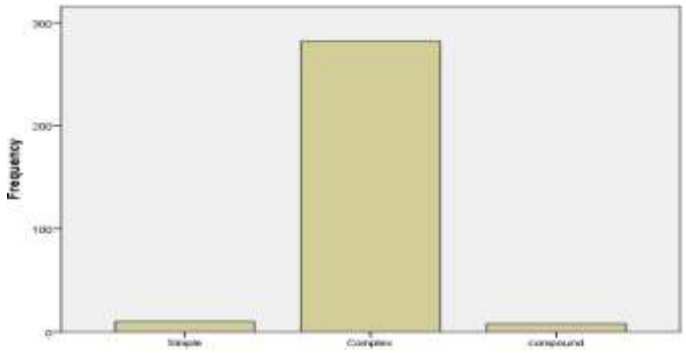


Figure 2: Frequency of types of sentences in acts

	Acts	
Type of sentence	**Frequency**	**Percentage**
Simple	10	3.3
Complex	282	94%
Compound	8	2.7
Total	300	100.0

Table 2: Frequency of complex sentences in Acts and in judgments

It is evident in the above Table-2 that frequency of complex is 282 in acts. This frequency of complex sentences is the highest in comparison to simple (10) and compound sentences (8).

4.1.1.4. Number of subordinate clauses in a complex sentence

Descriptive Statistics							
	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation
No.of sub cl	300	19.00	.00	19.00	1236.00	4.1200	3.39470
Ncl	300	4.00	.00	4.00	124.00	.4133	.76448
Rel cl	300	7.00	.00	7.00	201.00	.6700	1.09167
Adv F cl	217	7.00	1.00	8.00	479.00	2.2074	1.51790
Adv NF cl	148	12.00	1.00	13.00	274.00	1.8514	1.42076
Reduced Rel cl	300	5.00	.00	5.00	148.00	.4933	.78230
Verbless cl	10	1.00	1.00	2.00	12.00	1.2000	.42164
Nominal Rel cl	3	.00	1.00	1.00	0.00	1.0000	.00000
Valid N (listwise)	0						

Table- 3: Number of subordinate clauses in Acts

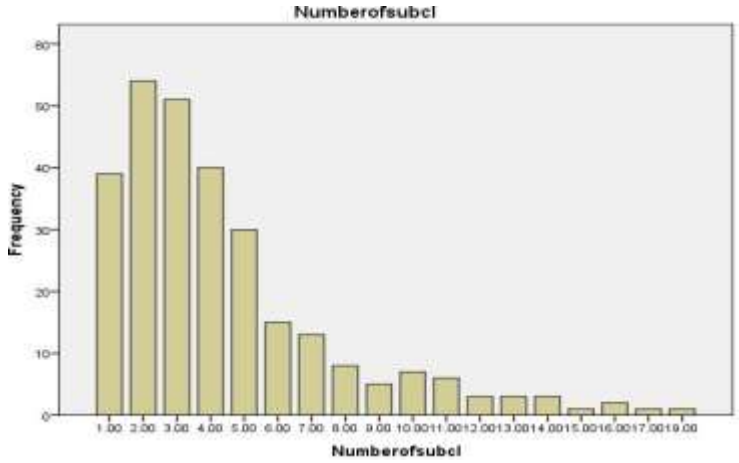


Figure 3: Number of Subordinate Clauses in Acts

It can be estimated from the above table 3 that approximately (679/ 282= 2.40) three subordinate clauses per complex sentences are used in acts. Apart from three subordinate clauses per sentence, the range, maximum and minimum frequency of their occurrence also draw our attention to noticeable fact that there are instances of 19 subordinate clauses in a sentence. During the present study, instances of 16, 15, 14 subordinate clauses in acts were found. The maximum frequency of adverbial and relative clauses indicates the elements that lengthen complex sentences.

4.1.1.5. Position of subordinate clauses

After sentence length and more number of complex sentences, position of subordinate clauses is another element that causes complexity in legal English. Quirk et al. (1985) define three positions of subordinate clauses; initially, medially and at the end of superordinate clause. Initial and medial positions refer to beginning of superordinate clause. The technical terms used for initial, medial and final position are left-branching, nesting and right –branching respectively. When subordinate clauses occur at the beginning of superordinate clause, they are known

as left-branching. The embedding of clauses in left-branching is limited to two degrees. Left-branching is more prominent feature in noun phrases than clauses.

Right-branching takes place when subordinate clauses are placed after superordinate clauses. In every type of writing, right-branching is commonly used. It is easiest to comprehend. When subordinate clause is placed in between superordinate clause it is called nesting. "Nesting or medial branching creates the most awkwardness, if nested clause is long and is itself complex. It also violates the principle of end-weight" (Quirk et al. 1985: 1037)

> ***Initial:*** *When you're ready, we'll go to my parents' place.*
>
> ***Medial:*** *We'll go, when you're ready, to my parents' place.*
>
> ***End:*** *We'll go to my parents' place when you're ready,*
>
> **(Quirk et al., 1985: 1037)**

In standard formal English left branching and nesting are not frequent rather they are avoided for maintaining the principle of end-weight. In the present study, there are 86 instances of left- branching in Acts data. In English there is a preference for placing long, complex, "heavy" constituents towards the end of the clause, because in this position they seem easier to understand or process. Legal English deviate from general English in this norm and violates the rule of end-weight by exerting more weight on left side of the sentence.

Example-

*{ **Adv cl F** (If the Chief Information Commissioner or a Information Commissioner in any way, concerned or interested in any contract or agreement **Red Rel cl** (made by or on behalf of the Government of India or participates in any way in the profit*

thereof) or in any benefit or emolument[Ncl] *(arising therefrom otherwise than as a member)and in common with the other members of an incorporated company, he shall, for the purposes of sub-section (/), be deemed to be guilty of misbehaviour.}*

(Companies Act 1956)

Levin and Garrett (1990) tested a hypothesis whether left- branching sentences are more formal than right-branching and center-embedding (or nesting). Their hypothesis proved positive. In their research, they found left-branching sentences more formal. Justifying the findings, they state:

> *Many theories of parsing claim that left-branching sentences (LB) are more difficult to produce and to understand than right-branching (RB) sentences. The reasons are variously thought to lie either in their load on memory required to process LB forms or the perceptual demands that LB sentences make.*
>
> **(Levin and Garrett 1990: 511)**

Gustafsson (n.d.) also reports the same reason in his article. Legal English uses highly formal sentences structure. There are considerable numbers of left-branching sentences found in acts. Tessuto (2008) remarks on disadvantage of left-branching:

> *while the majority of legal writers have advocated for the elements in a legislative sentence to be arranged as case/condition + legal subject+ legal action (e.g. Coode 1976; Thring 1902), variation in syntax of legislative texts has led linguists scholars to speak in terms of 'right-branching' (where conditional clauses are positioned initially) as opposed obviously to 'left-branching '. In syntax this means that a case description constitutes a subordinate adverbial clause linked to another hypotactically, and leads to very long, complex*

sentences. Hence the reader often cannot absorb the content of the provision; at least on a first reading.'

(Tessuto2008: 7)

Lehto (2012:16) states that 'the place where subordinated clauses are positioned in sentence also affect legibility since final subordination is usually more readable than initial or middle subordination.'

In most of case where subordinate clauses have occurred at the left side of the main clauses, we find that subordinate clauses occurring on the left are adverbial clauses beginning with 'if', 'where' and 'when'. Bhatia (1993: 110) finds it reasonable for case description in the form of adverbial clauses at initial position because there are a few legislative statements of universal application so drafters place them initially to have more focus on these statements. In this way, delaying of subject is done intentionally.

Where a notice is served by the Central Public Information Officer or State Public Information Officer, as the case may be, under sub-section (/) to a third party in respect of any information or record or part thereof, the third party shall, within ten days from the date of receipt of such notice, be given the opportunity to make representation against the proposed disclosure.

(Right to Information Act 2005)

4.1.1.6. Coordinated embedded clauses

As it is already known fact that 'embedding gives rise to the theoretical possibility of grammatical units having indefinite length' (Quirk et al. 1985: 44) and if embedded clauses are coordinated, it create a deadly combination of length and complexity of structure. Legal documents are full of these sorts of instances. Subordinate clauses are increased in the sentence in

two ways: one way is combined them with subordinate conjunctions and another way is to joint them with coordinate conjunction if clauses are of equal rank. If a subordinate clause embedded in another subordinate clause is combined with coordinate conjunction, it grows unnecessary complication and creates interruption in the flow of sentence. In the below example, consecutive four clauses : *that*-clause, relative clause and two adverbial clauses are embedded in one another and the second adverbial clause is coordinated with *or*.in this type of embedding sentence becomes remarkably complex.

> *For the removal of doubts it is hereby declared* ^{Ncl}*(that no commission shall be paid under clause (a) of sub-section (1) to any person on shares or debentures* $^{Rel\ cl}$ *(which are not offered to the public for subscription))*
>
> *Adv cl F (Provided that Adv cl F (where a person has subscribed or agreed to subscribe under clause (a) of subsection for any shares in, or debentures of, the company and before the issue of the prospectus or statement in lieu thereof any other person) or(persons has or have subscribed for any or all of those shares or debentures and that fact together with the aggregate amount of commission payable under this section) in respect of such subscription is disclosed in such prospectus or statement,) then, the company may pay commission, to the first-mentioned person in respect of such subscription.)}*
>
> *(**Companies** Act 1956)*

4.1.1.7. Majority of complex noun phrases

	Acts
Total No. of NP	221
Simple NP	35
Complex NP	186
Pre-modifiers	214
Post-modifiers	172

Table -4: Noun Phrase in Acts

The above table 4 shows high propensity for complex noun phrases. Out of 221 noun phrase in the data of 60 sentences has only 35 simple noun phrases (noun without its modifiers) while there 186 complex noun phrases. The high frequency of pre-modifiers indicates towards the fact that left-branching is common practice even in phrases. Apart from it; this is interesting to see that maximum range of pre-modifiers goes to 21 pre-modifiers in a sentence. Mean score which points towards occurrence of at least 3 noun phrases in a sentence consolidates Williams's (2005) remark that 'Legal documents tend to be 'nouny' rather than 'verby','The below table 5 shows which pre- modifier and post-modifier play prominent role in intensifying complexity in the texts. The study discovered that articles are major pre-modifier while in prepositional phrases have the highest frequency as post-modifiers. It also shows high frequency of complex noun phrases.

4.1.1.8. Syntactic discontinuity

There are 149 instances of syntactic discontinuity in acts in the present study. Syntactic discontinuity is outcome of an intentional and rational effort of structuring legislative writing. It occurs if two elements which would normally be situated beside each other in the sentence structure are formally separated by another expression or clause being inserted in between them. As a result of this, the two elements, which are both semantically

and structurally related, may end up distanced from each other in the structure of the sentence and the close semantic or structural relation between them may become less obvious. There is no doubt that to a certain extent syntactic discontinuity contributes to a greater complexity of the text and makes greater demands on the interpreter. Gibbons (1994) does not victimize complicated sentence structure (linguistic complexity) only factor of complexity rather he finds it one of the factors. As complexity is being associated to communication, Gibbons (1994: 287) lists out four sources of communication difficulty: specialized text structures and procedures used in the law, the extreme writeness of many legal documents, and technicality of legal discourse, power disparity and hyper formality.

Incidents of Syntactic discontinuity are at phrase and clause levels are abundant in legal English. Syntactic discontinuity is created in the text by using prepositional phrases, adverbial clauses and non-finite verb *–ing*. In the below example, we can see the subject *The committee* has been separated with its corresponding verb *commented on* by insertion of prepositional phrase.

> *The appropriate Government shall, if necessary, update and publish the guidelines referred to in sub-section*
>
> *(2) at regular intervals which shall, in particular and without prejudice to the generality of sub-section (2), include—*
>
> *(a) the objects of this Act;*
>
> *Right to Information Act 2005*

Syntactic discontinuity is more common in formal language than in informal language. In general English syntactic discontinuity appears to be a marginal feature of the syntax, in legislative writing it is one of the central syntactic features of the

text. Tessuto (2008: 12) affirms that syntactic discontinuity idiosyncratically characterizes English legislative writing.'

4.1.1.9. Self-embedding:

The medial subordination of one constituent within another constituent of the same kind is called self-embedding (Quirk et al, 1985: 1040). One layer of self-embedding is quite comprehensible but more than one layer makes comprehension difficult. In legal English, there are number of example of one layer of self-embedding. Self-embedding is more complex than multiple center-embedding of different clause types (Karlsson 2006). Relative clauses and *that*-noun clause functioning as complement is employed for self-embedding. The problem of comprehensibility arises when another subordinate clause is combined within self-embedding and creates a wide gap between subject of superordinate clause and verb.

> *The directions^{Ncl}(that may be given under this sub-section) may include a direction^{Ncl} (that one member of the company present in person or on proxy shall be deemed to constitute a meeting.)*
>
> ***(Companies Act 1956)***

4.1.1.10. Complex and archaic prepositional phrases

The most common stressed aspect characterizing legal language is the frequent use of prepositional phrases. In terms of the use of complex prepositions in formal document, according to Quirk et al. (1985:304) stated that *"Legal English is notable for complex prepositions, the following being among those found mainly in legalistic or bureaucratic usage: in case of, in default of, in lieu of, on pain of, in respect of* "One-word prepositions are called simple prepositions while two-word or three-word sequences are known as compound prepositions. A preposition is a word which generally comes after a noun or a noun phrase which is the object of the preposition. A preposition along with noun or noun phrases makes a prepositional phrase.

Prepositional phrases have syntactic functions of post nominal modifier of noun in a noun phrase, adverbial and complement of an adjective. Quirk et al. (1985) propose one more category of prepositional phrases: complex prepositions. They suggest subdivision of it as two- and three-word sequence. 'In two-word sequence, the first word (which usually is relatively stressed) is an adverb, adjective, or conjunction, and second word is simple preposition (usually *for*, *from*, *of*, *to*, or *with*).' The sequence of three-word is preposition + noun + preposition. (Quirk et al. 1985: 669)

The most ubiquitous feature of legal language is complex prepositional phrases. They have highest frequency of occurrence in all type of legal documents especially in Statutes or Acts. Legislation in English has more complicated and longer sentences to some extent because of large number of complex prepositions. Their high frequency is not alone responsible for the complexity but also their unstable position in the sentence. Mostly prepositional phrases are functioning as modifiers of nominal constituent in the present findings.

Vedralova´)´2008(concluded in his study of preposition in legislative text that sentence in legal texts are longer due to complex prepositions. Another observation he made that legal English uses more complex prepositions than literary texts. Prepositions such as *in conformity with*, *in line with in accordance with*, *accordance with*, *according to*, *pursuant to*, *in relation to*, *with regards to* convey reference. Preposition like *for the purpose of*, *with a view of*, *by means of* , *on the basis of* denote references.

Bhatia (1998) figures out the reason of high frequency of prepositional phrases. Bhatia logically illustrates that complex prepositions are used to signal textual authority in the legal text. Bhatia analysed the British Housing Act 1980 to investigate the reason for it usage. It is a kind of intertextual device used in the legal text (Bhatia ibid). Following Bhatia similar type of

examples were tried to find out in the present study and number of example were easily spotted due to its high frequency.

> *a. It shall be a constant endeavour of every public authority to take steps* ***in accordance with*** *the requirements of clause (b) of sub-section) (1) to provide as much information suomotuto the public at regular intervals through various means of communications, including internet, so that the public have minimum resort to the use of this Actto obtain information.*
>
> ***Right to Information Act 2005***

4.1.1.11. High frequency of finite adverbials clauses and phrases and their unusual positioning in the sentence

Adverbs are traditionally categorized as one of parts of speech in English grammar whose primary function is to modify verbs, adjectives, and adverbs. But other parts of speech also perform the same functions. In order to avoid confusion and to be more precise, Quirk et al. (1985) used the term "adverbial" to refer to all the elements of the clause performing the action of an adverb.

	N	Range	Minimum	Maximum	Sum	Mean	Std. Deviation
AdvF cl	217	7.00	1.00	8.00	479.00	2.2074	1.51790
AdvNF cl	148	12.00	1.00	13.00	274.00	1.8514	1.42076
Valid N (listwise)	116						

Table -5: Descriptive Statistics of adverbial clauses

Adverbial is fairly mobile sentence element in English (Gustafsson n.d.). Number of adverbial phrases and clauses in legal English is relatively high as legal English believes in pooling in all the factors and issues affecting the matter in

question. For this purpose, 'adverbials are placed where they will serve the demands of unambiguity' (Crystal and Davy 1969:203).Charrow and Charrow (1982) consider it quite absurd insertion of adverbial in the sentence and call it "misplaced phrases".

The high frequency of adverbial clauses is considered to be one of the characteristic features underwriting complexity of written legal language. Adverbial finite clauses are used to refer time, place, condition, concession, manner, contrast, and purpose (Bakshi 2000: 271-280). In finite adverbial clauses, subordinate conjunctions are used to create relation between matrix and subordinate clauses. The realization of non-finite adverbial clauses takes place in form of participle and infinitive. In the table 5, it can be seen that frequency of adverbial clauses is higher than other clauses; in fact they have highest frequency. As shown in table 5, When we compare adverbial finite and non- finite clauses, we find that finite adverbial clauses are higher than non-finite clauses in terms of their frequency (F=479, NF=
274) and percentage (72.3, NF= 50.7).

Here are the examples of adverbial clause which have exclusive feature of legal writing. The particular types of adverbial which have become distinctive feature of Acts are conditional and temporal adverbial finite clauses.These adverbial constructions are an essential part of legislative provisions, as their function is to establish the scope of application of legislative rules. Kurzon (1985:269) cites locative role of adverbial clauses and phrases beginning with *Where*…and *without*…*:*

This adverbial clause or phrase, often called 'case' in legal textbooks, sets down the circumstances in which the rule, the contents of the main clause, is to be applied (what is called in the Logic of Action 'the conditions of application'). Proviso is one of the components of Acts. It begins with conditional clauses beginning with *provided*. There are 112 conditional clauses

beginning with *if*, 56 beginning with *provided* and 31 beginning with *unless*.

4.1.1.12. Position of adverbial clauses

The occurrence of adverbial phrases can be seen in every possible place in the sentence. They can be seen before subject, after subject, between subject and verb, between auxiliary and main verbs, after main verb and after conjunctions. In the present study, most of the instances found of adverbial insertion between auxiliary and main verb. Placing them after main verb can create more scope for misunderstanding. Adverbial clauses and phrases in all ways contribute greatly in making legal English distinctive. Besides their frequency, their position is also noticeable feature. It is placed at unusual places in the sentences where adverbial are normally not placed. They are inserted in between helping verb and main verb, at some places they are placed just after the conjunctions.

> *Before taking any decision under sub-section (/),) the Central Public Information Officer or State Public Information Officer,*$^{Adv\ cl\ F}$*(**as the case may be,**) shall take into consideration the representation made by a third party under section 1*.
>
> **Right to Information Act 2005**

Kurzon (1985) offers two reasons for positioning of adverbial clause and phrases. The first reason is clarity of word order which makes reading comprehension easier and another meaning of clarity is clarity of interpretation of facts in court. The second reason is linguistic factors as historical tradition, theme and iconicity.

The Renton Committee (in Kurzon1985:270) locates the area of complexity in legal texts. The Renton Committee mentions positioning of adverbial clause and phrases before subject and between subject and main verb decreases clarity. Many pieces of psycholinguistic research (Bever 1970) have

proved that positioning of adverbial clause before subject is technically called 'left-branching' puts extra exertion on the memory. It has been decribed as 'scene-setter' by Kurzon (1985)

> *…the adverbial sets the scene for the subsequently mentioned event. In the case of legislative sentences, the initial adverbial clause or phrase sets the scene by laying down the circumstances for the application of the rule in the main clause. The circumstances in which the rule is to be applied may be divided into two types: (I) time and (2) text. By time, I mean that some event must take place before the rule comes into operation… By text, I mean reference to another statute or to another section of the same statute…*
>
> ***Kurzon (1985:274)***

This rationale is the result analysis of text by categorising its part into theme and rheme. Theme functions as link with previous text.

Conclusion

The stylistic features of acts extracted out from the present study are: long, complex sentences with multiple embedding of clauses and phrases, Middle English syntactic structure, considerable number of adverbial clauses, and high frequency of non-finite verb forms, syntactic discontinuity etc. Complexity in acts was found on both clausal and phrasal levels. Complexity of legal English has been major issue of discussion throughout this paper because it is one of the hurdles in the basic principle of communication. 'It is least communicative among all the variety of language (Crystal and Davy 1969:193). 'Communication difficulties' (Gibbons 1994: 286) arise due to complexity of legal English. In discussion of complexity, researcher feels that complexity of legal English is not only due to language alone rather there are some other reasons as understanding of legal principles and process too responsible for complexity of the

legal language. Therefore, the present study reveals that there are two layers of complexity in legal language. On the surface level, linguistic components as sentence structure and technical vocabulary create complexity and at the deeper level, legal concept are responsible for conceptual complexity.

References

Bakshi, Raj. Nath. (2000). *A Course in English grammar*. Hyderabad: Orient Longman.

Bever, T. G. (1970). The cognitive basis for linguistic structures. Cognition and the development of language.

Bhatia, Vijay Kumar. (1991). Prescription, description and explanation in ESP.In R. Amritavalli, A. (Eds).*A carnival of Grammarians*. Madras: McMillan Press Limited.

Bhatia, Vijay Kumar. (1998). Intertextuality in legal discourse. Retrieved on 13 July 2012 fromhttp://jalt publications.org/old_tlt/files/98/nov/bhatia.html

Bhatia, Vijay Kumar. (2004). *Worlds of written discourse: a genre-based view*. London: Continuum.

Candlin, Christopher, Vijay Kumar Bhatia, C. H. Jensen.(2002). Developing legal writing materials for English second language learners: problems and perspectives. *Department of English*, City University of Hong Kong, Hong Kong English for Specific Purposes 21 299–320 http://legalwritingcoach.com/legal writing materials article .pdf

Charrow, Robert P. and Veda R. Charrow (1979). Making legal language understandable: a psycholinguistic study of jury instructions.*Columbia Law Review* 79, 1306-1374.Retrieved on 5th July2014 fromhttp://www.jstor.org/stable/1121842

Crystal, D., & Davy, D. (1969). The language of legal documents. *Investigating English Style*, *Harlow: Longman*.

Crsytal, David. (1991). A dictionary of linguistics and phonetics. 3rd edition. Cambridge, MA: Basil Blackwell.

Cutts , Martin. (2009). *Oxford guide to plain English*. USA: Oxford University Press.

Devitt, Amy J. (2004). *Writing genres*. Carbondale: Southern Illinois University Press.

Eggins, Suzanne. (1994). An introduction to systemic functional linguistics. London: Cassell.

Freedman, Artemeva and Peter Medway. (1994). *Genre and the new rhetoric*. London: Taylor & Francis.

Gibbons, John., V. Prakasham, K. V. Tirumalesh, H. Nagrajan. (Eds.) (2004). *Language in the law*. Hyderabad: Orient Longman.

Gocheco, Paulina. (2011). An analysis of verb groups in legal discourse: implications for teaching English for specific purposes*TESOL JournalVol. 4(1), pp. 2-15* http://www.pdf-repo.com/pdf_1a/5b3g2o2em951f58768.html

Goutal, Jean Louis. (1976). Characteristics of judicial style in France, Britain and the U.S.A.*The American Journal of Comparative Law Vol. 24 No. 1 pp. 43-72* Retrieved on 29 March 2014 from http://www.jstor.org/stable/839167

Gustafsson, Martia., (1975).*Some syntactic properties of English law language*. Publications of the Department of English 4, Turku: University of Turku.

Gustafsson, Martia., (n. d.). What makes legal language complicated? Retrieved on 19 Jan 2014 from https://www.jyu.fi/hum/laitokset/solki/afinla/julkaisut/arkist o/33/gustafsson

Karlsson, Fred. (2006). Constraints on multiple center-embedding of clauses.*J. Linguistics* 43, 365-392. Retrieved on 2 March 2014 fromhttp://www.ling.helsinki.fi/~fkarlsso/ceb5.pdf

Kurzon, Dennis. (1984). Theme, hypertheme, and the discourse structure of British legal. *Text - Interdisciplinary Journal for the Study of Discourse*. *Volume 4*, *Issue 1-3*, Pages 31–56.

Lehto, Anu. (2012). Development of subordination in early modern English legal discourse.Journal of historical pragmatics. Volume 11, Issue 2 pp. 277–300 (24)http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-176.pdf

Levin, Harry and Peter Garrett (1990). Sentence structure and formality. Language in Society Vol. 19, No. 4 (Dec., 1990), pp. 511-520 http://www.jstor.org/stable/4168176

Mackinlay, Jana. (2004). Syntactic discontinuity in the language of UK and EU Legislation. Retrieved on 12th Aug, 2013 from http://www.espworld.info/Articles_7/Syntactic%20DiscontL egal.htm>

Masson, Michael. E. J. & Mary Anne Waldron. (1994). Comprehension of legal contracts by non-experts: effectiveness of plain language redrafting. *Applied Cognitive Psychology*, Vol. 8, pp. 67-85.

Mellinkoff, David. (1963). *The Language of the law*. Boston & Toronto: Little Brown and Company.

O' Barr, William. M. (1981).The langauge of law. In C.A. Ferguson & S. B. heath (Eds.), Language in the U.S.A. New York: Cambridge University Press.

Quirk, S., S. Greenbaum, G. Leech and J. Svartvik.(1985). *A comprehensive grammar of the English language*. London: Longman.

Swales, John M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tanner, Edwin. (2010). Legal language and the non-law research student. *Journal of the Australasian Law Teachers Association.pp77-86*

Tessuto, Girolamo. (2008). Drafting laws in UK settings: implementing plain language and discourse?Retrieved on 12

July 2103 from http://www.federalismi.it/focus/semplificazionenormativa/ApplOpenFilePDF.cfm?dpath=document&dfile=08012008030941.pdf&content=Drafting+laws+in+UK+settings:+implementing+plain+language+and+discourse?+-+dottrina+-+dottrina+

Tiersma, Peter. (1999). *Legal language*. Chicago: The University of Chicago Press.

Trosberg, Anna. (1997). Rhetorical strategies in legal language: discourse analysis of statutes and contracts.Germany : Gunter NarrVerlag Tubingen.

Vedralová, Lenka. (2008). Complex prepositions in EU legislation and their translation equivalents. Masaryk University Faculty of Arts. Retrieved on 5th July 2013 fromhttp://is.muni.cz/th/75440/ff_m/Text_diplomove_prace_IS.txt

Williams C. (2005). Tradition and change in legal English: verbal constructions inprescriptive texts. Peter Lang: Bern.

A Genderlect Study on Function Words Used in Suicide Notes

N. Nirmeen[1] and Dr. N. Vijayan [2]

(1) Research Scholar, Department of Linguistics
(2) Assistant Professor, Bharathiar University, Coimbatore

ABSTRACT

The choice of words in our day to day conversation is the mirror of our self. They are not just part of a language but they are the reflection of our ideas, thoughts, emotions and behaviour. Words are of two types; content and function words. The choice of words varies from gender to gender, religion to religion, region to region and so on. It is possible for a speaker/writer to restraint their choice of content words but function words aren't so. There are a deal of researches on the language of genders. These researches are constat that men and women use language differently. Studies suggest that victims of suicidal ideation and victims who complete suicide almost go through similar psychological pain. And psychologists believe that victims with specific psychological disorders or mental state share similar characteristics which includes language also. The present study investigates the language used by both men and women in suicide notes to conclude if their choice of function words depends on their mental state or gender.

Keywords: Function words, Gender language, Suicide note

1. Introduction

Words are more than a means of communication; they can shape our beliefs, behaviours, feelings and ultimately our actions. The power of words is beyond expression, it defines what the person wants to express and the style of expressing the thoughts. Word choice provides information about person's perception (*Semin & Fiedler*, *1988*). Words are divided into two types; Content words and Function words. Content words are categorised as being those lexical items which have a relatively 'specific or detailed' semantic content and as such carry the principal meaning of the sentence. They name the objects, events, properties, and locations and directions that are at the heart of the message that the sentence is meant to convey. As opposed to content words, function words have a more 'non-

conceptual' meaning and fulfil an essentially 'grammatical' function; in a sense they are needed by the surface structure to glue the content words together, to indicate what goes with what and how. Content words belong to open classes of words; that is, new members can freely be added to this class. Function words, on the other hand, have a fixed roaster, and once it has been learned, few other members are ever added. (*Mouton de Gruyter*, *2001*).

2. Function Words

Function words are also called as style words or junk words. The meaning of an utterance is reliable on the function words which include articles, demonstratives, pronouns, prepositions, conjunctions and auxiliary verbs. When people try to present themselves in a certain way, they tend to select what they think are appropriate nouns, verbs, but they are unlikely to control their use of articles and pronouns. In short, function words reflect linguistic style (*Pennebaker & King*, *1999; Pennebaker*, *Mehl*, *& Niederhoffer*, *2003*). Rather than reflecting what people are saying, these function words connote how they are speaking. These words not only aid in generating meaning but also track the thoughts and emotions of people as the thoughts and emotions of people are often translated into words. The use of function words depends on the context and the listener/reader, and it also varies according to the psychological state and biological health of the person. It is said that 55% of all the words that people speak, write, hear, and read are function words.

Content and function words tend to be processed in the brain very differently (*Miller*, *1995*). Function words reflect how we communicate and content words convey what we say. Function words are constructed depending on one's social and psychological worlds. Pennebaker's (*2001*) statistical analyses have shown that these small words may hint at the healing progress of patients and give insight into the personalities and

changing ideals of public figures, from political candidates to terrorists. He also says that low level words can give clues to large scale behaviours.

3. Genderlect and Suicide Notes

The term Genderlect refers to the language style of particular genders. Though genderlect studies include all the three genders, many researchers and the present study also focuses only on men and women. Men and women differ in their opinions, skills, thoughts, emotions, actions and so on. As mentioned earlier ideas, thoughts, emotions etc are translated into words. It is certain that their language style will not be similar as they have different views of the world. Studies on gender language have proved that men and women use language differently. Language of men and women are best understood as two distinct ways of communicating rather than assuming one to be superior.

According to Merriam Webster dictionary, the definition of suicide is "the act or an instance of taking one's own life voluntarily and intentionally". Suicide victims think their incapable of bearing their burden and decide to escape from it. It often happens due to forlornness, grim, anxiety, depression and other mental illness. Victims of suicide usually leave their final words to their loved ones through a letter, voice note or a video. Their choice of words and style of language differs from the people who are non suicidal and not being through such mental illness. Researchers claim that people with similar mental state as in suicide notes show similarities in language style. A deep analysis on the language of suicide notes can mirror their emotions, thoughts, behaviour and motivation.

To reduce the act of suicide, it is important to identify the risk factors among suicidal people and identify whether those risk factors varies among men and women. Language is one of the major clues they leave behind. To understand if the language

style categorised by genders correspond or diverge from each other in a particular context having similar mental state, it is necessary to analyse the language in depth. As mentioned above, 55% of our words are function words and it plays a major role in revealing our thoughts, emotions, behaviour and motivation.

Thus, the study analyses the function words used by men and women in suicide notes to identify what and how function words are used, and the similarities and variations on the usage of these words among them. Suicide notes written in English by men and women between the years 2010 to 2020 were extracted for the study from internet sources.

4. Analysis

The function words used by men and women were extracted to calculate the percentage of function words used in the suicide notes. It was found that 51.4% of function words were used in the suicide notes. The percentage of each function word used by both the genders was calculated to trace the graph of the particular word. And, the total percentage of particular function word was calculated to identify the function words which occurred frequently in the suicide notes.

FW	**W – SN (%)**	**M – SN (%)**	**Total %**	**FW**	**W – SN (%)**	**M– SN (%)**	**Total %**
Articles				**Prepositions**			
A	1.51	3.61	2.37	Of	2.70	3.25	2.96
The	3.52	8.67	5.64	For	2.39	0.72	1.70
Demonstratives				On	1.00	0.54	0.81
That	1.51	1.26	1.41	In	2.01	4.88	2.00
This	3.02	1.62	2.44	At	0.62	0.72	0.66
These	0.25	0.36	0.29	With	0.12	1.08	0.51
Those	0.37	0.18	0.29	To	5.54	7.41	6.31
Pronouns				About	0.37	0	0.22
I	12.59	5.06	9.50	From	1.13	0.54	0.89
Me	4.53	3.25	4.00	Out	0.12	0	0.07
My	5.65	3.61	4.23	Than	0.25	0	0.14

Myself	0.62	0.18	0.44	**Auxiliary verbs**			
Mine	0.12	0	0.07	Be	1.13	3.25	2.00
We	0.62	1.08	0.81	Is	1.25	4.33	2.52
Us	0.25	0.54	0.37	Am	2.39	2.16	2.30
Our	0.88	1.26	1.03	Are	0.25	1.08	0.59
You	9.94	4.15	7.57	Have	2.51	6.14	4.00
Your	1.25	1.44	1.33	Has	0.25	0.90	0.51
Yourself	0.12	0	0.07	Did	0.37	0	0.22
They	0.75	2.35	1.41	Was	1.38	0	0.81
Their	0.62	0.18	0.44	Were	0.62	0	0.51
Them	0	0.90	0.37	Had	0.37	0.36	0.22
It	3.02	2.71	2.89	Being	0.50	0	0.29
He	0.50	0.72	0.37	Been	0.88	0.18	0.59
She	0	0.72	0.29	Don’t	0.37	0.18	0.29
Him	0.12	0	0.07	Didn’t	0.50	0.18	0.37
His	0.25	0	0.14	May	0.25	0	0.14
Her	0	0.54	0.22	Might	0.25	0	0.14
Someone	0	0.18	0.07	Can	0.25	0.90	0.51
Anyone	0	0.36	0.14	Could	0.75	2.16	0.89
Something	0	0.18	0.07	Should	0.37	0	0.22
Nothing	0.62	0	0.37	Will	0.25	2.71	1.26
Others	0	0.72	0.29	Would	0.88	0.36	0.66
All	1.25	1.26	1.26	Can’t	1.38	0.54	1.03
Conjunctions				Couldn’t	0.37	0	0.22
And	5.28	7.41	6.16	Shan’t	0.12	0	0.07
But	2.64	0.72	1.85	With	0.12	1.08	0.51
Or	2.14	0	1.26	Without	0.37	0	0.22
Yet	0.25	3.79	0.14	Much	0.50	1.08	0.74
As	1.13	0.72	0.96	Both	0.25	0	0.14
If	0.88	0.72	0.81	Many	0.37	0.18	0.29
Once	0.12	0	0.07	Any	0.62	0.36	0.51
So	0.37	0.18	0.29				

FW- Function words, W – Women, *M – Men, *SN – Suicide notes Table 1.1

5. Findings

Articles and Demonstratives

Articles '*a'* and 'the' were identified in the notes. The use of indefinite article '*a*' was used to describe something unspecific. The definite article '*the'* was used more than '*a*' while they discussed about specific things, people and situation. The use of '*the*' more than '*a*' shows that they mention the things or people or situation in particular which is known to the reader or expects the readers to understand it. And it was identified that men used a higher percentage of articles than women.

Demonstratives '*this, that, these and those'* were identified in the writings of both men and women but women showed higher percentage than men. Women often made use of these demonstratives as they confessed about variety of memories and people. They either did not want to specify them or they expected the reader to understand. Men used lesser percentage of '*that and those'* as men focused less on past.

Pronouns

Pronouns are one of the useful linguistic elements that analyses intention and attention. Pronoun use is very important in showing the quality of a close relationship, because it shows how individuals are referring to each other (Pennebaker & Tausczik, 2010). Women frequently used pronouns than men such as *'I, you, he, it, this, that, those, him, his*, *my*, *myself*, *mine'* which were personal, reflexive and possessive pronouns. The percentage of first person and second person pronoun was higher among women. The use of such pronouns reflect being self- focused, self aware, and more self reflective. It also means that the person is insecure, torment, and depressed. Nirmeen and Vijayan *(2020)* analysed the language of suicidal and found that 81% of people with depression and suicidal ideation used first person pronoun. Simmons, Chambless, and Gordon (*2008*) found that use of second-person pronouns was negatively related

to relationship quality. The second person pronoun was used among women while blaming their partner or the person who was responsible for the distress which reflects the negative relationship. It was observed that only 15% of first person pronouns were used in the notes of men. Third person pronouns like '*she*, *her*, *it*, *they*, *them*, *others*, *someone*, *anyone*, *anything*, *something*' was observed in the writings of men which shows that they focused on others than self while recalling a memory which affected them predicting themselves as the victim. There was no much difference found in the use of words like '*we*, *us, our'*. A study says, these words reflect two very different uses of 'we'. One of which includes *'I* and *you'* which is very warm and reflects the relationship bonding which is highly found in the language of women. And the other includes not only '*I* and *you'* but also other people which is not warm and personal.

Auxiliary verbs

Auxiliary verbs like '*is*, *are*, *has, have, will'* which denote the present and future tense were observed frequently in the notes of men. This reflects that men often predicted their future and focused on the present situation which dragged them to the decision of suicide. And verbs like *was*, *had*, *were*, *did*, *would*, *may*, *might*, *should*, *could* were observed frequently in the notes of women. Auxiliary verbs in past tense denote that women often recall their unpleasant memories. Auxiliary verbs which may be called tentative auxiliaries or subjunctive auxiliaries like *may*, *might*, *can*, *could* etc reflect the uncertainty and insecurity. The word '*could'* was equally used by both the genders. The use of the word '*could'* indicate a failure to cope with problems and situations, a feeling of being defeated, a state that has been suggested as characterizing suicidal individuals (*Gilbert & Allan, 1998*; *Williams, 1997*). And the negations like *couldn't*, *don't*, *shan't*, *didn't*, *can't* were used in higher percentage among women which reflects the negative thoughts. People who use words reflecting negative thoughts are most likely to see

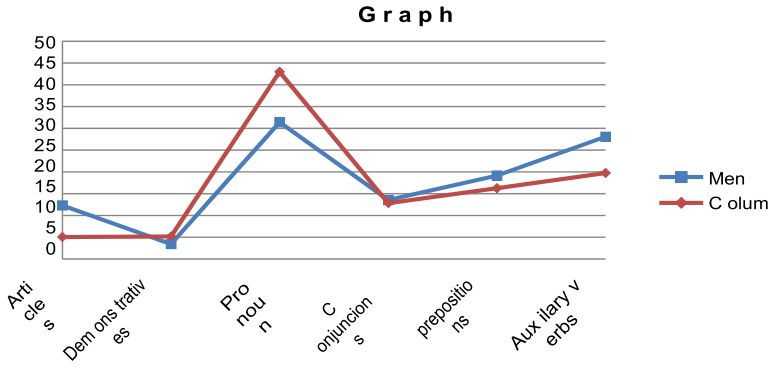
negative side of things and think of worst scenarios. Frequent use of negations reflects the stuck feelings of depression (*Nirmeen & Vijayan, 2020*).

Conjunctions

Inclusive and Exclusive conjunctions were identified frequently in the suicide notes. The use of inclusive conjunction *'and'* used by both men and women were equivalent. Women used more exclusive conjunctions such as *but, or, yet* than men. It was also observed that few women used the conjunction *'or'* as an inclusive. A study says, exclusive conjunctions are helpful in making distinctions. People use exclusion words when they are attempting to make a distinction. Women made use of these exclusives while expressing uncertainty, superior of self and while criticizing others.

Prepositions

Prepositions in the writings of both men and women had no much difference. Prepositions such as *of, for, from, in, on, to, than, after* etc were partially equivalent. The preposition *'with'* had a higher percentage in the writing of men and the prepositions *'about, out, without'* showed higher percentage in the writings of women. The use of *'with'* by men and the use of *'without'* by women reflect the negative thoughts of women. Though men were also going through similar emotions as women, they used lesser negatives than women did.



6. Conclusion

The study attempted to identify the function words present in suicide notes and to analyse if the use of function words varied among men and women in suicide note. 51.4% of the words in suicide notes were function words. Articles, prepositions and auxiliary verbs showed higher percentage in the writings of men and pronouns showed higher percentage in the writings of women. The use of demonstratives, conjunctions were almost similar among men and women. Comparing the results of the study with few studies related to the language of genders and suicide notes, it was observed that the language used by both men and women were different in suicide notes also. Though men and women go through the same trauma in life, their way of expressing and view of their world was different. It can be concluded that men and women use language differently even in the same context and mental state. Function words alone cannot decide the language style of suicide notes but it can corroborate with other linguistic features to determine the language style.

References

Gilbert, P., & Allan, S. (1998). The role of defeat and entrapment (arrested flight) in depression: An exploration of an evolutionary view. *Psychological Medicine*, *28*, 585–598.

Miller, G. (1995). The science of words. *Scientific American Library: New York*.

Nirmeen, N., & Vijayan, N. (2020). Voice of suicidal minds. Working Papers on Linguistics and Literature, Department of Linguistics, Bharathiar Univeristy: Coimbatore,
Vol 13, 301 – 307.

Pennebaker, J. W., Francis, M. E. & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC: 2001
.
Erlbaum, *Mayway*.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*, 1296-1312.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer. K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547-577

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558-568.

Simmons, R. A., Chambless, D. L., & Gordon, P. C. (2008). How do hostile and emotionally over involved relatives view relationships? What relatives' pronoun use tells us. *Family Process*, *47*, 405-419.

Tausczik Y. R., & Pennebaker J. W. (2010). The psychological meaning of words: LIWC and computarised text analysis methods, *Journal of language and social psychology*, *29(1)*, 24-54

Williams, M. (1997). Cry of pain. *Penguin: New York*.

https://www.merriam-webster.com/dictionary/suicide

The Language, Gender and Power in The Print Media and The Electronic Media in India– The Brief Case Study

Srijani Mondal

PhD, Eighth Semester, Department of Linguistics, Jawaharlal Nehru University, New Delhi – 110067, India.

ABSTRACT

Media is one of the most important means in communication. Media are diversified m edia technologies that are intended to reach a large audience through mass communication. The technology through which this communication takes place varies. The study reveals the interaction between language, gender and power in the print media and the electronic media in India and their influence on each other. With a boon coming in the world of print communication with the availability of printing press, telegraph, the publishing industry holds the print media as a rising sensation. Electronic media are media that use electronics or electromechanical energy for the end user [audience] to access the content. The paper focuses its study on the newspapers in the print media and films among the electronic media. Gender refers to the roles and responsibilities of men and women that are created in our families, our societies and our cultures. Gender roles and expectations are learned. They can change over time and they vary within and among cultures. The concept of gender is vital, because when applied to social analysis it reveals how women's subordination [or men's domination] is socially constructed. The media's role in affecting language in wider society. The media language reveals as a mirror of the wider society and culture. Language is biased, in that it is male dominated and actually favours men [Spender 200]. Words associated with males more often have positive connotations; they convey notions of power, prestige, and leadership. In looking at the many ways in which women are defined, Mills (1989) discovered categories of words that expose how women are stereotyped and consequently become more fragmented.

Keywords: Gender, Print media, Electronic media, Power, Language.

1. Introduction

One of the most important forms of media text that is the newspaper has an important role in human's life because it presents various local, national and international information and

events. One of the most obvious properties of media news, ignored or neglected in both traditional and more recent approaches to media reporting is that news reports, whether in the press or on TV constitute a particular type of discourse. Newspapers are a local, mass medium. Media carry messages to or from a targeted audience and can add meaning to these messages.

1.1. Language in the Print Media

The study of language in the media has much to offer to the different disciplines. It touches – linguistics, sociolinguistics, discourse analysis, semiotics, communication studies, sociology and social psychology. The kind of samples that are used to gather the data ranges from recordings of many months of broadcast news, or stories from a hundred different countries. The accessibility of media is a source of data for some language feature we want to study interest in the way the media use some language feature, also found in ordinary speech. Interesting is media's role in affecting language in wider society. Interest in what media language reveals as a mirror of the wider society and culture. Interest is in the way the language of the media affects attitudes and opinions in society through the way it presents people and issues. Focusing on the availability, media language is easier to collect than conversation. Most nations lack a true national daily press because of the difficulties of nationwide distribution. There is one feature of researching the media which still manages to surprise even experienced practitioners: the porcupine reaction of media personnel. If there is any one institutional disease to which the media of mass communication seem particularly subject, it is a nervous reaction to criticism, a reaction that puzzles us, for it is the media themselves which so vigorously defend principles guaranteeing the right to criticize. Sometimes the reaction can be extreme enough to jeopardize the conduct of publication of research. The likelihood of reaction should not drive media language researchers away, but rather

forewarn us on how to approach media organizations and present our research to them with care. Broadcasters operate from moment to moment with no assurance that they even have an audience. Although Feedback is not absent from the mass communication process, in few cases is the audience member on equal terms with the communicators. The audiences exercise their main influence on the media just through being the audience or by deciding to be someone else's audience. Mass communicators have only the haziest concept of what kind of people make up their audience. Isolation from the audience is a characteristic of mass communicators. Ironically, the more mass the medium, the greater the isolation. The inter-speaker or social dimensions have been correlated with differences in the measurable social characteristics (age, gender, social class) of a person – the speaker. Speakers are often primarily responding to their audience in the language they produce. But they also on occasion take more initiative and use language to redefine their relationship to their audience. If the style does not shift to suit the audience, the audience will shift to style that does suit. Variable editing rules thus function to shift the style of the input text closer to the style, which the station deems suitable for its kind of audience showing the way in which a number of divergent styles can be derived from a single text and how different media outlets achieve different linguistic styles.

1.2. Power and the Print Media

The power of the media is generally symbolic and persuasive, in the sense that the media primarily have the potential to control to some extent the minds of readers or viewers, but not directly their actions. Psychological and sociological evidence suggests that despite the pervasive Political Communication in action symbolic power of the media, the audience will generally retain a minimum of autonomy and independence, and engage more or less actively, instead of purely passively, in the use of the means of mass

communication. It also involves a study of the mental representations, including so-called social cognitions such as attitudes and ideologies, shared by groups of readers or viewers. Indeed, it may well be in the best interests of the elites that such public understanding be minimal. Language is a powerful tool especially when used consciously and deliberately. It can be used in many different ways in order to reinforce and manipulate a message. The print media is a linguistic medium, and one of the most important vehicles for the exchange of ideas. Thus it is important to be aware of the fact that the output of the print media is related to the need to make a profit; to have external relations with other industries, with financial institutions and with official agencies [Fowler, 1991:20]. Most newspapers are owned by people or companies which are commercial enterprises [Fowler, 1991:121], which means that printed media can be influenced by the owner or by the advertisers in its representation of events, thus potentially compromising its freedom of expression. Social power is a property of the relationship between groups, classes, or other social formations, or between persons as social members. Although we may speak of personal forms of power, this individual power is less relevant for our systematic account of the role of power in discourse as social interaction. At an elementary but fundamental level of analysis, social power relationships are characteristically manifested in interaction. Social power is usually indirect and operates through the minds of people, for instance by managing the necessary information or opinions they need to plan and execute their actions. The resources that socially enable the exercise of power, or the application of sanctions in case of noncompliance usually consist of socially valued, but unequally distributed attributes or possessions, such as wealth, position, rank, status, authority, knowledge, expertise, or privileges, or even mere membership in a dominant or majority group. Power is a form of social control if its basis consists of socially relevant resources. Besides this form of power distribution, which also

involves various forms of power sharing, there is the important dimension of resistance: Dominated groups and their members are seldom completely powerless. Under specific socioeconomic, historical, or cultural conditions, such groups may engage in various forms of resistance. The exercise and maintenance of social power presupposes an ideological framework. This framework, which consists of socially shared, interest-related fundamental cognitions of a group and its members, is mainly acquired, confirmed, or changed through communication and discourse. Power must be analyzed in relation to various forms of counter power or resistance by dominated groups [or by action groups that represent such groups], which also is a condition for the analysis of social and historical challenge and change.

1.3 Gender and the Print Media

Language is the cultural, not a biologically inherited, function, futility of interjectional and sound imitative theories of the origin of speech. Language is the human capacity of acquiring and using complex systems of communication and language is an instance/example of such a system. Estimated numbers of languages in the world vary from 6000 to 7000. Language, as a system of phonological, syntactic and logical structures and rules, is not inherently sexist or 'man-made' in Dale Spender's sense. Linguistic systems, however, serve as the basis for the production and interpretation of sets of related utterances – discourses or newspapers – which effect and sustain the different categorizations and positions of women and men. It is on the newspaper discourses and not on language in general or linguistic systems, that feminist analysis have to focus. It is one thing to formally implement non-sexist language change, that is, to recommend, encode and document new or alternative linguistic item. Questions of actual usage of non-sexist language and new progressive forms are something else again. Evidence is sparse but not non-existent. Quantitative analysis of change in

discourses would be impossible, given that a discourse cannot be 'seen', only its 'traces'. Use of 'gender-inclusive' language has been documented in speech and writing [Bate 1978, Cooper 1984, Hellinger and BuBmann 2001–02, Pauwels 1998].Cooper and his seminar group, for example, conducted a 'manual' study of 'generic he' and 'generic man' in US newspapers between 1971 and 1980, and reported a decline. One of the most influential figures as regards context, language use, and interpretation is Dell Hymes. Unlike Chomsky, who was interested in isolated sentences, Hymes was interested in the 'rules of speaking 'within a community, and the associated 'speech situations', 'speech events', 'speech acts, 'newspaper reports and 'communicative competence'.

1.4. The Language Power Gender and the Electronic Media

The art of cinema is an expression of human thought and emotions through the medium of communication. A film is created by photographing actual scenes with a motion picture camera, by photographing drawings or miniature models using traditional techniques. Films are cultural artefacts created by specific cultures. They reflect those cultures, and, in turn, affect them. Film is considered to be an important art form, a source of popular entertainment, and a powerful medium for educating or indoctrinating citizens. The visual basis of film gives it a universal power of communication. Some films have become popular worldwide attractions by using dubbing or subtitles to translate the dialog into the language of the viewer. Some have criticized the film industry's glorification of violence and its sexist treatment of women. Language is the basic means of communication. There exist gender differences in talk to discourse and the social/textual/linguistic construction of gender. It highlights the interdisciplinary nature of gender and language, different theoretical approaches conversation analysis, feminist, poststructuralist discourse analysis and discursive psychology. Since the early 1970s the study on Gender and language

developed creating an expanding paradigm which draws on a wide range of disciplines and theoretical and methodological approaches. Several studies identified sex-exclusive linguistic features that are features used only by women or only by men, within a given speech community. Sex-Exclusive uses of language occur rarely and contrast with the common sex-preferential uses. These refer to differential tendencies in which women and men tend to talk differently from each other in a given context. Sex-Preferential phonetic, intonation, lexical, syntactic and wider interactional tendencies have been identified. Different phases of Feminism can be seen as the driving force behind the male dominance and cultural difference approaches to the study of gender and talk. Feminism in general and feminist theory in particular also drove the subsequent critique of dominance and difference as a single approach. Gender relations carry the potential not just for differentiation and differential empowerment, but also dominance, disadvantage, and economic, educational and political inequality. Some studies, however, have taken a different approach by looking not so much at power in mixed-sex interactions as at how same-sex groups produce certain types of interaction.

1.5. Research Methods

The study is based on the deductive approach. The following are some steps involved in conducting studies, following by specific aspects associated with analyzing the newspaper and the films –

1.5.1. Selecting a certain groups, worlds, topics, issues, or phenomenon, and then find different representations of the topic/phenomenon in a film or TV.

1.5.2. Noting patterns in these representations in terms of similarities portrayals/images instances of stereotyping or essentializing categories.

1.5.3. Noting value assumptions in terms of who has power, which solves problems, how problems are solved.

1.6. Data Collections Findings and Discussions

Word forms that reflect newspaper discourse in general such as Is, Has, Who and Says. Word forms that belong to what McLachlan and Reid (1994, 104) refer to as part of the 'circumtext' of a given article, such as 'author' [name of the author], 'paper', 'title', 'Guardian' and 'Times' [newspaper title], 'London' and 'England' [place of publication], 'date' [of publication], 'section' [rubric], 'page', 'copyright', as well as terms deriving from set phrases such as 'newspapers', 'limited', 'all', 'rights' and 'reserved'. Terms relating to recent technological innovations such as the World Wide Web which appeared considerably more frequently in LANGCORP than in the BNC due to the difference in construction dates. The BNC was completed in 1993, which predated the upsurge in new media phenomena and terminology, such as 'WWW', 'Google', and '.com', from the mid-1990s onwards. Word forms which only occurred [in disproportionately high numbers] in one single article or type of article that was considered statistically and conceptually insignificant [e.g. lists of election results after the general elections in Britain in May 2005 and the United States in November 2004, personal columns]. Sinclair refers to texts that differ radically from the others in a corpus and, hence, threaten to distort the data for the sake of overly objective coverage, as rogue texts. To guarantee accountability, he advocates the documentation of such exclusion procedures.

1.6.1. New York Times Headline – 163,000 jobs added in july :

Unemployment rate rose to 8.3%

1.6.2. Fox News Headline – The Jobless rate rises to 8.3

percent, hiring picks up but still falling short.

☐ The same news headline is presented in different newspapers differently –

1. In the first headline, the tense is referred to in the past, where is in the second sentence is in the present continuous tense
2. The sentence structures of both the sentences differ from each other. There are uses of complex phrases that organize the whole sentence.
3. The meaning of both the sentences is also slightly different. The first sentence provide the meaning of the action already being completed in the past but in the second sentence it provides the sense of the action still being continued. In the second sentence, the tonal variation is included. The sentence is said with a falling-rising tone.
4. Sentence structures and use of articles differ.

The Grammar of Headlines –

Parts of speech, Articles, Auxiliary verbs, Possessive adjectives, Verb forms, Present simple, Past participles and Infinitives.

If stereotypical perceptions of gender traits persist, it jeopardizes the development of human potential and capabilities, given that people will be confined to develop within the straight- jackets of socially defined roles and gender-appropriate behavioral patterns. This also reinforces the segregation and rigidity of gender roles and thus social statuses.

1.7. Conclusion

The manifestation of ideal body image as a predominant image portrayed in the media.The investigation of the nature of the portrayal of a person appearing in the visual media and check whether his age is a factor for his likelihood. Showing the dominance of the issue of identity crisis in the media texts. The reflection of power in the cultural and linguistic identities through the language in the media.

References

L. Miriam Rachel. (2007). Research into the Representation of Gender and Body Image in the Press: A Visual and Textual Analysis Examining the Presentation and Portrayal of Gender and Body Image in British National Daily Newspapers. pp.10, 30-33.

E. Penelope. (2005). Linguistic Variation as a Social Practice – The Linguistic Construction of Identity in Belten High.

K. Scott, S. Prem.N T. Ellen.E. 2004. Exploring Language Effects in Ethnic Advertising: A Sociolinguistic Perspective – The Journal of Consumer Research, Vol.20, pp-575-585

The Language Rights of Persons Belonging to National Minorities Under the Framework Convention. (2012). Thematic Commentary No.3.

P. Kirtikumar.R. (2011). Sociolinguistic aspects of media language with special reference to internet. Vol-1. Issue-5.

E. Penelope. (2005). Linguistic Variation as a Social Practice – The Linguistic Construction of Identity in Belten High.

E. Penelope. C. Sally.Mc -Ginnet. (1992). Communities of practice: Where language, gender, and power all live.

Sentiment Analysis of Persian children's bedtime stories

Ms. Sanaz Montazer[1] and Dr. Srikanth Muddam[2]

(1) Department of Linguistics Osmania University
(2) Department of Linguistics, Osmania University

ABSTRACT:

In this century, data have been counted as the major asset to different types of companies around the world. We are living in an era where text information is growing rapidly. The webs, blogs, emails, social networks provide millions of data in seconds and all of these resources can be a useful base for extracting knowledge. As the number of reviews grows, it is necessary to develop an efficient Sentiment analysis method that is capable of extracting product aspects and determining the sentiments for these features. The majority of researches have been done on English and few on the Persian language and its challenges. The following paper discusses the issues related to the sentiment in Persian children's bedtime stories and develops a tool to recognize the sentiment in the stories automatically.

Keywords: Sentiment analysis; Persian language; opinion mining

1. Introduction

Persian is spoken by more than a hundred million speakers throughout the globe and it is the official language of Iran, Tajikistan, and Afghanistan. Persian is a challenging language for Sentiment analysis and the resources and tools for Persian text processing are quite rare. Sentiment analysis is an automated process that identifies positive, negative, and neutral opinions from the text by using Artificial Intelligence. Sentiment analysis is also known as opinion mining explains the analysis of the feelings (i.e. attitudes, emotions, and opinions) by using NLP. The present study focuses on Sentiment analysis of Persian children's bedtime stories.

In this age, data is a valuable property of enterprises around the world. Websites allow customers to share their ideas and feedback about the products and services. All of these resources can be used for extracting knowledge. It has called the

power of text analysis. The information on social media can be used to improve products and services. For example, if someone wants to buy a TV, the buyer will look at the quality of screen resolution, the quality of the sound, and the connection boards and also compare it with other brands. The automatic analysis of such information, however, requires a thorough understanding of natural language. [1] In other words, the internet has become the main measure of connection nowadays and without it, the connection is lost. Technology has made life easier, besides customers are free to give their feedback online, which is viable for everyone. The producer and seller can find out about how happy their client is with their product by their written ideas on the website. Customers' reviews play an important role in the selling rate of a product. As the number of reviews grows, it is necessary to develop an efficient Sentiment analysis model that is capable of extracting product aspects and determining the sentiments for these aspects. [1]

Apart from the English language, Sentiment analysis has been done in the Persian language as well. This language is an Indo-European language and has a strong potential for deep research in any aspect of NLP, especially Sentiment analysis. The majority of researches have been done on English Sentiment analysis and a few on Persian sentiment analysis. Persian is a challenging language for Sentiment analysis and the resources and tools for Persian text processing are quite rare. [2] In this paper, firstly, we will explain the Sentiment analysis in the English language, its applications, and data resources and then we discussed the Persian language and its emulate. Ongoing work includes the solutions to the problems.

2. Sentiment Analysis and its Applications

As mentioned above, Sentiment analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text, to find out whether the writer's view towards a particular subject, product, and service is positive, negative, or

neutral. It is widely used for getting insights from social media comments, surveys, responses, and products reviews to help a business to understand the social sentiment of their service, brand, or product while monitoring online conversations. Currently, all big companies such as Amazon, Google, Face book, most of the applications like Twitter and Instagram, popular brands, online stores, hotels, and airlines are using Sentiment analysis to improvise on their products and services. This analysis applies a mix of statics, natural language processing, and machine learning to identify and extract subjective information from a text file. Sentiment analysis is a complex process that involves five different steps to analyse sentiment data. These steps are:

Data collection: the first step of Sentiment analysis consists of collecting data from user generated content contained in blogs, forums, social networks. These data are disorganized, expressed in different ways by using different vocabularies, slang, the context of writing, etc. Manual analysis is almost impossible. Therefore, text analytics and natural language processing are used to extract and classify

Text preparation: consists of cleaning the extracted data before analysis. Non-textual contents and contents that are irrelevant for the analysis are identified and eliminated.

Sentiment detection: Sentiment detection is often also called Sentiment analysis, Sentiment classification, or Opinion analysis/mining. The extracted sentences of the reviews and opinions are examined. Sentences with subjective expressions (opinions, beliefs, and views) are retained and sentences with objective communication (facts, factual information) are discarded. It automatically identifies emotions in textual data and enhances business intelligence applications. Generally, Sentiment detection considers a machine learning algorithm divided into the supervised method and unsupervised (statistically) one. The lexicon-based feature sentiment detection

approach is based on the co-occurrence of a sentiment lexicon that might include a list of emotional orientations. Usually, a sentiment lexicon consisted of a set of words in a special language, containing emotion weight and annotated with some orientations or particular rating score. In this context, when the lexicon is richer in words besides using a multi-orientation emotion spectrum, compared to an easy binary polarity method, can enrich the classification accuracy in sentiment detection. [3]

Sentiment classification: in this step, subjective sentences are classified in positive, negative, good, bad; like-dislike, but classification can be made by using multiple points.

Presentation of output: the main objective of Sentiment analysis is to convert unstructured text into meaningful information. When the analysis is finished, the test results are displayed on graphs like pie charts, bar charts, and line graphs. Also, time can be analysed and can be graphically displayed constructing a sentiment timeline with the chosen value (frequency, percentages, and averages) over time.

2.1. Approaches to sentiment analysis [4]

There are various methods or approaches in sentiment analysis research. Following are some of them.

1-Machine learning approach: Machine learning is a data analytics technique that tells computers to do what comes naturally to humans and animals: learn from experience. Machine learning algorithms use computational methods to "learn" data directly from information without depending on a predetermined equation as a model.

2-Lexicon- based methods: This is the basic approach to automatic sentiment categorization. The lexicon may be developed manually or semi-automatically, to organize the affective lexicon, it is necessary to have a model for the effective knowledge representing emotions, moods, attitudes, and traits.

3- Natural Language Processing: It is a field in machine learning with the ability of a computer to understand, analyse, manipulate, and potentially generate human language.

4-Decision Tree Learning: Decision Trees are a non-parametric supervised learning method used for both classification and regression tasks. The aim is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

5- Techniques of Information theory and Coding: It is the scientific study of the quantification, storage, and communication of digital information.

6- Semantic Orientation Approach: It is an opinion on a feature of states whether the opinion is positive, negative, or neutral.

7- Hybrid Approaches: A Method to improve the expected output of Semi-Structured Sequential Data. The hybrid approach of sentiment analysis exploits both statistical methods and knowledge-based methods for polarity detection. It inherits high accuracy from the machine learning (statistical methods) and stability from the lexicon-based approach

Applied method: Among the above, the Lexicon-based methods were chosen for this research, because it has a good scope controlled linguistic analysis.

3.1. Lexicon-based methods

This is the basic approach to automatic sentiment categorization. This technique calculates the sentiment orientations of the whole document or set of sentence(s) from semantic orientation of lexicons. Semantic orientation can be positive, negative, or neutral. The lexicon may be developed manually or semi-automatically, to organize the affective lexicon, it is necessary to have a model for the effective knowledge representing emotions, moods, attitudes, and traits.

Using this framework, we should be able to identify a large number of affective concepts, organize them in a hierarchical structure and connect them with lexicon. The drawback of using a lexicon is that words can have multiple meanings and senses, and those which are common in one domain may not be similar in another. Furthermore, words that are not generally considered sentiment-bearing can imply sentiments in specific contexts. Given a sufficiently large training corpus, a machine learning model is expected to outperform a lexicon-based model. [5]

4. Data and Resources

Review sites

Blogs

Forums

Dataset

s

Micro-blogging

Google Play Android Application Store

Social media such as Twitter, Facebook, Instagram, etc.

[6] The following is the sample children's bedtime

story of

Farsi along with the translation.

شب وحشتناک آقا گوسفنده و خانم گوسفنده

آن روز، از صبح تا شب، آقاگوسفنده و خانم گوسفنده و بزچالاقه تنها از آغل بیرون نرفتند. در آسمان باز شده بود و باران تند و تیز می بارید. آقاگوسفنده و خانم گوسفنده و بزچالاقه با عجله خشکی که توی آغل داشتند، شکمشان را

.سیر کردند و شب که شد، هر کدام گوشه ای خوابیدند

نصفه های شب بود که همه شان از خواب پریدند و دیدند آب باران از گوش های راه پیدا کرده و کف آغل را خیس

کرده است. خوابیدن روی زمین گل آلود و نمناک امکان نداشت.

خانم گوسفنده به آقاگوسفنده گفت: «چه کنیم حالا؟ کجا بخوابیم؟»

آقاگوسفنده گفت: «نمی دانم. راه چاره ای به نظرم نمی رسد. یا باید روی همین زمین گل آلوده بخوابیم یا تا صبح بیدار

بمانیم و ببینیم صبح که شد، چه کسی به دادمان می رسد.

خانم گوسفنده گفت: «روی گل و لای بخوابیم یا تا صبح بیدار بمانیم؟ این هم شد راه چاره؟ نمی دانم تو با این عقل و

هوشت، چرا مشکل گشای مردم دنیا نشده ای! بلند شو مرد! کاری بکن.

آقاگوسفنده گفت: «چه کار کنم؟ تو که دست تمام عاقلان دنیا را از پشت بسته ای، بگو که توی این تاریکی شب، چه می

تو انم بکنم؟ «

خانم گوسفنده گفت: «چه می دانم؟ برو ببین آب از کجا به آغل راه پیدا کرده. راه ورود آب را ببند.

آغلمان را آب برداشت آقا! عجب! من که نمی توانم تا صبح سر چهار تا دست و پاهام بیدار بمانم. «

آقاگوسفنده گفت: «بستن راه ورود آب، بیل می خواهد، کلنگ می خواهد، زور بازو می خواهد که هیچ کدامش را ما

گوسفندها نداریم. «

خانم گوسفنده گفت: «این را که خودم هم می دانم که بیل و کلنگ نداریم. اینکه دیگر گفتن ندارد. «

آقاگوسفنده که از غرغر کردن خانم گوسفنده عصبانی شده بود، صدایش را بلند کرد و گفت: «مثل اینکه تو هم جز مسخره کردن من، کار دیگری بلد نیستی. فکر می کنی خودت خیلی می فهمی؟ این راه حل که باید راه ورود آب بسته

» .شود، به فکر هر بره ای هم می رسد.

خانم گوسفنده که دید آقا گوسفنده هم حق دارد و هم خیلی از دستش عصبانی شده، دیگر حرفی نزد و بغض کرد و حالا

.گریه نکن، کی گریه کن.

بزچلاقه با اینکه کنار آقاگوسفنده و خانم گوسفنده ایستاده بود، حرفهای آنها را نمی شنید. از همان اول هم فهمیده بود که نه امکانات و قدرت بستن راه ورود آب را دارد و نه می تواند شب تا صبح روی چهار لنگه دست و پایش بایستد و بیدار بماند. با این که خوابش می آمد، فکر کرد و فکر کرد. ناگهان، انگار که به راه حل درستی رسیده باشد، جستی زد و به طرف کپه علوفه های خشک رفت. آقاگوسفنده و خانم گوسفنده وقتی حرکت یکباره بزچلاقه را دیدند، از گریه

ودعوا دست برداشتند، ببینند دوستشان چه می کند. بزچلاقه با سر به جان بسته های علوفه های خشک افتاد. به آنها شاخ می زد و سعی می کرد جا به جایشان کند.

آقاگوسفنده و خانم گوسفنده چشم دوخته بودند به بزچلاقه و نمی دانستند می خواهد چه کار کند. بزچلاقه آنقدر سرش را به کپه های علوفه خشک کوبید که توانست آنها را کنار هم قرار دهد و با کنار هم گذاشتن آنها، تختخواب بزرگی که شب وحشتناک آقا گوسفنده خیلی هم از زمین نمناک فاصله داشت، بسازد. و خانم گوسفنده کارش که تمام شد، بدون اینکه به دوستانش چیزی بگوید، به بالای کپه علوفه های خشک جست خواب بعد از خستگی دلنشین بود. آقاگوسفنده و خانم گوسفنده نگاهی به یکدیگر انداختند و گفتند که چرا این راه حل به فکر ما دو تا نرسید . بعد، از تختخواب علوف های بزچلاقه بالا رفتند و روی آن دراز کشیدندکه بخوابند. خانم گوسفنده گفت: «اگر بزچلاقه نبود، ما چه کار می

کردیم.

Terrible night of Mr. Sheep and Mrs. Sheep

On this day, from morning till night, Mr. Sheep and Mrs. Sheep and Goats could not get out of the cage. It was open in the sky, And it rains a little. Mr. Sheep and Mrs. Sheep and Goats fed their bellies with the dry fodder they had in the cage, and that night, they each slept in a corner. It was midnight when all three of them woke up and saw rainwater come from the ears and wet the floor of the cage. It was impossible to sleep on muddy, damp ground. "What shall we do now?" Said Mrs. Sheep to Mr. Sheep. Where do we sleep? "I do not know," said Agha Gosfandeh. I do not think there is a solution. "I do not know," said Agha Gosfandeh. I do not think there is a solution, We either have to sleep on this muddy ground or stay up until morning, and let's see who comes to us in the morning.

"Shall we sleep on the mud or stay up until morning?" Said the sheep. Was this also the solution? I do not know why you have not solved the problems of the people of the world with this wisdom and intelligence! Get up, man! Do something. "What should I do?" You who have tied the hands of all the wise men of the world from behind, Tell me, what can I do in this night? "What do I know?" Said Mrs. Sheep, go and see where the water got to the cage, Block the entry of water. The water was taken away from us, sir! move one! I cannot stay up until four in the morning with my arms and legs. "Closing the water intake requires a shovel, a pickaxe," said Agha Gosfandeh, which requires arm strength, none of which we have. "I know I do not have a shovel or a pickaxe," said the sheep. He has nothing more to say. Mr. Gosfandeh, who was angry with her muttering raised his voice and said, "As if you, too, except you do not know how to make fun of me. Do you think that you can handle it? This solution, which should close the way for water to enter, comes to the mind of every lamb. Ms. Gosfandeh, who saw that Mr. Gosfandeh was both right and very angry with her, said nothing. And she started crying continuously.

The goat did not listen to them, even though he was standing next to Mr. Sheep and Mrs. Sheep. He knew from the beginning that he does not have the facilities and power to block the entry of water nor can he stand on his legs and stay awake all night long. Even though he was asleep, he thought and thought. Suddenly, he looked as if he had found the right solution. He jumped and went to the pile of dry fodder, Mr. and Mrs. Sheep saw the goat moving suddenly, they stopped crying and fighting to see what their friend was doing. The goat fell head over heels in the dry fodder package. He was horning at them and trying to move them. Mr. and Mrs. Sheep were staring at the goat and did not know what he wanted to do. The goat bangs his head so hard on the piles of dry fodder that he could put them together and by arranging them, he made a big bed that Mr. Sheep's terrible night was too far from the damp ground, and Mrs. Sheep, when she was done, without saying a word to her friends, went to the top of the pile of dry fodder to sleep as she got tired. Mr. and Mrs. Gosfandeh looked at each other and said why this solution did not come to our minds. Then, goats got out of bed and lay on it to sleep. "If we were not goats, what would we do?" Said the sheep.

5. Data and Resources of this Research

The corpus contains seventeen children's stories which have been collected from various sources including short stories. Based on this corpus some of the lexical items related to happy and sad were collected. These words were listed in a file without any morphological analysis, i.e. word as it is. This lexicon contains nouns, verbs, adjectives, etc. which have happy or sad concepts in them. This lexical database will be searched in the input data and mapped words are sorted out based on the number of occurrences. The frequency of these words will be calculated and the sentiment of the text is identified.

The basic algorithm for analysis of the stories has been done in Python and is as follow:

The Algorithm of the

System. The input text is

called.

This input is compared with the database.

Then it extracts the words which have the meaning related to happy or sad.

These extracted lexical items will be calculated.

Based on the number of occurrences the engine decides whether an input text contains happy or sad sentiment.

Using the above algorithm, a tool was developed in a python programming language. Following is the sample python code:

```
i=0 a=0
while
True:
#for i in range (0,3,1):
try:
ages1= re.findall (teldicss[i],
indaa) print (ages1) print
(len (ages1)) a= a+len
(ages1) i=i+1 except for
Type Error:
                    brea
k except Index
Error:
 break print(a)
 ratio =
a/indaa1*100 print
(ratio)
```

6. Conclusions

To summarise, nowadays technology plays an important role in every aspect of the business.

Everyone checks the review of the product or services before going for it. Due to the huge amount of generated data daily on e-commerce and social media, companies need to apply

a method to get their client's feedback in order to improve their product or services. Sentiment analysis is the fastest way of understanding the customer's feedback about a product and a service in order to improve the quality. It is a field of natural language processing, which attempts to identify and extract subjective information from textual resources. Generally speaking, SA aims at extracting topics that are discussed in a review or comment and then determining the attitude of the writer towards the extracted topic. Due to its efficiency and rapid growth in textual resources, it has a wide range of applications including customer review monitoring, political affiliation extraction, and market movement prediction. In other words, applications of sentiment analysis are becoming more popular every day. In order to extract subjective information, most researchers have focused on automatically determining the polarity (e.g. positive or negative) of the terms in the given document. Every day, a vast number of Persian internet users around the globe share their ideas and feelings about a product, service, or political issue. Due to the challenges of Persian Sentiment Analysis, there are good scopes for researchers in this area. Overall, sentiment analysis is the process of analysing the feeling of the buyer of a product or the reader of a text towards it. The challenges are about the plurality of the lexicon and the written form in the Persian language.

The output of this tool has been given around 60% efficiency. Since this percentage is not sufficient to use the tool, it needs to be improved further by testing it on a bigger scale. More output need to be collected; also, Lexical semantics has to be examined extensively. Moreover, sophisticated algorithms have to be developed. The Persian language has to be compared in terms of lexical and conceptual similarity and divergence if any is available.

References

Dashtipour, Kia, et.al. "A Hybrid Persian Sentiment Analysis Framework: Integrating Dependency Grammar Based Rules and Deep Neural Networks", Cognitive Computation, 10(6) (2018) 1152-1166.

Bagheri, Ayoub, Saraee, Mohammad and de Jong, Francisca," Care more about customers: Unsupervised domain-independent aspect detection for sentiment analysis of customer reviews" Knowledge-Based Systems Volume 52, November 2013, Pages 201-213.

Faiz Alasmari, Salma and Dahab, Mohamed, Sentiment Detection, Recognition and Aspect

Identification, International Journal of Computer Applications (0975-8887), Volume 177No.2, November 2017

Supriya B. Moralwar, Sachin N. Deshmukh "Different Approaches of Sentiment Analysis" International Journal of Computer Sciences and Engineering Open Access Review Paper Volume-3, Issue

Khoo, Christopher S.G.; John khan, Sathik Basha, Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons. Journal of Information Science, in press.

Heydari, Mohammad. "Sentiment Analysis Challenges in the Persian Language." 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT),2016, pp.628-632.

Rhetorical Devices in 'Antarani Vasantam'

Dr. K. Susheel Kumar

Assistant Professor, Department of Linguistics, Osmania University, Hyderabad

ABSTRACT

In literature the writers or the speakers use rhetorical devices considering its precedence to lend beauty, variety, vividness, force, and power to the language. These devices are the key elements which takes the pivotal role in successfully producing a desired or intended result in a literary discourse. The present paper attempts a study of 'Simili' and 'Personification' in the Telugu novel 'Antarani Vasantam'(Untouchable Spring) written by G.Kalyana Rao. All the examples in this paper are provided in Telugu language to focus on the style of the author in the novel.

Keywords: Simili, Personification, Style, Vividness

Introduction

The word rhetoric has been derived from Latin 'rhetorica' or 'rhetorice', the word originally taken from Greek 'rhetorike' meaning the technical art of rhetor an orator or the public speaker. Aristotle, in his first book Rhetoric (322-320B.C), has defined rhetorical discourse as the art of discovering all the available means of persuasion in any given case. Rhetorical devices serve as a tool to lend beauty, variety, vividness, force, and power to the language and literature. There are so many rhetorical devices presented by the author in the novel, however in this paper the most commonly used rhetorical devices like Simili and Personification used in 'Antarani Vasantam' are discussed in the following analysis.

Simile

Simile is the root notion of tropes. It is a comparison of two things derived from likeness. The comparison is made explicit by means of such constructional elements as like, as, so, etc. Simile and metaphor differ only in degree of stylistic refinement. In the words of Leech (1969, p.156) Simile is an

overt and Metaphor a covert comparison. This means that for each metaphor, we can devise a roughly corresponding simile, by writing out tenor and vehicle side by side, and indicating (by "like" or some other formal indicator) the similarity between them. The ship ploughs the waves', a stock classroom metaphor, may be translated into a simile as follows: "The ship goes through the waves like a plough ploughing the land". The examples illustrates that a metaphor can be converted into a simile and can again be compressed into a metaphor. The simile, in which a comparison is made directly between two objects thus belongs, in the words of Read (1963, p.23), to an earlier stage of literary expression. It is a correspondence, often persuaded for its own sake. But a metaphor is the swift illuminating of an equivalence. Two images, or an idea and an image stand equal and opposite; clash together and respond significantly, surprising the reader with a sudden light.

- అచ ుş అలYˊ oుoపç ఎ>ˇgల aిటśలYˊ అరwడs. (10)
- జకపకమ ుంæ కo*Y* లaాల 6ాo*నటśc వç ుంటc ుంo*. (21)
- మ ుంట చ ుటśe వలయ ుంల కeరcşనg జన ుం. (23)
- అద ృశ్çaˇÛన aˇలcతvరc పరcచ ుsింటe మc ుంద ుకc వ3Šª ుంo*. eస్టీæటªwెట. మధçలˊద ు. ఊo* బయట మ ుంటల వç ుంo*. (23)
- శబTల మధç aˇలcతvర.c పగcలcతvనg aˊఘల మధç పరుగులుæకో aˇరcపçల. వç>ge. (23)
- oేaాలయల.ª aˊo*ఐª న ృత్ç ుంల అగ.వoşాలc oచ ుsినg ఉరcమcల ఆæ>çస ుం. (25)
- 9ళ్gలªsª wేపల జo*6ª æన9 పటścsమ9 అయ్ç9 అ&ªYˊaాడs. (45)
- 9జ9sª ృ% మªన ుంYా లˊద ు. 9కటªYా లˊద ు చ ుక4ల., oా%. æణcగcరcల., షYY*>ిప& ే

- జ¤లcకc ఎద◌ు౦*ో◌ే ◌ృ% తనsశస◌ం ఙ¤న◌ు స◌ృఊ ◌ిం◌చ◌ుsశనg wే◌పల.¤,పఙ 9టªజ¤లc స◌ంY*త◌ంల¤
- ఎప9టªల¤Y´ వç◌ంo*. (54)

- ఎ◌ండలª a˘o*>´ పచş9 బ◌ంY ర◌ం %గల¤ వç◌ంo*. కళvg మ◌ండsతvనg శ%ఙలవç>¤ge. (105)
- w¤ల¤మ◌ంo* మ¤ల¤మ¤o*గల.¤ &వయç s˘Ó.సక వçడయ¤çడs. (131)
- జ¤z6eo˘&ª గc>´gన◌ుగcల¤ వç>¤gడs. (148)
- sశత sశయట¤9sª >◌ిద్a˘ûన ప◌ంట కతక vల.¤ వç◌ంట¤e. (152)
- గ◌ంగల¤ ఉ9sªన క9gళ.v వç>¤ge. a˘>˘gల.¤ కco*>◌ిన ఆన◌ంo¤లc వ్gఙe. (160)
- చ◌ంద◌ు ఙఞª a˘>˘g%g చ◌ు>´ 9త.src&ª హ◌ృదయ◌ం ల¤◌ంటª జక¤పక◌ం. (163)
- e3 4o*e´తv యeo¤కc ఙళgకc ఙే&¤ ల´ద>¤gడs. (209)
- o ళg మ¤టలc గco*a`టకªన బ¤ఙ¤ల.¤ వç◌ం&ేఙ. (217)
- మడవఙద s లca`టకª 6 రa`s˘%క న స◌ుభద. ¾దలc 9ళgలª wే◌పల¤ ఙఙఙక%>◌ి6ª eన రe9o¤s. (228)

Personification

Personification or "actualization" is another figure of speech related to metaphor. It is a symbolic endowment of inanimate objects or qualities with personality or human attributes. "In ordinary language (the kettle sings) and in literature, inanimate objects are sometimes talked of as they were alive" Turner (1973, p.126). Like metaphor, as Read (1963, p.31) explains, the terms of comparison are suppressed, or identified with the object to which it is compared. Personification also has wider uses than to enliven abstract ideas. It usually incendies human qualities to non-human subjects to give a sense of life to it. Though this is entirely appropriate to poetic expression, a sustained process of it is also

realized to give vitality to the story of the Novel, as seen in those examples.

బª మqలc ఏడs3Š౽ℯ. బª మqలc నవçgæ¤ℯ. (19)

అక4డ ఆమంట చుటకe పళయం నంŠ*3ªŠ ంo*. (23)

సంంoె a˘లcగcలc 9వgo* నవçgలc నవçgతv>¤gℯ. (120)

ఆ చంక4ల.ª ఏ చంక4 బ¤ల ంow¤9sª పo*చయం

ఏసంనం ¾టk¾దట షuే>ింo*.

a౽రమq a˘తvకcæ౽ంo*. (136)

జŠ¤పకం ఏడs3Šª ంo*.

జŠ¤పకం గcంo&ె%g aిండsæ౽ంo*.

జŠ¤పకం ఆ 9డ aˇo*. ఆక.ందన9

æ993Šª ంoరeతvకc ℯప9టªsª

స9షkంoY౽ æ99ంowే నవçgలc. జ˘>ీ7

నవçgలc.

6ª నgంoY* aిటk నవçgలc (204)

తన 9డ. ℯమ¤qనంℯ´లc భeææ౽ మ¤ట.¤&ª వçంంట¤డs.
9జంoY౽>´ అoే wేa౽డs. భeææ౽ మ¤ట.¤&¤డs. ఏo* wెaి9
వçంటcంo*. భeæ. బª ల´డs రహ3౽çలc wెaి9
వçంటcంo*. 9తçం తన9 నమcqs౽9 తన9 3ªంతం
wేసంs౽ల´9 తన 9డ ల కథలc wెaి9 వçంటcంo*. కల.
¤ంo Y*ంజల బతvకcలc wెaివçంటcంo*. అంæే>¤. తన9
తన 9డ లకc దక4కcం&¤ wెరలª వçం9న భe3౽gమcల
మోసం కe&¤ wెaి9 వçంటcంo*.

Conclusion

From the above discussion it is clear that any author or a speaker's use of rhetoric in a literary discourse is not just the subject or the situation to consider how thoughts are expressed in a literary discourse. However rhetoric itself shows its impact on the author or the speaker's think in different ways and

modifies their thoughts in any shape or form in a discourse, to attain this goal the author or a speaker needs to have a very serious consciousness. The rhetorical devices used in a literary discourse helps us in deconstructing meaning in various discourses. The rhetorical devices like the 'Simili' and 'Personification' are consciously used in the novel 'Antarani Vasantam' by the author to bring beauty to the novel and also to impress the readers with his style.

References

Leech G. N. (1969). 'A Linguistic Guide to English Poetry', London & Harlow: Green & Company., Ltd. p.156.

Read H. (1963). English Prose Style. London: G. Bell & Sons, Ltd. p.23, 31.

Turner G.W. (1973). Stylistics. Harmondsworth: Penguin Books. p.126

Kalyana Rao G. (2000). 'Antarani Vasantam', Vijayawada. Published by Viplava Rachayitala Sangaham, Andhra Pradesh.

Kennedy.G. (1980). Classical Rhetoric and its Christian and secular Tradition from Ancient to Modern Times, Chapel Hill: University of North Carolina Press, p. 25-26.

Simpson. P. (2004). Stylistics, New York: Routledge, 2004, p.27.

Sonority Profiles of Initial Consonant Clusters of Telugu Words

Vidyamrutha Imandi[1] and C.S. Swathi[2]

(1) Research Scholar, Department of Linguistics, Osmania University
(2) Asst. Professor, Department of Linguistics, Osmania University

ABSTRACT

Sonority of a sound produced has been defined as acoustic intensity, relative loudness, carrying power, degree of openness, etc. by many linguists. Its hierarchy is often debated to be universal cross-linguistically, but a few linguists (Sherwin 1999, Jane et al. 2007) claim otherwise that require language specific sonority scale. The aim of this study was to analyze the application of Sonority Sequencing Principle (SSP, having 5-point scale) as proposed by Clements (1990) to Telugu Language, or is there a need for Language Specific Scale. Analyzing the sonority profiles of 50 words with initial consonant clusters taken from Telugu Academy publication's Telugu-English dictionary, it was observed to have six different onset cluster patterns (ON, OL, OG, NG, OO and GG). The findings of this descriptive study showed that majority of the words had rising sonority profiles (ON, OL, OG and NG) as per Clements proposal, however, some of them showed plateaus (OO and GG). Words with sonority plateaus were then analyzed using a Language Specific Sonority Hierarchy Scale (LSSHS, having 13-point scale). It was observed that using this scale, more number of words were adhering to SSP. For example the sonority profile of /kṣa:mam/ - 'famine, scarcity' has a sonority value of (1-1-5-2-5-2), with the initial values (1-1) creating a sonority plateau in the primary onset cluster of the word when the scale is used as proposed by Clements. But when considered from the LSSHS and by dividing the obstruents into stop and fricative, the values change to (2-5-13- 10-13-10) which do not violate the SSP. Thus, LSSHS has been applied to sonority plateaus resulting in minimal violations like sonority reversals and plateaus.

Keywords: Sonority, Sonority Sequencing Principle, Telugu, Onset Consonant cluster

1. Introduction

Telugu, one of the languages from the Dravidian family, is the language of focus of this paper. With a wide range of L1 and L2 speakers, this language is slowly reaching a billion mark (2011 census). Over the long process of development, this

language has gone through a lot of changes due to the influence of various factors, and has borrowed a lot of words from Sanskrit, Prakrit, Perso-Arabic and English languages. According to the paper 'Illustrations of the IPA – Telugu' written by Peri Bhaskararao and Arpita Ray (2016), our phonemic system had been upgraded with new set of sounds due to borrowing. Sounds that are absent in native phonemic system like aspirates /pʰ, bʱ, tʰ, dʱ, ʈʰ, ɖʱ, ʧʰ, ʤʱ, kʰ, gʱ/ and fricatives like
/ʃ, ʂ, h/ are borrowed from Sanskrit. Just like the phonemes above, /f/ is borrowed from English and Perso-Arabic phoneme inventory. Sounds that were not directly borrowed were rendered to a closer sound of the Telugu phoneme inventory. For example, the phoneme /ɵ/ in English, which is taken as /tʰ/ in Telugu (Bhaskararao and Ray, 2017). Similar adaptations of the phonemes are also seen from Perso-Arabic borrowed phonemes. Many linguists are also now focusing on developing language specific words to enrich its vocabulary instead of just borrowing loan words. The purpose of this paper is to identify whether the universal sonority scale applies to Telugu or should it opt for language specific sonority scale (LSSH) and help future researchers make a definite step regarding the use of sonority in syllabification.

Sonority

The topic of sonority has been garnering attention from as early as the 19th Century when Sievers (1881) spoke about Sonority Sequencing Principle (SSP) in his book, which became a pioneer of later works about sonority. Heffner (1950), in his work 'General Phonetics' spoke about sonority as the degree of audibility of the voice and by more or less equating it with acoustic energy. Just like the word Jurisprudence in Law that has been defined by many jurists each contributing their own opinion, sonority has also been defined by many linguists like Ladefoged, Parker, Collins and Mees, etc. who put forth few more definitions which brought out its importance. Sonority has

been defined as a linear regression equation derived from intensity measurements (Parker 2002), relative resonance (Clements 2006), relative loudness and carrying power (Collins and Mees 2008), degree of opening of the vocal tract during speech articulation plus sound propensity for voicing (Yavas 2011), etc. of a sound. Though these definitions have been both accepted and debated by majority of the linguistic community, sonority has never been denied to the fact that it plays a major role in arrangement of phonological segments in a syllable i.e. syllabification in general. The arrangement of phonemes in a syllable is in such a way that more sonorous the sound, the closer it is to the nucleus.

Sonority Hierarchy

Sonority Hierarchy or the Sonority Scale is the grading of sonorous and non-sonorous segments in a language. Often, it is debated to be universal or language-specific in nature in its application cross-linguistically. Hogg and McCully (1987) proposed the sonority levels of sound segments of a language based on two important aspects namely, Voicing and loudness. According to them, in terms of voicing, voiced sounds have higher sonority and in terms of loudness, sounds produced with greater amplitude are more sonorous.

Sounds	Values
Low Vowels	10
Mid Vowels	9
High Vowels	8
Flaps	7
Laterals	6
Nasals	5
Voiced Fricatives	4
Voiceless Fricatives	3
Voiced Stops	2

Voiceless Stops	1

Table 1: Sonority Scale proposed by Hogg and McCully

After two years, Clements (1990) proposed a scale which was rather accepted by many world languages. The significance of this scale is that he simplified the scale by dividing the speech sounds into five major natural classes:

Natural Class	Sonority Value
Vowels(V)	5
Glides (G)	4
Liquids (L)	3
Nasals (N)	2
Obstruents (O)	1

Table 2: Sonority Scale proposed by Clements (1990)

Lagdefoged (1993) also proposed a scale which is exactly similar to Hogg and McCully's with an exception of missing flap sounds in the sounds list. Other linguists that proposed this scale after Clements have more or less a similar indexing of the speech sounds which bought the theory of universality of this scale as there was not much of a difference in terms of sonority values in terms and phonemes and allophones.

Sonority Sequencing Principle

Selkirk (1984) defined Sonority Sequencing Principle (SSP) as a syllable that consists of a sonority peak and is preceded and/or followed by a sequence segments of progressively decreasing sonority value. Ideally, the peak is occupied by a vowel which is the most sonorous of all the segments.

1. *(a)* */pra.ti/* 'every'
   (b) */se:.va/* 'service'

Clements (1990) in his work 'Role of sonority in Core Syllabification' proposed that only core syllables follow the route of high sonorous sounds marking the nucleus or peak while the borders end with low sonorous sounds. Anything in violation of the regular norm i.e. low sonorous sound in nucleus position or high sonorous sounds in the borders are expected to be in violation to Sonority Sequencing Principle (SSP).

But, there are cases where this nucleus or peak position is occupied by consonants. These are called '*syllabic consonants'* or '*vocalic consonants*'. They form their own syllables in absence of a vowel. For example *(Parke,r 2011)*

Slovak

2. (a)	*/kr̲.vi̲/*	'blood'
(b)	*/vl̲.ka̲/*	'wolf'

Tachelhit

3.(a)	*/r̲ .gl̲t/*	'lock'
(b)	*/tz̲.dm̲t/*	'gather wood'

In 2, the data shows that the /r/ and/l/ sounds acted as syllabic consonants. They are of the liquids class. But in 3, the data shows other than /r/ and /l/, /z/ and /m/ also took the nucleus position of a syllable. Here, /m/ is of nasal class and /z/ is of obstruent. The reason for such an occurrence is sonority of the segment itself. Considering the first syllable /tz/ of 3(b), the segment /t/ is a voiceless alveolar stop and /z/ is a voiced alveolar fricative. According to the Hogg and McCully's (1987) or Ladefoged's (1993) scale, /z/ is more sonorous than /t/ and as these are the only two sounds that constitute a syllable here, the most sonorous one out of the two i.e. /z/ occupied the nucleus or peak position. Clements (1990) on the other hand, classified SSP of a syllable in to three types:

i. Sonority Peak - The sound segments have ascending sonority values towards the peak or nucleus.
ii. Sonority Plateau - The neighboring segments have same level of sonority values.

iii. Sonority Reversal – The segments have descending values towards the nucleus.

He suggested that only those with sonority peaks in the syllable are following the SSP and the other two types are in violation of SSP. He even denied syllabic consonants in the nucleus position. But considering the data in (1), (2) and (3), it can be noted that the sonority hierarchy can be both universal and language-specific which in turn makes the SSP be applicable both cross-linguistically and language–specifically in some languages.

2. Review of Literature

Telugu has a total of 56 alphabets but not all those are available in the currently used alphabet chart. Now-a-days, the alphabet chart shows only 51 alphabets. This elimination is due to the reduced use of those alphabets by the populace. Among the 56 alphabets that existed originally, quite a few sounds were borrowed from Perso-Arabic, Prakrit and Sanskrit (Bhaskara rao and Ray, 2016). The sounds from these other languages are assimilated in to the Telugu language to fill the gaps of Telugu phonemic inventory that occurred due to borrowing. Bhaskara rao and Ray discussed in their paper about how these aspirated phonemes and few fricative sounds do not occur in native language with the exception of /d^h/ which is only used in the careful pronunciation of two words in the compound numerals,

4. /padd^{h}enimidi/ ‘eighteen’
/padd^{h}na:lugu/ ‘fourteen’

They gave an IPA illustration of Telugu phoneme sounds by clearly mentioning that the sounds shown were based on the educated Telugu speakers’ speech of an eastern dialect area because only they tend to fully pronounce the newly added or extended phonemes in public domain areas like public lectures, over the radio, places of worship, etc., in the Telugu speaking states before bifurcation.

Clements' (1990) sonority hierarchy has initially been accepted by a lot of world linguistics, as a universal scale as the natural classes division is applicable to many languages. It was not until a difference in observed by Sherwin (1999) in the Hindi language of the many languages that she chose to study with. The presence of phonemic aspirated stops and non-sibilant fricative /f/ being ranked low on the sonority hierarchy made Sherwin to deny the universality of Clements' scale. Later in 2006, Clements explained that the sonority scale is for understanding the segmentation of syllables based on sonority and the ranking of segments rely not only on audibility but also on physical and perceptual properties at times.

Jany, Gordon, Nash and Takara (2007) also examined universality of the acoustic basis for the sonority scale found that Mongolian, one of the few languages they studied is leaning towards Language Specific Sonority Hierarchy because when the grading is done in-depth like between voiced and voiceless sounds, aspirated and unaspirated sounds, etc. the concept of universality loses its cross-linguistic applicability.

3. Methodology

This descriptive study looks into the sonority profiles of word initial onset consonant clusters in Telugu. Initially, 137 words with two consonant clusters were collected from the Telugu-English dictionary published by Telugu Academy. These words were selected from each type of initial consonant clusters that were available in the dictionary. These words were further reduced to 50 words through 'Random number Generation' tables. The words were not categorized according to word classes and include both the native and loan words. The final words selected were transcribed as per the IPA symbols and the profiles were drawn according to the sonority scale given by Clements (1990) as this is generally a universally accepted scale for languages. Profiles with SSP violations at the initial

consonant cluster were reanalyzed with LSSHS to re-evaluate the sonority adherence before concluding the final result.

4. Data Analysis and Results

The aim of the study was to analyze whether initial two consonant clusters (2CC) of Telugu were adhering to grading the speech sounds as per Clements sonority scale (1990), or was there a need for a Language Specific Sonority Hierarchy Scale (LSSHS). In accordance with the aim of this study, in section 4.1., all the words with 2CC were graded based on Clements 5- point Scale. A list of words not adhering to the scale were noted. In section 4.2., the scale was reorganized to accommodate all the sounds of Telugu Language along with their features into a Language Specific Hierarchy Scale with 13-point scale. In section 4.3., the words that were not adhering to the Clements Scale, were graded based on the LSSHS, to check the applicability. In section 5., the feasibility of application of LSSHS has been discussed.

4.1. Application of Clements 5-point Scale to Telugu Language and Data

Applying the Clements 5-point Scale the sounds of Telugu Language can be listed as follows.

Grade	Class of Sound	Telugu Sounds
5	Vowels	/a/, /a:/, /I/, /i/, /u/, /u:/, /e/, /e:/, /o/, /o:/
4	Glides	/j/ and /ʋ/
3	Liquids	/l/ and /ɭ/ and /r/
2	Nasals	/m/, /n/, /ɳ/, /ɲ/, /ŋ/
1	Obstruents	Stops: /p/, /t/, /ʈ/ /k/, /b/, /d/, /ɖ/, /g/, /pʰ/, /tʰ/, /ʈʰ/ /kʰ/, /bʰ/, /dʰ/, /ɖʰ/, /gʰ/ Affricates: /tʃ/, /dʒ/,/tʃʰ/, /dʒʰ/ Fricatives: /f/, /s/, /ʂ/, /ś/, /h/

Table 1: Grading of Telugu Sounds based on Clements 5-point scale

It can be noted that, not all sounds of Telugu get represented in physical form. For example nasals and glide sounds can only be heard as a part of colloquial speech production and do not get represented in the written form. However, as this aspect was not the aim of this study, we are not considering this in our analysis.

Based on the grading given in table 1, the onset consonant clusters were analyzed using the sonority hierarchy scale. For the 50 randomly selected words, type-token ratio was obtained. It was observed there were 29 tokens, which could be classified into 6 patterns of cluster combinations. The list of the onset clusters is given below.

S.no	Onset clusters	Combination	Scale	Value
1	/pr/	Obstruent + Liquid	Rising Sonority	1-3
2	/ʊj/	Glide + Glide	Sonority Plateau	4-4
3	/kl/	Obstruent + Liquid	Rising Sonority	1-3
4	/bʰr/	Obstruent + Liquid	Rising Sonority	1-3
5	/gr/	Obstruent + Liquid	Rising Sonority	1-3
6	/st/	Obstruent + Obstruent	Sonority Plateau	1-1
7	/tr/	Obstruent + Liquid	Rising Sonority	1-3
8	/kr/	Obstruent + Liquid	Rising Sonority	1-3
9	/gɲ/	Obstruent + Nasal	Rising Sonority	1-2
10	/kʂ/	Obstruent + Obstruent	Sonority Plateau	1-1
11	/skʰ/	Obstruent + Obstruent	Sonority Plateau	1-1
12	/dʊ/	Obstruent + Glide	Sonority Plateau	1-1
13	/gʰr/	Obstruent + Liquid	Rising Sonority	1-3
14	/nj/	Nasal + Glide	Rising Sonority	2-4

15	/sn/	Obstruent + Nasal	Rising Sonority	1-2
16	/sp^{h}/	Obstruent + Obstruent	Sonority Plateau	1-1
17	/śr/	Obstruent + Liquid	Rising Sonority	1-3
18	/st^{h}/	Obstruent + Obstruent	Sonority Plateau	1-1
19	/dr/	Obstruent + Liquid	Rising Sonority	1-3
20	/śʊ/	Obstruent + Glide	Sonority Plateau	1-1
21	/ʤʊ/	Obstruent + Glide	Sonority Plateau	1-1
22	/śl/	Obstruent + Liquid	Rising Sonority	1-3
23	/sm/	Obstruent + Nasal	Rising Sonority	1-2
24	/sʊ/	Obstruent + Glide	Sonority Plateau	1-1
25	/hr/	Obstruent + Liquid	Rising Sonority	1-3
26	/tʊ/	Obstruent + Glide	Sonority Plateau	1-1
27	/dj/	Obstruent + Glide	Rising Sonority	1-4
28	/dd^{h}/	Obstruent + Obstruent	Sonority Plateau	1-1
29	/ʤj/	Obstruent + Glide	Rising Sonority	1-4

Table 2: Combination of Onset Clusters in Telugu

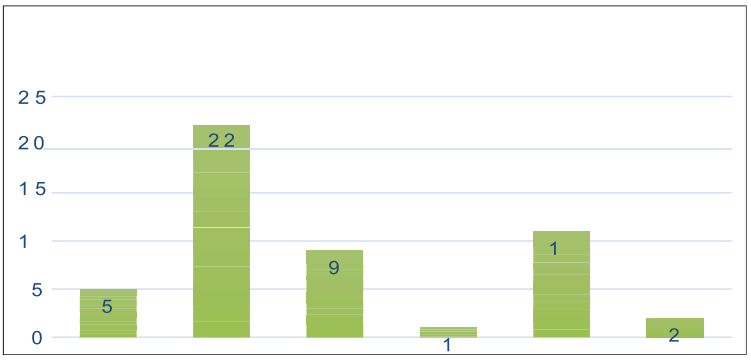
As shown above, there are about 29 different consonant clusters found in the 50 words. Out of these, 74% of the words containing onset consonant clusters were of rising sonority (i.e. adhering to SSP) whereas the remaining 26% were plateaus, which were non-adhering. The following cluster combinations can be listed based on the above table -

• Obstruent + Nasal (ON)	• Nasal + Glide (NG)
• Obstruent + Liquid (OL)	• Obstruent + Obstruent (OO)
• Obstruent + Glide (OG)	• Glide + Glide (GG)

The number of words falling under each cluster type is depicted in the table and graph below.

	Type of Cluster	Number of words
Adhering SSP	ON	5
	OL	22
	OG	9
	NG	1
Non- Adhering SSP	OO	11
	GG	2

Table 3: Type-token Ration of 2 Consonant Clusters



Graph 1: Type-token Ratio of Cluster Combinations

The following examples show the cluster combinations that are violating SSP as per Clements scale.

- Obstruent + Obstruent (OO)

	Telugu	gloss	Sonority
Value 5. (i)	*/stu:.pam/*	'pillar'	[1-1-5-1-5-2]
(ii)	*/spha.ʈi.kam/*	'crystal'	[1-1-5-1-5-1-5-2]

Glide + Glide (GG)

	Telugu	gloss	Sonority value
6 (i)	*/v̯ju:.ham/*	'strategy'	[4-4-5-1-5-2]
(ii)	*/v̯ja:.ʤjam/*	'law suit'	[4-4-5-1-4-5-2]

4.2. Language Specific Sonority Hierarchy Scale (LSSHS)

Earlier attempts were made to propose LSSHS for Telugu Language. For example, Swathi (2014) proposed a scale for grading Telugu sounds based on Sonority Hierarchy, in which she proposed 8-point scale, however, aspirated sounds were not separately listed in this scale. Hence, proposed here is a 13-point Language Specific Sonority Hierarchy Scale (LSSHS) taking into considerations all the features of sounds of Telugu Language as proposed by Bhaskara rao and Ray (2017).

Natural Class		**Value**
Vowels		13
Glides		12
Liquids		11
Nasals		10
Voiced Affricate	Unaspirated	9
	Aspirated	8
Voiced Stops	Unaspirated	7
	Aspirated	6
Voiceless Fricatives		5
Voiceless Affricates	Unaspirated	4
	Aspirated	3
Voiceless Stops	Unaspirated	2
	Aspirated	1

Table 4: Language Specific Sonority Hierarchy Scale (LSSHS) for Telugu Language

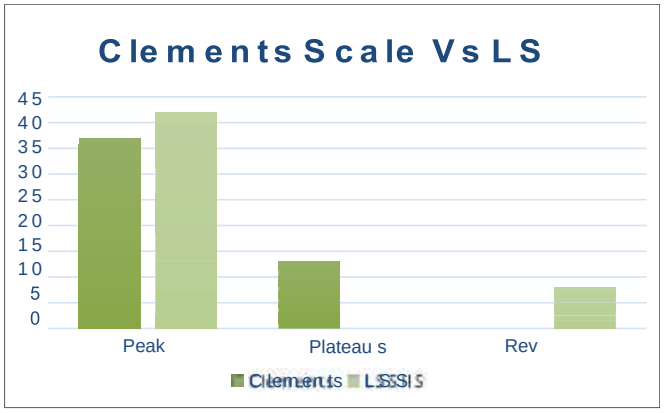
4.3: Application of Language Specific Sonority Hierarchy Scale

Application of Language Specific Sonority Hierarchy Scale (LSSHS) to clusters violating sonority sequencing principle found that the number of clusters violating the SSP has decreased. The application of LSSHS to the 50 words now results in 42 sonority peaks (adhering to SSP) and 8 reversals (non-adhering to SSP). The clusters that were initially found to

have sonority plateaus have now changed to SSPs with rising sonority peaks. This was because the fricatives, affricates and stops have been differentiated in the scale in terms of voicing and aspiration features.

Comparison between Clements and LSSHS

Using Language Specific Sonority Hierarchy Scale 84% of its onset clusters seem to adhere to SSP by having a rising sonority peak. This was 10% higher to applying the Clements universal sonority scale. Though using LSSHS resulted in more number of Sonority plateaus turning into SSP adhering peaks, it also resulted in 8 instances of Sonority reversals.



Graph 2: *Comparison of SSP types in Clements Scale and LSSHS*

5. Conclusion

The profiling of Telugu speech sounds using sonority scale resulted in interesting findings. Initially, the scale developed by Clements was applied to initial two consonant clusters of Telugu language. This scale could account for only 74 % of the data, while obtaining 26% plateaus. Thus, to account for more data, a Language Specific Sonority Scale was developed. Using this scale 84% of the data could be accounted for but it also resulted in reversals. A probable reason for this could be the length of the

scale. However, as the aim of this study was to account for as many syllables as possible, a LSSHS can be preferred.

The most common clusters observed from the data for initial 2 consonant clusters were CCV or CCVC syllable type. In case of sonority peaks, the C_1 was usually occupied by the sounds /p/, /t/, /k/, /d/, /g/, /b^h/, /g^h/, /s/, /ś/, /h/, /ʋ/, /ʤ/, /n/ i.e. the obstruents and nasals. Glides or liquids are not seen in C_1 position of a consonant cluster. In case of C_2 position in clusters,
/p^h/, /t^h/, /k^h/, /t/, /n/, /m/, /ɲ/, /l/, /r/, /j/, /ʂ/ sounds are predominantly seen. It is also observed that sibilant sounds /s/ and /ś/ are not seen in C_2 position. The reversals are seen in clusters such as /st/, /st^h/, /sk^h/, /sp^h/ where initial consonant C_1 is always /s/ - a voiceless alveolar fricative. The second consonant C_2 is always seen to be a voiceless stop which is either aspirated or unaspirated. As such, Telugu language is both - mostly adhering to SSP and rarely violating the SSP.

References

Bhaskararao, P., & Ray, A. 2017. Telugu. *Journal Of The International Phonetic Association*, Vol 47(Issue 2), 231-235. Doi :10.1017/S0025100316000207. Retrieved On 08/10/2021 From Www.Cambridge.Org/Core/Journals/Journal-Of-The- International-Phonetic-Association/Article

Clements, G. N. 1990. The Role Of Sonority In Core Syllabification. In J. Kingston & M.E. Beckman (Eds.), *Papers In Laboratory Phonology I: Between The Grammar And Physics Of Speech*. New York: Cup. 283-333. Doi: Https://Doi.Org/10.1017/Cbo9780511627736.017. Accessed On 20 November 2021 From Http://Www.Ai.Mit.Edu/Projects/Dm/Featgeom/Clements90 .Pdf

Jany, Gordon, Nash and Takara. 2007. How Universal Is The Sonority Hierarchy? A Cross-Linguistic Acoustic Study. Icphs16. 1401-1404.

Joel, A. 2015. *Sonority Profiles of Swahili Words*. Unpublished Master's Thesis, Osmania University, Hyderabad.

Nezhad, M. & Miri, M., 2013. Sonority Sequencing Principle In Initial Consonant Clusters In Sistani Dialect. Language In India. Vol 13:11. 331-343.

Parker, S. 2011. Sonority. In Oostendorp, Van Mark , Ewen, J.C., Hume, E. and Rice, K. (Eds.) *Blackwell Companion to Phonology*, Vol-II, 1160-84. https://onlinelibrary.wiley.com/doi/book/10.1002/9781444 433 5262

Sailaja, P. 1999. *Syllable Structure in Telugu*. Icphs99. San Francisco. 743-746

Sherwin, S. 1999. The Sonority Sequencing Principle in Inter-Language Phonology. Working Papers in Linguistics.

Swathi, C.S. 2014. Phonological Errors in Telugu Aphasics: The Role of Sonority. In Vasanta, D. (Ed) Clinical Applications of Phonetics and Phonology in Indian Languages, ISHA Monograph, Vol 14, No.1, 121-158.

Vidyamrutha, I., & Swathi, C.S. 2021. *Sonority Profiles of Reduplicated Words In Telugu*. Presented online at 43rd International Conference of the Linguistic Society of India, 21-23 December.

Venkateswara Sastry, J. 1994. *A Study of Telugu Regional and Social Dialects – A Prosodic Analysis*, Mysore: Central Institute of Indian Languages. 59-112.

Syntactic performance of Telugu speaking school going children with Hearing Impairment

Dr.S.Santhi Prakash[1] , Mrs.V.J.Vinila[2]

(1) Reader Spl.Edn (HI), Ali Yavar Jung National Institute of Speech & Hearing Disabilities (Divyangjan), Regional Centre, Secunderabad.
(2) ASLP, Ali Yavar Jung National Institute of Speech & Hearing Disabilities (Divyangjan), Regional Centre, Secunderabad.[2]

ABSTRACT

"I am just as deaf as I am blind", wrote Helen Keller.

The study was aimed to investigate differences in Syntactic performance between children with normal hearing and children with hearing impairment. 240 children were selected and divided into 2 groups of 120 children with normal hearing and 120 children with hearing impairment. The purposive random sampling technique was adopted in selection of the sample. The tool i.e "The Test of Language in Telugu" developed by Prakash & Prakash (2006) was used to study the syntactic performance of Telugu speaking children with hearing impairment who are studying 7th Standard.

Results revealed that children with normal hearing showed better syntactic performance in terms of reception and expression when compared to children with hearing impairment. Among children with hearing impairment, children in Non-Government Special Schools showed better performance on syntax as compared to children studying in Government Special Schools. In terms of Gender girls performed better than boys and children following Oral communication performed better than total communication method. The implications for both educators and researchers are to understand the syntactic development of Telugu speaking children with hearing impairment and to design treatment programs to improve syntactic and semantic relational skills. The administrators can also plan and develop curriculum for improving the syntactic performance and also develop relevant teaching aids for children with hearing impairment.

1. Introduction

"I am just as deaf as I am blind", wrote Helen Keller.

Deafness is a misfortune, and its problems are intensive and complex. It has been recognized as a handicap for many centuries not only because it can affect personality so adversely, but also because it prevents children from developing speech and

language skills. In these individuals most vital stimulus the sound of a voice that plays a major role in language development to keep the individual in the intellectual company of a man is lost (Stevens & Warshofy, 1971). A hearing-impaired child does not have adequate speech and language skills. Deaf child is unable to express their feelings and thoughts through language. So the acquisition of language skills in hearing-impaired children is very essential for academic and social development. Language plays an important role in the development of the personality of the child in all his/her effective transactions in day-to-day life situations.

Language is an acquired system of structured but arbitrary vocals, signs, and symbols that provide meaning by cataloging and representing people, places, things and feelings, and other abstract concepts. According to Bloom (1973) "Language is a code whereby ideas about the world are represented through a conventional system of arbitrary symbols for communication".

Language has different parts called components of language. These parts deal with what to say (content), when to say (use), and how to say a word or a sentence (form). We will be able to communicate efficiently by making use of these components. The major components are further sub-divided into
(a) Phonology (b) Morphology (c) Syntax (d) Prosody (e) Semantics (f) Pragmatics. The various levels of linguistic analysis from phonology to syntax and semantics exist entirely in the form of abstract mental constructs inclusive of cognitive mechanism, which is said to regulate the acquisition of linguistic constructs. The sensory, motor and perceptual mechanisms regulate the acquisition of the skills needed to transmit and receive the linguistic signals at the physiological level. Disturbance in any of these mechanisms interrupts normal language acquisition.

Auditory channel is the major input device for normal language acquisition. It is the controlling element of servo

system in language acquisition. Hearing helps in the storage of auditory information that a child receives. When this controlling device is damaged, the whole servo system gets disrupted, affecting its smooth and automatic functioning. Language input of Hearing Handicapped individuals gets restricted both by external and internal sources. The hearing loss directly restricts the organism's ability to process speech input at the level of sensation. Despite the acoustic amplification, speech intake generally remains fragmentary and incomplete, since hearing aids and lip-reading cannot compensate for most severe to profound hearing loss. This results in language handicap either in the form of insufficient language or faulty language use.

Apart from its direct influence on language acquisition, hearing handicap also has an indirect influence on the communicating environment. This may be manifested in the form of reduced language input by the parents, family members, and the social group with whom a hearing handicapped individual interacts. Most communication between a deaf child and his immediate family members takes place by iconic and situation-bound gestures. Hearing-impaired children need vigorous training. Despite intensive training, a good number of hearing-impaired individuals may reach a ceiling in their syntactic performance.

Need for the Present Study

The field of language learning and language pathology is relatively recent in India. There is a paucity of research studies in the Indian language on the acquisition of language by normally hearing children and children with hearing impairment. A few studies that are available on normal language acquisition are of Master's dissertations or Doctoral theses (Sridevi, 1976; Subramanyaiah,1978; Prema, 1979; Roopa, 1980; Vijayalakshmi, 1981;. Studies on the language of hearing- impaired are even fewer (Lakshmi, 1990; Rashmi,1992; Suma, 1992; Uma,S, 1994; Kusuma, 2002).

Studies on the language of the hearing-impaired children agree on the point that the language of hearing impaired is far below that of normal hearing children. However, not much information is available on the nature of syntactic difficulties experienced by them in learning certain syntactic structures in any of the Indian languages. Since such knowledge is essential in planning intervention strategies, the current study was undertaken.

Objectives of the Study

The following are the objectives of the present study:

I. To study the difference in Syntactic Performance between children with normal hearing and children with hearing impairment, boys with normal hearing and boys with hearing impairment and girls with normal hearing and girls with hearing impairment.

II. To study the difference in Syntactic Performance of children with hearing impairment studying in Govt and Non-Govt special schools, between boys and girls, between hearing aid users and non-hearing aid users and exposed to total communication and oral communication methods.

Hypotheses

For the purpose of present study, the following hypotheses have been formulated

1. There would be significant difference in terms of Syntactic Performance between (a) Children with normal hearing and children with hearing impairment, (b) Boys with normal hearing and boys with hearing impairment and (c) Girls with normal hearing and girls with hearing impairment.

2. There would be significant difference in terms of Syntactic Performance of children with hearing impairment, (a) Studying in Govt and Non-Govt special schools, (b) between boys and girls, (c) Between hearing aid users and non-hearing aid users and (d) Exposed to total communication and oral communication methods.

2. Review of Literature

Quigley and his colleagues conducted an extensive investigation on comprehension and production of several syntactic structures by deaf children and adults (power and Quigley, 1973; Wilbur, Quigley and Mentanelli, 1975; Quigley, Montanelli and Wilbur, 1976). The research was done to investigate the order of difficulty of certain syntactic structures, establishment of syntactic rules, stages of development of syntax and the presence of distinct syntactic structures. The findings indicate that deaf children had similar order of difficulty in the acquisition of syntactic structures as normal hearing children. Hearing impaired children were found to impose subject-verb- object pattern in inappropriate situations.

Karanth(1984) tested eight children with moderate hearing loss the results showed poor performance by the hearing- impaired children. She also inferred that the syntactic development is severely affected when compared to normal hearing children within the same age range. Their phonological and semantic developments were better than syntax.

Thirumalai and Gayatri (1988) described the verbal speech characteristics of hearing impaired children at length. She reported that the sentence length of the hearing impaired was shorter characterized by broken sentences generally not connected by any grammatical markers.

Lakshmi (1990) studied the return samples of 17 hearing impaired children with profound sensorineural hearing loss.

Theirmean age was 18 years two month. The performance of sharing impaired was compared with the written language of sharing students with a mean age of 14 years 2 months. The tasks are administered with storing narration through pictures picture description the title of the picture depicted being farm and essay writing on My house in Kanada and analysis of the test result revealed that hearing impaired children used short and simple sentences. The sentences had more nouns than other words. Very few syntactic markers were used by the hearing- impaired children. Incorrect usage of nouns and verbs were noticed, and a number of incorrect sentence constructions were seen.

The overall picture that merged supported the prevailing view that the language of hearing impaired was inferior to that of the hearing children.

Kusuma (2002) studied semantic and syntactic aspects of language in Hindi speaking hearing impaired children and reported that there is a significance difference between semantic and syntactic receptive abilities in hearing impaired children and semantic expressive abilities are better than that of syntactic expressive abilities.

However, the knowledge in this area is far from complete, not much information is available on the nature of syntactic difficulties experienced by children with hearing impairment in learning syntactic structures especially in Telugu language. Hence the present study aimed to investigate the Syntactic Performance of Telugu Speaking Children with hearing impairment studying 7thclass.

3. Methodology

The following method was adopted for the study.

3.1 : Participants

A total of 240 children studying class 7th in both normal schools and special schools for hearing impairment with Telugu as medium of instruction were selected for the study.

The subjects are divided in to two groups.

Group I consist of 120 children with normal hearing (60 boys and 60 girls) studying in 3 Govt. Schools.

Group II consist of 120 children with severe to profound hearing impairment age matched with children with normal hearing studying in Govt and Non-Govt special schools for the hearing impairment were selected for the study. Out of 60 children (30 boys and 30 girls) from 3 Govt. Schools for the hearing handicapped located at Hyderabad, Karimnagar and Nalgonda districts. 60 children (30 boys and 30 girls) from 5 Non-Govt schools for hearing handicapped located at twin cities of Hyderabad and Secunderabad.

3.2 : Tool

The "Test of Language in Telugu" developed by Prakash and Prakash (2006) was used. This tool was developed to measure Semantic and Syntactic abilities among children with hearing impairment. There are 12 categories in Semantics and 10 categories in Syntax. The tool was standardized by establishing reliability and validity. The present study focuses on Syntactic abilities.

3.3 : Procedure

The test was administered to the selected sample of children with hearing impairment and children with normal hearing studying form class 7th spread in 11 schools selected. The test is conducted to each child separately and time limit of approximately half an hour is allotted for each child.

3.4 : Scoring

Each category has equal number of items. Each item under each category was scored on the basis of 'correct', 'emergent' and 'no response'. Child's response on each item is given a score of 1 for correct response, half for correct response but not mentioned in the list (emergent behavior) and 0 for wrong or no response. Total scores of syntax reception, syntax expression and a c combined score of syntax i.e Scores of both syntax reception and expressions are obtained for both children with normal hearing and children with hearing impairment.

3.5 : Analysis

Statistical analysis is done by finding the Means and SD's for all the variables studied with respect to syntactic performance. Statistical comparisons of the variables with respect to syntax between children with normal hearing and children with hearing impairment were one using 't'test.

4. Results

The collected data was analyzed and interpreted the scores with respect to syntactic performance of children with normal hearing and children with hearing impairment with the following findings and observations.

Syntactic performance between the children with Normal Hearing and children with Hearing Impairment

- Children with normal hearing and children with hearing impairment

S.No	Category	N	Mean	SD	t value	df	Level of Significance
1	Children with normal hearing	120	58.35	1.03	17.17	238	000

2	Children with hearing impairment	120	40.05	7.84			

☐ Boys with normal hearing and boys with hearing impairment

S.No	Category	N	Mean	SD	t value	df	Level of Significance
1	Boys with normal hearing	60	58.26	1.00	13.95	118	000
2	Boys with hearing impairment	60	38.40	7.74			

☐ Girls with normal hearing and Girls with hearing impairment

S.No	Category	N	Mean	SD	t value	df	Level of Significance
1	Girls with normal hearing	60	58.26	1.00	13.95	118	000
2	Girls with hearing impairment	60	38.40	7.74			

Tables showing Difference in Mean scores of Children with normal hearing and Children with Hearing impairment.

There is a significant difference in terms of overall syntactic performance between

- Children with normal hearing and children with hearing impairment.

- Boys with normal hearing and boys with hearing impairment.
- Girls with normal hearing and girls with hearing impairment.

Observations

- Children with hearing impairment showed delayed and poor performance when compared to children with normal hearing.
- Boys and girls with hearing impairment performed poorer than the boys and girls with normal hearing.

Syntactic performance of children with hearing impairment

☐ Children studying Govt and Non-Govt special schools

S.No	Category	N	Mean	SD	t value	df	Level of Significance
1	Govt. Spl Schools	60	35.21	4.43	8.57	118	000
2	Non-Govt Spl Schools	60	44.89	7.53			

☐ Boys and Girls with hearing impairment

S.No.	Category	N	Mean	SD	t value	df	Level of Significance
1	Boys	60	38.40	7.74	2.35	118	.862
2	Girls	60	41.70	7.64			

☐ Hearing aid users and hearing aid non-users

S.No.	Category	N	Mean	SD	t value	df	Level of Significance
1	Hearing aid users	63	41.14	8.15	2.65	118	.001

2	Hearing aid non-users	57	37.54	6.50			

☐ Total Communication and Oral Communication methods

S.No.	Category	N	Mean	SD	t value	df	Level of Significance
1	Total Communication	98	37.47	6.76	4.31	118	.070
2	Oral Communication	22	42.13	7.27			

Tables showing Difference in Mean scores of overall syntactic performance of Children with Children with Hearing impairment in terms of Govt and Non-Govt Special schools, Boys and Girls, hearing aid users and hearing aid non-users and exposed to Total communication and Oral communication methods.

There is a significant difference in terms of overall syntactic performance between

- ☐ Studying in Govt and Non-Govt special schools
- ☐ Between boys and girls
- ☐ Between hearing aid users and non-hearing aid users
- ☐ Exposed to total communication and oral communication methods

5. Discussion

The children with hearing impairment studying in Non-Govt Special Schools performed better in overall syntactic performance than children studying in Govt.Special Schools. This may be due to better educational facilities provided by NGOs such as use of teaching aids, monitoring regular usage of hearing aids, conducting parental counseling sessions, motivation of teachers, Special attention given to the children.

Overall syntactic performance of girls is better than boys. Acquisition of syntax includes the usage of categories such as plurals, tenses, PNG markers, case markers, conditional clauses, transitive, intransitive causatives, sentence types, conjunctives/quotatives, comparatives and participle construction which are difficult for the hearing-impaired individuals to acquire due to their hearing loss. But this acquisition is better in girls as they have good social stigma i.e getting opportunities to get exposed to the environment to learn language and to interact with their peers and other members in the society.

Overall syntactic performances of children with hearing impairment using hearing aids are better than children non hearing aid users. This may be due to the fact that usage of hearing aids regularly by children with hearing impairment maximizes the reception of speech through auditory and/or visual channels. Therefore, it provides a means of amplification to them to hear and learn the different concepts and the categories that are essential for developing language easily. Thus, the hearing aid provides a tool for a child to acquire the linguistic development. Apart from only usage of hearing aids, providing only auditory and auditory visual learning experiences for the children, care and maintenance of hearing aids, successful aural rehabilitation and motivation by the teachers yields better performance by the hearing aid users when compared to non-users.

Overall syntactic performances of children with hearing impairment exposed to oral communication performed better than the children exposed to total communication. This difference may be attributed to the fact that oral communication method which is a unisensory approach relies on residual hearing almost exclusively and speech reading occurs only in the natural context of daily living and also children who lean language through oral method receive input through the

combination of speech reading and amplification of sound while expressing themselves through speaking. And also that language is readily learned through audition and that only oral education can give children access to auditory/articulator codes on which language is based when compared to total communication. These results are in support with findings by Ling (1984) and Gatty (1987) who reported that there is much individual variability in children's ability to acquire speech by using oral approach where the most children with hearing impairment were successful. Therefore, it can be concluded that the oral communication method is the best method in which the hearing-impaired individuals understand the concepts for development of syntax.

Implications of the Study

The findings of the study have the following implications for Audiologists, Speech and Language Pathologists, Educators and Researchers.

- Act as a precursor for the educators in designing the training programs to improve semantic-syntactic relational skills of children with hearing impairment.
- Help in planning and developing curriculum for improving syntactic performance.
- Help in netter understanding of syntactic performance of Telugu speaking children with hearing impairment.
- Teacher empowerment.
- Further can help to develop relevant teaching aids to improve the syntactic performance in children with hearing impairment.
- The outcome of the study may also appraise the educational administrators for reviewing the facilities and concessions given to the children with hearing impairment.

Limitations

- The sample size is relatively small.

- Variables such as interference of gesture/multilingual factors were not controlled.
- There is a need to include children with hearing impairment studying in integrated/inclusive schools.
- There is a need to study other components of language such as semantics and pragmatics.
- The following areas need special attention to promote efficient language acquisition in children with hearing impairment.
- Early identification and early intervention- to enable the hearing handicapped child to acquire natural language.
- Demonstration classes for the parents to guide them in teaching language to their children with hearing impairment.
- Following the aural-oral approach of language teaching from the very early stages.
- Teachers should also have adequate knowledge of language acquisition in children with hearing impairment.
- Setting integrated schools where the children with hearing impairment will have a wide exposure to peer language use.
- Public awareness program to educate the people on the importance of early identification.

6. Conclusion

Education is the basic tool for the development and reconstitution of a humanbeing.It helps an individual to realize his inherent qualities and develop his entire personality. This in turn has a direct hearing on the economic and social progress of the nation as a whole. Education for All (EFA) is, therefore, seen as a process of investment in human development, which empowers people to contribute better to their own self and family and to the society in general.

Although this study is a very small but sincere effort have been made to investigate the Syntactic performance of school

children with Hearing impairment and suggest programs for accelerating the growth and development of language acquisition, in hearing-impaired at the crucial stages of early childhood. The researcher is aware of the limitations of this study and proposes to study it in detail even further in the future. Underlying this is an earnest endeavor to seek the means of providing equality of opportunity to every child in the larger context of education for all in India.

References

Kusuma, (2001).' Semantic And Syntactic Aspects Of Language In Hearing Impaired Children - A Comparative Study. Unpublished Master's Dissertation, University Of Bombay.

Lakshmi, (1990). 'Syntax In The Written Language Of Normal Hearing And Hearing Students Of The 8thstd'. A Comparative Study. Unpublished Master's Dissertation. University Of Mysore.

Ministry Of Law, Justice & Company Affairs (1996) The Persons With Disabilities (Equalopportunities, Protection Of Rights And Full Participation) Act, 1995 (No.1 Of 1996) New Delhi; The Gazette Of India P.24.

NSSO (2002) National Sample Survey Organization, Disabled Persons In India. Report No.485: 58th Round, New Delhi, Department Of Statistics, Government Of India.

Prakash And Prakash (2006): "Test Of Language In Telugu" Manual Of AYJNISHD, Mumbai Project.

Prema, K.S. (1979). Some Aspects Of Syntax In 5-6-Year-Old Children. A Descriptive Study In Kannada. Unpublished Master's Dissertation. University Of Mysore.

Rashmi, P. G. (1992). Story Comprehension, Recall, And Generation In The Hearing Impaired. Unpublished Master's Dissertation: University Of Mysore.

Sreedevi, N. (1976). The Aspects Of Acquisition By 2-Year-Old Child Children. Unpublished Master's Dissertation, University Of Mysore.

Subbarao, T.A. (1990). Manual on Developing Communication Skills In Mentally Retarded Children.

Subramanaiah, M. G. (1978). Development Of Some Morphological Categories In Kannada. A Study Of Children 6-8 Years Age Range. Unpublished Master's Dissertation. University Of Mysore.

Suma, N. (1992). Sentence Structure And Reading Comprehension In The Hearing Impaired. Unpublished Master's Dissertation, University Of Mysore.

Uma.S. (1994). 'Language Acquisition by Hearing-impaired children in Kannada'. Unpublished Doctoral Thesis, University of Mysore.

Vijayalakshmi, A. R. (1981). “Development of a test in Kannada for assessing Language acquisition in children. Unpublished doctoral thesis, University of Mysore.

Aesthetic Importance of Onomatopoeia in 'Antarani Vasantam'

Dr. Kalloji Susheel Kumar

Assistant Professor, Department of Linguistics, Osmania University, Hyderabad

ABSTRACT

Onomatopoeia in literature is the phenomenon that is used to imitate natural sounds. These sounds are made an echo to the sense. Onomatopoeia is the formation or making of a word by imitating the sound of the noise or action designated or made such as twitter, hiss, water- fall etc. The present paper deals with the relation between onomatopoeia and arbitrariness and also it analyses how the poet uses onomatopoeic words in the novel 'Antarani Vasantam' referring to its rhetoric importance. The use of onomatopoeia in the novel is clarified with the examples used by the author G.Kalyana Rao. As the analysis is on Telugu novel, all the examples are provided in Telugu without translation in order to retain onomatopoeic effect of the words, and author's style in the novel.

The paper concludes with the inference that onomatopoeia in literature is a beautiful phenomenon which is found in all the languages. However it focuses on the author's use of onomatopoeia which can be perceived at the Visual, Auditory, Audio-Visual, and Physical to bring the aesthetic effect and beauty to the novel and to create vehemence among the reader of 'Antarani Vasantam'

Key words: Onomatopoeia, Antarani Vasantam, Aesthetic , Rhetoric

Introduction

The very interesting characteristics of language by comparison with other codes or communication are its flexibility and versatility. Language has various properties, and arbitrariness is one among them. The meaning of the term arbitrary is defined as something like 'inexplicable in terms of more general principle' (Lyons;1990:19). In generally there is no linguistic connection between a linguistic form and its meaning. The arbitrary relationship between the linguistic signs and the objects used to indicate the forms of human language is called arbitrariness. (Yule:1985:18-19). Onomatopoeia words are

considered a part of the natural sounds 'theory of language' origin (Ibid). There are sporadic instances in all languages which are traditionally referred to as onomatopoeia. Leonard Bloomfield (1933:156) distinguishes between words which are actually imitative of sounds, like meow, moo, and bow-wow, and differs from language to language.

Onomatopoeia can be defined as the formation of words whose sound is imitative of the sound of noise or action designated such as hiss,buzz,and bang. (Harper;2010:p1) such words are used for poetic or rhetorical effect. (Ibid:1). Also it can be defined as it uses to imitate actual sounds as the ringing bells, the singing birds, or voices animal. In a broader sense it refers to any combination of the imitative sound and rhythms that are used to reinforce the sense or moods of a passage of poetry or prose. (Ibid:7).

Onomatopoeia: Phonological Echoism

As the style of the language with which we are concerned, has to be investigated in all the varieties of its function, this paper explore some of the additional aspects of the phonological patterns which are determined not mainly by the collocation of sounds in a series of words as in "anuprasa" or by the patterns of equivalent vowels and consonants but of something more, which deals with some part of the expressive functions of the language and which goes even a litter further than the arbitrariness of the language. These words create an imitation of the thing talked about, and strengthens the cognitive meaning of the word. However, in as much as these words which imitate the natural sounds are comparatively few and are even conventionalized in different languages, they however do not invalidate the general principle of the arbitrary relation between a word and its referent. The imitation of these can be perceived at the Visual, Auditory, (both visual & auditory put together) and also at the physical & senses of a human being. To create this impart in the reader the writer has used this device, along with the story.

Visual

- 9కటª చ ంక4లc ఆణcకc ఆణcకc మ ంటc>gꝏ. (54)
- ꝏపç9&ే oాaాల ఈ 9లcగc ˘gల. (113)
  చ ంక4ల a>

Auditory

- త% c ఆెంచ ంsి9 ఆి>´ పçటకడs. (13)
  aగ s<వçgనక
- వలవల ల aాన9 జలజల ꝏ´టª కద%క9. ( )
  కco*>´ ఢg 6ాo´ >ల
- స ా%sª ఆకc (89)
  ం Y*ర.cన %oన
  &Y
- ఖం0&.ªర9 ఉఆq మo* వwşనం. (95)
- ఆెల.o*జమcన >˘పç9లc sాడ మ%g
  ఆYాꝏ. కరకర 6ª దŢ
  ¾దలయçꝏ. (196)
- జ˘>ీ7 wెపç9 అంటª Y*ర.cన %oY* తల ఆత 6ª టఆద a`టª sాళvg >నమq
  ఆద aటడs. (203)

Audio-Visual

- oేaాలయం ధgజసక మంటe ఆo*Y*ప&ªన శబŢం (34)
  ంభం ఆఆ
- మటª9 9లcşకcంటª ª (133)
  9లcşsింటª మట9 ఆos వ3క ª
  oo
- a´ల హ3కాలc, o<పo<పల&ే లª జం0&.
  Yింతvలc, జం0&లc జీ7w%
  aల

  రe9 wే%లª జం0&.(213)

Physical

- అనgం aటక>క´ ఆవçoావçరcమ9 %>gడs. (26)
- దగ.రకc %సంsింట´ బª రcన క వçం ాళvg.
  ఏ&şడs. (26)
- ఒsి43ాo* కరణం Yిడ.sిటకం దగ.ర నటc ద>a
  గబగబ ఫ(49)

- &వయ్ç ఆ మ◌ాటల◌ª ఉ%sª4ప&ˌ◌ాడs. (119)
- జ◌ామcs◌&ª మ◌ా తఀ◌ృ%.ప&ª చ◌ా◌ాª◌ాడs. (146) కe>ి◌ి◌oo. ośz
- aి◌ల. లc sˇవçgన s´కa´a◌ాrc. (195) (204)
- a´o* ఉ%sª4ప&ª ఆ మ◌ాట అనg ఆఁకి చ◌ా>ి ◌oo.

Conclusion

Poets or the writers use of onomatopoeia to access the reader's auditory sense and create rich sounds capes. It is one of many rhetoric devices used by the author in the novel . Finally, onomatopoeia helps the readers hear the sounds the words reflect. Hence the reader can enter the world created by the poet or author with the aid of these words. The beauty of onomatopoeic words lies in the fact that they are bound to have an effect on the readers senses whether they are understood or not. Moreover, a simple, plain expression does not have the same emphatic effect that conveys an idea powerfully to the readers. The use of onomatopoeic words helps to create emphasis in literary discourse.

References

Bloomfield. (1933). Language. New York: Holt, Rinehart and Winston, p-156

Harper, Douglas. (2010). Origin and Meaning of Onomatopoeia by Online Etymology dictionary. onomato……<word<www.etymoline.com

Kalyana Rao G. (2000). 'Antarani Vasantam', Vijayawada. Published by Viplava Rachayitala Sangham, Andhra Pradesh.

Lyons, J. (1977). Semantics. VolI. London and New York: Cambridge University Press.

Yule, George. (1985). The Study of Language. Cambridge University Press pp18-19.

Morphophonemic processes of Banjara Language

K. Balu Naik

Asst. Professor, Department of Linguistics, Osmania University, Hyderabad

ABSTRACT:

Banjara is an Indo-Aryan language is spoken at Telangana has undergone many influence in lexeme and grammar there are many cultural and linguistic borrowing that happened in Banjara a very less work has been done on the Morphophonemic processes of Banjara. Hence this paper will discuss the Morphophonemic changes in Banjara involve assimilation, loss and the addition of phonemes with suitable examples. Since it is in contact with Telugu at times, the differences are also discussed in this paper.

Introduction:

Lambada language is one of the Indo-Aryan languages. The language continued its existence in the form of oral communication as its script has not been developed. In case of necessity of written communication, the script of local predominant language is used. Hence, in Telangana also, Telugu script is used for such kind. Because of the influence of regional predominant language, most of the Lambdas are generally either bilinguals or multi-linguals. Banjara is nomadic tribe in Telangana as well as India; especially in Central India. The tribe is known with different names viz. Banjari, Wanjari, Brinjari, Lambani, Labani, Sugali etc.According to LSI, Grierson states that the name Banjara is derived from the Va:nijyaka:raka 'a merchant'. Later in the Prakrits it has it has changed to Vanijjarao 'a trader'. Banjaras claim that they have descended from the Mota and Mola who descended form the cows' guardian of Lord Krishna. The original home was Rajastan, because the folk tales narrates that their original home is Rajasthan. They have long history of migration. Another story also depicts that these people have come as assistants to the Mogul kings Calvary. Since Moguls were doing wars all over India, they also spread and settled in most of the states of India. The language belongs to Central group of Indo-Aryan family of

languages. Many works on the language had been carried out either cultural perspective or anthropological perspective. Very few works were carried out on the language especially in Telangana.

Keywords: Morphophonemic, Banjara Language, Assimilation

Review of Literature

For this kind of interdisciplinary research, a researcher needs to have orientation in multiple fields of knowledge. Hence, certain books and research works in the fields of Lambada history, culture, literature, and language are studied as background knowledge for this study.

Uma Shankar Upadyay (1975) explained the grammar of Lambada language. The book included phonology, morphology, syntax and vocabulary. B. Cheeniya Naik (1998), in his book 'Banjara charithra samskruti pragathi', described various aspects such as history, tradition and culture of Lambada community. This book is a comprehensive description in Telugu language about life and living of Lambadas. Apart from these books, there are some studies to focus in this background. Most of these are written in Telugu language.

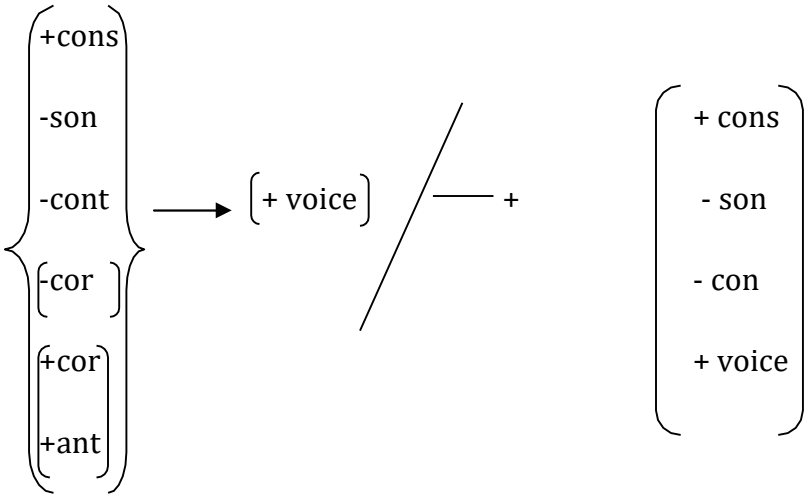
Methodology:

For the work, the data is being collected from the Nalgonda district Telangana. Scheme for Protection and Preservation of Endangered Languages (SPPEL) questionnaire designed by Central Institute of Indian Languages by a group of Experts is used for collecting data. Male female informants are selected from the village who have less experience of travelling to other places. Fifteen male and fifteen female informants were interviewed and elucidated the data. Along with the informants, I myself speak the language and corrected wherever necessary.

Morphophonemic changes in Banjara involve assimilation, loss and the addition of phonemes.

Regressive assimilation

Voiceless final consonants /p/, /t/, /k/ and /c/ of the verb stems become voiced before voiced initial consonants of the following stems.

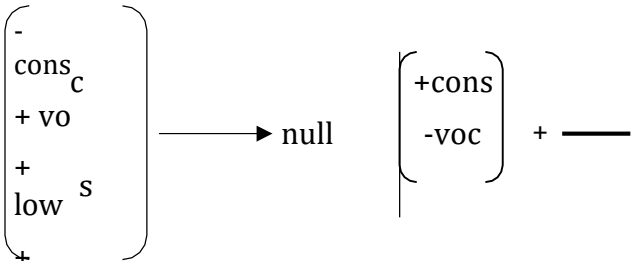


Example:

git	'song'	*gid bol*	'sing'
dhok	'pray'	*dhog de*	'worship'
puc	'reach'	*puj go*	'reached'
ba:p	'father'	*ba:be:Ta*	'father and son'

Loss of phonemes

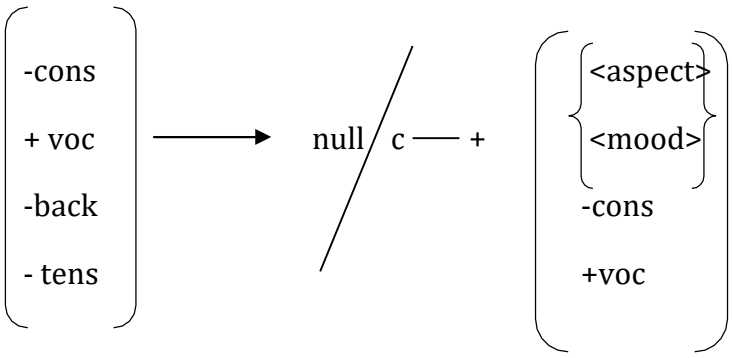
Initial -a is lost after stem final vowels



Example:

gawDi	'cow'	an	'and'	ghoDo	'horse'	gawDinghoDo	'cow and horse'
kothri	'female dog'	an	'and'	kothra	'male dog'	kothrinkothr	'female dog and male dog'
be:Ti	'daughter'	an	'and'	be:Ta	'son'	be:Tinbe:Ta	'daughter and son'

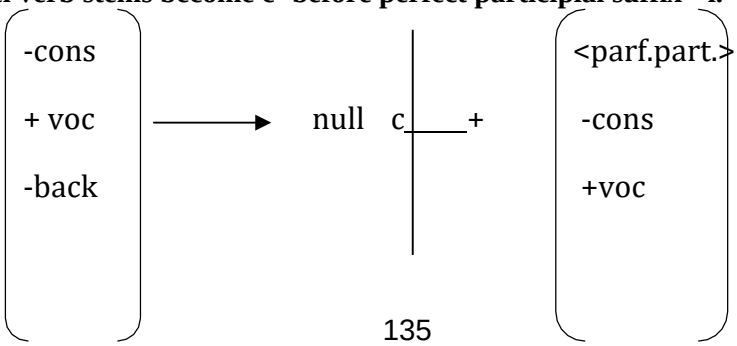
In ce verb stems ce becomes c before vowel initial aspect/mood markers.



Example:

le + ũ - lũ '(I) take'
de + i- ũ - diũ ' I will give'
we + i-e - wie 'will happen'

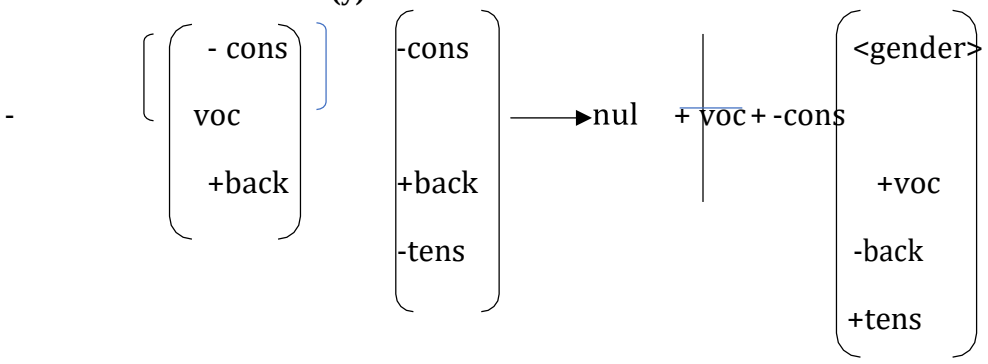
ci verb stems become c- before perfect participial suffix –i.



+ tense + tense

Example: pi 'drink + i pi 'having drunk'

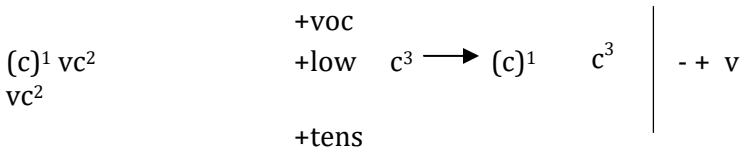
Stem final – (y) a/o are lost before feminine suffix-i.



Example:

nankya	'small' + i	*nanki*	'small' (fem)
chora	'boy' +i	*chori*	'girl'
ghoDo	'horse'+I	*ghoDi*	'mare'

$(c)^1$ vc^2 ac^3 becomes $(c)^1vc^2$ c^3 followed by any vowel.



Example
:

nikaL	'come out'	*nikLa*	'to drive out'
pakaD	'hold'	*pakDo*	'caught'
aTak	'obstruct'	*aTka*	'to obstruct'

Addition of phonemes

'come' a – becomes 'aw' before verb initial g- and imperfect -t.

a 'come + gi 'went' awgi 'she came'
a 'come + te - ne awtene 'to coming one'

Miscellaneous

In allegro speech the following changes occur.

D + i becomes LL and D + n/N becomes NN.

SD:
$$\begin{pmatrix} +\text{cons} \\ -\text{son} \\ -\text{cont} \\ +\text{cor} \\ -\text{ant} \\ +\text{voice} \end{pmatrix} \quad \left\{ \begin{pmatrix} +\text{voc} \\ +\text{cons} \\ +\text{lat} \\ +\text{ant} \end{pmatrix} \; \begin{pmatrix} +\text{voc} \\ +\text{son} \\ +\text{cor} \end{pmatrix} \right\}$$

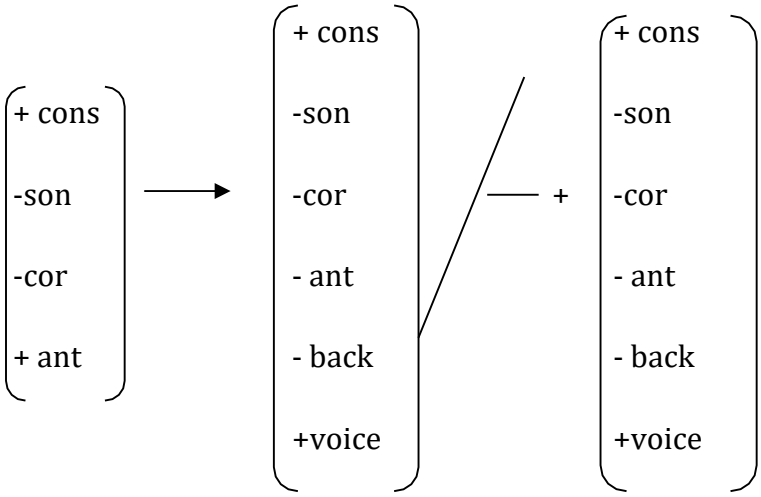
SC: 1
$$\left\{ \begin{pmatrix} +\text{voc} \\ +\text{cons} \\ +\text{lat} \\ -\text{ant} \end{pmatrix} \; \begin{pmatrix} +\text{voc} \\ +\text{son} \\ +\text{cor} \\ -\text{ant} \end{pmatrix} \right\} \longrightarrow 2 \left\{ \begin{pmatrix} +\text{voc} \\ +\text{cons} \\ +\text{lat} \\ -\text{ant} \end{pmatrix} \; \begin{pmatrix} +\text{voc} \\ +\text{son} \\ +\text{cor} \\ -\text{ant} \end{pmatrix} \right\}$$

Example:

toDlen	toLLen	'having
broken' kaDnak		kaNNak
	'open'	

kaDno kaNNo 'should open'

Phonemes /t/. /d/ and /s/ are assimilated to the following /j/.



Example:

mat	'not'	+	jo	'go'	majjo 'do not go!'
kud	'kump'	+	jo	'go'	kujjo '(you) jump!'
dhas	'runs'	+	jo	'go'	dhajjo 'run away'

Conclusion:

Morphophonemic changes in Banjara involve assimilation, loss and the addition of phonemes. Voiceless final consonants
/p/, /t/, /k/ and /c/ of the verb stems become voiced before voiced initial consonants of the following stems. Initial -a is lost after stem final vowels. In ce verb stems ce becomes c before vowel initial aspect/mood markers. ci verb stems become c- before perfect participial suffix – i. 'come' a – becomes 'aw' before verb initial g- and imperfect -t. D + i becomes LL and D + n/N becomes NN. Phonemes /t/. /d/ and /s/ are assimilated to the following /j/.

References

Umashankar Upadyay ph.D thesis. (1975). Department of Linguistics, A Descriptive study of the Banjara Language grammar texts &vocabulary.

Krishnamurti, Bh. (1957). Sandi in colloquial Telugu. Indian Linguistics17, 178-180.

_____and J.P.L .Gwaynn.(1985). A Grammar of modern Telugu.Delhi:Oxford University press.

Naik, D. B. (2000). *The art and literature of Banjara Lambanis*: a socio-cultural study. Abhinav Publications

Krishnamurti.Bh. (2003) .The Dravidian Languages.Cambridge: CUP.

Nagamma Reddy.K. (1979). problems of syllable-division in Telugu.work in pogress12.135-140. Linguistics Department, Edinburg University.

Urdu Verb Agreements with Personal Pronouns Within Distinct Cases

Karishma Khatoon

Research Scholar, Department of Linguistics, AMU Aligarh.

ABSTRACT

The paper demonstrates Urdu cases with a personal pronoun and their verb agreements. Here, an attempt has been made to know how verb agreement varies within different cases along with different pronouns. In this paper, mainly four cases have been studied that are nominative, accusative/dative, ergative, and genitive case. This aspect is crucial to know to learn a new language (Urdu).

The case indicates the grammatical function of a noun and pronoun. Agreement refers to variation in verb forms; its variation not only depends upon the subject but distinct grammatical cases which show little changes in verb agreements. It is not necessary that verb agreement always keep changing, rather it remains the same in few grammatical cases. It has been briefly explained in this proposed paper. The paper intends to demystify how verb agreement varies with personal pronouns within different cases.

Keywords: Verb agreement, Case, Grammatical function, Nominative, Accusative, Ergative, Genitive, etc.

Introduction

The paper has been written by keeping in mind the difficulties of English speakers who learn the Urdu language as a second or foreign language. Urdu grammar differs from English grammar in various ways. Urdu is a lingua Franca of Pakistan. And one of the official languages of India. Verb agreement refers to a system in which the form of the verb reflects the person and number.

A language applies two important devices case and agreement to identify grammatical relations within a sentence. The purpose of this paper to demystify where and how verb agreement changes in a sentence within different cases in consonance with a personal pronoun. There are such cases where conjugation remains the same.

Noun and Pronoun function as a subject of a sentence and the subject interprets various roles within clauses like the role of subject, direct object, or possessor, etc. In other words, the case expresses relation upon one word to another within a sentence. Case grammatically describe the function of a noun phrase that noun phrase can be a noun and pronoun.

There are six cases found in the Urdu language such as nominative (ø), ergative (ne), accusative (ko), genitive (k-), instrumental (se), and locative (me, par, tak, se, etc.). In this paper, the focus is limited to four cases that are nominative, ergative, accusative, and genitive. With a noun, the case marker is usually written as a separate word but with a pronoun, the case marker is tagged.

Here, in this paper only the present form of verb agreement is shown, it may also change into past and future tense and in the subjunctive mood by adding affixes. However, the focus is on the present form of verb agreement that agrees with the pronoun in person and number and for the negative construction need to add the „nəhı‟ form before the main verb.

Review of Literature

The concept of case is used to characterize the interaction between verbal lexical-semantic, grammatical relation, and word order.

Blake (2001:1) states that *‘‘case is a system of marking dependent nouns for the type of relationship they bear to their heads’’*. As cited in Butt, M. (2005). p.4.

Richards & Schmidt (2013) asserted that *‘‘ case is a grammatical category that shows the function of the noun or a noun phrase in a sentence’’.*

Verb agreement refers to inflection in verb forms to express person, number, tense, aspect. To know the person and number, in more than one tone verb must agree otherwise,

sentences will seem to be ambiguous for the listener or reader. For the learners, it is inevitable to know that when verb inflection should change and where it will be the same, and with which component of a clause or sentence verb will agree.

Lehmann (1988 Haspelmath; 2002:65ff) states*: "the term agreement refers to describe a situation in which the grammatical feature of a noun/noun phrase determines the morphological shape of a word that is syntactically related to the N or NP someway"*

However, verb agreement varies according to the subject of the sentence and the subject plays different roles in pursuance of different cases.

Methodology

The learners of this proposed study are American English speakers learning Urdu at AIIS (educational institute). The sample size is 40 from a homogenous group. Only a few examples and exercises have been selected to explain verb agreement within Nominative, Accusative/Dative, Ergative, and Genetive cases. It is based on text data and day-to-day class observation of learners‟ performance.

1. Nominative case

The nominative case is used for the subject (whether noun or pronoun) of a verb. It is crucial to mention that the paper is restricted to pronouns.

Kachru (1980) demonstrates that *"Nominative direct case is phonologically null as it does not take any clitic. It can appear in the subject position as well as in the object position* (Kachro, 1980 in Islam, Akhtar & Bukhari, 2009 as cited in Iqbal, M., Mangrio, R. A. and Mustafa, R. E. 2016. p. 138).

The thing here we have to note that, the verb agreement varies according to person and number within the nominative case as stated above and mentioned below in the table.

	Nominative		Verb Inflection	
	Singular	**Plural**	**Singular**	**Plural**
1st Person	maı	hum/humlog	h	haı
2nd Person	tuːm/aːp/tuː og	tuːm/tuːml /aːp/aːplog	ho/haı/haı	ho/ haı
3rd Person	yəh/vəh	ye/yelog/voː/voːlog	haı	haı

Here a few examples of how the nominative case and its verb agree with the subject of a sentence in the Urdu language has given. In the nominative case, the verb directly agrees with the subject of the sentence (pronoun). However, verb inflection also changing just as personal pronouns changing in person and number.

a. maı kuːtte se dərta hːː.
b. tuːm school jate ho.
c. aːp mɔːsıqı sunte haı.
d. hum ahːː nəhiː rehte haı.
e. ye jəgəh bəhuːt achchiː haı.

2. Accusative/Dative Case

The accusative case is identical to the dative (ko). According to Mahajan (1990), Davison

(1998) *"It has been observed that in Urdu/Hindi lack accusative case and (ko) is an inherent dative case".*

Islam Akhtar & Bukhari (2009) explain: *"Urdu ko dative, as well as accusative, is homophonous. It is very complex in its usage".* (as cited in Iqbal, M., Mangrio, R.A., and Mustafa, R.E. 2016. p.133).

In the accusative/dative case verb agreement remains the same with all pronouns because its subject is attached with the postposition "ko/e" and the postposition blocks the direct agreement with the subject. It marks a goal or specifies an experiencer.

	Accusative		**Verb Inflection**	
	Singular	**Plural**	**Singular**	**Plural**
1st Person	mujh+ko/e	hum+ko/	haɪ	haɪ
2ndPerson	tujh ko/e	tum ko tumh	haɪ	haɪ
		a:pko	haɪ	haɪ
3rd Person	is/us/+ko/e	in/un ko inh/unh	haɪ	haɪ

There is a slight difference between the accusative and dative cases; the accusative case has a direct object of the verb and the dative case has an indirect object of a verb. Dative case simultaneously takes (_e and _ko) markers.

Accusative and dative case markers do not use with the demonstrative pronoun. There is a difference between the use of (_e and _ko) (_ko) can be used with noun and pronoun preceded by object but in case (_e) can only use with pronoun preceded by an object.

a. mujhe do qələm cha:hɪye.
b. a:pko bukha:r haɪ.
c. tumh e kurta pəsənd haɪ?
d. humko fu:tball khelna pəsənd haɪ.
e. ɪse kɪta:b de do.
f. mujhe tumha:re ba:re me ma:lum haɪ.

3. Ergative Case

An ergative case has (ne) marker that always comes with the subject. It is tagged with a pronoun and stands alone with a noun. In the ergative case, the verb does not agree with the

subject of the sentence rather it agrees with the object in person and number. The ergative case is mostly used in the perfect tense and with both transitive and intransitive verbs.

	Ergative		**Verb Inflection**	
	Singular	**Plural**	**Singular**	**Plural**
1st Person	maı+ ne	hum+ne	-	-
2nd Person	tum/tu/aːp ne	tum/aːp ne	-	-
3rd Person	is/us/+ne	inh /unh +ne	-	-

Here we can see in these examples a few sentences are in the present perfect tense and some of the simple past. If sentences will be in simple past, there is no use of verb agreement. If sentences in the present perfect tense, they must agree with the object. Therefore, the verb will agree with the object of the sentence in Person, number, and gender. In the simple past, the main verb will agree with the object in person, number, and gender and end with "aː e ı" we can see in (d) and (e). The (ne) mark blocks the direct agreement with the subject.

a. unh ne əwaːm se vaːda kıya haı.

b. maıne photoz khıːnche haı.

c. tumne coffiː bənaːi.

d. humne fılm khətəm nəhı kı.

e. aːpne saːdı khərıdiː?

f. mene khaːna nəhı khaːya haı.

4. Genitive Case

Here one thing important to note that in the genitive case pronoun is always preceded by a noun hence, they stand together as a subject of a sentence (ie. meri kitab). In this case, the verb agrees with the subject in number as well as the main verb or adjective agrees with the object by adding (_i) with feminine, (_a) with masculine, and (_e) with plural form whether

masculine or feminine. If the subject of the sentence will be in the singular form like (meri kitab) verb inflection will be in the singular form. The verb inflection depends on the noun which is followed by a pronoun. If it is in singular form verb must be in the singular form, if it is in the plural form then it must be in the plural form.

	Genetive		**Verb Inflection**	
	Singular	**Plural**	**Singular**	**Plural**
1st Person	mer+a+i+e	humar+a+i+e	haı/haı	haı/haı
2nd Person	tumhar+a+i+e	tumahar+a+i+e	haı/haı	haı/haı
	a:pk a i e	a:pk a i+e	haı/haı	haı/haı
3rd Person	isk+a+i+e	unk+a+i+e	haı/haı	haı/haı

Sometimes, in the genitive case, the subject can require such terms as meri ma, mere bade Bhai, uski Ammi, and so on under such condition verb inflection change same as plural forms. Here (ka) is inflected like an adjective to agree with governing nouns in gender and number.

In the genitive case, one more pronoun “apna” is used for 1st person singular “mera”. It is used when the action of the verbs refers to the subject of the sentence i.e. (vo apni kita:b pədhta hai).

a. merı kıta:b phətı hai.
b. humare usta:d Hındusta:n ke ba:re me bətate haı.
c. mere bəde bha:ı ke do bache haı.
d. uski ammı bi:mar haı
e. tumha:ra ghar chota hai.

Findings

In nominative case there is no postposition attached with pronoun therefore verb must agree with the subject in number and person. In all three cases (accusative/dative, ergative, and genitive) subject must have a postposition so, the verb doesn't need directly to agree only with the subjects of a sentence. It is crucial to point out that in the Urdu language there are various honorific terms used in daily conversation whether noun or pronoun. The honorific terms require a plural form of verb inflection even if it is in a singular form. In Urdu, we have various honorific terms in 2nd person singular and plural like a:p a:plog consist of same verb inflectional form as mentioned above. On the other side in the accusative/dative case verb agreement is blocked so verb inflection remains the same with all pronouns. When the subject strat with mujhe hume inh , unh
, etc. there must be an inherent "ko".

The ergative case demonstrates that the verb agrees with the object. If the object is in the plural form verb inflection change in plural form. The genitive case is a little bit complex in the use of verb agreement as compared to other cases. Here verb agrees with the subject and the subject is formed by using a pronoun followed by a noun.

Conclusion

Based on the above discussion it has been concluded that If the subject is attached with any postposition the verb will not agree with the subject the postposition block the direct agreement with the subject in such cases verb agree with an object and if the subject is not attached with postposition verb must agree with the subject.

Verb agreement is a crucial part of a language if learners use it incorrectly it will be caught by the language teacher at once. Urdu is structurally very different from English so, it is difficult for Urdu learners but the presented study tried to

demystify the verb agreement with pronouns within distinct cases.

As mentioned above learners are native speakers of English so they found it difficult to learn verb agreement within different cases. In the matter of difficulty their native language interfere in other words they use just two or three verb inflection with all pronouns which lie in different cases. It is based on the observation when the teacher guides them to use the correct verb agreement in simple conversation they follow it with all pronouns as well as cases. However, the paper elaborates and clarifies how verb agreement varies within different cases.

It has been found in this paper that verb agreement does not remain the same it requires changes according to case so, by keeping in mind the use of distinct cases learners would use it correctly in an attempt to communicate.

References

Kroeger, P. R. (2005). Analyzing Grammar: An Introduction. New York: Cambridge University Press.

Iqbal, M., Mangrio, R. A., and Mustafa, R. E. (2016). Punjabi *nuN* and Urdu *Ko*: A Morpho- Semantic Analysis. *Pakistan Journal of Languages and Translation Studies*. Vol. (4). 133-155.

Butt, M. (2006). Theories of Case. New York: Cambridge University Press.

Ahmad, M. (2013). Gender System in Urdu Language [Power Point Slide].

Ilyas, M. and Shibani, M. A. (2019). Asymmetrical Subject-Verb Agreement in Standard Arabic and Urdu Languages: A Comparative Study. *International Journal of Innovation*, *Creativity*, *and Change*. Vol. (7). p. 228-243. Issue 5, 2019.

Koul, O. N. (2008). Modern Hindi Grammar. United State of America: Dunwoody Press.

Zeb, F. (2019). Nominative and Accusative Case in English and Urdu: A Comparative Study. Pakistan Journal of Humanities and Social Science. Vol. (7). No. 4. p. 365-377.

Butt, J. M. (1993). The Structure of Complex Predicates in Urdu. Leland Standford Junior University: United State Press.

Plural Formation in Lambadi Language

Krishna Banothu[1] and C.S. Swathi[2]

(1). Research Scholar, Department of Linguistics, Osmania University,
(2). Asst. Professor, Department of Linguistics, Osmania University.

ABSTRACT

Lambadi is a nomadic tribal community which is spread all over India. Lambadi language belongs to Indo-Aryan language family and is spoken in Telangana, Andhra Pradesh, Karnataka, Rajasthan and Maharashtra. Although it belongs to Indo-Aryan language family, it is also widely spoken in Dravidian family geographical areas and has been influenced by these languages. Pluralization is the way of transforming the words singular nouns into plural forms and it is a part of noun morphology. The aim of this study is to describe the various processes of plural formation in Lambadi language spoken in Telangana. In this study words ending with Obstruents and vowels were only considered. For this purpose, the data was collected from a native speaker of Lambadi, using word list from CIIL, Mysore. The data obtained was verified by the first author to validate it. The descriptive analysis of the data revealed that in Lambadi language, pluralization is formed by suffixation. Nominal words ending in Obstruents (stops, fricatives and affricates) and Laterals, the plural marker /e/ is being added. For the singular nouns ending with the front vowel /i/ take the back vowel /o/ as a plural suffix. For words ending with the low vowel /a/, back high rounded vowel /u/ is suffixed to form plural. It is also observed that if gemination occurred in the prefinal position, while ending in 'o', it took the plural suffix as /e/. In conclusion, it can be noted that plural suffixes are added to the singular nominal bases. The plural suffixes are principal inflectional nominal suffixes and can be added to all the nominal bases.

Keywords: Lambadi Language, Pluralization, Suffixation

1. Introduction

Lambadi is a nomadic tribal community which has spread all over India. Lambadi language belongs to the Indo-Aryan language family and is spoken in Telangana, Andhra Pradesh, Karnataka, Rajasthan and Maharashtra. Although it belongs to the Indo-Aryan language family, it is also widely spoken in geographical areas where Dravidian languages are spoken and as a result, it has been influenced by these languages. The

descriptive analysis of the data reveals that in Lambadi language, plural nouns are formed by suffixation and substituting vowels.

2. Review of Literature

There are some studies related to different aspects of the concept of plural formation.

Krishnamurthy and Gwynn (1985) in his work 'A Grammar of Modern Telugu' discusses various rules of plural formation in Telugu that primarily includes the syllable [lu/Lu] to convert singular stems into plural nouns, e.g. [a:wu] 'cow' changes to [a:vulu], [baNTu] 'soldier' changes to [baNTlu], [paNDu] 'fruit' changes to [paNDLu] 'fruits' etc. Vidyasagar (2003) also carried out an extensive research on 'Plural in Telugu Nouns' where he discusses how singular stems change their forms when plural morpheme '-lu' is added e.g. [ceppu] – [ceppulu], [ma:ta]-[ma:talu] etc. There have been multiple researches carried out by various linguistic levels in Lambadi, but no extensive research has been done on plural formation in Lambadi. Therefore, the researcher felt the need to study on plural formation in Lambadi. This paper emphasizes on the various processes of pluralization by suffixation.

3. Methodology

For this purpose, the data was collected from a native speaker of Lambadi, using word list from CIIL, Mysore. The data obtained was verified by the first author to validate it.

4. Pluralization by Substituting Vowels

When a singular stem ends with the vowel (i.e., /o/), it is substituted by the vowels like /a/ and /e/ in Lambadi.

In the following data, the vowel /o/ in singular changes to /a/ in plural forms.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[kanabɖo]	'temple'	[kanabɖa]	'temples'
	[ʈ̪apa:ɭo]	'forehead'	[ʈ̪apa:ɭa]	'foreheads'
	[a:nkiɖaɭo]	'eye'	[a:nkiɖaɭa]	'eyes'
	[munɖo]	'face'	[munɖa]	'faces'
	[ma:ɭo]	'nest'	[ma:ɭa]	'nests'
	[goɖo]	'knee'	[goɖa]	'knees'

Rule-1: In the words ending with the vowel '/o/' in singular, it changes to /a/ form into plural counterparts.

In the following data, the vowel /o/ in singular changes to /e/ to form plural noun.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[ma:ʈ̪ʈ̪^{h}o]	'head'	[ma:ʈ̪ʈ̪^{h}e]	'heads'
	[b^{h}edʒdʒo]	'brain'	[b^{h}edʒdʒe]	'brains'
	[puʈ̪o]	'back'	[puʈ̪e]	'backs'
	[kãnkko]	'armpit'	[kãnkke]	'armpits'
	[poʈ̪ʈ̪o]	'grandson'	[poʈ̪ʈ̪e]	
		'grandson'		

Rule-2: Words ending with the vowel '/o/' in singular and preceded by geminated consonants take the plural marker /e/ to change to plural forms.

5. Plural formation by suffixation

There were different patterns observed for words ending in vowels and consonants. There are plural suffixes in Lambadi and these are '-e', '-u' and '-o'. These are predominantly morphologically conditioned.

5.1: Words ending in Vowels /a/ and /i/

In the following data, When the singular nouns end with the low vowel /a/, /u/ is suffixed to form a singular noun into a plural noun in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[haɖka]	'bone'	[haɖkau]	'bones'
	[ʈaɭva]	'foot'	[ʈaɭvau]	'feet'
	[ki:ɖa]	'insect'	[ki:ɖau]	'insects'
	[mi:na]	'month'	[mi:nau]	'months'

Rule 3 : In Words ending with the low vowel /a/, back high rounded vowel /u/ is suffixed to obtain the plural forms.

The following data shows that the singular nouns ending with the front vowel /i/ take the back vowel /o/ as a plural suffix to change them into plural nouns in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[ɖa:ɖi]	'beard'	[ɖa:ɖio]	'beards'
	[musɖi]	'gum'	[musɖio]	'gums'
	[k^huɳi]	'elbow'	[k^huɳio]	'elbows'
	[lãnɖi]	'neck'	[lãnɖio]	'necks'
	[a:ŋgɭi]	'finger'	[a:ŋgɭio]	'fingers'
	[haʈɭi]	'palm'	[haʈɭio]	'palms'
	[muʈʈi]	'fist'	[muʈʈio]	'fists'

Rule - 4: Words ending in [i] – the suffix [o] is added to obtain the plural forms.

5.2. Words ending in Obstruents

The following data shows that the singular nouns ending with the voiceless bilabial stop /p/ and /b/ take the front vowel
/e/ as a plural suffix to turn into plural nouns in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[sa:p]	'snake'	[sa:pe]	'snakes'
	[ba:p]	'father'	[ba:pe]	'fathers'
	[dʒi:b]	'tongue'	[dʒi:be]	'tongues'
	[d^ha:b]	'well'	[d^ha:be]	'wells'

As per the data given below, the singular nouns ending with the voiceless dental stop /t̪/ take the front vowel /e/ as a plural suffix to change into plural nouns in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[d̪ a:nt̪]	'tooth'	[d̪ a:nt̪e]	'teeth'
	[ha:t̪]	'hand'	[ha:t̪e]	'hands'
	[dʒa:t̪]	'clan'	[dʒa:t̪e]	'clans'
	[ra:t̪]	'night'	[ra:t̪e]	'nights'
	[gammat̪]	'joke'	[gammat̪e]	'jokes'
	[ka:nd]	'shoulder'	[ka:nde]	'shoulders'
	[go:d]	'lap'	[go:d]	'lap'

As per the data given below, the singular nouns ending with the voiceless dental stop /t/ and voiced retroflex stop /ɖ/ take the front vowel /e/ as a plural suffix to form plural nouns in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[ho:t]	'lip'	[ho:te]	'lips'
	[po:t]	'luggage'	[po:te]	'luggage'
	[tʃo:t]	'beat'	[tʃo:te]	'beats'
	[mo:t]	'elder'	[mo:te]	'elders'
	[gaɖ]	'fort'	[gaɖe]	'forts'
	[dʒaɖ]	'root'	[dʒaɖe]	'roots'

As per the data given below, the singular nouns ending with the voiceless velar stop /k/ and voiced velar stop /g/ take the front vowel /e/ as a plural suffix to change them into plural nouns in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[no:k]	'nail'	[no:ke]	'nails'
	[hi:k]	'chest'	[hi:ke]	'chests'
	[pa:k]	'ripen'	[pa:ke]	'ripens'
	[g^{h}o:g]	'throat'	[g^{h}o:ge]	'throats'
	[dʒo:g]	'leech'	[dʒo:ge]	'leeches'
	[ta:ŋg]	'leg'	[ta:ŋge]	'legs'
	[dʒa:ŋg]	'thigh'	[dʒa:ŋge]	'thigh'

Words ending in both voiceless and voiced velar stops, suffix [e] is added to obtain the plural forms.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	Gloss
	[kass]	'armlet'	[kasse]	'armlets'
	[ve:s]	'banjara dress'	[ve:se]	'banjara dresses'

Words ending in voiceless fricative, suffix [e] is added to obtain the plural forms.

According to the below data, the singular nouns ending with the voiceless post-alveolar affricate /tʃ/ and voiced post-alveolar affricate /dʒ/ take the front vowel /e/ as a plural suffix to change to plural nouns in Lambadi.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[pãntʃ]	'dhoti'	[pãntʃe]	'dhotis'
	[ka:ntʃ]	'mirror of the blouse'	[ka:ntʃe]	'mirrors of the blouse'
	[ri:ntʃ]	'bear'	[ri:ntʃe]	'bears'
	[ga:dʒ]	'mushroom'	[ga:dʒe]	'mushrooms'
	[ra:dʒ]	'king'	[ra:dʒe]	'kings'
	[ma:dʒ]	'enjoyment'	[ma:dʒe]	'enjoyments'
	[mo:dʒ]	'measurement'	[mo:dʒe]	'measurements'

Words ending in both voiceless and voiced affricates - [e] suffix is added to obtain the plural forms.

Rule 5: All words ending with Obstruents, the plural marker /e/ is suffixed, to form the plural words.

5.3: Words ending in /l/ and /ɭ/

The singular nouns ending with /l/ and /ɭ/ are suffixed with /e/ to form plural forms.

e.g.	**Lambadi**	**Gloss**	**Lambadi**	**Gloss**
	[ga:l]	'cheek'	[ga:le]	'cheeks'
	[ma:l]	'grocery'	[ma:le]	'groceries'
	[wa:ɭ]	'bend'	[wa:ɭe]	'bends'

Rule - 6: Words ending in lateral approximant, [e] suffix is added to obtain the plural forms.

Conclusion

The aim of this paper is to discuss the process of pluralization in Lambadi language and it is observed that the system of pluralization is very distinct and simple in Lambadi. Basically, the plural suffixes and substitution play the major role in transferring from singular to plural at word level. The plural marker /e/ is added to nominal words ending in Obstruents (stops, fricatives and affricates) and Laterals. In words ending in
/i/, /a/ suffix is added and in words ending in /o/, /o/ becomes /a/, and in words ending with /a/, it changes to /u/. It is also observed that if gemination occurred in the prefinal position, while words ending in 'o', it takes the plural suffix as /e/.

References

Krishnamurthy, Bh. & Gwynn (1985). A Grammar of Modern Telugu. Oxford University Press. New Delhi.

Hunter, W.W. (1976). A Comparative Vocabulary of the Languages of India and High Asia. New Delhi: Cosmo Publications.

Lal, B. (2012). Historical Facts about Migration of Banjaras to Hyderabad State.

Masica, C.P. (1999). *The Indo-Aryan Languages*. Cambridge University Press.

Naik, D. B. (2000). The art and literature of Banjara Lambanis: a socio-cultural study. Abhinav Publications.

Ronald L.Trail. (1970). *The Grammar of Lambani*. Published by Summer Institute of Linguistics in the University of Oklahoma, Norman.

Vidyasagar, J. (2003). 'Plural in Telugu Nouns'. Osmania Papers in Linguistics.

Language and Power: The Role of Language in Creating Political Demagogues

Gabriel Vinod Kumar Dolla

Faculty, OHS, ECISD, Texas

ABSTRACT

"Language is power, life and the instrument of culture, the instrument of domination and liberation" -Angela Carter

The study of the relationship between "language, power" and social relations has clearly shifted from proving the existence of this relationship, to probing and understanding its underpinnings and implications. Most linguists now agree that the relationship between "language and power" is a mutual relationship. Powerful institutions and individuals use language as both a means to construct their power and as a way to maintain it. Language thus becomes necessary for the maintenance of power, and the power and effect of language in turn rely on the power of individuals and institutions themselves.

Language is delineated as "a social practice" (Fowler 61), by which power relations are established and sustained. This supports Norman Fairclough's (1989) view that power is not only built and sustained via coercive means (by force), but also via indirect ways (the use of language). Besides institutional power, or the power exercised by entities that are overtly recognized for holding a position of authority (the police, for example), there also exist other types of power relations: between family members, between educated and uneducated people, and so forth. Individuals and groups in this category of power relations use language as their main tool for maintaining status and power.

A father in a phallocentric family, for example, would not forcibly have to resort to force in order to impose his authority over the other members of his family. Consider the following example taken from a short story entitled A Meeting in the Dark, by African writer Ngugiwa Thiong'o:

> *Sit down. Where are you going?'*
>
> *'For a walk, Father,' he answered evasively.*
>
> *'To the village?'*
>
> *'Well-yes-no. I mean, nowhere in particular.*

We can clearly see in this short conversation that the father is exerting his authority on his son. The father here does not resort to force while talking to

his son, something that we can deduce from the absence of any textual reference by the author to the use of power by the father. On the contrary, the father resorts to a straightforward, strict style of address with his son ('Sit down. Where are you going?').

The style of speech deployed in the example is clearly loaded with power and authority, bearing in mind that the context in which this conversation takes place is within a patriarchal family, in which the father, who is a fervent preacher, is the member in control. The example above supports Fairclough's (1989) view that power relations are not class-bound, and can be manifested though the use of language during social interactions between individuals.

Other scholars, namely Fowler (1985) and Kramarae, Schultz and O'Barr (1984), have supported the view that power is developed and maintained via interaction (the social practice of language). These scholars argue that language or discourse serve to construct and manipulate concepts of power in society. The discursive construction of power by institutions and individuals can also lead to changing social practices and realities (Mayr 4). The way institutions and individuals use language to consolidate and promote their ideological interests, or simply to preserve their authority over other institutions and individuals, changes social relationships (between married couples, for example) or social practices (e.g., the way one does his/her work).

For instance, an uneducated person who is exposed daily to a media content that uses language to promote patriarchal principles is likely to be influenced by such ideas, and might consequently put those principles into practice within his or her own marital life. The media in this case serves certain ideological purposes that might pertain to a dominant community. Hence, these types of media use language in such a way they change people's perceptions of and attitudes towards their social practices and relationships.

Fairclough (1989), Fowler (1985) and Kramarae, Schultz and O'Barr (1984) all agree that language is "a social practice," and that power is constructed and developed via social interactions marked by hierarchy and asymmetry. However, we can still draw a distinction between two different uses of language in the context of power relationships: language as public discourse, and language as private discourse. The former refers to the language used by powerful public institutions through the mass media. The power exercised via this public use of language usually stems from to governments and political parties. Those powerful institutions use language or "public discourse" to construct and promote their dominance, by producing knowledge about society and advertising a given social (Mayr 3).

The press, for example, uses a variety of techniques to change the way people perceive of and think about a certain social phenomenon. Fowler (1979) presents a "checklist" that students of Critical Linguistics can refer to in their analysis of the functions that discourse serves in the context of power relations.

Keywords: Language, Power, Transitivity, Media, Demagogues

Introduction

Transitivity is one main feature included in Fowler's (1979) checklist of categories of structure. It generally refers to how meaning is represented in a clause. In the print media, for instance, choices of transitivity structures are not arbitrary. They serve a certain ideological purpose that is clearly manifested in how different newspaper headlines, for example, report the same events, but with dissimilar focuses and interests. Below are two headlines from two different American daily newspapers reporting the same piece of news:

The New York Magazine

Iraq forces suspension of U.S. surveillance flights

USA Today

U.N. Withdraws U-2 Planes

It is clear that each headline reflects the same events. As far as transitivity structure is concerned here, the ways participants in each clause are structured, as well as the words chosen to report the occurrence, are illustrative of how language is used differently to promote or reflect different interests. The New York Magazine structured its headline in such a way that the reader would understand that the agent of action in the happening is "Iraq forces." On the other hand, USA Today structured its headline dissimilarly, in such a way that the agent of action is rather the U.N., rather than 'Iraq Forces,' as contrarily reported by The New York Magazine. Though the two newspapers are reporting the same event, this discrepancy between the headlines is not accidental. This difference of focus

demonstrates Foucault's definition of discourse, which is the way knowledge and reality are constructed. Such discrepancy in the way discourse constructs reality is illustrative of the "power of language."

Other than public discourse, which is one way by which institutions and individuals construct and maintain power, private discourse is another way of using language for the same purpose. Private discourse refers to the language used by individuals in their interactions with one another. Like public discourse, private discourse is another way to maintain power relations among individuals.This type of discourse retrieves its power and legitimacy from the social roles individuals play in their society. One example of such manifestation of private discourse in social relationships is doctor-patient interactions. Consider the following example cited in Allen and Guy (1989:46):

Doctor: *You have diabetes? Your pancreas metabolism is impaired and its ability to manufacture insulin, and therefore, your digestive processes cannot utilize your intake of glucose.*

Patient: (to Nurse): *What did he say other than I have diabetes?*

Nurse: *Diabetes results when an organ in your body called pancreas is not able to manufacture insulin. Without insulin in adequate amounts, your body can't handle all the sugar you eat daily. The doctor plans to supply your body with insulin it needs so that the sugar you eat will be used up and not wasted.*

Patient: *So why didn't he say that in the first place?*

This extract from a doctor-patient interaction exemplifies how private discourse is used to maintain a certain power relation between interlocutors. D (Doctor) uses a highly technical jargon that is characteristic of his field of expertise to explain to P (Patient) about his/her health condition. Perhaps D

could have used the same discourse used by N to explain P's health condition. Rather, D favors a discourse not accessible to P, as it is clearly deducible from P's responses that P's knowledge of the medical field (D's area of expertise) is mediocre compared to D. D's use of language in the extract above is not random. D could at least consider the fact that P is a patient, and not his or her colleague in medicine. D's personal use of language denotes asymmetry and an interest in maintaining a doctor-patient distance (power relationship).

In contrast, N (Nurse) successfully infers that D's discourse is undecipherable for P, and thus uses a discourse that is more comprehensible for P. Therefore, we can say that N is relatively not interested in maintaining any power relationship between herself and P, as she resorts to "a communicative" use of language, which is aimed at producing understanding rather than maintaining power.

Analysis:

On January 15, 2022 the Chief Minister of Assam Dr. Himanta Biswa Sarma was criticized for "misbehaving" with an IAS Officer in public. The news channels showed Mr. Sarma standing on the highway with one bus and a truck waiting in front of him and many officials along with his personal security officer surrounding him. At one point he was heard saying " *SP ko bulao* [call the SP]" and then telling the Deputy Commissioner, " *Arre DC saab*, *ye kya natak hain? Kyu gaari rukwaya hain?* [DC sir, what's this drama? Why have you stopped the vehicles?]" " *Koi raja-mahara aa raha hain kya?* [Is any king coming?]" he added. When Mr. Hivare tried to say something, Mr. Sarma was heard saying: " Hatt*! Aisa mat karo aage! Logo ka kasht ho raha hain!* [Don't do this in future! People are suffering!)]"

Irrespective of the intent of the Chief Minister, using an unparlimentary word like "hatt" with a respected civil servant is

not just wrong but also sends out a wrong message to the entire bureaucracy. This is where the intermingling of language and power takes place: the use of an uncouth word like “hatt” places the user of the word in a higher position in the hierarchy of power and the receiver in a far lower position.

In the run-up to the 2014 Indian General Elections, Narendra Modi's speeches revealed his cunning communication skills that helped conceal his broader purpose. This is comparable to demagogues throughout history. The fascist project is carried out by securing the public's vote and their active political engagement, just like in other histories of fascism in other parts of the world. Fascism grows as a result of a leader's popular and catalytic language in the political public sphere, which manipulates popular opinion among the common people. The fact that so many Germans backed Hitler during the Third Reich is evidence of how popular and powerful the fascist ideology is.

Dictatorships have historically been dedicated to the goals of growth, progress, and development. The level of economic advancement attained under Stalin's administration was highly regarded. Russia under Stalin was a prime example of a powerful industrial state. But he qualifies as a mass killer due to his act of violence. The conflict of whether Stalin was a "villain" or a "hero" is something many Russians are still finding difficult to deal with.

Language, which turns into a site of power and violence in the political public domain, facilitates the agenda of demagogues. The language was the foundation for the crimes of fascism and stalinism. According to Roland Barthes, fascism forces speech rather than prohibiting it. Demagogic rhetoric combines exclusive nationalism and patriotism with the language of growth. They mention the poor and poverty frequently. They have a lot to conceal. They stifle options, diversity, and variety.

Modi depicts inflation and the struggling poor in vivid detail. The hearth is not lit, or "*Chulha nahin jalta*". He solely seems to discuss interpersonal difficulties. His conception of growth is elitist and superficial. His use of the metaphor of light alludes to a development lamp that fuels a consumerist ideology and culture. He uses Vivekananda's plainly spiritual rhetoric to legitimise his prejudiced notion of progress and a worldwide India combined with patriotism, which he defines via blood, sacrifice, and sweat. However, when it comes to addressing the social injustices prevalent in India, Modi's vision does not seem to take that into account. Modi's decision to revive Shastri is an attempt to temper his political language by using the kisan (peasant) only as a frame.

The former US President Donald Trump used many rhetorical linguistic devices to tap into the raw emotions of potential voters that helped him win the 2016 presidential elections. Here are some of the language mechanisms he used:

Braggadocio

To elicit an emotional response to his lofty promises, Trump frequently used superlatives. The size of the things was "yuuuuuge," not big. Not only would he make things great again, but "wonderful," "tremendous," and "the best." He wasn't just going to prevail. He promised that he would "win so much that you'll be sick and tired of winning".

Binaries

Trump employed the divisive "we and them" language to its most destructive potential. Either you support the jihadist, destructive Islam, or you support the Judeo-Christian west. You are either a "criminal alien" or a diligent, mistreated American. You are either "legal" or "illegal." But by the time he called for unity in his victory address, the "we and them" language from the campaign trail appeared to have lost some of its lustre.

Hyperbole

In addition to boasting about himself, Trump was also prone to dramatisation and exaggeration. When talking about one nation, Syria, Trump would say, "The whole world is burning apart." Three times in two lines during his victory speech, Reince Priebus was referred to as a "superstar"—a strong assertion for a man who is mostly unknown outside of the Washington, D.C.

Euphemisms

Trump had the ability to "play things down as well as up". Of course, his most notorious euphemism was "locker room banter," which he allegedly used to justify confessing to "grabbing women by the pussy" without their permission.

Conclusion:

Noam Chomsky, a political scientist, details how language is the currency of power in elections in his book *Language and Politics*. Voter persuasion depends on effective communication and persuasive speechwriting, and in democracies, the system demands that the populace buy into what politicians are saying in their campaign speeches. Live or televised debates have historically been used by politicians to influence the electorate, albeit social media is becoming a significant tool in their toolbox for doing so. They also employ language to develop catchphrases that can energise the electorate and take on the form of a chant or battle cry, such as MAGA (Make America Great Again), or *Bharat Mata ki Jai (*Hail Mother India). Political discourse is indexical, which means that every word used either reflects a political viewpoint overtly or implicitly. Even anything as small as an accent or how individuals are addressed could indicate this. Political discourse constantly aims to engage the audience through interruption, discussion, and negotiation. Political rhetoric also has a tendency to be ambiguous, almost like the words of a horoscope, leaving it up to the electorate to interpret it anyway they see fit. Therefore, it

is the responsibility of the listeners (receivers) to tread very carefully as the language of the politicians could eventually lead to unprecedented pandemonium.

References:

Dowling, Lee H. *Rocky Mountain Review of Language and Literature*, vol. 38, no. 1/2, 1984, pp. 96–97. JSTOR, https://doi.org/10.2307/1347169. Accessed 20 Oct. 2021.

Fairclough N. (1989). *Language and power*. Longman.

Fowler, Roger. *Language and control*. Routledge, 1975. Print: 61-62.

Dijk, Teun A. "Discourse, Power and Access." *Discourse and Power*. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan, 2008. Web: 84

Mayr, Andrea. *Language and Power: An Introduction to Institutional Discourse*. London: Continuum, 2008. Web: 4-8.

Thiong'o, Ngugi Wa. *A Meeting in the Dark*. 1674. Print: 3.

PTI. 2022, January. "Assam CM Himanta Biswa Sarma scolds Nagaon DC publicly for halting traffic for CM's convoy". *The Hindu*. https://www.thehindu.com/news/national/himanta-scolds- nagaon-dc-publicly-for-halting-traffic-for-cms- convoy/article38276686.ece

Barthes, Roland. *The Neutral: Lecture Course at the Collège de France*, 1977-1978.

Datta, Nonica. 2013, May. "The Language of Narendra Modi". *Economic & Political Weekly*. https://www.epw.in/journal/2013/18/web-exclusives/language-narendra-modi.html

Nunn, Gary. 2016, November. "Winning words: the language that got Donald Trump elected". *The Guardian*. https://www.theguardian.com/media/mind-your-

language/2016/nov/11/winning-words-the-language-that- got-donald-trump-elected

Chomsky, Noam. *Language and Politics*. 2004, March. A K Press.

Phenomenon of Code-Mixing in News18 Urdu TV Channel

Khalid Mulla

Research Scholar, Department of Linguistics, Osmania University, Hyderabad

ABSTRACT

Code-Mixing occurs when lexical items and grammatical features of two or more languages exist in the same sentence (Muysken,2000). Code mixing is a phenomenon that is part of multilingual society, which produces mixing of two or more languages. It has already been found that the phenomenon of code-mixing not only exists in conversations among language users but also occur in sentences used in the news and political debates in electronic media. As such, the study concerns the identification of English lexical items that were code-mixed into Urdu language used in News 18 Urdu news channel.

In this paper a maximum of 24 sentences that were code-mixed in the language used in News18 Urdu news channel were collected and analyzed. English nouns and verbs are most frequently code-mixed in telecasted news bulletins, they are also analyzed and discussed as well. News readers, anchors often mix English lexical items in their Urdu language, during interviews, debates, discussions and while reading the news, which brings about the popularity of code-mixing in News 18 Urdu channel.

Keywords: Code-mixing, English, Urdu, Hindi, News, News18, TV, Broadcast, Multilingual society

Introduction

Code-mixing is a phenomenon that frequently occurs in multilingual and pluralistic societies like India. It has been found that the Urdu code-mixed language used in the News18 TV Channel is due to exposure and good command over Urdu and English by the news readers, anchors and desk staff. The news anchors and journalists, political analysts of News 18 Urdu use code-mixed language not only while reading the live news bulletins but even during interviewing the public, politicians, artists, business people and sports persons etc. indulge in code- mixing. Code mixing can also be identified in the language of guests who are invited to the news channel for debates and

discussions on contemporary social issues, like politics, government policies, community issues etc. Most of the programs that are aired in News 18 are reported from what has been gathered through interviews and conversations and news reports coming from field staff and news agencies like ANI, FPI etc. In these broadcasted news bulletins the most noticeable thing is the prevailing application of English lexical items that are often code-mixed with the Urdu language used by the anchors, journalists, copy editors of the New18 Urdu channel.

Review of Literature

According to Myers-Scotton(1993), Boeschoten(1998) and Azuma (1998)(cited in HO,2007), code-mixing is the change of one language to another language in the same oral or written text. It is a common phenomenon in the society where two or more languages are used together. Studies of code-mixing enables us to understand the constraints of language.

Nik(1998) defines code-mixing as a situation where language users code-mix two or more languages. McLaughlin(1984, cited in Hoffman,1991) claims code-mixing occurs within the sentential level and usually involves the use of lexical items.

According to Kachru(1978,p.28), code-mixing refers to "The use of one or more languages for consistent transfer of linguistic units from one language into another and by such a language mixture developing a new restricted or not so restricted code of linguistic interaction."

In his research work (Chen,1999) identified non-sino and non-alien mixed words which are attributed to the need of doing business. For example in Xi an selling earthenware pots to foreign tourists by saying 'lianf yuan one' (two dollars each one) and 'San dollar two'(three dollars for two), this kind of mixed language is really the language user's verbal communication

tool, which is likewise the specific embodiment of the lively language diversification under such circumstances (Chen,1999).

Lan (2007), studies the phenomenon of code-mixing in Chinese business advertisements in Malaysia. His research shows that English lexical items used in Chinese business advertisements are predominantly nouns adjectives and verbs. He has also found that the majority of proper nouns in English are maintained and that they are not translated in the form of sounds and meanings in Chinese. A good number of English verbs and adjectives involved are not only short, precise and easy to pronounce, but also regularly used in conversations among the users of the Chinese language in Malaysia. In addition, some English adjectives such as 'in', 'hit', 'hot' and 'cool' that are deliberately inserted into Chinese business advertisements to exert some sense of modernity and youthfulness.

Purpose of the study

The study focuses on code-mixing as well as examining the phenomenon of code-mixing that exists in various telecasts and programs which are telecasted in the News18 Urdu Channel which include news bulletins, special news-based programs on contemporary issues, documentaries etc. this research paper is hoping to identify the English and Hindi words that have been code-mixed in the sentences used in the News18 channel. The research question that this study hopes to answer is what are the English and Hindi lexical items that have been code-mixed in the News18 news, discussions, current affairs programs and why?

Methodology

For the analysis 10 Sentences that are code-mixed into Urdu in News18 channel, which include news were selected. The sentences were then analyzed and further discussed focusing on the reason why code-mixing takes place in news media? In

order to achieve the objective of the present research, the English lexical items inserted into the sentences of Urdu language used in the News 18 were observed, focusing on the code-mixing speech of the news readers to find out whether the news readers are using code-mixed language spontaneously or with a consciousness.

Language of news channels

As mentioned before, language of entertainment news is not of the same genre used in academic, religion, education, administration, medicine, science and law. Hence, lexical items, sentences and the style of the language used to form sentences do not necessarily conform to the standard variety. In other words, code-mixing allows the news readers to incorporate some 'differences' into their telecasted news and this will help to 'attract' the attention of the viewers.

The concept of adding the lexis of another language into news debates of News18 Urdu channel is not a rare phenomenon because it has been accepted by the majority of viewers in India. As a matter of fact, observations indicate that code-mixing is a distinctive feature of the genre used in News 18 Urdu channel in multilingual society like India. This shows the genre that is aired in News18 is dissimilar from other genres.

The following sentences provide the code-mixing

sentences. Examples:

The following sentences provide the code-mixing sentences used in news bulletins in News 18 Urdu in first week of December 2018.

Examples:

1. سپلائی میں کمی اور مانگ میں اضافہ ہوا ہے۔

/səˈplaɪ meiə̃ : kʌmiː aur mɑːŋg meiə̃ : izɑːfa huwa hai/

There is a shortage in supply and a surge in demand.

2. بازار میں انتظامات کو لے کر میونسپل کارپوریشن کی لاپرواہی سے پریشانی

/bazɑːr ke ɪŋθzɑːmɑːθ ko lekər mu(ː)ˈnɪsɪpəl ˌkɔːrpəˈreɪʃən kiː laparwahiː se pareʃɑːniː/

The negligence of Municipal Corporation in managing the market has caused difficulties.

3. یہ مانا ہے لکھنؤ کے خواجہ معین الدین چشتی اردو عربی فارسی یونیورسٹی کے وائس چانسلر کا

/je mɑːnnɑ hay lʌknaʊ ke kwaʤa mɔːɪnuddɪn ʧɪʃθɪ ˈʊrduː ˈarəbɪ fɑːrsiː <u>ˌyuːnɪvarsɪti</u> ke vaɪs ˈʧɑːnsəlar ka/

The Vice Chancellor of Khwaja Moinuddin Chishti Urdu Arabi Farsi University thinks so.

4. نروڈا پاٹیا معاملے میں آج ہائی کورٹ میں آسکتا ہے فیصلہ

/narɔːda patɪyə mɑːmle meiं : ɑːʤ haɪ kɔrt meiं : ɑː sakta hay faɪsla/

A decision may come today in the case of Naroda Patiya.

5. ملزم ایم ایل اے کلدیپ سینگر کی سیکیورٹی میں کمی

/mʊlzɪm ɛm-ɛl-ey kʊldiːp sɪŋ seŋɑr kɪ se:<u>kyʊərɪti</u> meiं : kamɪ/

Accused MLA Kuldeep Sengar’s security has been reduced.

6. فائر ٹینڈر کی بارہ گاڑیاں موقع پر پہونچی ہیں.

/ˈfaɪə ˈtɛndə kɪ bɑːrɑːh gɑːdɪyं : maʊqe pɑːr pʌhuːं :ʧɪ hay/

Twelve fire tender vehicles have reached the spot.

7. ان مجرمین نے سزا کو گجرات ہائی کورٹ میں چیلینج کیا ہے.

/ɪn mʊʤrimiːn neɪ sʌzʌ ko gʊʤarɑːt haay eɪʧkort meiं : ˈʧælɪnʤ kɪyʌ hay/

These culprits have challenged the punishment in Gujarat High Court.

8. پسماندہ ذات کے افراد نے ریزرویشن کا مطالبہ کیا

/pʌsmɑːnda zɑːt ke afrɑːd ne ˌrɛzəˈveɪʃən ka mʊt ɑːlaba kɪy ʌ/

Backward caste people have demanded the reservation.

9.ایمرجنسی میں جوانوں کی سرجری کی ڈرل کی گئی۔

/ɪ'mɜːʤənsi mei◌̃ : ʤawano kɪ 'sɜːʤəri kɪ drɪl kɪ gayi:/
A surgery drill was conducted in emergency among the army men.

10.جج نے کرائم سین کا معائنہ کیا۔

/ˈʤʌʤ ne kraɪm siːn ka mʊa:yna kɪyʌ/

The judge has inspected the crime

scene. 11. کی قیادت میں ایک اہم میٹنگ ہے۔
کانگریس

/ˈkɒŋgrɛs kɪ qayadat mei◌̃ : ek əhəm ˈmiːtɪŋ hay/
An important meeting is scheduled under the leadership of Congress.

12. میٹنگ پارلیمنٹ میں غلام نبی آزاد کے آفس میں ہوگی۔

/ˈmiːtɪŋ
ˈpɑːrləmənt mei◌̃ : gʊlam nabɪ ɑːzɑːd ke ˈɒfɪs mei◌̃ : hogi:/
Meeting will be conducted in the office of Ghulam Nabi Azad in Parliament.

13.چھوٹا سا بریک لے رہے ہیں۔

/ʧoʈa sʌ bre:k le rʌhe hai◌̃
:/Taking a short break!

14. ذرائع کے مطابق سناتن سنستھا کے چیف سے کی جا سکتی ہے پوچھ تاچھ
-

/zʌra:ye ke mutʌ:bɪq sʌnaθan sʌnsθa ke ʧiːf se kɪ ʤʌ sʌkθɪ hʌɪ puːʧθɑːʧ/

According to the sources the chief of the Sanatan Sanstha may be quizzed.

15. ہندستان میں کمپنی کا کارپوریٹ آفس قائم کرنے کی دی ہدایت۔

/hɪndustɑːn me kʌmpəni kʌ ka:rporet ɒfɪs qʌ:yʌm kʌrne kɪ di: hɪdɑːyʌt/

to open a company corporate office in India.

Instructed ریلائنس فاؤنڈیشن نے 50 کروڑ کا ریلیف میٹریل
16. دیا۔

/rɪˈlayɪəns faʊnˈdeɪʃən ne 50 karod kʌ rɪˈliːf materərɪəl dɪjʌ/

Reliance Foundation has provided the relief material worth Rs.50 crores.

mʊta:bɪq 17. ستارا پولیس اسٹیشن میں درج ریکارڈ کے مطابق ...

Sa:ta:ra: po:li:s sʈe:ʃn mɛːi᷉ d̪ar d͡ʒ re:korɖ ke:

According to the records in Saatara Police Station…

18. چاہے لنچنگ کا معاملہ ہو۔

t͡ʃa:he: li:n͡tʃi:ng ka: ma:mʔla: ho:

Even if it’s the case of lynching…

19. جانکاری ساجھا کرنے کے لیے کوئی میکانزم کی ضرورت ہے۔

d͡ʒɑːŋkɑːrɪ sa:d͡ʒʱa: karne: ke: lɪje: koi: mɛːkjanɪʒm ki: zaru:rat hay:

There is a need to build some kind of a mechanism to share information.

20. خطیر رقم چیف منسٹر ریلیف فنڈ میں جمع کی گئی ہے۔

xat̪i:r raqm t͡ʃi:f mɪnɪsʈər rɪli:f fən᷉d m᷉n d͡ʒɑma: ki: gəɪ həe:

Huge amount has been collected in Chief Minister’s Relief Fund.

21. فاؤنڈیشن کی چیئرپرسن نیتا امبانی موجود تھیں۔

faʊnˈdeɪʃən kɪ tʃɛrpərsən ˈpɜːsən ni:ta: əmba:ni: maʊd͡ʒu:d̪ t̪ʱ᷉

Chairperson of Foundation Neeta Ambani was present.

بازار کے انتظامات کو لے کر میونسپل کارپوریشن کی لاپرواہی سے
پریشانی 22.

/bazɑːr ke ɪŋθzɑːmɑːθ kəʊ lekər mju(ː)ˈnɪsɪpəl ˌkɔːpəˈreɪʃən kiː laparwahiː se pareʃɑːnɪ

23. اے ٹی ایس، سی بی آئی اور مقامی پولیس کا مشترکہ آپریشن

ˈeɪ tiː əs ˌsiːˌbiːˈaɪ ˈɔːr maqa:mi: pəˈliːs ka: mʊʃtarɪka: ˌɒpəˈreɪʃən

A joint operation of ATS, CBI and local police…

سب سے چھوٹے بکرے کا وزن کم از کم 150 کلو ہے.

24.

sab/sie/ /tʃʰoʈɛː/

bəkre/ka:/wazn/kəm/sie/kəm/150/kilo/hai/ The weight of the smallest goat is at least 150 kgs.

Conclusions

From the above two dozen sentences used in the News18 channel it was found that various English words are mixed into the sentences in Urdu. It is observed that some Hindi words also are code-mixed in news bulletins. The most commonly used parts of speech in the sentences are common nouns, proper nouns, followed by verbs. The most commonly code-mixed words are the nouns and next to the nouns are the verbs. The code-mixing phenomenon in TV 18 Urdu Channel is due to the fact that it had a national rather international audience which required the channel broadcasters to use the language which conveys the news to a very wide range of TV viewers across the spectrum. Hence the channel staff including news readers are tempted to use code-mixed words in daily news over the pure Urdu words borrowed from Persian, Turkish, Arabic etc.

References

Muysken, P. (2000), Bilingual Speech: A Typology of Code-mixing, Cambridge: Cambridge University Press.

Bhatia, T. K. & Ritchie, W. C. (1999). Language Mixing and Second Language Acquisition: Some Issues and Perspectives. In Klein, E. C. & Martohardjono, G (Eds.), The Development of Second Language Grammars: A Generative Approach. Armsterdam: John Benjamins Publishing Company, pp. 241-267.

Asmah, H. O. (1992). The Linguistics Scenery in Malaysia, Kuala Lumpur: Dewan Bahasa dan Pustaka.

Awang, S. (2004), Bahasa Rojak (Rojak Language), [Online] Available:

http://www.pts.com.my/modules.php?name=News&file=article &sid=192 (July 15, 2006).

Ho, J. W. Y (2007), Code-mixing: Linguistic Form and Socio-cultural Meaning. [Online] Available: The International Journal of Language Society and Culture, http://www.educ.utas.edu.au.users/tle/JOURNAL/issues/2007/21-2.pdf (May 18, 2010).

Maschler, Y. (1998), On the transition from code-switching to a mixed code, In Auer, Peter (Ed.), Code-switching in Conversation: Language, Interaction and Identity, London: Routledge, pp. 125-149.

Lau, S. K. (2007). Fenomena Percampuran Kod dalam Iklan Perniagaan Bahasa Cina di Malaysia. Unpublished Research Report of Master of Modern Languages Studies, University of Malaya, Faculty of Languages and Linguistics.

Liu, P. (2008), Code-switching and Code-mixing, GRIN Verlag.

Hoffman, C. (1991), An Introduction to Bilingualism, London:
Longman.

Chen, J. M. (1999), Zhongguo Yuyan he Zhongguo Shehui.
Guangzhou: Guangdong Jiaoyu Chubanshe.

Kachru, B. B. (1978), Toward Structuring Code-Mixing: An Indian Perspective. International Journal of the Sociology of Language. 16, 27-46.

Book Review

Madhumita Barbora and Gautam K Borah (eds.) (2016). *Aspects of Modern Assamese*. Guwahati, Assam: Bhabani Books in association with Department of English and Foreign Languages and the Registrar, Tezpur University (Under UGC's SAP-DRS II, 2015 – 2020). 214 pages.

Bipasha Patgiri
Assistant Professor (Linguistics)
Department of English, Tezpur Central University, Assam

Assamese is the easternmost Indo Aryan language spoken by approximately 15 million people of Assam. This collection of papers provides a comprehensive view of the theoretical aspects concerning Modern Assamese mostly from a synchronic perspective. Many important issues of current interest in linguistics, especially on Syntax-Semantics interface are covered in *Aspects of Modern Assamese*which contains eleven scholarly papers by senior professors and research scholars of linguistics of Tezpur Central University.

The Foreword of the book is by Prof Kārumūri V. Subbārāo, acclaimed linguist and author of *South Asian Languages: A Syntactic Typology* (2012). He mentions how the location of Linguistics wing in Tezpur Central University is advantageous to conduct research on innumerable lesser-studied languages from at least five different language families – Indo- Aryan, Tibeto-Burman, Mon-Khmer, Dravidian and Tai-Kadai. As Prof Subbārāo rightly points out, "the study of any unexplored language provides deep insights into the structural and functional aspects of language". He also mentions the need to work on the rich agreement patterns, the incorporation of various elements in the verb, deixis, noun modification strategies, control structures, structure of the NP, compound formation etc. among other interesting phenomena. He opines that the current volume has insightfully picked up some

intriguing aspects of syntax and semantics and the thread of such work must continue with inclusion of works of enthusiastic students. The editorial gives a gist of the papers outlining the salient features. The book contains papers with topics ranging from the language's socio-historical background, bare noun indeterminacy, agentive-ergative marking, ingressive- progressive aspect, genitive subject, causative constructions to scope of negation etc. The editors state, "Assamese has a very long and distinguished literary career, yet it is not yet sufficiently studied in the light of the latest theoretical developments". The book intends to be fill this gap and initiate enthusiasm among young researchers.

In the first paper **"Assamese: Socio-Historical Background"**, Prof M M Sarma sketches linguistic scenario of Assam from both diachronic (the only paper to deal with historical aspects) and synchronic perspectives. He divides the paper into three sections – Assamese in its formative period (4th to 13th Century) and 'Kamrupi Prakrit' with a discussion on influences of Non-Aryan elements on Assamese language. In the second section, he presents the history of Assamese script, Assamese prose and subsequent standardization of the language. The third section deals with a discussion on the varieties of Assamese and their classification. The paper includes some fundamental questions pertaining to the language's development, uses, contact situation and varieties. For the first time in Assamese linguistics, Prof Sarma lists some morphophonemic changes and changes in classifiers and idiomatic expressions in the standard variety.

The second paper **"The Indeterminacy of Bare Nouns in Assamese and its Determination"** by Prof Gautam K Borah explains some interesting observations concerning noun determination. He argues for a 'classifier phrase' (following Greenberg, 1974 and Allan, 1977) instead of a conventional 'noun phrase' in Assamese and also states that classifiers are not

always definitives but has a crucial role to play for 'reference'. The paper also explains with proper justification how Assamese has 'two separate sets of inherently indefinite and definite quantifiers' for non-numerical quantification. The paper also elaborates on 'absence of articles', 'number forms', 'definite collective markers used only for humans', 'generic references' in Assamese etc. The paper is a good read for scholars working on syntactic and semantic features of NPs in general.

In **"The Agentive Ergativity in Assamese"**, Prof M Barbora and D Nath draw reader's attention to a oft-cited yet controversial aspect of morphosyntax, i.e., case marking on subjects. Following an extensive discussion on the nature of 'e- marked' subjects, they conclude that instead of merely terming such subjects in Assamese as ergative, a more appropriate categorization could be 'agentive ergative'. The paper also discusses 'split ergativity', 'sematicization of '-e'', 'agentive light verb' and the inherent relationship between 'feature checking and case marking'.

Another paper by Prof G K Borah, **"The Ingressive Progressive Aspect in Assamese"**, throws some lights into the existence of an ingressive aspect within the usual progressive aspect. The ingressive aspect focuses on the beginning of an action rather than the progression of the action. In his following paper **"The Fuzzy Boundary between Projected and Potential Realities and *iba*"**, author Prof Borah shows that embodiment of future within modality is fuzzy in many respects. He then argues that the conventional glossing of the marker *iba*for future tense is somewhat misleading as it actually involves elements of both futurity and modality.

In the paper **"Genitive Subjects in Assamese"**, Dr D Nath explores the distinction between experiencer subjects and experiencer objects of South Asian languages and examines the nature of possessive subjects which get genitive case marking. He states that genitive subjects in Assamese are 'inherently case

marked' and can 'occur with unaccusaive verbs to form a complex predicate'. The paper discusses some other motivationsbehind classifying possessive subjects as genitive subjectsin Assamese such as 'goal theta role', 'requirements from certain predicate types', 'feature checking' and 'subjecthood of genitive case' etc. Among the specific types of predicate types which require a genitive case-marked NP, the author lists subjects ranging from 'state of consciousness', 'state of sickness' etc. among others. The paper is a good read to understand the variable nature of Assamese subject NPs.

In her second paper **"Causative Construction in Assamese"**, Prof M Barbora deals with an important aspectconcerning derivational relations. The paper is divided into three main sections –an overview on causatives and its relationship with the concept of valency, range of transitivity in Assamese and morphological causatives. In the last section, she discusses three types of 'causative clauses' based on transitivity and their thematic hierarchy. She also gives two interesting schematic representations – one for morphological causative types in Assamese and the other for demotion/addition of arguments for each causative type. In her next paper **"The Particle *ze*: a Clitic, a Wh-operator or a Complementizer"**, Prof M Barbora explore the finite complement clauses in Eastern Indic languages and the syntactic roles played by the complementizer *ze*in Assamese. She also distinguishes between the clausal-initial particle *ze*and clausal-final particle*buli*. Interestingly, the particle *ze*can either be internal (like Bangla) or peripheral (like in Oriya).It also functions as a Wh-operator in Assamese.

The paper**"(Non-Finite) Clausal Complementation and Assignment of Case in Assamese"** by Dr D Nath discusses the distinction between 'finite and non-finite clauses', tow types of non-finite clauses - 'infinitival clauses' and 'participle clauses' and also 'ECM (Exceptional Case Marking) constructions' and

their feature checking in real examples from Assamese. In fact, this is the first work on Assamese on ECM constructions. The author rightly argues that in ECM, the matrix verb assigns accusative case to the embedded subject NP.

The paper **"The Scope of Negation in Assamese"**, Dr K Sarma presents a review of literature dealing with this issue and then discusses scope of negation in Assamese. He shows that unlike in English where only the object is within the scope of negation, both the subject and the object are within the scope of negation in Assamese.

In **"The Meaning of *ma:ne*: a Conversational Filler, a Structural Necessity and a Cognitive Endowment"**,Dr S Barua deals with the cognitive aspects of conversational fillers. The author argues that *ma:ne*in Assamese is much more than a conversational filler. She initiates the discussion by exploring whether *ma:ne*is a discourse complement, a conversation facilitator or cognitive endowment. The paper also discusses the rhetorical constituents of discourse – exigency, audience and constraints adopting Bitzer's analysis. She also presents extracts from two talk-in-interactions from young Assamese native speakers. The paper concludes with remarks that such conversational fillers are multifunctional and act as hedges. The paper is a good read for researchers working on discourse analysis within cognitive framework.

> *Aspects of Assamese thus is an ensemble of in-depth articles presenting a much-neededoverview of major grammatical categories and particles in Assamese pertaining to their syntax-semantics interface. Such works must continue to motivate scholars to enhance research in these areas so as to contribute to the larger domain of Assamese linguistics in particular and world's languages in general.*

Implementation of UNL Based Machine Translation system for Telugu Language

Dr. G Renuka Devi*

Abstract

*The purpose of the present paper is to discuss the **U**niversal **N**etworking **L**anguage (UNL) representation of Developing Dictionaries from **N**atural **L**anguage (NL) to Universal Networking Language and vice versa for Telugu Language. Universal Networking Language acts as an artificial language or medium in converting one language to any other languages through UNL framework. As UNL is a language independent platform, it captures information and finds its major claim in Machine Translation (MT). This is an attempt to introduce the readers to the implementation of the UNL framework and explain how it is being used to find out various linguistic issues involved in Telugu to English and vice versa.*

1. Introduction

Languages differ greatly in their "surface structures", they all share a common "deep structure"; hence came the idea of creating a universal representation capable of conveying this deep structure while enjoying the regularity and predictability natural languages lack. Although inter-lingua is a promising idea, the number of interlinguas created is still very limited. UNL, has succeeded in standardizing its tools, tagset and methodology as well as rely on meaning as an intermediate representation (Al Ansary 2009).

UNL is a kind of mark-up language which represents the core information of a text. The UNDL Foundation, the founder of UNL, has created a wrapper application for development of various UNL tools and applications (Martins 2012, Martins and Avatesyan 2009). All engines, resources and tools are available through the UNL web (www.unlweb.net) that contains many tools designed for linguists, computational linguists as well as non- professionals. These tools are used in analysing and generating natural languages. IAN, the Interactive ANalyzer, it employs the analysis grammar rules to analyze input and finally generate its corresponding UNL expressions. It operates semi-automatically; word-sense disambiguation is still carried out by a language specialist, nevertheless, the system can filter the candidates using an optional set of disambiguation rules. Various syntactic aspects are addressed with the help of a fully automatic engine known as the dEep-to-sUrface natural language GENErator (EUGENE) which functions as DeConverter by using language grammar rules. Firstly Network Processing (NN) is performed to generate a UNL

* Dept. of Linguistics, The English and Foreign Languages University, Lucknow.

graph that is pre- processed by the NN rules in order to become a more easily manageable semantic network. Second Network-to tree (NT) processing is performed resulting in network structure being converted by the NT rules into a syntactic structure. Third Tree-to-Tree (TT) processing is performed to transform deep syntactic structure into a surface syntactic structure by the TT rules. Fourth Tree-to-List (TL) processing is performed to convert surface syntactic structure to various other changes according to the TL rules to generate a NL structure. Finally List Processing (LL) processing is performed to get a natural language sentence by the LL rules.

EUGENE (the dEep-to-sUrface natural language GENErator) is a fully automatic engine; it simply uses the target language grammar rules in order to decode the incoming UNL document and generate it in natural languages.

Interactive Analyser (IAN) and Eugene use two types of Natural language dictionaries; enumerative and generative. The enumerative dictionary of IAN contains all inflected word forms of a language together with their corresponding Universal Words (concepts) and a set of linguistic features covering different linguistic levels. The generative dictionary, on the other hand, is the same as the 'enumerative' one but it contains all lexemes of language as bases together with a morphological paradigm number that controls the generative morphological behaviour (e.g. agreement and inflected forms) of words in natural language (Martins and Avetisyan 2009). It might be a fact that all languages have classical reference grammars in grammar books. Such a reference grammar may be defined as a description of the grammar of a language, with explanations of the principles governing the construction of words, phrases, clauses, and sentences.

It is designed to give someone a reference tool for looking up specific details of the language. In Natural Language Processing, computers should also learn a language in order to give a comprehensive and objective test-bed that enables us to evaluate, compare and follow up the performance of different grammars.

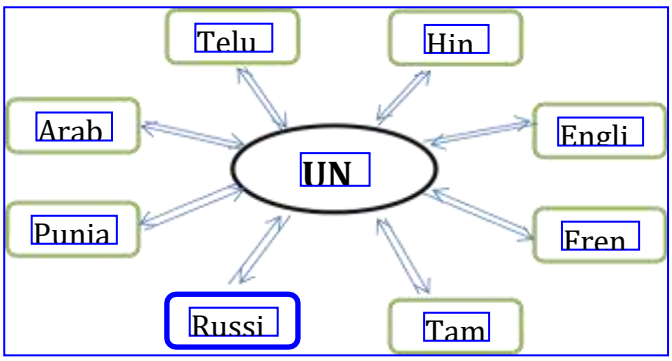


Fig 1: The UNL system

A formalized reference grammar is needed in order to synchronize different languages; the UNL is initiating this idea as it utilizes a standardized environment. The current study is limited to Telugu to UNL and UNL to Telugu only; it is organized as follows:

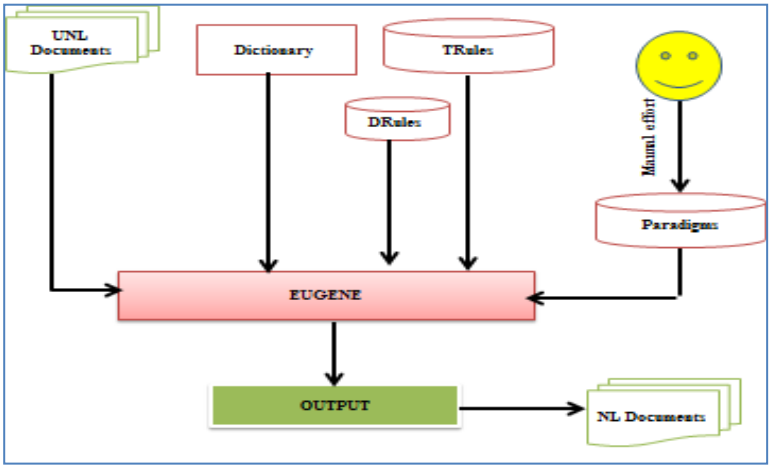


Fig 3: NLization of UNL Document

The above flowchart describes the process of NLization of UNL document. The output generated from the IAN tool becomes the input of the EUGENE tool. Only base forms are extracted from Tel – UNL dictionary to develop UNL – Telugu dictionary and added inflectional paradigm number to the words containing inflections. Similarly, Transformation grammars and Disambiguation grammars and Inflectional grammars are developed and saved in a plain text document with .txt extension. Finally, UNL sentences, Dictionary, T-rules and D-rules and Inflectional rules are uploaded into the EUGENE tool to generate Telugu sentence out of UNL sentences.

Reference corpus

Corpora are considered essential language resources necessary when building grammars. A reference corpus has been compiled as an experimental English corpus in order to prepare the initial version of analysis and generation grammars. Telugu parallel Corpus has been compiled by translating the English reference corpus into Telugu, this corpus consists of (100+300+250) 650 sentences and a short story collected from UNDL. It is supposed to cover the basic and common linguistic phenomena between all languages that may be encountered in the process of building grammars within the UNL framework such as: temporary entries (e.g. URLs, symbols etc.), words that are not found in the dictionary (a grammar in NLP may face a set of words that might not be found in the dictionary), numerals, determiners, prepositions, conjunctions, noun phrase structures, expressions of time, verb forms, pronouns and sentence structures. The English reference corpus is manually annotated to make a

standard version of UNL reference corpus. Both versions; English reference and UNL corpora, are available on the UNL web (http://www.unlweb.net/wiki/Corpus-UCA1, UCA2, UGOA1, AESOP).

In this study, the IAN tool was used for UNLization and two Experimental NL corpora, UCA1 and UCA2 were used for the study. The former corpus was of initial level and consisted of 100 sentences and the latter was the second level, which comprised 300 sentences. In this paper, the challenges faced by the investigator while UNLizing the sentences from these corpora. In order to discuss the issues involved in UNLization of the corpus in question, Categorization of all the sentences[1] in the corpus as follows.

Table 1: Different Parts of Speech in UCA1 and UCA2 corpus

S.No.	UCA1-Types	No. of Sent.	UCA2- Types	No. of Sent.
1	Temporary	05	Cardinals numbers	70
2	Determiners	10	Ordinals numbers	50
3	Pronouns	10	Fractions	20
4	Prepositions	20	Other numerals	10
5	Conjunctions	10	Determiners	35
6	Noun Phrase	20	Prepositions	20
7	Verbs	10	Noun Phrase	15
8	Sentence structures	15	Time	15
9			Verbs	10
10			Pronouns	25
11			Sentence structures	30
	Total sentences	100	Total sentences	300

Process of UNLization

To UNLize the UCA1 and UCA2 corpora[2], a Telugu-UNL dictionary and Telugu grammar containing Telugu Transformation Rules (T-Rules) were prepared. We present them in this section.

[1]UNL sentences are the basic unit of representation inside the UNL framework. UNL sentences or expressions are the hypergraphs generated out of nodes (Universal words) interlinked by binary semantic universal relations which are modified by universal attributes.

[2]UCA1 and UCA2 contain experimental data in English. For the purpose of the research, we had translated the entire data into Telugu. This data was used as the input for UNLization.

Telugu – UNL Dictionary

Telugu – UNL Dictionary is an analysis dictionary that is used for linking Natural Language entries to Universal Words (UWs) along with its corresponding semantic features and attributes. This has been developed according to the UNL dictionary specifications provided by UNDL Foundation. For UCA1 corpus, there were 141 dictionary entries and default dictionary entries and for UCA2 corpus 710 dictionary entries and default dictionary entries.

Syntax of each Dictionary entry is as follows:

[NLW] {ID} "UW" (ATTR , ...) < FLG , FRE , PRI >;

Here, NLW refers to Natural Language Word; ID refers to Unique Identifier; UW refers to Universal Word; ATTR refers to Attributes; FLG refers to three-character Language Code; FRE refers to Frequency; PRI refers to Priority. A sample entry from Telugu-UNL dictionary is given below;

[పుస్తకం]{}"book" (LEX=N,POS=NOU,NUM=SNG)<tel,0,0>;

Snapshot of Telugu default dictionary is shown below.

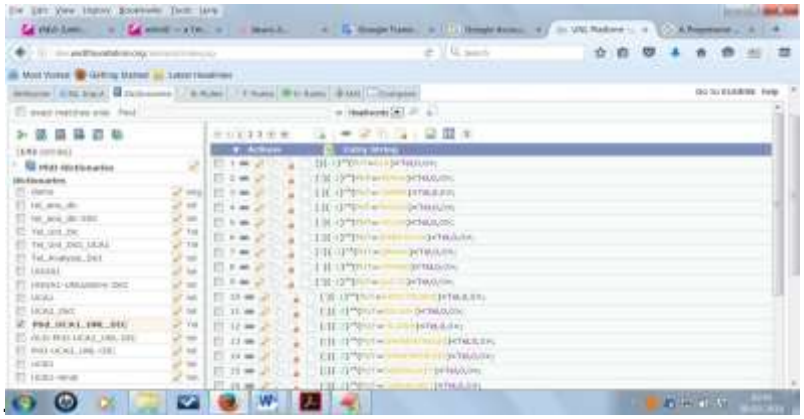


Fig 4: Snapshot of Telugu Default Dictionary

3. Building the UNLization (analysis) Grammar

UNLization is the process of representing the content of a natural language structure using UNL. In order to UNLize any Natural language text, the UNLization (analysis) grammar for that natural language should be, first, developed.

The UNLization reference grammars for UNL and Telugu reference corpora have been already built to represent the content of both corpora. UNL and Telugu grammars have common modules such as; the tokenization, numeral, attribute, syntactic and syntax-semantic modules.

3.1: The Tokenization module

The tokenization algorithm is strictly dictionary-based; the

system tries to match the strings of the natural language input against the entries existing in the dictionary. In case it does not succeed, the string is considered a temporary entry. There are no predefined tokens: spaces and punctuation marks have to be inserted in the dictionary in order to be treated as non- temporary entries. The tokenization algorithm goes from left to right trying to match the longest possible string with dictionary entries, and it assigns the feature **TEMP** (temporary) to strings that are not found in the dictionary.

For instance, any URL such as "www.undlfoundation.org" should be considered TEMP; however, it is tokenized according to the entries found in the dictionary as [www] [.] [u] [nd] [l] [foundation] [.] [or] [g], which is incorrect since we expect the whole string to be treated as a single temporary entry. In order to avoid that, a disambiguation rule applies to consider any string a single node if followed by blank space or a full stop (is at the end of the sentence). The tokenization algorithm blocks the segmentation of tokens or sequences of tokens prohibited by disambiguation rules. Disambiguation rules are not only responsible for the segmentation of the input, but also responsible for choosing the word senses most appropriate to the context. For instance, “you” have two different realizations in the dictionary; the singular second person pronoun and the plural second person pronoun. In the sentence “you love yourself”, disambiguation rules should prevent the choice of the plural pronoun, thus, causing the engine to choose the singular pronoun if the verb is followed by a singular personal possessive pronoun.

3.2: The Numerals Module

Numerals in UNL are temporary UWs and should be represented in UNL as Digits between quotes. There are two cases in Numerals; they may be present in the input as digits in which case the engine will consider them as TEMP automatically, or, they may be written in letters, in the latter case the numerals module is activated. In order to handle numerals in both English and Telugu, both a dictionary and analysis rules are required. The numerals module is part of both the English and Telugu grammars. There are 4 types of numerals to be covered in this module: cardinal, ordinal, partitive and multiplicative. To examine the cardinal numbers first as they constitute the base for other types of numerals. There are many subsets of cardinal numbers such as units, tens, hundreds, thousands, millions…etc. The first step towards analyzing them is compiling a small dictionary that will enable rules to convert numbers in both English and Telugu into digits. Some cardinal numbers will be inserted in the dictionary as is such as the numbers from one to nineteen. Other numbers will be inserted incomplete in the dictionary to be later completed by rules; for example, tens are inserted without their zeros, “twenty” is inserted as “2” etc. The second step is to develop the required rules; units and numbers from ten to nineteen are retrieved from

the dictionary without any modification by rules. Tens starting from the number twenty have two possibilities in analysis: the first is adding tens to units; for instance in the case of "twenty one", "twenty" which is stored in the dictionary as "2' and "one" which is stored as "1" will be joined by a rule and will be treated as a single number "21". The second is not adding tens to units as in "twenty", a zero will be added to "2" and joined together by a rule to become "20". The analysis of partitive numerals depends on their existence in the dictionary. In ordinal and multiplicative numbers; after converting the numbers in letters into digital numbers, an attribute "@ordinal" will be assigned to the number. If the number is followed by the word "times" such as "four times", the attribute "@times" will be assigned to "4"to be "4.@times".

3.3: Attributes module

In UNL, attributes have been used to represent information conveyed by natural language grammatical categories (such as tense, mood, aspect, number, etc). The set of attributes, which is claimed to be universal, is defined in the UNL Specs (http://www.unlweb.net/wiki/Attributes).

The attributes module can handle determiners, pronouns, prepositions and verb forms. It is responsible for substituting certain words or morphemes with attributes, as in the case of quantity quantifiers) "a lot of", "several", "few", "all", "any…etc. (which will be deleted and substituted by the attributes "@multal, @paucal, @any, @all. Etc." to be assigned to the following word. In UNL, pronouns are "empty concepts" represented semantically as "00". The person, number and gender of the pronoun are described by UNL attributes.

3.4: Syntactic module

After assigning the necessary attributes, the syntactic module should start drawing the syntactic trees for noun phrases, verb phrases and sentence structures that are part of the corpus, according to the X- bar theory (http://www.unlweb.net/wiki/X-bar_theory). The syntactic module is divided into two phases; the list–to-tree phase and the tree-to-tree phase.

The list-to-tree phase

In this phase, rules are used to parse the tokenized input sentences into a tree structure. It starts by composing small trees for the small phrases in the sentence and then combining these small trees together to form a bigger tree. List-to-tree rules are responsible for building the trees for language structures; ordering of rules is required; rules for building noun phrase trees should be followed by rules for building verb phrase trees.

The tree-to-tree phase

The constituents will be mapped onto their syntactic roles. The methodology used in this phase is to start decomposing constituents from the biggest to the smallest. Ordering of rules is required; rules for decomposing verb phrase trees should be followed by rules for decomposing noun phrase trees. VP is the last constituent to be composed,

Hence, it will be the first one to be decomposed. A key assumption of X-bar theory is that branching is always binary, thus, the decomposition of any constituent will affect the tree. A constituent is decomposed into a syntactic role between a node and the head of the adjacent constituent to make the binary relation.

3.5: Syntax – semantic module

In this module, rules have been built to derive the semantic network from the syntactic graph. Orders of rules in this module are not necessary since the semantic features assigned to nodes from the list-to-tree phase, constrain the rules enough to be carried out in their context.

4. Building the NLization (Generation) Grammar

This section discusses the NL-ization of the reference corpus from the inter-lingua representation (UNL) into both Telugu and UNL. To achieve this purpose, Telugu and UNL linguistic resources have been developed. These resources are Telugu and UNL specialized dictionaries in addition to Telugu and UNL NL-ization grammars. The process of generation may be seen to some extent as a mirror image of the analysis process; generating well-formed sentences has to pass through a set of grammar modules which are: the semantic-syntactic module, the syntactic module, the attributes module, and the numerals module.

4.1: The Semantic-Syntactic Module (Network-to-Tree Phase)

This module is responsible for mapping the semantic relations onto their syntactic equivalents.

4.2: The Syntactic Module

The syntactic module is the second module of the NL-ization grammar, it is responsible for transforming the deep syntactic structure generated from the semantic-syntactic module into a surface syntactic structure. The Syntactic module is divided into two phases; the tree-to- tree phase and the tree-to-list phase. The tree-to-tree phase is responsible for gathering individual syntactic relations and forming higher constituents while the tree-to-list phase is responsible for linearizing the surface tree structure into a list structure.

The tree-to-tree phase

In the tree-to-tree phase, rules are responsible for building the

surface syntactic structure of the sentence by building the intermediate constituents (XBs) which are combined to form the maximal projections (XPs) and finally combined to form the sentence structure.

The tree-to-list phase

In the tree-to-list phase, rules are responsible for transferring the surface syntactic structure into a list structure and also adding the required spaces. Thus, the syntactic tree will be transformed into the natural language list.

4.3: The Attributes Module

This module is responsible for converting the attributes represented in the inter-lingua into the suitable natural language words or affixes.

4.4: The Numeral Module

This module is responsible for converting digital numbers onto their counterpart natural language string (Telugu or UNL). There are four types of numerals to be covered in the numerals module; cardinal numbers, ordinal numbers, partitive numbers and multiplicatives. For cardinal numbers, the basic conversion mechanism is converting individual digits from (0 to 9) directly onto the counterpart natural language string, and then converting multiple digits by combining the converted individual digits to form bigger numbers. The numerals module handles also ordinals, multiplicatives and partitive numbers.

5. Evaluation

The output of the UNLization process for both Telugu and UNL languages has been evaluated based on a corpus that is annotated manually semantically in order to figure out the quality and accuracy of the automatically generated semantic networks. The output of the NLization process has been evaluated based on a manually translated corpus. The F-measure (F1-score) is used to measure of the grammar accuracy, according to the formula: F-measure = 2 x ((precision x recall) / (precision + recall)). Finally, the success rate of the output of the UNLization and NLization by calculating F- measure is satisfactory.

For example, the F-measure grammar accuracy for UCA2 corpus is 0.901, which is reflected in the snapshot below.

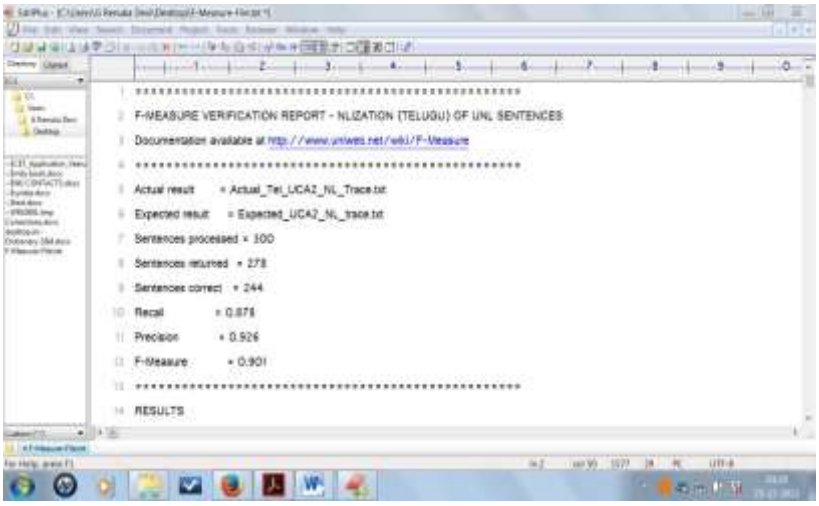


Fig. 5: The F-measure grammar accuracy for UCA2 corpus is 0.901

NLization of UCA1

The F-measure grammar accuracy for UCA1 corpus is 0.806, which is seen in the snapshot below.

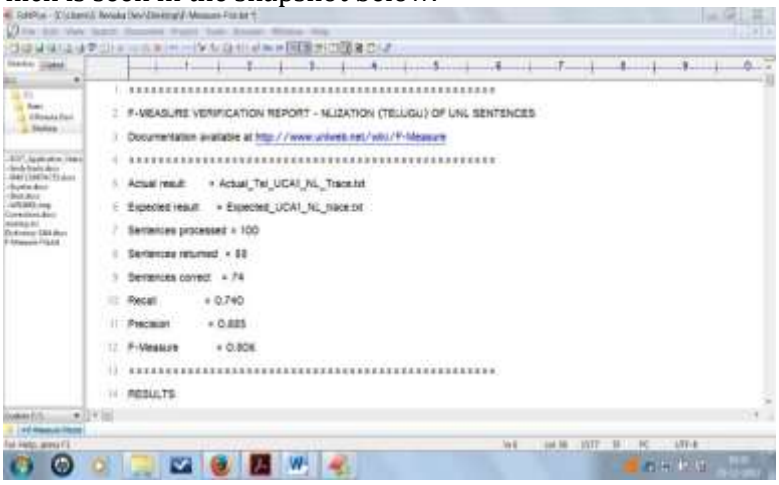


Fig.6: The F-measure grammar accuracy for UCA1 corpus is 0.806

6. Findings

This section presents all the findings of the research done on UNLization and Nlization of Telugu corpora. Both UNLization and NLization of the Telugu corpora needed to address certain issues. UNL is more like English and therefore writing UNL dictionaries and grammar for conversion of English data into UNL or UNL to English is comparatively easier. However, writing UNL dictionaries and grammar for conversion of Telugu data into UNL or UNL to Telugu involved writing new grammars.

This raised certain linguistic issues, which are as follows:

a. Telugu does not have a separate indefinite article as in English. A noun is indefinite when it is preceded by the numeral ఒక[oka] 'one'. That is, in Telugu this numeral is ambiguous and

it can be disambiguated only in context. While UNLizing, the Telugu numeral 'one' is mapped to either the numeral 'one' or indefinite article 'a' in UNL.

b. Similarly, there is no definite article in Telugu. The absence of any indefinite marker like the numeral 'one', will make the noun definite. Hence, in this research, we used NULL before a noun, when it is definite for UNLization and a BLANK is generated for NLization.

c. There is no negative determiner 'no' in Telugu. As a result, an English noun phrase with a negative determiner is translated as a sentence with a negative existential verb **లేదు** [leedu] 'be-not'. Both in UNLizing and NLizing, a lexical approach was followed. However, this has not really solved the problem. T-Rules changing the structure need to be written for mapping a negative NP to a negative existential sentence.

d. UNLizing cardinal numbers involved a few issues. As in English, double-digit numbers can be derived by using certain T-rules as they have a common pattern – XY, where X is a constant and Y can range from one to nine. For example, ఇరవై - ఆరు, [iruvai-aaru] 'twenty-six' and ఇరవై - ఏడు, [iruvai-eeDu] 'twenty-seven'.

Non-human nouns are modified by cardinals as they are without any change while human nouns are modified by cardinals (human) which are derived by affixation of – గురు,[–guru] in the case of cardinals from three to seven, మంది,[maMdi] in the case of cardinals from seven onwards. The analysis dictionary contains only cardinals and T- rules are used to derive the cardinals which modify human nouns.

e. Ordinals in Telugu are derived by suffixing either –వ /va/, or –-ో /O/ to cardinal numbers. That is, a T-rule to derive ordinals is written. However, in some cases, [యో- *yoo*] is added to the cardinal numbers, For example *[*ఇరవయో– *twentieth]*. Such cases are taken care of by adding specific T-rules to the grammar.

f. Another area of difficulty with UNLizing and NLizing Telugu data was personal pronouns particularly with first person plural prnoun. In Telugu there are two first person plural pronouns (i),[manamu] and మేము [meemu] 'we'[exclusive] and (ii) మనము [manamu] 'we'[inclusive]. That is, English 'we' is ambiguous. While translating the first person plural pronoun from English to Telugu, the ambiguity cannot be removed.

g. In English the reciprocal pronoun consists of two words, due to which the two words are represented as two words in UNL and wrongly translated into Telugu, which contains one word reciprocal. However, we need to recognize that Telugu reciprocal also has two reduplicated words., but function as one.

h. Telugu adjectives have only form and they are not inflected for either Person, number and gender features nor comparative/superlative features. In other words, Telugu

adjectives have only one form. However, when an adjective is used as a predicative, i.e., after a be-verb as complement, the adjective is suffixed with 'waaDu' in the case of males and 'adi' in the case of females. In addition there is agreement marker which shows agreement with the subjects. For example,

(i) neenu manchi-waaDi-ni

I good –waaDu+1st person singular 'I am good.'

(ii) aame manchidi

she good-adi 'She is good.'

Also notice in the above sentences, there is no copula. These are verb-less sentences. Hence the difficulty is to T-rule to change an adjective into nominal in Telugu.

The second issue with regard to adjectives is that there are no comparative and superlative forms of adjectives. There is only the use of the counterpart of the word 'very' [chaala] and 'than' [kanTe]. [kanTe] is used in comparative sentences. Equivalent word of 'chaala', 'atyanta' is used as premodifier of the adjective to give superlative meaning.

i. In Telugu, phrases and sentences are joined by lengthening the final of the vowel of the first conjoint. Conjunction 'mariyu' or 'leedaa' are used for joining sentences. Commas do not act as conjunctions as in English. Therefore the problem is to add a conjunction for every comma after every conjoint if we are UNLizing English and while NLizing into Telugu, and 'conjunction' needs to be translated into an affix or a proper conjunction.

For example, "Paper, pencil and eraser" needs to be translated into UNL where each comma representing a conjunction and while translating this UNL version into Telugu each conjunction needs to be translated into a conjunction marker (long final vowel of each conjunct). The strategy we used is to add comma equivalent to conjunction in the dictionary. Though our data did not contain commas for the purposes other than conjunction, the accuracy was less due to this ambiguity of a comma.

7. Limitations and Scope of UNL-based Translation

a. The study described in this paper is limited to the UNLization and NLization of Telugu data provided by UNDL foundation. This data is only an experimental data and consists of words, phrases and a few sentences.

b. UNDL foundation has provided some other data too. However, there are no UNL documents exist for the remaining data. As a result, the investigator is constrained to work only on the data where UNL documents available.

c. The research data is restricted only to sentence level. When the program is implemented on a large data, more accurate results will be obtained.

8. Conclusion

In this study various issues in detail are discussed that are involved in Telugu to UNL generation and UNL to Telugu and English. The presented issues of certain simple and multiple sentences and the ways that are improvised the output. There are several issues, of changing importance, are still to be addressed to accept the UNL based Machine Translation for representing knowledge and information of various languages. Hence the UNL information needs to be provided without any ambiguity by strengthening the dis-ambiguity or simply D- rules. Hence UNL should give a way to UNLize ambiguous information to handle machine translation.

References

Adly, N., & Al Ansary, S. (2009). Evaluation of Arabic machine translation system based on the universal networking language. In *Natural Language Processing and Information Systems* (pp. 243-257). Springer Berlin Heidelberg.

Arden, A. H. (1905). *A Progressive Grammar of the Telugu Language with Copious Examples and Exercises*. Society for promoting Christian knowledge.

Bhattacharyya, P. (2012). Natural language processing: A perspective from computation in presence of ambiguity, resource constraint and multilinguality. *CSI Journal of Computing*, *1*(2), 1-13.

Martins, R., Avetisyan, V. (2009). Generative and Enumerative Lexicons in the UNL Framework, the seventh international conference on computer science and information technologies (CSIT 2009), 28 September – 2 October, 2009, Yerevan, Armenia.

Martins, R. (2012). Le Petit Prince in UNL, the 8th international conference on language resources and evaluation (LREC'12), 23-25 May 2012, Istanbul, Turkey.

Uchida, H., Zhu, M., & Della Senta, T. (1996). UNL: Universal Networking Language–An Electronic Language for Communication, Understanding, and Collaboration. *Tokyo: UNU/IAS/UNL Center*.

Uma Maheshwar Rao G & Prameshwari, K. (2010). On the description of morphological data for morphological analyzers and generators: A case study of Telgu Tamil & Kannada. *Knowledge Sharing Events, LDC-IL, CIIL, Mysore*.

Viswanatham, K. (2006). Structure of Telugu Phrases. *Central Institute of Indian Languages, Mysore.*

Website References:

Calculation of F-measure 2015:
http://www.unlweb.net/unlarium/index.php?tools=fmeasue
Introduction to UNL 2014:
http://www.unlweb.net/wiki/Introduction_to_UNL
NLP Applications:
*http://www.doc.ic.ac.uk/*nd/surprise_97/journal/vol1/hks/intro.html*
Uchida H, Zhu M and Senta T D 1999 The UNL, A gift for a millennium. Tokyo: Institute of Advanced Studies, The United nations University.
UNDL Foundation:
http://www.unlweb.net/wiki/index.php/UNLization
UNDL Foundation 2016:
http://www.undlfoundation.org/undlfoundation/
http://lekhini.org/nikhile.html (transliteration tool)
http://en.wikipedia.org/wiki/Part-of-speech_tagging.
*UNDL Foundation:*www.unlweb.net/wiki/For#For *UNDL Foundation:*www.unlweb.net/ wiki/F-measure *UNDL Foundation:*www.unlweb.net/wiki/Tagset *UNDL Foundation:*www.unlweb.net/wiki/Grammar