Mapping of Semantic Class of Nouns in Urdu Conjunct Predicates to the English WordNet

by

Farhat Abdullah

(ISB-F-181764)

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

DEPARTMENT OF ENGLISH, FACULTY OF SOCIAL SCIENCES

AIR UNIVERSITY, ISLAMABAD

NOVEMBER, 2023

Mapping of Semantic Class of Nouns in Urdu Conjunct Predicates to the English WordNet

by

Farhat Abdullah

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

SUPERVISOR at Air University Islamabad: **Dr. Uzma Anjum**

SUPERVISOR at Nanyang Technological University:

Dr. Francesco Perono Cacciafoco

AIR UNIVERSITY, ISLAMABAD

CERTIFICATE

Department of English

It is hereby certified that Farhat Abdullah (ISB-F-181764) has successfully completed her thesis.

Dr. Uzma Anjum

Associate Professor

Air University

Supervisor

Dr. Sham Haider

Associate Professor

Air University

APA References

Dr. Sadia Irshad

Assistant Professor

Air University

Thesis Coordinator

Dr. Farzana Masroor

Associate Professor

Air University

Chair Department of English

Prof. Dr. Munawar Iqbal Ahmad

Professor

Air University

Dean Faculty of Social Sciences

Prof. Dr. Zafarullah Koreshi

Professor

Air University

Dean Faculty of Graduate Studies

Dr. Francesco Perono Cacciafoco

Senior Lecturer

Nanyang Technological University (NTU), Singapore

Supervisor

DECLARATION

I declare that all material in this thesis is my work, and that which is not my own work has been mentioned as such, and that no material from this work has previously been submitted or approved for the award of a degree by this or any other university.

Signature: _____

Author's Name: Farhat Abdullah

Dated: November, 2023

It is certified that the work presented in this thesis is carried out and completed under my supervision.

Signature: _____

Supervisor's Name: **Dr. Uzma Anjum**

Associate Professor

Department of English

Air University, Islamabad

Dated: November, 2023

Author's Declaration

I, Farhat Abdullah, hereby state that my PhD thesis titled: Mapping of Semantic Class of Nouns in Urdu Conjunct Predicates to the English WordNet, is my own work and has not been submitted previously by me for taking any degree from this **AIR UNIVERSITY** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my Graduation, the University has the right to withdraw my PhD degree.

Signature:_____

Author's Name: Farhat Abdullah

Dated: November 2023

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled: Mapping of Semantic Class of Nouns in Urdu Conjunct Predicates to the English WordNet, is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the **HEC** and **AIR UNIVERSITY** towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/ cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the University reserves the rights to withdraw/revoke my PhD degree and that HEC and Air University has the right to publish my name on the HEC/ Air University Website on which names of students are placed who submitted plagiarized thesis.

Student/Author Signature:

Name:

Publications

Abdullah, F., Ahmed, T., & Anjum, U. (2021). Urdu conjunct predicates (N+ V) inventory from Urdu Universal Dependency Corpus. *Corprorum: Journal of Corpus Linguistics*, *4*(1), 15–29. https://journals.au.edu.pk/ojscrc/index.php/crc/article/view/265/214

Ahmed, K., Zahra, F. T., Abdullah, F., & Habib, M. A. (2021). Voices from the world: An analysis of discourse fragments from world leaders on Covid-19. *Linguistica Antverpiensia*, *6*(3), 4315–4345.

Zahra, F. T., Saleem, T., Abdullah, A. F., & Khan, M. A. (2021). Testing Matrix Language Framework Model On Urdu-English online news entity: A creative approach. *Multicultural Education*, *7*(1), 265–274. https://doi.org/265–274. 10.5281/zenodo.4460350

Ruba, A., Abdullah, F., Ahmed, K., & Basharat, A. (2021). Online learning experience and challenges of undergraduate students during COVID-19. *Journal of English Language, Literature and Education*, *3*(01), 1–28. https://doi.org/10.54692/jelle.2021.030164

Zahara, F., Saleem, T., Joiya, N., & Abdullah, F. (2020). Mayers-Scotten's 4-M Model: A Qusai-experimental study of Pashto-English morphological ability. *Amazonia Investiga*, *9*(34), 8–16. https://doi.org/10.34069/AI/2020.34.10.1

Abdullah, F. (2013). The Use of E-contents for teaching language subjects. *Paper Proceedings of Conference on Language, Literature & Linguistics 2013*, 41–60. https://uniqueca.com/archives/pdf/2013/Proceeding%20LLL%202013.pdf#page=42

Zafar, A., Mahmood, A., Abdullah, F., Zahid, S., Hussain, S., & Mustafa, A. (2012). Developing Urdu WordNet using the merge approach. Proceedings of the Conference on Language and Technology, 55–59. https://www.cle.org.pk/Publication/papers/2012/8.pdf

Acknowledgments

When I chose this spring midnight to write this last part of my thesis, it gave me a melancholy feel. Concluding and saying bid the work you put in efforts to nurture, is not easy. Now, when you see it growing and taking shape, it is natural to think of further shaping it to the next level, but the time constraint does not allow it, and compels you to take a pause here.

If I reflect, I can clearly see that accomplishing this research task was not possible without the relentless support and continued mentoring of Dr. Baber Sultan Khan. May it be an early Sunday morning or after business hours, whenever I sent him a naïve question in my state of bewilderment, I always found him passionate to respond with a scholarly eagerness which has been a beacon for me to see my pathway. As a linguistic giant, he has been tolerant of my limited knowledge and all I learned is because of scholarly and challenging discussions with him. I am indebted to him for his detailed feedback on the discussion part of this dissertation. He always inspired me to think and come up with a reasonable argument. Sitting in his office, I not only improved my linguistic knowledge but also saw him heading the department with enthusiasm and utmost concern for every issue being discussed with him. I am the one who witnessed his empathetic behavior with both faculty members and students.

Dr. Uzma Anjum lives in my heart for her unconditional motherly support. She always supported me and allowed me to experiment with my research ideas. In the low phases of this doctoral journey, she has been a friendly counselor who extended her supervision to keep my wavering emotions on track and pushed me to complete it on time.

I cannot be honest without mentioning Dr. Mirriam Butt when I met her at Conference on Language and Technology (2020) in UET, Lahore. Her passion for complex predicates in Urdu is evident in her prolific and continued publications on the complex predicates. In a quest for a research gap, I many times tried to frame different hypotheses which ultimately got a definite shape after a brief but fruitful discussion with her during the tea break.

At the same conference, I spotted Dr Tafseer Ahmed whom I requested guidance on the data collection. He is the one who mentored me in selecting and extracting data on

Urdu conjunct predicates from Urdu Universal Dependency Treebank Corpus which could not have been nearly possible for me without any computational background.

Faheem, my husband, has been very supportive of my aims and ambitions. His trust in my abilities encouraged me to follow my goals. To do all this, the energy and euphoric motivation has been drawn from my three children: Ibrahim, Nafia, and Sofia. I cannot think of a more hardworking and disciplined son than Ibrahim. He has been my monitor and whenever I stuck in research somewhere and wandered mindlessly on social media, he is the one who kept on reminding me of my Ph.D. timeline. Nafia, a tech-savvy and grammar-nazi teen, has always been a great support since her early childhood, and her understanding of my academic deadlines always makes me emotional and indebted. I highly value her support of my academic journey even when she was a toddler. Sofia, exactly like her name, has greatly impressed me with her deep mystic and philosophical thoughts. To explain my research venture to her, I made use of an analogy of a quest for a pearl lying deep in the ocean. She kept asking me about my success to find the pearl until one day when I told her that I have found it.

I, hereby, acknowledge the Higher Education Commission of Pakistan which selected me for the prestigious Faculty Development Ph.D. Scholarship Award, and enabled me to complete Ph.D. research program at Nanyang Technological University (NTU), Singapore. Dr. Francesco Perono Cacciafoco, my foreign supervisor, not only appreciated my work ethics but also entrusted me with the teaching opportunity at this top 13th university in the world. He helped me in conducting an official research seminar at the Department of Humanities, NTU, Singapore http://soh.school/lms-farhat. Credit goes to Dr. Perono for letting me share my work with eminent professors including Dr. Randy LaPolla and Dr. K. K. Luke. Special thanks to Dr. Randy LaPolla who with his questions gave me a chance to elaborate on the knowledge related to my work.

During my Ph.D. Research program, I have been privileged to be hosted by a Pakistani family in Singapore. Akhtar Bhai and Nadia with their kids Aulia and Zahra have been greatly generous and warmhearted. Humming nursery rhymes with Aulia and Zahra has been therapeutic and made me less homesick and allowed me to focus on my work. Both Akhtar Bhai and Nadia are great cooks as I still remember the scrumptious food with which they have been treating me. The learning part from traveling came from them accompanying me to Santosa Beach, Marina Bay Sands, Gardens by the Bay, The Jewel, and the famous Mustafa, and especially Aulia who has been my guide on MRTs and introduced me to the Kiaso culture of Singapore. Nadia has been a great walking

companion to the Bedok community library, one of my favorite places where I used to spend on average ten hours daily for my research work. I am so happy that because of my Ph.D., I got connected with these wonderful people.

I am thankful to Dr. Irfan Abbas who looked at my reference list and identified things to be improved. I must admit that I struggled to meet the formatting expectations of my institution, and if you see my work as correct in formatting and referencing, it is because of Dr. Sadia Irshad and Dr. Sham Haider who meticulously reviewed my draft.

Mapping of Semantic Class of Nouns in Urdu Conjunct Predicates to the English WordNet

by

Farhat Abdullah

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Linguistics

Abstract

This study aims to develop an inventory of Urdu conjunct predicates which helps to draw semantic classes of Urdu nouns, and to investigate compatibility patterns between Urdu nouns and light verbs such as argument structure, case marking, and semantic role of the subject argument influenced by Urdu host nouns in the clause, and to develop a tool to identify true Urdu conjunct predicate. The complex predicate and its subtypes are considered as single constituents where a light verb is hosted by a verb (V_1+V_2), a noun (N+V), or an adjective (Adj+V) and known as compound predicate and conjunct predicate respectively. Urdu Conjunct predicate (N+V), a type of complex predicate, has not been fully explored to date for its semantic and syntactic information. The lack of adequate lexical resources related to Urdu conjunct predicates has been a challenge for Natural Language Processing applications. To develop an Urdu N+V inventory, two Urdu Corpora i.e., Universal Dependency Urdu Treebank Corpus and Urdu WordNet Wordlist 0.1 are used. To draw semantic classes of Urdu nouns, semantic senses of Urdu nouns are mapped onto twenty-five semantic noun primes in English WordNet. To elaborate on combinatory restrictions between Urdu nouns and light verbs, a theory of correlation between semantic and syntactic context has been employed (Levin, 1993). The resultant Urdu N+V inventory mostly included abstract noun classes with very low instances of tangible noun classes which are compatible with light verb. A tool is developed to identify true conjunct predicates so that these can be inserted into Urdu electronic resources such as Urdu WordNet. Achievement of these research objectives may contribute to the digitally enabled status of the Urdu language which is less resourced to date.

Thesis Supervisor: Dr. Uzma Anjum

Title: Associate Professor of Linguistic

Table of Contents

Chapter 1: Introduction 1

1.1 Brief Overview 2

1.2 Background of Urdu Complex Predicate 3

1.3 Statement of the Problem 9

1.4 Research Objectives 10

1.5 Research Questions 10

1.6 Significance of the Study 10

1.7 Delimitation of the Study 11

1.8 Chapter Breakdown 12

1.9 Summary 14

Chapter 2: Literature Review 16

2.1 Complex Predicate 19

2.2 Conjunct Predicate: A Type of Complex Predicate 25

2.3 Incorporation of Noun in Conjunct Predicate (N+V) 29

2.4 Constituency Test, including Movement, Conjunct question, and Coordination Tests 37

2.5 Semantic Contribution of Nouns in Conjunct Predicate 40

2.6 Semantic Contribution of Light Verbs in Conjunct Predicate 52

2.7 Compatibility between Noun and Light Verb in Conjunct Predicate 56

2.8 Influence of Light verbs in a Conjunct Predicate Construction on the Subject Case-Marking 56

2.9 Argument Structure of Conjunct Predicate 59

2.10 Semantic Roles in Conjunct Predicate Constructions 62

2.11 Tense/Aspect 64

2.12 Different Theoretical Frameworks used to Study Different Complex Predicates 65
2.12.1 Lexical Functional Grammar 65
2.12.2 Semantic Componential Approach 67
2.12.3 Relational Semantics 67
2.12.4 Lexical Semantics 71
2.13 Complex/Conjunct Predicates and English WordNet 73
2.14 WordNets in Different Languages 74
2.15 Mapping of Urdu Nouns onto English WordNets 80
2.16 Summary 83
Chapter 3: Research Methodology 84
3.1 Research Design 84
3.2 Theoretical Framework 86
3.3 Choice of Urdu Corpora 89
3.4 Data Collection Procedures 92
3.5 Data Analysis Procedures 95
3.6 Semantic Orientation of the Light Verbs 102
3.7 Semantic Compatibility of Urdu Noun with Light Verbs 102
3.8 Mapping of Urdu Conjunct Predicate onto English WordNet 103
3.9 Summary 105
Chapter 4: Results 107
4.1 Development of Urdu N+V Inventory 108
4.2 Semantic Classes of Nouns and Light Verbs in Urdu Conjunct Predicates (N+V) 129
4.2.1 Urdu Nouns 129
4.2.2 Light Verbs 131
4.3 Noun Act 133
4.3.1 Noun Act: Semantic Orientation 133

4.3.2 Noun Act: Compatibility with the Light Verbs 137
4.4 Noun Communication.....145
4.4.1 Noun Communication: Semantic Orientation.....145
4.4.2 Noun Communication: Compatibility with the Light Verbs 148
4.5 Noun Cognition.....153
4.5.1 Noun Cognition: Semantic Orientation 153
4.5.2 Noun Cognition: Compatibility with the Light Verbs.....155
4.6 Noun Attribute 157
4.6.1 Noun Attribute: Semantic Orientation.....157
4.6.2 Noun Attribute: Compatibility with the Light Verbs 158
4.7 Noun Artifact 160
4.7.1 Noun Artifact: Semantic Orientation.....160
4.7.2 Noun Artifact: Compatibility with the Light Verbs 160
4.8 Noun Process 161
4.8.1 Noun Process: Semantic Orientation.....161
4.8.2 Noun Process: Compatibility with Light Verbs 162
4.9 Noun State.....162
4.9.1 Noun State: Semantic Orientation 163
4.9.2 Noun State: Compatibility with the Light Verb 163
4.10 Noun Feeling.....165
4.10.1 Noun Feeling: Semantic Orientation 165
4.10.2 Noun Feeling: Compatibility with the Light Verbs.....166
4.11.1 Noun Phenomenon: Semantic Orientation 168
4.11.2 Noun Phenomenon: Compatibility with the Light Verbs.....168
4.12 Noun Time 168
4.12.1 Noun Time: Semantic Orientation.....168
4.12.2 Noun Time: Compatibility with the Light Verbs 169

4.13 Noun Event 170
4.13.1 Noun Event: Semantic Orientation 170
4.13.2 Noun Event: Compatibility with Light Verbs 171
4.14 Noun Group 172
4.14.1 Noun Group: Semantic Orientation 172
4.14.2 Noun Group: Compatibility with the Light Verbs 173
4.15 Noun Possession 173
5.14 4.15.1 Noun Possession: Semantic Orientation 173
5.15 4.15.2 Noun Possession: Compatibility with Light Verbs 174
4.16 Noun Person 175
4.16.1 Noun Person: Semantic Orientation 175
4.16.2 Noun Person: Compatibility with the Light Verbs 175
4.17 Noun Relation 176
4.17.1 Noun Relation: Semantic Orientation 176
4.17.2 Noun Relation: Compatibility with the Light Verbs 176
4.18 Summary 177
Chapter 5: Discussion 178
5.1 Inventory of Urdu Conjunct Predicate (N+V) 180
5.2 Semantic Classes of Nouns in Urdu Conjunct Predicate (N+V) 182
5.3 Semantic and Syntactic Connection between Nouns and Light Verbs in Urdu Conjunct Predicates (N+V) 191
5.4 Noun Act 192
5.4.1 Noun Act: Semantic Orientation 192
5.4.2 Noun Act: Compatibility with the Light Verbs 193
5.4.3 Noun Act: Case Marking 194
5.4.4 Noun Act: Argument Structure 195
5.4.5 Noun Act: Semantic Roles 196

5.5 Noun Communication 197
5.5.1 Noun Communication: Semantic Orientation 197
5.5.2 Noun Communication: Compatibility with Light Verbs 197
5.5.3 Noun Communication: Case Marking 198
5.5.4 Noun Communication: Argument Structure 200
5.5.5 Noun Communication: Semantic Roles 200
5.6 Noun Cognition 201
5.6.1 Noun Cognition: Semantic Orientation 201
5.6.2 Noun Cognition: Compatibility with Light Verbs 201
5.6.3 Noun Cognition: Case Marking 202
5.6.4 Noun Cognition: Argument Structure 203
5.6.5 Noun Cognition: Semantic Roles 204
5.7 Noun Attribute 204
5.7.1 Noun Attribute: Semantic Orientation 204
5.7.2 Noun Attribute: Compatibility with Light Verbs 205
5.7.3 Noun Attribute: Case Marking 205
5.7.4 Noun Attribute: Argument Structure 207
5.7.5 Noun Attribute: Semantic Roles 207
5.8 Noun Artifact 208
5.8.1 Noun Artifact: Semantic Orientation 208
5.8.2 Noun Artifact: Compatibility with Light Verbs 208
5.8.3 Noun Artifact: Case Marking 209
5.8.4 Noun Artifact: Argument Structure 210
5.8.5 Noun Artifact: Semantic Roles 210
5.9 Noun Process 211
5.9.1 Noun Process: Semantic Orientation 211
5.9.2 Noun Process: Compatibility with Light Verbs 212

5.9.3 Noun Process: Case Marking....212
5.9.4 Noun Process: Argument Structure....213
5.9.5 Noun Process: Semantic Roles....214
5.10 Noun State....214
5.10.1 Noun State: Semantic Orientation....214
5.10.2 Noun State: Compatibility with Light Verbs....215
5.10.3 Noun State: Case Marking....215
5.10.4 Noun State: Argument Structure....217
5.10.5 Noun State: Semantic Roles....217
5.11 Noun Feeling....217
5.11.1 Noun Feeling: Semantic Orientation....218
5.11.2 Noun Feeling: Compatibility with Light Verbs....218
5.11.3 Noun Feeling: Case Marking....218
5.11.4 Noun Feeling: Argument Structure....219
5.11.5 Noun Feeling: Semantic Roles....220
5.12 Noun Phenomenon....221
5.12.1 Noun Phenomenon: Semantic Orientation....221
5.12.2 Noun Phenomenon: Compatibility with Light Verbs....221
5.12.3 Noun Phenomenon: Case Marking....221
5.12.4 Noun Phenomenon: Argument Structure....222
5.12.5 Noun Phenomenon: Semantic Roles....223
5.13 Noun Time....223
5.13.1 Noun Time: Semantic Orientation....223
5.13.2 Noun Time: Compatibility with the Light Verbs....223
5.13.3 Noun Time: Case Marking....224
5.13.4 Noun Time: Argument Structure....225
5.13.5 Noun Time: Semantic Roles....226

5.14 Noun Event 226
5.14.1 Noun Event: Semantic Orientation 226
5.14.2 Noun Event: Compatibility with Light Verbs 227
5.14.3 Noun Event: Case Marking 227
5.14.4 Noun Event: Argument Structure 229
5.14.5 Noun Event: Semantic Roles 230
5.15 Noun Group 230
5.15.1 Noun Group: Semantic Orientation 230
5.15.2 Noun Group: Compatibility with Light Verbs 231
5.15.3 Noun Group: Case Marking 231
5.15.4 Noun Group: Argument Structure 232
5.15.5 Noun Group: Semantic Roles 232
5.16 Noun Possession 232
5.16.1 Noun Possession: Semantic Orientation 232
5.16.2 Noun Possession: Compatibility with Light Verbs 232
5.16.3 Noun Possession: Case Marking 233
5.16.4 Noun Possession: Argument Structure 235
5.16.5 Noun Possession: Semantic Roles 235
5.17 Noun Person 236
5.17.1 Noun Person: Semantic Orientation 236
5.17.2 Noun Person: Compatibility with Light Verbs 236
5.17.3 Noun Person: Case Marking 236
5.17.4 Noun Person: Argument Structure 237
5.17.5 Noun Person: Semantic Roles 237
5.18 Noun Relation 238
5.18.1 Noun Relation: Semantic Orientation 238
5.18.2 Noun Relation: Compatibility with the Light Verbs 238

5.18.3 Noun Relation: Case Marking 238
5.18.4 Noun Relation: Argument Structure 239
5.18.5 Noun Relation: Semantic Roles 239
5.19 Insertion of Urdu Conjunct Predicate in Urdu WordNets 240
5.19.1 Categorization of Urdu Conjunct Predicate 241
5.19.2 Connection between the Argument Structure, Semantic Sense and Case Marking 242
5.20 Identification of True Urdu Conjunct Predicate 243
5.21 Significance of Research Contribution 246
5.22 Summary 250
Chapter 6: Conclusion 253
References 262

List of Figures

FIGURE	TITLE	PAGE
1.1	Size and Vitality of Urdu Language	4
2.1	Types of complex predicate	22
2.2	Noun-Centric Analysis	62
2.3	Multi-word Analysis	62
2.4	Concept of a Lexical Matrix	75
3.1	Universal Dependency Urdu Treebank Corpus	92
3.2	Semantic and Syntactic Compatibility of Nouns with the Light Verb	97
3.3	Urdu Keyboard Used in an Online Urdu-English Dictionary	99
3.4	Multiple Senses of Noun ‘Memory’ in Princeton English WordNet	100
5.1	Semantic Orientation of Urdu Conjunctive Noun Classes	185

List of Tables

TABLE	TITLE	PAGE
2.1	List of 25 Unique Beginners for Noun in WordNet	49
2.2	Common Urdu Light Verbs	55
2.3	Urdu Case Markers and their Functions	58
2.4	Urdu Pronominal Forms in Nominative and Accusative /oblique Case	59
2.5	Categories of Tenses	66
3.1	Noun Classes and their Lexicographer File Codes	102
4.1	Urdu Noun + Verb Inventory	109
4.2	15 Unique Conjunctive Urdu Nouns	130
4.3	Urdu Light Verbs	131
4.4(a)	Mapping of Semantic class and Argument-taking Ability of Urdu Light Verbs	132
4.4(b)	Mapping of Semantic Class and Argument-taking Ability of Urdu Light Verbs	132
4.5	Mapping of Urdu Conjunctive Noun Act	134
4.6	Compatibility Pattern of Noun Act with Light Verbs	138
4.7	Mapping of Urdu Conjunctive Noun Communication	144
4.8	Compatibility Pattern of Noun Communication with Light Verbs	147
4.9	Mapping of Urdu Conjunctive Noun Cognition	151
4.10	Compatibility Pattern of Noun Cognition with Light Verbs	153
4.11	Mapping of Urdu Conjunctive Noun Attribute	155
4.12	Compatibility Pattern of Noun Attribute with Light Verbs	156
4.13	Mapping of Urdu Conjunctive Noun Artifact	157
4.14	Compatibility Pattern of Noun Artifact with Light Verbs	158
4.15	Mapping of Urdu Conjunctive Noun Process	158
4.16	Compatibility Pattern of Noun Process with Light Verbs	159

4.17	Mapping of Urdu Conjunctive Noun State	160
4.18	Compatibility Pattern of Noun State with Light Verbs	161
4.19	Mapping of Urdu Conjunctive Noun Feeling	162
4.20	Compatibility Pattern of Noun Feeling with Light Verbs	163
4.21	Mapping of Urdu Conjunctive Noun Phenomenon	164
4.22	Compatibility Pattern of Noun Phenomenon with Light Verbs	165
4.23	Mapping of Urdu Conjunctive Noun Time	165
4.24	Compatibility Pattern of Noun Time with Light Verbs	166
4.25	Mapping of Urdu Conjunctive Noun Event	167
4.26	Compatibility Pattern of Noun Event with Light Verbs	167
4.27	Mapping of Urdu Conjunctive Noun Group	169
4.28	Compatibility Pattern of Noun Group with Light Verbs	169
4.29	Mapping of Urdu Conjunctive Noun Possession	170
4.30	Compatibility Pattern of Noun Possession with Light Verbs	171
4.31	Mapping of Urdu Conjunctive Noun Person	171
4.32	Compatibility Pattern of Noun Person with Light Verbs	172
4.33	Mapping of Urdu Conjunctive Noun Relation	172
4.34	Compatibility Pattern of Noun Relation with Light Verbs	172
5.1	Semantic Inventory for Noun Classes in Urdu Conjunct Predicate (N+V)	180
5.2	Noun Act and Di/transitive Light Verbs	192
5.3	Noun Communication and Di/transitive Light Verbs	196
5.4	Noun Cognition and Di/transitive Light Verbs	200
5.5	Noun Attribute and Di/transitive Light Verbs	203
5.6	Noun Artifact and Di/transitive Light Verbs	207
5.7	Noun Process and Di/transitive Light Verbs	210
5.8	Noun State and Di/transitive Light Verbs	213
5.9	Noun Feeling and Di/transitive Light Verbs	216
5.10	Noun Phenomenon and Di/transitive Light Verbs	219
5.11	Noun Time and Di/transitive Light Verbs	222
5.12	Noun Event and Di/transitive Light Verb	226

5.13	Noun Possession and Di/transitive Light Verb	231
5.14	Noun Person and Di/transitive Light Verb	233
5.15	Noun Relation and Transitive Light Verb	236
5.16	Categories of Urdu Conjunct Predicates	237

Abbreviations

Abbreviations and Acronyms

Complex Predicate	CP
Agentive Ergative Nominative	AEN
Experiencer Dative	ED
Syntactic Compound Verbs	SCpdVs
Lexical Compound Verbs	LCpdVs
Masculine Singular	M.SG.
Noun +Verb	N+V
Adjective +Verb	Adj+V
Natural Language Processing	NLP
Expanded Graded Intergenerational Disruption Scale	EGIDS
First language	L1
Second Language	L2
Universal Dependency Urdu Tree Bank	UD_URDU_UTB
Universal Dependency	UD
Verb + Noun Phrase	V+NP
Semantic Structure	SEM STR
Argument Structure	ARG STR
Grammatical Function Structure	GF STR
Grammatical Category Structure	GC STR
Lexical Functional Grammar	LFG
Argument predicate	ARG-PRED
Noun Act (A class of conjunctive nouns)	N_{act}
Intransitive Light Verb	$V_{intrans}$

Transitive Light Verb	V_{trans}
Ditransitive Light Verb	$V_{ditrans}$
Direct Object	DO
Indirect Object	Indirect Object
Light Verb 'Do/*Kar'*	LV_{kar}
Perfective Aspect	PERF
Perfective Aspect	+PERF
Imperfective Aspect	-PERF
Imperfective Aspect	IMP
Noun Communication (A class of conjunctive noun)	N_{com}
Noun Cognition (A class of conjunctive noun)	N_{cog}
Noun Attribute (A class of conjunctive noun)	N_{att}
Noun Artifact (A class of conjunctive noun)	N_{art}
Noun Process (A class of conjunctive noun)	N_{proc}
Noun State (A class of conjunctive noun)	N_{state}
Noun Feeling (A class of conjunctive noun)	N_{feel}
Noun Phenomenon (A class of conjunctive noun)	N_{phen}
Noun Time (A class of conjunctive noun)	N_{time}
Noun Event (A class of conjunctive noun)	N_{eve}
Noun Group (A class of conjunctive noun)	N_{grp}
Noun Possession (A class of conjunctive noun)	N_{pos}

Noun Person (A class of conjunctive noun)	N_{per}
Noun Relation (A class of conjunctive noun)	N_{rel}
Adverb of Time	Adv+Time
Semantic Role Labeling	SRL
Multiword Expressions	MWEs
Universal Part of Speech	UPOS
Part of Speech	POS
Center for Language Engineering	CLE
Unicode Transformation Format	UTF
Lexico-Syntactic Pattern Extraction	LSPE
Urdu Web	UrduWaC
Semantic Role Label	SRL
Multiword Expression	MWEs
Universal Part of Speech	UPOS

Chapter 1: Introduction

This chapter introduces the background of Urdu conjunct predicate as a subtype of complex predicate which is a common phenomenon in the Urdu language. This section gives an overview of studies related to complex predicates and identifies the need to work on Urdu conjunct predicate as there is not much research conducted in this domain. Inadequacy of proper lexical Urdu resources leads to problems in developing Natural Language Processing applications such as information retrieval, machine translation, semantic sense disambiguation (Saeed et al., 2019).This chapter aims to introduce the need to develop an Urdu conjunct predicate (N+V) inventory which can be used as a basis to draw semantic classes of host nouns which are yielded by mapping Urdu noun sense to the nouns senses found in English WordNet. The chapter highlights that the conjunct predicate is an abundant feature of the Urdu language. Though Urdu has attained an institutional status and has a large speech community, it is still a less-resourced language (Eberhard et al., 2019). This introductory part also addresses the different semantic and syntactic interpretations of incorporated Urdu nouns followed by a light verb which arise due to this abundant presence of this feature in the Urdu language. Furthermore, this section tries to present the background knowledge on the complex predicate and its subtypes specifically conjunct predicate, and how different previous works have contributed to its ontological perspective and addressed the linguistic ambiguities involved in the semantic and syntactic contribution of N+V as a single constituent.

This chapter also highlights the problem statement and necessitates the development of an N+V Urdu inventory required to fill the knowledge gap of absence of semantic classes of Urdu nouns in the conjunct predicate. Four clear research objectives work in progression as the development of work depends on the accomplishment of the prior achieved research objectives. Development of Urdu N+V inventory is the first step to figure out the semantic classes of Urdu nouns compatible with light verbs. It later moves on to elaborate on the compatible syntactic conditions required for an Urdu N+V instance. As a final tangible task in this research thesis, a tool for the identification of true Urdu conjunct predicates is developed as all the Urdu N+V instances may not be termed as conjunct predicates. This section also highlights the significance of the research study to improve the less-resourced status of the Urdu language which can

increase the probability of formulation of required lexical resources of Urdu used in Natural Language Processing applications for Urdu.

1.1 Brief Overview

Conjunct predicate, a type of complex predicate, is a common linguistic construction found in most Indo-Aryan languages including the Urdu language where a noun or an adjective host is incorporated into a light verb. Conjunct predicate has been a part of important research studies in Bantu, Romance, Urdu, and Hindi languages as a category of complex predicate (Alsina, 1993, 1997; Butt, 1995; Hook, 1974; Mohanan, 1994, 1997). Interesting but not completely revealed linguistic nature of conjunct predicate incited many research studies to explicitly describe its semantic and syntactic knowledge. In the world of digitization, where there is an abundance of Natural Language Processing applications available in the well-resourced languages such as English and Chinese. The presence of lexical resource related to conjunct predicate is quite significant to digitally empower the Urdu language in which these conjunct predicates are found abundantly.

Urdu which has a speech community of more than 280 million people all over the world. Its standard dialect has been selected and consequently adequate literature has been codified and documented. Furthermore, as a national language of Pakistan, it functions have been elaborated in the education system, media, bureaucracy, judiciary, etc. Urdu has also been accepted as a national language according to the constitution of Pakistan (1973). Its status has been declared as 'institutional' (Eberhard et al., 2019). Despite all this prominence, Urdu is a digitally less resourced language (Ahmed & Butt, 2011).

All three types of complex predicate are present in Urdu: compound predicate (V+V), conjunct predicate (N+V) and conjunct predicate (Adj+V). Conjunct predicate (N+V) is the focus of this work, where a noun host is incorporated and followed by a light verb. These conjunct predicates are abundantly found in the Urdu language. Neither all the possible instances of Urdu N+V are known nor can be anticipated due to the continuous influx of nouns as an open class category and their several combinations with different light verbs. Therefore, not an exhaustive knowledge about the semantic classes of nouns in Urdu conjunct predicate is available (Ahmed & Butt, 2011). Absence of this knowledge leads to inadequate information regarding semantic and syntactic combinatory restrictions of host noun and light verb in Urdu conjunct

predicates. In addition to this, not all instances of Urdu N+V instances cannot be termed as true conjunct predicate which necessitates a tool to recognize the true Urdu conjunct predicate. Development of such a tool may facilitate the insertion of Urdu conjunct predicate to digital lexical tool which may be beneficial to improve Urdu as a digitally enabled language.

To fill this gap, this work aims to develop an inventory of Urdu N+V instances which in turn further facilitates the discovery of the semantic classes of nouns in Urdu conjunct predicates. Detection of the conjunctive semantic noun classes was yielded through mapping the semantic senses of Urdu nouns onto noun senses found in English WordNet. Compatibility patterns of nouns with light verbs help to explicitly describe the semantic and syntactic combinatory restrictions. This knowledge of sets of combinatory principles supports the development of a tool for the identification of true conjunct predicate. Resultantly, the developed lexical resource provides prospects to help natural language processing experts to develop language applications adaptable to the Urdu language; at the same time, improves the efficacy of already existing language Natural Language Processing (NLP) programs.

This chapter introduces the need to work on Urdu conjunct predicate highlighting the problem it aimed to solve. Moreover, it also provided a brief introduction to the existing ontological knowledge about conjunct predicate in different languages, and the need to elaborate the knowledge explicitly. It clearly mentions the research objectives step by step which gradually helped to achieve broader research goal of empowering Urdu as a digitally enabled language.

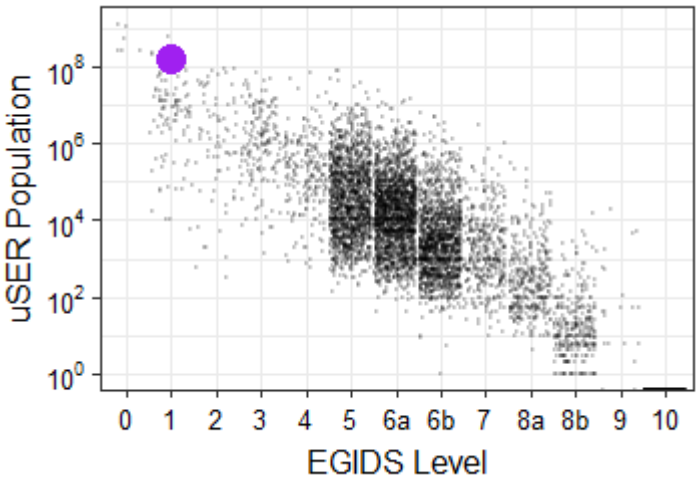
1.2 Background of Urdu Complex Predicate

Urdu is an Indo-Aryan language which has a large speech community in the world with the status of National language in Pakistan. As reported by Ethnologue (https://www.ethnologue.com/language/urd) on March 03, 2021, user population of Urdu is 164,000,000 in Pakistan. According to the census of 2018, there is a population of 15,000,000 people with Urdu as their first language in Pakistan; whereas, 149,000,000 people use it as their second language. Total users of the Urdu language in all countries are 230,052,270 (as L1: 69,006,470; as L2: 161,045,800). The Expanded Graded Intergenerational Disruption Scale (EGIDS) for the Urdu language is 1 (See Figure 1.1) which means that it has gained institutional status. It has been

developed to the extent that it is standardized and institutionalized for education, media, judiciary, and other domains of life specially in Pakistan (Eberhard et al. 2019).

Figure 1.1

Size and Vitality of Urdu Language (Eberhard et al. 2019)



The Urdu language possesses flexible word order, like most languages of Indo-Aryan origin, which mostly display Subject Object Verb (SOV) pattern (Bukhari, 2009). It is categorized as one of the 20 most spoken languages of the world (Eberhard et al., 2019). It is scarcely resourced for developing required computational programs to match the rapid pace of digitalization of other languages. This study may be a step towards digitally empowering Urdu and making it a technically enabled language compatible with English language which is used as a second language in Pakistan (Warsi 2004).

Review of the previous studies on complex predicate provides a sound foundation to move forward (Ahmed & Butt, 2011, Ahmed, 2010; Butt, 1993, 1995, 2010, Butt & Geuder, 2001; Butt & Ramchand, 2005; Ehsan & Butt, 2020; Kiani, 2013; Mohanan, 1994, 1997). The review of related studies also helps in identifying the knowledge gap. Types of complex predicate in the Urdu language consist of a variety of main predicational elements i.e., noun, main verb, or adjective compatible with a light verb. In all these complex predicates, light verb behaves as a syntactic head of this constituent (Butt, 1995). Morphological inflexions of the light verbs in complex predicates may

form causatives. Light verbs and the main predicational element do not always form a single syntactic category; therefore, some tests may be applied to identify true complex predicates (Butt & Geuder, 2001). Some complex predicate constructions in Urdu/Hindi have been identified:

i. Aspectual Complex Predicate

Light verb which is followed by a main verb shows the completeness (see Example 1) or spontaneity (see Example 2) of the action (Butt, 1995).

1.

Mussafir ne Panni Pii Liya
Passenger-M.F.ERG water-M.NOM drink Take- M.SG PERF
'Passenger drank the water.'

2.

Mussafir Chilla Utha
Passenger-M.SG.NOM. shriek rise-M.SG. PERF
'Passenger shrieked out.'

ii. Permissive Complex Predicate

Main infinitive verb is followed by a finite light verb, and it reflects the permission granted by the subject as an argument (see Example 3). This argument is semantically shared both by the finite light verb and the infinitive main verb (Butt, 1995).

3.

Malli ne Mussafir ko Phool Tornay Dia.
Gardener-M.SG.ERG Passenger-M.F.SG.ACC flower-M.SG.NOM Pluck INF.OBL give-M.SG.PERF
'Gardner let the passenger pluck the flower.'

iii. Causatives Complex Predicate

Urdu causative verbal predicates are overtly marked by two causative Urdu morphological inflexions (as shown in Examples 4 and 5) '-aa' and '-vaa' (Alsina, 1997; Butt, 1995).

4.

Akhtar ne ghar ban-**aa**-ya
Akhtar-M.SG.ERG house-M.SG.NOM made-CAUS.M.SG.PERF
'Akhtar built a house.'

5.

Akhtar ne mazdoroon se ghar ban-**vaa**-ya
Akhtar-M.SG.ERG laborer-M.PL.INST house-M.SG.NOM Made-CAUS.M.SG
'Akhar had a house constructed by laborers.'

iv. Urdu Conjunct (N+V) Predicate

Light verb is preceded by a noun (see Example 6). Light verb agrees to the left most unmarked argument (Mohanan, 1994). Nouns show compatibility with only selected light verbs. N+V combination is highly productive in the case of the Urdu language and could not be fully documented to date due to the diverse, individual, innovative, and creative use of language choice.

6.

Talib-e-Ilm ne sabaq yaad kia
Student-M.SG.ERG lesson-M.SG.NOM memorize- F.SG do-M.SG.PERF.
'Student memorized the lesson.'

v. Urdu Conjunct Predicate (Adj+V)

Light verb is preceded by an adjective (see Example 7) to form a conjunct predicate.

7.

Mazdoor ne raasta saaf kia
Laborer-M.SG.ERG passage-M.SG.NOM clean-M.SG do-M.SG.PERF
'Laborer cleaned the pathway.'

The light verb combined with different types of stem word either a main verb, noun or an adjective form a single syntactic constituent. Both the light verb and the stem word contribute to the semanticity of the verbal predicate (Butt, 1995).

Ahmed (2010) identified Urdu main verb classes based on their compatibility with the most frequent light verbs in Urdu complex predicates (V_1+V_2). It involved a compatibility pattern of V_1+V_2 using Urdu native speaker's intuition about the

acceptability of the collocation of main verb followed by a light verb. The work revolved around checking the acceptability of main verbs with three most frequent Urdu light verbs: 'Give/*Dia'*, 'Take/*Lia'* and Go/*Jaa'*. The work contributed to disambiguating the semantic senses of polysemous Urdu verbs. This work did not focus on the identification of true Urdu complex predicate based on any syntactic agreement. Ahmed (2010) followed Levin's (1993) proposition of a connection between semantic features and syntactic contexts, but instead of using alternations as not all exist in the Urdu language, he made use of acceptability of light verb with the main verb. This notion of acceptability is described as compatibility in the present work.

The same theoretical framework of Levin (1993) is followed to draw semantic classes of Urdu conjunct predicate (N+V) based on the syntactic pattern and lexical choices (Ahmed & Butt, 2011). The number of arguments present in the sentence, case markers on the subject argument and the compatibility with different light verbs depend on the semantic features of noun in Urdu conjunct predicate (N+V). The identification of semantic features of Urdu nouns and then categorization of Urdu conjunct predicate may contribute to develop a lexical resource which is claimed to be used for Natural Language Processing tools to improve the less resourced nature of the Urdu language. It is an effort to provide lists of argument taking nouns and verbs with subcategorization frames. Development of subcategorization frame of Urdu main verbs is doable in a considerable finite time, but Urdu Noun + Verb is a highly productive phenomenon and all Urdu N+V instances are difficult to predict. Absence of possible list of Urdu conjunct predicates is a hindrance in the development of verbal subcategorization for Urdu conjunct predicates. It creates a gap of knowledge regarding the combinatory restrictions between Urdu noun and the light verb. This knowledge gap leaves the Urdu language left behind in the era of digitization.

For the last three decades, though a need to technically empower the Urdu language has already been felt, it is still an under resourced language. Lack of relevant lexical resources has been a hindrance in developing Natural Language Processing tools for the Urdu language. Natural Language Processing programs need explicit records of conjunct predicates as single semantic constituents to develop computational tools for a language. A lack of adequate lexical resources may pose difficulty in developing useful computational programs which can be facilitated by building a systematic semantic standard for identifying conjunct

predicates. The UD_URDU_UTB corpus which is a universal dependency Urdu corpus (Bhat et al. 2017; De Marneffe et al., 2021) is used for this study. Universal Dependency annotated linguistic resources have become a preference of computational linguists in the recent decade for advanced computational linguistic programs. Furthermore, to incorporate a local dialect, Pakistani Urdu corpus is also used to elicit the conjunct predicate (N+V) pattern (Urdu WordNet 1.0 Wordlist, 2013)**.**

The present work aims to focus on drawing the semantic classes of nouns in Urdu N+V conjunct predicates. It explores the compatibility of different noun classes with varying morphological forms of light verb lemmas. It is an effort to develop the inventory of Urdu N+V collocations. Here, it is very important to announce that not every N+V instance can be labelled as conjunct predicate due to un/availability of incorporation of its nominal host. In a true conjunct predicate, the nominal part (N) is incorporated with the light verb, and it does not agree with it in terms of its phi features such as person, gender, and number. In fact, light verb in a conjunct predicate agrees with its covertly unmarked leftmost object argument. In this research many Urdu N+V instances are investigated for its semantic and syntactic compatibilities which provides an adequate data to interpret a comprehensive tool to recognize the true conjunct predicate. In this study, an agreement test is employed to verify the noun incorporation to light verb. This research provides a sizeable repository of Urdu N+V and provides a tool to recognize the true conjunct predicate. It not only layers out the combinatory restrictions but also highlights the phenomenon in which the noun is incorporated and becomes a part of verbal constituent. In this work for an ease of referring, N+V instances are written as conjunct predicate; however, it focuses on deducing a theory related to the identification of true conjunct predicate towards the end after getting the relevant semantic and the syntactic information related to Urdu noun and light verbs.

Levin (1993) classic proposition of verb classification is used as a theoretical framework for the current study. It is guided by the notion that syntactic context of verb in terms of its expression and realization of arguments is mainly dependent on its semantic sense (Dukes, 2000). Her work revolves around the notion that through the syntactic constructions of a verb, its meaning can also be elicited. This research reported here paves the way towards the evolution of lexical knowledge of Urdu conjunct predicates. Highly productive and sometimes unprecedented instances of N+V in the Urdu language drew the attention of many researchers. Different

paradigms have been employed to find the connection between the syntactic and semantic features of complex predicates (Butt 1995; Kiani 2013; Mohanan 1997), but conjunct predicates remained a challenge for both theoretical and computational linguistics. The development of an exhaustive lexical tool will bridge the knowledge gap which has been a bottleneck to develop Natural Language Processing programs for the Urdu language. Princeton English WordNet plays a role of gold standard in determining the sense and further its mapping onto the respective Urdu noun with the similar semantic sense (Fellbaum et al., 2010).

1.3 Statement of the Problem

Conjunct predicate N+V is an abundantly found instance in Urdu where it represents a complex predicate consist of two semantic heads but mostly one verbal constituent. Ability to recognize and read conjunct predicate accordingly is integral to develop Natural Language Processing programs in Urdu. However, not all possible instances of Urdu conjunct predicate N+V could be documented due to the prolific and productive combinatory nature of host noun and light verb. Continuous influx of Urdu noun to the lexicon is another challenge to have a pre-decided inventory of Urdu N+V. Furthermore, not a related knowledge is available in the recent related literature about the semantic classes of host nouns and their combinatory syntactic restrictions with light verbs in Urdu conjunct predicate. Lack of such lexical tools has been a bottleneck in the development and efficacy of Natural Language Processing (NLP) applications available in Urdu which makes it a less resourced language. To bridge this knowledge gap, a good sizeable inventory of Urdu N+V is needed which can be used as a basis to further study the semantic classes of host nouns and help in explicitly draw compatibility restrictions between Urdu nouns and light verbs. The nouns included in the Urdu N+V inventory are extracted from two diverse Urdu corpora i.e., Universal Dependency Urdu Treebank (De Marneffe et al., 2021) and Urdu WordNet 1.0 Wordlist (2013) which have diverse contextual, multi generic and diachronic Urdu data. Identification of true conjunct predicates in Urdu is much needed phenomenon to systemize the insertion of Urdu conjunct predicates to electronic linguistic resources used in different computational linguistic programs. Absence of such tools and linguistic resources have been a hindrance in elevating the status of Urdu as a fully resourced language. An agreement test is devised as a tool to identify true Urdu conjunct predicates.

1.4 Research Objectives

The research objectives of this study are

1. To develop an inventory of Noun + Light Verb instances from Urdu corpora
2. To investigate the semantic classes of nouns in Urdu N+V instances, and map them onto the existing semantic senses in English WordNet
3. To interpret the range of possible syntactic constraints on the combinatory possibilities of noun and light verb in Urdu N+V instances
4. To formulate the tool for the identification of true conjunct predicate based on semantic and syntactic data interpretation

1.5 Research Questions

This research work focuses on the following research questions:

1. What are the possible instances of Urdu nouns followed by light verbs?
2. How can semantic classes of nouns in Urdu (N+V) instances be drawn while mapping them onto the semantic senses in English WordNet?
3. How the connection between semanticity and syntactic context be drawn by exploring the compatibility patterns between host noun and light verb in Urdu N+V instances?
4. How can a comprehension tool for the identifying true Urdu conjunct predicate (N+V) be regulated based on semantic and syntactic combinatory restrictions?

This study plans to identify significant observable restrictions on Urdu N+V instances as conjunct predicate. The research product came out as a categorized and codified semantic inventory which can be used as a lexical resource for building Natural Language Processing programs. The inventory of Urdu conjunct predicates will also specify the semantic classes of nouns that can possibly merge with light verbs. Development of inventory and declaration of semantic classes of host nouns help in highlighting the syntactic conditions which facilitate the compatibility of nouns with light verbs in a conjunct predicate.

1.6 Significance of the Study

This work aims to develop a systematic semantic inventory of Urdu N+V which can provide enough data to explore syntactically and semantically compatibility patterns

between host noun and light verbs in Urdu N+V collocations. This inventory is much needed to understand the correlation between the semantic class of Urdu nouns, argument structure of Urdu light verb, case marking and semantic roles of subject arguments in Urdu conjunct predicates. Finding the semantic and syntactic connection between Urdu nouns and the compatible light verbs is very importance to identify N+V as true conjunct predicate and to insert them in English WordNet. Development of an exhaustive lexical tool to identify true Urdu conjunct predicate N+V will bridge the knowledge gap which has been a barrier to develop Natural Language Processing programs for Urdu.

Unique, highly productive, and sometimes unprecedented instances of conjunct predicate in the Urdu language drew attention of many researchers who used different paradigms to study the intervening syntactic and semantic factors, but semantic classes of nouns in conjunct predicate have not been fully explored to date (Butt, 1995; Kiani, 2013; Mohanan, 1997). During the accomplishment of decided research objectives in this research, it is realized that Urdu nouns in different semantic relations such as synonymy, antonymy, and hyponymy can be linked to reveal their similar syntactic contexts. Some data evidence has been presented to confirm the semantic and syntactic connection between synonyms, antonyms and hyponyms though more can be done in a future study to further validate the presence of such a correlation.

1.7 Delimitation of the Study

This study does not claim to find all syntactic and semantic limitations to ensure the perfect semantic inventory of Urdu N+V. I used Urdu corpora to find the comprehensive data of N+V instances. First, a sizable, annotated Urdu corpus was required to elicit the possible N+V collocation to study the semantic and syntactic compatibilities of nouns with different light verbs. For this purpose, it needed multilayered annotated Urdu corpus so that conjunct predicate can easily be mined. I have chosen Universal Dependency Urdu Treebank Corpus (UD_URDU_UTB) which is a multilayer annotated corpus for almost all possible grammatical and syntactic information of the entries included in the sentences (Bhat et al., 2017; De Marneffe et al., 2021). The selection of this Universal Dependency Urdu Treebank Corpus for Urdu may benefit the study in terms of making it more compatible for future computational

work on Urdu because Universal Dependency annotated linguistic resources have become a preference of computational linguists in the recent decade. Furthermore, to incorporate a local dialect, Urdu WordNet Wordlist 1.0 is also crawled to elicit the unique Urdu nouns. This corpus was not annotated for its grammatical and syntactic categories but a list of 5000 words where the parts of speech are elicited using the intuition of native speaker of Urdu. The developed inventory of Urdu N+V instances contained all naturally occurring light verbs found in the Universal Dependency Urdu Treebank corpus. This inventory contained Persio-Arabic script just to maintain the original shape of the lexical tool in the target language (Urdu) which may not be readable for all the readers of majority readers, but these N+V instances are later on transliterated and translated as well in the later parts of the thesis just to enhance the understanding of the data. List of 280 unique Urdu Nouns is formed using two Urdu corpora. The nouns in this list could not be tested for their compatibility with all naturally occurring light verbs due to the time constraint allowed for the work. I limited this study to investigate ten most frequently occurring light verbs: 'Do/*Kar',* 'Become/*Hu*', 'Be/*He',* 'Put/*Rakh',* 'Come/*Aa*', 'Give'/*Dia,* 'Go/*Ja*', 'Take/*Lena*', 'Remain/*Rah',* and 'Hit/*Laga*'. 'Be/*He'* is the inflectional variation of the 'Become/*Hu',* but it may also serve as a copula that is why it is treated as a different verb in the list of ten light verbs.

More light verbs can be explored to see their compatibility with Urdu nouns. Due to the limited time available for the present study, it can be venture in a future advanced study in the same area. Urdu Adjective followed by light verbs also form conjunct predicate. Time constraints did not allow to include information about Urdu Adj+V instances, though it can be ventured in a future study.

1.8 Chapter Breakdown

First chapter introduces the overall field of conjunct predicate as a type of complex predicate. After a brief overview of the background studies, it highlights the research problem. The first chapter aims to clearly state what is the work about, what was the problem, why it is important to solve the problem, and how the research work would be executed. Research objectives and research questions are stated in an arrangement of the research flow and interlinked with the accomplishment of formerly mentioned stated research objective mentioned in that order. Each research question is the clear

indicator of research objective and arranged according to the course of action. This section of the study emphasized on how the research product may be directly linked to the industry, and what change it could bring after filling the mentioned knowledge gap. Second chapter comprises a comprehensive survey of the related literature on complex predicate especially conjunct predicate in different languages. Beginning from the general topic of complex predicate, it becomes specific to discuss conjunct predicate (N+V) and its components. Semanticity of nouns and light verbs is much debated and different categorizations are highlighted. It addresses how nouns in different languages are classified. The second chapter provides information on types of complex predicates in Urdu, argument structure, semantic classes of Urdu conjunct predicate. It also presents a taxonomy of 25 semantic noun primes which is used a theoretical support to categorized Urdu nouns in conjunct predicates. Among different theoretical frameworks to deal with conjunct predicate, Levin's (1993) proposition of correlation between the semanticity and syntactic context of an entity is highlighted to be used as a guiding theoretical framework for the present work. It debates on how host noun and light verb semantically and syntactically contribute to the holistic nature of conjunct predicate. Here both semantic and syntactic levels are elaborated for their probable connection. Furthermore, it is analysed that how these semantic and syntactic layers interact together to create a holistic effect to the sentence construction which include the argument structure, case marking and semantic roles of the subject argument. The second chapter also elaborates on different tests used to detect the single constituency status of complex predicates. It also necessitates on the identification of true conjunct predicate in the Urdu language so that these can be inserted in Urdu WordNet as single constituents.

Methodology section, the third chapter, clearly shapes out the research design with deliberate and clearly stated research method. It included explicit description of the theoretical framework employed in the present work. It answers, 'how Urdu corpora are chosen?', 'what was the rationale behind choosing specific Urdu corpora for the study?', and 'how noun and light verb instances are picked up and further checked for their probable compatibility with the light verbs?'. Compatibility of Urdu noun and light verbs is investigated in an elaborated way. Types of light verbs and their impact on case markers and the semantic role of subject argument were tabulated in a systematic way. This detailed inquiry helped to see the emerged pattern of semantic

classes of noun in Urdu conjunct predicate. Conjunctive Noun classes are emerged as the probable semantic classes of host noun in Urdu conjunct predicate.

Result section, fourth chapter, contains the extracted N+V instances from the Urdu corpora and the manually configuration of mapping of semantic senses of Urdu nouns to the English WordNet. Once the mapping of the noun data is completed, it gave rise to a pattern of compatible conjunctive Urdu noun classes. The data related to the semantic classes of Urdu nouns is presented in separate tabulated forms. In the second round of result compilations, Urdu nouns are checked for their compatibility with different light verbs. Th discussion chapter attempts to draw patterns from the compiled data. It discusses the nature of Urdu N+V inventory. It evaluates the emerged semantic classes of Urdu conjunctive nouns, provided an overview about the general composition of Urdu conjunctive nouns. It analysed the achieved research objectives one by one:

- The development of Urdu N+V inventory which can be used for future research studies.
- Theory related to the semantic classes of nouns in Urdu conjunct predicates.
- Explicit documentation of combinatory restrictions of noun and light verb in N+V instances.
- A tool is developed to identify true Urdu conjunct predicates.

Finally, the research is concluded by clearly stating down the research outputs. In addition to this, this section contains clear and specific recommendation for future studies to advance in the discipline.

1.9 Summary

This chapter highlights the need to work on Urdu conjunct predicates. It also necessitates that the development of a good sizeable inventory of Urdu N+V is important to further investigate semantic and syntactic information related to Urdu conjunct predicates. For example, it is needed to draw semantic classes of Urdu nouns after mapping them onto English WordNet. Furthermore, this section presented adequate background knowledge of the conjunct predicate as a type of complex predicate and its semantic-syntactic interpretations. It illuminated the research problem by highlighting the absence of any relevant lexical resource such as an inventory of Urdu conjunct predicate, knowledge about the semantic classes of Urdu nouns and a set of combinatory restrictions along with an absence of a tool to identify true Urdu conjunct predicates. This chapter also elaborates on how these research gaps can be

filled to contribute to the status of the Urdu language and make it a well-resourced language. The overall research is planned to discover the correlation between the semantic and syntactic contexts of components of Urdu conjunct predicates.

Upcoming chapter reviews the sufficient literature on the complex predicate and its subtypes emphasizing the conjunct predicate. As the debate progresses, it finds its niche to develop the rationale to work on developing the Urdu N+V inventory to explore more information about the semantic and syntactic correlation of N+V in an Urdu sentence.

Chapter 2: Literature Review

This chapter is about the related literature on complex predicate and its major categories. It has got vital and detailed information regarding the types of languages including Urdu and other Indo-Aryan languages in which the complex predicates are found. Enough literature is provided to elaborate on the taxonomy of complex predicates specifically in the Urdu language: compound/verbal complex V_1+V_2, nominal conjunct predicate (N+V), and adjectival conjunct predicate (Adj+V). This section highlights the noun categorization which helps to draw semantic class of nouns in Urdu conjunct predicate. Components of complex predicates contribute complexity to the semantic and syntactic structure of the constructions. The contribution of the stem (main verb, noun, or adjective) and the light verb to the argument, syntactic structure and semantic structure of the clause is analyzed. Owing to the diversity and productivity of the combinatory variations of complex predicates in such languages, it poses a hinderance in developing computational linguistic programs which results in less-resourced status of such languages including the Urdu language which is the focus of current research. This knowledge gap necessitates the exploration of the ways to deal with the inadequacy of resources related to complex predicate and its types from eclectic linguistic perspectives following different theoretical frameworks. The semantic contribution of complex predicates is interpreted differently in different languages which includes prepared/unprepared mind, language user's knowledge and belief state, completion of task and volitionality, etc.

A complete and detailed overview of complex predicates smoothly navigates the reader to the focus of the study i.e., conjunct predicates (N+V) in the second main part of this chapter. Both the components of conjunct predicates i.e., noun and the light verb contribute to the argument structure, case marking, semantic roles in one way or the other. The focus of the work is to figure out the combinations patterns and set of restrictions related to compatibility between different noun hosts and light verbs. High productivity of N+V instance results in large number of compatibilities which are difficult to scale owing to the open class of its host noun. Semantic class and semantic value of the noun host in Urdu conjunct predicate (N+V) is said to have a connection with the argument structure of the clause and the case marking on the subject argument. Light verbs in Urdu conjunct predicate influence not only the argument structure of the construction, but also affect the semantic role and case marking of the subject argument

in the sentence. Subsequent third main section includes the literature on how a noun host in Urdu/Hindi conjunct predicate is examined as an incorporated activity (Hook, 1972, 1978). Different tests such as Adjectival Modification, Wh-questions, Relativization, Conjoining, Addition of accusative marker to the noun host, Movement, and Agreement are explained with examples which may elaborate the incorporated nature of the nominal host in N+V instances. This section highlights the Urdu Aspectual and Permissive complex predicates proposed by Butt (1995), and it also contains five types of Hindi compound predicates proposed by Bhattacharyya et al., (2007) because of application of different tests. Identification of true conjunct predicate aims to resolve the issue of their insertion in a lexical resource like WordNet which may improve the less-resourced status of the Urdu language. Fourth main heading of this chapter is about the semantic contribution of nouns which are the predicational components in conjunct predicates. Governing features of noun taxonomy are explained. Noun primes of English WordNet and noun classifiers in other languages are discussed in detail. It is elicited that some nouns host i.e., 'Reliance/*bʰarosaa*' etc. are said to have influence on the argument structure of the construction. Case marking which is usually licensed by the light verb in conjunct predicate is also said to have been influenced by the semantic features of noun in N+V instance. Semantic value of noun host 'Attention/*dyhaan*' clearly influences the case marking of object argument. Verbs are comparatively more polysemous than nouns which reflect their semantically more flexible nature which also depend on their compatibility with the host nouns. Then there is relevant literature on how nouns are categorized in different languages. This noun classification follows different rules and norms which tell us the way the whole world is categorized. This section also talks about verb classification. Classification of verb is also linked with the physical properties or the position of an adjoining object noun argument. Most importantly, this section lists and reviews the twenty-five noun primes/classes suggested by Miller et al., (1990) to construct Princeton English WordNet. In the current research work, these noun classes are used as a main reference and a theoretical framework to map the senses of Urdu nouns in N+V instances.

Next main heading is about the semantic contribution of light verbs in conjunct predicate (N+V). This section contains specific knowledge about different terms used for main verb and light verb in (V_1+ V_2) instance. Furthermore, it also tries to differentiate light verbs from modal and auxiliaries in terms of no restriction of

compatibility with other non-finite verbs, and no contribution to the argument structure respectively. The phenomenon of compatibility issues is their ability to license arguments in the sentence are limited to light verbs only. Here some recognized light verbs in Urdu, Hindi, Punjabi and Gojri have also been reviewed. Furthermore, it elaborates on how a light verb contributes to an element of volitionality and subsequently licenses the case marking of the subject argument. Then the phenomenon of licensing of case marking by light verb is considered, and it is seen that how different Urdu light verbs influence case marking on the arguments in the construction. There are some transitive light verbs which assign agentive role to the subject argument.

Notion of predictability of compatibility is illuminated between noun host and light verb by examining the semantic classes of compatible nouns, and the transitivity of light verbs is discussed in this section which may validate the semantic and syntactic relevance hypothesis (Levin, 1993).

Taxonomy of Urdu cases, their markers and functions are presented in a detailed way to enable a new reader to navigate through examples presented in the thesis. Furthermore, Urdu pronominal forms in nominative and accusative/oblique case have been described explicitly.

The argument structure of conjunct predicate, which is its syntactic valency, is elaborated in the following part. Conjunct predicates can be categorized based on the semantic classes of Urdu nouns and argumenthood of light verbs i.e., transitivity, diatransitivity or intransitivity. Two approaches to examine arguments influenced by light verb construction are presented i.e., Noun-centric Analysis and Multiword Analysis. Assignment of the semantic role to the subject argument is another ability of light verbs which is elaborated with examples in a detailed manner. To elaborate on semantic roles of 'Agent' and 'Experiencer', their connection with the argumenthood of the light verbs in conjunct predicates is analyzed with ample examples.

After discussing the significant issues related to conjunct predicate, literature on different encountered theoretical frameworks is presented keeping their research focus in mind. For instance, Lexical Functional Grammar (LFG) which presents language in interrelated structures has been a frequently implied theory in dealing with complex predicates. LFG was often refereed to address the mismatch between syntactic and semantic information related to verbal complex phenomenon.

How a word can be broken down into its semantic primitives is dealt with in Componential Analysis though it remained inadequate to address issues related to

conjunct predicates. How lexical entities are semantically defined using Relational Semantics can also be a building block of discovery of some patterns in the noun classes of Urdu conjunct predicates. Last but not the least, Levin's (1993) model of English verb taxonomy is elaborated using her examples of English alternations. As all alternations are not found in the Urdu language, compatibility between different noun classes and light verbs is used to draw the semantic and syntactic connections between them.

Last section of literature review talks about WordNet as a useful electronic lexical resource which augmented the digital status of English language. Princeton English WordNet is also used as a 'Gold Standard' to build WordNets in other languages as well. This section previewed the semantic approaches involved in the development of English WordNet. Like other languages, a WordNet in Urdu is also built though it is in its nascent stage. Linguistic forms like complex predicates and their subtypes including conjunct predicate are found abundantly in the Urdu language. Not all instances of V_1+V_2, N+V and Adj+V are complex predicates and conjunct predicates respectively, and therefore all of them cannot be inserted into a WordNet as single constituents. This issue can be resolved by mapping the Urdu nouns in an N+V instances onto the related semantic senses of nouns in English WordNet. It may enable us to draw semantic classes of Urdu nouns in conjunct predicate based on which some compatibility pattern can be drawn which can afterwards subject to the identification of true Urdu conjunct predicate. These true Urdu conjunct predicates may be inserted in Urdu WordNet as single constituents. The discovery of compatible noun classes of Urdu conjunct predicate and the development of a tool to identify Urdu true conjunct predicates may likely to improve the less-resourced status of the Urdu language and augment the existing linguistic tools to enhance the efficacy of Urdu natural language programs.

2.1 Complex Predicate

Jespersen (1949) is said to be the first person who coined the term 'complex predicate' which was primarily applied to English V+NP collocation such as *have a nap*, *take a step*, *give a bath* etc. (Jespersen, 1949). Since the beginning of 1990s, the term

'complex predicate' has been getting constant attention in linguistics especially in the areas of semantics and syntax.

Known with different names, complex predicate, conjunct predicate, or compound verb is a commonly observed linguistic phenomenon in most Indo-Aryan languages such as Urdu, Hindi, Gujrati, Marathi and Punjabi. This linguistic feature provokes researchers to explore its semantic and syntactic interpretations from different linguistic perspectives with diverse approaches. Different theoretical frameworks have been employed to explore the syntactic and semantic aspects of complex predicates. Among the very sound works on complex predicates in Hindi-Urdu, Hook (1978), Mohanan (1994; 1997) and Butt (1993, 1994a; 1994b) are very comprehensive and explanatory studies which were further explored in several studies (Butt 2010; Butt & Geuder 2001; Butt & Ramchand 2005). These works have been dealing with diverse phenomenon such as noun incorporation, particle verbs, auxiliarization, serialization, phrasal verbs, causatives, denominal verbs, etc. (Manetta, 2019). In general, if we try to talk about complex predicates, it is about some linguistic items such as words or phrases which behave like single syntactic constituent but morphologically may consist of two or more lexical items. This specific feature calls for the existence of different levels of representation such as argument structure, phrase structure and functional structure (Alsina, 1993; Bowern, 2008; Butt, 1995; Folli et al., 2005; Godard & Samvelian, 2021; Mohanan, 1997). Whereas the other school of linguists made use of different theoretical backgrounds to deal with the complex predicate through frequently used syntactic complementation, morphology and head movement etc. (Amberber et al., 2010; Baker & McCallum, 1998; Davies & Rosen, 1988).

Khailna (to play), and *Chamkna* (to shine) are simple verbs in Urdu. *Maar Daalna* (to kill)and *Yaad Karna* (to remember/memorize), *Kush karna* (to please) are V+V, N+V and Adj+V constructions respectively which are different forms of complex predicate (see Examples 8, 9 and 10).

N+V and Adj+V have usually been categorized in literature as conjunct predicate, whereas V_1+V_2 is termed as compound predicate.

8.

Usne chor ko **maar daala**
S/he-M/F.SG.ERG theif-M.SG.DAT kill pour
'S/he killed the thief.'

9.

Usne sabaq **yaad kia**
S/he-M/F.SG.ERG lesson-M.SG.NOM memorize-F.SG do-M.SG.PERF
'S/he memorized the lesson.'

10.

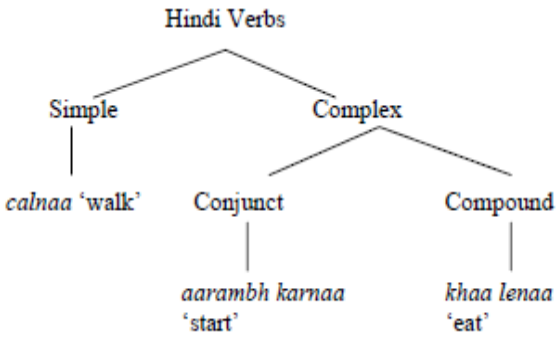
Usne apnay afsar ko **kush kia**
S/he-M/F.SG.ERG his/her officer-M.SG.DAT please do-M.SG.PERF
'S/he pleased his/her officer.'

Example number 8 is V_1+V_2 compound predicate where V_1 is the main verb and V_2 is the light verb. Examples 9 and 10 are categorized as conjunct predicates where nouns and adjectives are followed by a light verb. Butt (2005) is of the view that these nouns, adjectives, or main verbs are the main predicative elements of a complex predicate whereas the light verbs are usually the syntactic head. This light verb does not carry its distinctive or strong semantic sense but interacts with main verb or noun/adjective to convey the complete sense. She adds that light verbs do not always form a single syntactic category, and these can easily be distinguished from auxiliaries.

Bhattacharyya et al. (2006) presented the taxonomy of Hindi complex verbs as follows in the Figure 2.1 given below:

Figure 2.1

Types of Complex Predicate



(Source: Bhattacharyya et al., 2007)

Compound verb is one in which the main verb is followed by another verb. The second verb may be a light verb, modal or aspectual. Light verbs exhibit the meaning of completeness and suddenness. Complex predicate and its types consist of two lexical items but syntactically behave as a single unit. Overall semantic features are contributed by both the stem and the light verb. Compatibility of stem and light verb and their overall semantic orientation is easily understandable by native speakers, but at the same time, it requires deliberate and explicit information to run a computational program with linguistic data containing complex predicates and its other types.

Butt (1994a) asserts that the structure of a complex predicate is composed of a sequence of predicates that formulates a single predication. Furthermore, the constituents of complex predicate share the same tense, aspect, and mood as they form single sequential entity without any syntactic gap between them. Butt (1994a) elaborates those multiple functions performed by complex predicate and their semantic value which depends on the predicate semantic classes, sentence structure and other contextual assumptions. There have been two types of complex predicates identified in oceanic language: nuclear juncture and core juncture (Bril, 2007). These are further categorized into symmetrical and asymmetrical complex predicates. According to this classification, core juncture symmetrical complex predicate carries out the sequential purposive actions, whereas another specialized form of complex predicate i.e., nuclear asymmetrical gives information related to adverbs of manner (Van Valin, 1990; Van Valin & LaPolla, 1997).

Two processes of complex predicate formation have been investigated: merger and coindexation (Baker, 2010). Merged complex predicates are formed by merging the same kind of lexical and semantic predicates; whereas coindexation produces a different variety of complex predicates whose senses are difficult to express by simple predicates. One of the general rules of serial verb construction is that finite verbs are always light verbs which carry information of tense, aspect, mood, and agreement (Bhatt, 2005, 2008). Similar findings regarding light verbs have also been reported by Ahmed and Butt (2010).

The phenomenon of complex predicate in Bardi, a Nyulnyulan language spoken in North Australia, has also been studied with claims that there are three verbal types: raising verbs, restructuring predicates and light verbs (Bowern, 2008).

Pandharipande (1993) employs different syntactic phenomenon such as *passivization, participialization* , *agreement* and *causativization* to talk about the semantic compatibility between V_1 and V_2 in Marathi compound predicates (V_1+V_2). Passivization, participialization and agreement are only applied on the light Verb whereas causativization is applied to both verbal elements in a compound predicate (V_{1+} V_2) construction (Bhattacharyya et al., 2007).

Semantic compartmentalization of compound verbs of Kalasha based on 'prepared' and 'unprepared mind' is said to have categories related to the 'knowledge and belief state' of the speaker (Bashir 1993).

Two sorts of complex predicate i.e., lexical, and syntactic complex predicate have been presented by Williams (1997). In English, only lexical complex predicates are found. English lexical complex predicates have structures like Verb + Adjective (make clear, wipe clean) or Verb + Preposition (put together, kick over) (Cattel, 1984; Williams, 1997). On the other hand, both kinds of lexical and syntactic complex predicate are found in French language. The clear distinction of causative predicate drawn by Williams (1997) states that it is lexically complex if the embedded verb is intransitive, and it is syntactically complex if the embedded verb is transitive.

Another approach claims about the complex syntactic structures of all verbal predicates which are constructed by atomic components and licensed by unique syntactic principles (Hale & Keyser, 1997). English denominal verbs are studied to explain the blocking of certain morphological derivations in terms of violations of some syntactic principles which poses a challenge in syntax (Hale & Keyser, 1997). Responding to their argument about denominal verbs, Paul Kiparsky (1997) argues that only syntactic approach will not be adequate to explain the systematic blockage of certain interpretation of denominal especially location verbs. He continues that the conceptual knowledge is based on the semantic theory of Bierwisch (1983) with a clear differentiation between levels of semantic form and conceptual structure (Wunderlich, 1997). The case of denominal verbs, action mostly refers to the conventional use of the noun (Kiparsky, 1997).

Complex predicates in Hungarian language are syntactically separable but they do not allow morphological derivation. Hungarian particle verbs are considered as lexical constituents which may consist of different syntactic category (Ackerman & LeSourd, 1997).

Three levels of grammatical representation i.e., argument structure, functional structure and phrase structure as per Lexical Functional Grammar have been used to categorize aspectual and permissive complex predicate (Butt, 1995). She uses a term 'Elaborated Argument Structure' for the interconnectedness of Urdu aspectual light verb and the case marking properties of resultant complex predicate; therefore, she used the semantic theory of Jackendoff (1992) to explain this linguistic phenomenon.

Four different levels of syntactic representation: argument structure, semantic structure, grammatical function structure and grammatical category, have been used to explain Hindi predicate with a special focus on case marking and agreement (Mohanan, 1997). Semantic structure (SEM STR) here, elaborates semantic aspects associated with the arguments of the predicate which are related to the morphological and syntactic information. Argument structure (ARG STR) carries the information about the valency which is the information about the number of arguments a predicate takes. Grammatical function structure (GF STR) gives the information about grammatical arguments of the predicate in terms like subject, object, and other grammatical features such as case, number, person, and gender. Grammatical category structure (GC STR) reflects the information about lexical entities as noun, verb, etc.

At the lexical level, the grammatical function, grammatical features, and category associated with the arguments of predicate are probable from their SEM STR and ARG STR detail. However, it is claimed that the reverse is not applicable because of asymmetric interdependence of these structures (Mohanan, 1997). This semantic structure is also termed as Lexical-conceptual structure in different studies. Simple argument structure in Hindi N+V complex predicate links with complex semantic structure and complex grammatical category structure as well (Mohanan, 1997). This linguistic phenomenon of grammatical complexity is handled by Alsina (1997) who adopted the Lexical Functional Grammar like Butt (1995) and explained that causatives in Bantu and Romance language behaved in the same way at the level of argument but differently at the phrase structure level.

In polysynthetic languages, there are obligatory verbal morphemes to cross-reference nouns with a grammatical function i.e., subject, object, and indirect object along with frequent phenomenon of noun incorporation. Two features of complex predicates have their roots in the head-marking feature of such polysynthetic languages: firstly, complex predicates are always single constituents, and secondly, causatives are only constructed from unaccusative verbs (Baker, 1997).

Verb serialization occurs in different morphosyntactic types of languages with similar properties (Durie, 1997). Durie (1997) employes a functionalist approach to explain conditions on serial verb formation. He also explains a process of argument structure merging within the domain of lexical-conceptual structures suggested by Jackendoff (1992). Lexical-conceptual semantic structures of Jackendoff (1992) are used to examine complex predicate in Yimas which cannot be fully explained using the semantic roles such as agent, theme etc. (Butt, 1995; Durie, 1997; Foley, 1997).

Studying the complex predicate among polysynthetic languages continued by Evans (1997) who studied noun incorporation in Mayali complex predicate, a language of Australia. His work focuses on explaining which argument of a multivalent verb used in complex predicate will undergo incorporation. In contrast to previous research (Baker, 1997) who made use of phrase structure configuration to explain the linguistic condition of an argument to incorporate. Only the argument which denotes prototypically inanimate referent will be incorporated (Evans, 1997). Thus, animacy, semantic feature of referent, rather than a property of roles or phrase structure formation is the decisive factor. This finding proved to be a concluding remark in the theory of argument structure. A comprehensive overview of complex predicates enables the reader to dive into specific details related to its only one type i.e., conjunct predicate in the next section.

2.2 Conjunct Predicate: A Type of Complex Predicate

In conjunct predicate, a noun is followed by a light verb to describe the action. In a conjunct predicate, information related to the number of arguments, their semantic sense, and the case markers is either determined by a noun or the light verb (Mohanan, 1994). The focus of the above-mentioned study is investigating the connection between the semantic structure of complex predicate (which will entail conjunct predicate) and their syntactic context possibilities. Furthermore, Mohanan (1994) emphasized the need to identify two concepts of 'lexicality' based on which the incorporation construction can be concluded. As Urdu and Hindi are flexible word order languages, light verbs can be separated from their nominal host, and can proceed in CP construction and occur initially (V+N) e.g., *'*to remember/*Kia yaad'*. Therefore, if we follow '*lexical integrity hypothesis'* which does not support syntactic rearrangement within the lexical categories, conjunct predicate does not support rearrangement of nominal host. As it

has been investigated in different studies that the argument structure of complex predicate is controlled jointly both by noun and a light verb. Then according to the principle of '*direct syntactic encoding*' which propagates lexical alternation of argument-grammatical function association, conjunct predicate is a lexical unit (Mohahan, 1994). Here, I see Mohanan (1994) navigating smoothly into different levels of structures which she suggests resolving this conflict of incorporation by representing complex predicate at different levels of structures i.e., Grammatical constituent structure, (GC STR), argument structure (ARG STR) and Grammatical Function Structure (GF STR).

Conjunct predicate has been under debate for various other reasons too. Such as the 'Noun' which is a part of the predicate along with the light verb, it not only affects the valency of the construction, but also correlates with the meaning and case markers on the arguments (Mohanan, 1994). In addition to this, the nominal part is an argument along with other arguments of the sentence or clause which may passivize and may also agree with the light verb in the N+V construction. Mohanan (1994) presents the solution to deal with the dual nature of the 'Noun' by expressing its multiple association with an argument (ARG) and Argument predicate (ARG-PRED) at Argument structure (ARG STR).

Influence of the semantic class and semantic value of nominal part in N+V construction on the arguments in the construction and the case marking on the subject can best be studied by looking at the examples 11 and 12 (a) and (b):

11.

a. Mazdoor ne ajnabi per bhrosa kia
   Laborer-M.SG.ERG stranger-M.SG.LOC trust-M.SG do-M.SG.PERF
   'Laborer trusted the stranger'.

b. Mazdoor ne apna kaam kia
   Laborer-M.SG.ERG self work-M.SG do-M.SG.PERF
   'Laborer did his work.'

12

a. Maan ne bachay ko phool dia
   Mother-F.SG.ERG child-M.SG.DAT Flower give-M.SG.PERF
   'Mother gave a flower to child'.

b. Maan ne pachay per dihaan dia (CP)
   Mother-F.SG.ERG child-M.SG.LOC attention-M.SG. give-M.SG.PERF
   'Mother paid attention to the child.'

Example 11(a) and (b) have light verb 'Do*/Kar*' which is a transitive verb can also be called as dyadic verb. The nominal part of N+V instance in 11(b) is 'Work/*Kaam*' which refers to some tangible and concrete outcome of an action, but in 11(a), the nominal part is 'Trust/*Bhrosa'* which refers to an abstract cognitive activity as a 'noun cognition'. In 11(a) and (b), 'Laborer/*Mazdoor'* has got the thematic roles as an agent; however, here in 11(b) there is another argument 'stranger' which could not be licensed by the verb 'Do*/Kar*'.

This may best be explained in terms of the effect of semantic association of 'Trust/*Bhrosa'* which licensed this argument 'Stranger/*Ajnabi'*. So, a theory may be propagated that the semantic class of the nominal part of the N+V instances contributes to the argument and case marker in the construction.

If 12(a) and (b) are studied, both constructions carry two similarities; 'Mother/Maan' as giver and a verb 'Give*/Dia*' which is a ditransitive verb. It is a well-recognized notion that the 'Givee' is usually an indirect object (the goal of giving) takes dative case (Mohanan, 1994). However, it is Locative in case of 12(b) with a marker 'per'. The locative case marker can better be interpreted through its systematic association of the semantic configuration, and it can clearly be put forward that this semantic configuration is not part of the SEM STR of 'Give*/Dia*' (Mohanan, 1994). Hence the ditransitive light verb 'Give*/Dia*' may not license the locative case marker 'Per'. So, 'Attention*/Dihaan*' can be interpreted as a noun cognition which contributes to the semantic configuration associated with an argument and consequently the case marker as well.

Thus, it can now be said that the semantic class of noun in N+V instance influences the number, meaning and the case marking of the arguments in the sentence or clause. This ability is a main characteristic of predicate. Therefore, all such nouns which control the arguments of the clause whether in terms of semanticity or the case marking will be predicates though a noun as grammatical category. Such N+V instances will be treated as conjunct predicates.

The influence of noun in a conjunct predicate over the meaning, and case marking of the argument in the clause does not come with a notion of void participation of light verb. It can best be understood by considering the examples with the verb 'Become/*Hua*' and 'Do*/Kar*' in the following examples 13 (a), 13(b), 14(a) and 14(b):

13.

a. Ijlaas ki shruaat hui
   Meeting-M.SG.GEN beginning-F.SG become-F.SG.PERF
   'Meeting began.'
b. Sadar ne ijlaas shru kia
   President-M.SG.ERG meeting-M.SG.NOM begin-M.SG do-M.SG.PERF
   'President began the meeting.'

14.

a. Ijlass multavi hua
   meeting-M.SG.NOM delay-M.SG.NOM become-F.SG.PERF
   'Meeting got delayed.'
b. Sadar ne ijlass multavi kia
   President-M.SG.ERG meeting-M.SG.NOM delay-M.SG.NOM do-M.SG.PERF
   'President delayed the meeting.'

In examples 13(b) and 14 (b), 'Do*/Kar*' assigns an agentive semantic role to the subject argument; on the other hand, 'Become*/Hu*' in 13 (a) and 14 (b) is not capable of exerting the same influence over the subject argument in the construction. So, from the detailed analysis of the data, it may be said that light verbs also have an impact on the argument structure of conjunct predicate.

This notion has been proved in contradiction with the licensing effect of Japanese verb 'Suru' (Grimshaw & Mester, 1998).

Light verbs in N+V instances also influence the argument structure of conjunct predicate which is related to the semantic structure of the arguments in the construction. Being very productive in nature, it is very difficult to come up with an exhaustive list of all conjunct predicate (N+V) in a language. In conjunct predicate (N+V), semanticity of noun affects the case marking on subjects, the number of arguments and the compatibility with light verbs (Ahmed & Butt, 2011). Identification of semantic

criteria of Urdu nouns for conjunct predicate is the first step towards developing a lexical resource for nouns in the Urdu language. Three classes of conjunct predicate have been identified based on syntactic and lexical choices (Ahmed & Butt 2011).

i. Class A: Full Range

These conjunct predicates have nouns which are compatible with light verbs 'Do/*Kar'*, 'Become/*Hu'* and 'be/*he'* where subject argument gets ergative and dative case marker with agentive and experiencer roles respectively.

ii. Class B: Exclusion of Dative Subjects

Class B contains two subtypes. One where nouns in these conjunct predicates are felicitous with 'Do/Kar' and draws ergative case on subject argument. Compatibility of noun with 'Become/*Hu'* does not license dative case to the subject but it draws nominative case to the subject in case of intransitive state of the light verb 'Become/*Hu'*. The second subtype is realized by an accusative case of direct object in case of 'Do/*Kar'* as light verb, and this accusatively marked object alternates as a dative subject in case of host noun combination with 'Become/*Hu'*.

iii. Class C: Exclusion of Light Verb 'Become/*Hu'*

Noun hosts in these conjunct predicates are compatible with 'Do/*Kar'* and 'be/*he'* but not with 'Become/*Hu'*.

This identification is based on syntactic analysis of the N+V predicate. Their research asserts that semantics of nouns in the conjunct predicate (N+V) influences the choice of the light verb which in turn provides information about the case marking on subject, agentive vs. experiencer thematic role of subjects, tense/aspect and agreement information. Conjunct predicate N+V can further combine with more light verbs. Related research is conducted by Bowern (2008) in which he asserts that light verbs can be categorized depending on its preceding noun, adjective or verb in case of conjunct predicate (Adj/N+V) and complex predicate (V+V) respectively. He elaborates on the use of two productive light verbs in Turkish i.e., '*dur*/Stop*'* and '*et*/Do'. When verb '*dur*/Stop' is combined with gerund, theta roles assignments are controlled by the gerund.

2.3 Incorporation of Noun in Conjunct Predicate (N+V)

Verb in N+V instance agrees with the left most unmarked argument subject in the clause. If an argument is marked with an overt case, in that condition predicate will agree with overtly unmarked argument. Verb agrees for its grammatical features such

as gender and number with the highest unmarked nominative argument (Mohanan, 1994). This theory plays an important part in identifying noun incorporations in the case of real conjunct predicate which clarifies the distinction between a categorial word and a function word. In conjunct predicate N+V instances, when the arguments of the clause are marked with case, light verb agrees with the host noun which in that case will be an argument. In the case of simple predicate, when all arguments are overtly marked with case, the verb is inflected with default masculine singular perfective (-aa).

15.

a. Shikaari ne sanp ko maraa

   Hunter- F/M.SG.ERG snake-M.SG.ACC kill-M.SG.PERF

   'Hunter killed the snake.'

b. Shikarri ne chirya ko maraa

   Hunter- F/M.SG.ERG sparrow-F.SG.ACC kill- M.SG.PERF

   'Hunter killed the sparrow.'

In the above example 15 (a) and (b), the morphological realization of gender is covert for 'Hunter/*Shikari*'; therefore, same lexicographical representation is used for both female and male. That is why it is reflected in the glossing mentioned with examples here i.e., 'Hunter-F/M.ERG'. In these examples, simple verb 'Kill/*Mara*' took the default perfective masculine singular inflection (-aa) in the presence of all overtly marked arguments.

Apart from applying this agreement test, Mohanan (1994) mentions other tests to reflect on the incorporated nature of the host noun in N+V instances such as '*Adjectival Modification'*, and '*Conjoining*', '*Wh-Questions'*, '*Relativization*'. As a result of incorporation, Mohanan (1994) calls the host noun in N+L instance as an internal incorporated noun which does not allow adjectival and numeral modification as reflected in the following examples 16 (a) and 16(b) respectively.

16.

a. Akhtar-ko (* intihaigh) gussa aayaa.

   Akhtar-M.SG.DAT unrestrained anger-M.SG come-M.SG.PERF

   ' Akhtar got uncontrollably angry'.

b. Maryam-ne Raneem-ko (*ek) kabar kiyaa.

Maryam-F.SG.ERG Raneem-F.SG.DAT one news-M.SG do-M.SG.PERF
'Maryam informed Raneem.'

Cojoining of incorporated nominal host in a (N+V) conjunct predicate is also not possible as evident from the examples 17(a) and 17 (b) below (Mohanan, 1994).

17.

a. *Raneem-ko kahaanii yaad aur pasand aaii.
Raneem-F.SG.DAT story-F.SG.NOM memory-F.SG and-CC liking-F.SG come-M.SG.PERF
'Raneem remembered and liked the story.'

b. *Akhtar-ne Raneem ko-muaaf aur qabool kiyaa.
Akhtar-M.SG.ERG Raneem-F.SG.ACC pardon-F.SG and acceptance-M.SG do-M.SG.PERF
'Akhtar pardoned and accepted Raneem.'

Internal or let us say an 'incorporated noun' in (N+V) conjunct predicate does not support the production of wh-questions. Thus, the following examples will be unacceptable as shown in the examples. See examples 18 (b) and 19 (b) given below.

18.

a. Raneem -ne Maryam-par bharosaa kiyaa.
Raneem-F.SG.ERG Maryam-F.SG.LOC reliance-M.SG do-M.SG.PERF
'Raneem relied on Maryam.'

b. *Raneem-ne Maryam-par kyaa kiyaa?
Raneem-F.SG.ERG Maryam-F.SG.LOC what do-M.SG.PERF
(* what did Raneem do on Maryam?)

19.

a. Raneem-ne kahaanii-par dhyaan diyaa.
Raneem-F.SG.ERG story-F.SG.LOC attention-M.SG give-M.SG.PERF
'Raneem paid attention to the story.'

b. *Raneem-ne kahaanii-par kyaa diyaa?
Raneem-ERG story-LOC what give-M.SG.PERF
(*what did Raneem give to the story?)

The incorporated nominal host in conjunct predicate (N+V) cannot be replaced by the gap of a relative clause. This characteristic may be better understood by studying the following examples 20 (a), 20 (b), 21(a) and 21(b) as used by Mohanan (1994):

20.

a. [vaheksarsais [jo raam-ne __ kiyaa]]nahut muskil thaa.
That exercise-F.SG.NOM that ram-M.SG.ERG do-M.SG.PERF very difficult be-PA
'The exercise that ram did was very difficult.'

b. [vah kitab[jo niinaa-ne raam-ko __dii]] mere pass hai.
That book-F.SG.NOM that Nina-F.SG.ERG Ram-M.SG.ACC give-F.SG.PERF I-G near be-PR
'The book that Nina gave Ram is with me.'

21.

a. *[vah bharosaa[jo raam-ne mohan -par _ kiyaa]]…
That reliance-M.SG.NOM that Ram-M.SG.ERG Mohan-M.SG.LOC do-M.SG. PERF

b. *[vah dhyaan [jo niinaa-ne kahaanii-par __diya]]…
That attention-NOM that Nina-ERG story-LOC give -PERF

Agreement of internal nominal host of a conjunct predicate with the light verb is a paradoxical situation and called a theoretically puzzling situation where the noun host will be treated as an argument which can undergo passivisation (Mohanan, 1994). She aims to demystify this situation by representing the dual nature of structural representation in conjunct predicate.

Usually, a verb agrees with its unmarked nominative subject argument, and the nominal host in conjunct predicate does not show any case marking. In the case of overt case

marking of arguments of conjunct predicate construction, there would be an agreement between light verb and its nominal host.

Let us take examples 22 (a) and 22 (b) from noun feeling 'Worship*/Ibadat*' with the transitive light verb 'Do/*Kar'*:

22.

a. Admi Khuda kii Ibadat karta hay
   Man-M.SG.NOM God M.SG.GEN Worship-F.SG Do-M.SG.PR Be-PR
   'Man worships God.'

b. Aurat Khuda kii Ibadat karti hay
   Woman-F.SG.NOM God-F.SG.GEN Worship-F.SG do-F.SG.PR Be-PR
   'Woman worships God.'

The verb 'Do/*Kar'* agrees with the male and female Nominative subjects in the examples 22(a) and 22(b) respectively. Another information contained in these examples is the agreement between the gender information related to genitive case markers 'Ki' of object argument 'God/*Khuda*' with that of host noun in conjunct predicate.

When subject argument is not nominative, light verb will agree with the nominal host as given below in Examples 23 (a) and 23 (b).

23.

a. Admi ne Khuda ki Ibradat ki
   Man-M.SG.ERG God-M.SG.GEN Worship- F.SG Do-F.SG.PERF
   'Man worshiped God.'

b. Aurat ne Khuda ki Ibadat ki
   Woman-F.SG.ERG God-M.SG.GEN Worship- F.SG Do-F.SG.PERF
   'Woman worshiped God.'

The same notion has been true for the examples 24(a) and 24 (b) for noun cognition 'Sense/*Andaza*' with intransitive light verb 'Become/*Hua*' including the agreement between host noun and light verb as well as the agreement between the genitive case marker of an object argument.

24.

a. Larkay ko Mushkil ka andaza hua
   Boy-M.SG.DAT Difficulty-M.SG.GEN sense-M.SG Become-M.SG.PERF
   'The boy sensed the difficulty.'

b. Larki ko Mushkil ka andaza hua
   Girl-F.SG.DAT Difficulty-M.SG.GEN sense- M.SG Become-M.SG.PERF
   'The girl sensed the difficulty.'

Let us take another set of examples 25(a) and 25(b) from noun feeling 'Love/*Piyar*' with the intransitive light verb 'Come/*Aa*' although the locative case of object argument 'Child/*Bachaa*' is not morphologically inflected with feminine information:

25.

a. Maan ko bachay per piyar aya
   Mother-F.SG.ACC child-M.SG.LOC Love-M.SG come-M.SG.PERF
   'Mother loved the child.'

b. Baap ko bachay per piyar aya
   Father-M.SG.ACC child-M.SG.LOC love-M.SG come-M.SG.PERF
   'Father loved the child.'

The same pattern is true for noun feeling 'Hate/*Nafrat*' with the transitive light verb 'Do/*Kia*' with the instrumental case marker 'Se' on the object argument.

26.

a. Chirya ko Tufaan se nafrat hui

Sparrow-F.SG.DAT Storm-M.SG.LOC Hate-F.SG Become-F.SG.PERF
'Sparrow hated the storm.'

b. Ustaad ko jhoot se nafrat hui
Teacher-M.DAT lie-M.SG.LOC Hate-F.SG.NOM Become-F.SG.PERF
'Teacher hated the lie.'

It is reflected from the above examples 26(a) and 26 (b) that there is an agreement between the nominal host of a conjunct predicate and light verb.

Depending on the earlier discussion it is evident that verb agrees with the left-most overtly unmarked argument. Then the nominal host of a N+V instance will be an argument and not an incorporated entity. This is the situation called '*structural paradox'* by Mohanan (1994). Agreement of nominal host with the light verb is only possible when the subject is non-nominative. In that case, the agreement of nominal host a N+V conjunct predicate with the light verb may only treated as an agreement between a verb and an argument.

Here are some N+ V instances where the subject is non-nominative and does not show an agreement with light verbs shown in Examples 27(a) and 27 (b).

27.

a. Ustaani ne larki ko pasand kia
Teacher-M/F.SG.ERG girl-F.SG.ACC Like-F.SG do-M.SG.PERF
'Teacher liked the girl.'

b. Maan ne bachii ko yaad kia
Mother-F.SG.ERG child-F.SG.ACC Miss-F.SG do- M.SG. PERF
'Mother missed her child.'

28.

Larkay ne takleef ko bardasht kia
Boy-M.SG.ERG hardship-F.SG.ACC bear-F.SG do-M.SG.PERF
'Boy borne the hardship.'

When the light verb fails to agree with the nominative nominal host, it is a true conjunct predicate. In examples 27 (a/b) and 28, nominative noun hosts which did not show an agreement with the light verb, belong to noun class 'Noun Cognition'. The

general category of noun cognition is an abstract noun. To forfeit this finding, I want to put forward the coinciding information here that the nouns in N+V instances mentioned by Mohanan (1994) are mostly abstract nouns.

The term 'serial verb' is used for compound predicate (V_{1+} V_2) which express language user's intention, belief, and attitude (Kachru, 1993). According to her, some of the serial verbs have regular monoclausal or biclausal derivations; therefore, concluded that there was a need to distinguish serial verb from ($V_{1+}V_2$) construction.

The sequence of noun followed by a verb is quite prolific that all instances are quite difficult to predict and list in different Indo-Aryan languages. Therefore, there is a need of some phenomenon to distinguish between mere N+V construction and conjunct predicate as a single syntactic unit.

Bhattacharyya et al. (2007) have contributed to the distinction of complex predicates from a series of verbs. They made use of three empirical tests to investigate the incorporated nature of nouns in Hindi (N+V) construction which they are called conjunct predicates. Firstly, they checked it by adding an accusative case marker to the noun. Secondly, different constituency tests such as movement, conjunct question and coordination test. Thirdly, modifiers were added to the noun phrases. Intuition of native speaker of Hindi is used to see the acceptability of added accusative case marker to the noun. When a noun is overtly marked with an accusative marker 'Ko' and it is unacceptable to the native speaker, it would be interpreted as an incorporated noun and resultantly it would be called as a true conjunct verb.

The three empirical tests contributed by Bhattacharyya et al., (2007) are used to differentiate between the Hindi noun incorporation in form of conjunct predicate and mere N+V sequence. After the tests, once their status is confirmed as true conjunct predicate, they can be added to lexical resource as a single entity.

These tests are as follow:

1. Addition of the accusative case marker to the noun:

It is based on the intuition of a native speaker of Hindi to see the acceptability of an overtly marked noun with the accusative case marker. An overt argument allows accusative case marking whereas an incorporated noun does not entail this sequence N+V as a conjunct predicate. They presented an example of 'to take tea/*caae lii*' and 'to yawn/*jamhaaii lii.* Accusative case marker '*Ko'* is added to the noun in these two

N+V instances. *Caae ko lena* is acceptable to native Hindi speakers whereas *jamahaaii ko lena* is not. It shows that Yawn/*jamahaaii* is incorporated in N+V instance as a single syntactic unit. With various examples, it has been explicitly stated that conjunct predicate does not allow addition of accusative case marker to the incorporated noun.

When a modifier is added before an N+V instance, it modifies only the noun in case of an argument whereas it modifies the whole N+V phrase in case of an incorporated noun and a true conjunct predicate. In 'Taking a lot of tea/*Bohut Chai Lena*' the modifier 'a lot/*Bohut'* of is only modifying 'Tea/*Chai*'. In 'Push hard/ *Zor se dhaka Maarna*', this modifier 'Hard*/Zor se*' modifies the whole N+V construction as a true conjunct predicate.

2.4 Constituency Test, including Movement, Conjunct question, and Coordination Tests

Three tests have been used to analyze the internal structure of conjunct predicate. Object in Hindi allows movement to the other positions in sentences. Whereas an incorporated noun resists such movement which shows its single lexical entity.

Constituent response test is another tool to see whether noun is an overt argument or incorporated into verb. When questions are formed for *jamahaaii lena* and *chalaang marna* , it is 'what did you do/*apnay kia kia'* instead of 'what did you take or beat/ *apnay kia leya or apnay kia mara*'. This very characteristic talks about the holistic meaning of N+V as single syntactic entity which is different from its constituents. Identical syntactic constituents allow coordination. For instance, 'Take tea and snack/ *caae or namkeen lena*' is possible but 'Sleep and yawn*/neend or jamachaaii lena* ' is not suitable for a native speaker of Hindi (Bhattacharyya et al., 2007).

Non-incorporated nouns are free to occur in different positions other than conventional. If a noun in N+V construction did not allow its relocation to some other position in the sentence, this N+V will be a true instance of conjunct predicate (Bhattacharyya et al., 2007). For example, in Taking tea/ *Chai lena*, tea/*Chai* can occur to any unconventional position in the sentence and still gives the same meaning and allows acceptability by a native speaker. On the other hand, in Taking/*Lena* shape, shape/*Roop* cannot be reallocated in the sentence. So, the later N+V instance would be considered as a true conjunct predicate.

Coordination occurs only between uniform syntactic constituents, and an incorporated noun in N+V does not allow such coordination. Means a true conjunct predicate will

not allow coordination between its nominal stem and another noun. For instance, taking tea and snack/ *Chai our namkeen lena* will be acceptable whereas taking sleep and yawn/ *neend our jamaii lena* will not be acceptable to the native speaker of Hindi.

2. Addition of modifier to the noun phrase:

Adjectives, determiners, quantifiers modify overt arguments but in case of conjunct predicate they modify the action (N+V) as a whole. For example, 'Drank a lot of tea/ *bohut caae pi*' means drank a lot of tea; whereas ' Push hard/*zorka dhaka diya*' does not mean heavy push only but it intensifies the whole action.

So, the following rules can be summarized in the light of tools suggested by Bhatachrayya et al. (2007) to test conjunct predicate:

i. Conjunct predicate (N+V) resists inclusion of accusative case marker to the noun.
ii. Conjunct predicate (N+V) resists free movement of the noun.
iii. Constituents of conjunct predicate i.e., Noun and verb behave differently in constituent response test when an interrogative sentence is formed to inquire about the action. Conjunct predicate gives a holistic meaning which is different from the literal meaning of its constituents.
iv. Conjunct predicate N+V) does not permit coordination between an object and an incorporated noun.
v. When a modifier is added, it is applied to the whole constituent N+V as a whole and not the only noun.

Polar verb (V_1) is inflected for agreement, whereas the vector verb (V_2) remains uninflected (Raza, 2011). For example, '*Khana* ***Khati*** (V_1. F.SG) *ja*(V_2) *rahi thi* (F.SG) and *khana* ***khata*** (V_1.M.SG) *ja*(V_2) *raha tha* (M.SG) she kept eating and he kept eating' respectively. Some constituency tests have also been applied on V_1+V_2 to see their status as a single syntactic constituent (Bhattacharyya et al., 2007). Based on which they devised five types of compound predicates:

i. V_1 inf-e+V_2: *likhne* (V_1 inf-e) *lagnaa* (V_2). Here V_1 is an infinitive from, and V_2 is usually the verb '*lagnaa/*attach, literally*'*. V_2 carries the information of number, tense, gender, etc. '*Likhne lagii* F.SG.PAST/(She) started writing' VS '*likhne lagaa* M.SG.PAST/(He) started writing'. Here, *Lagna* (V_2) would be treated as a modal auxiliary.

write'.

ii. V_1–kar+V_2: In this compound predicate, -kar is attached to V_1. For example, 'took and went/*lekar gayaa'*. V_2 here also reflects the information regarding number, gender, tense, etc. Both V_1 and V_2 can be modified with adverbs which shows that this V_1+ V_2 instance is not a single constituent.

iii. V_1 stem+V_2: V_1 occurs in the stem form without any inflection, and V_2 gets all the inflections of number, gender, tense, etc. For example, 'Kill-put/*maar* (V_1) *Daalna* (V_2)'; 'Write put/likh (V_1) dalnaa (V_2)'; 'Write take/likh (V_1) lenaa (V_2)'. Here, (V_2) is diminished in its meaning, but does not lose it all, and adds some meaning to the whole compound verbal sequence.

iv. V_1 inf+V_2: V_1 takes the infinitive form and V_2 is usually '*Parnaa*/literally fall' which gives the meaning of compulsion and a forced action. For example, 'Compelled to write *liknaa paraa*'; 'Compelled to go/ *janaa paraa*'; and 'Compelled to eat/*khana paraa*'.

v. V_1 inf-pp+ V_2 stem: V_1 is an infinitive verb immediately followed by a preposition 'for/ *ke lie*' which is shortened to 'To/ *ko*' in some cases. For example, 'Asked to eat/ *khane ke lie kahaa*'; 'Asked to sit/ *bethne ke lie kahaa*' and 'Asked to leave/ *jane ke lie kahaa*'.

Some diagnostic tests are used by Butt (1993) to identify compound predicate V_1+V_2 as a single constituent. These tests include scope of adverbs; scope of negation; nominalization; passivization; causativization and movement. An exhaustive application of these tests on different types of compound predicate may reveal their status as single constituency which may facilitate their insertion in the lexical resources as a unit. Their addition to WordNet aims to augment this lexical resource for computational and Natural Language Processing (NLP) purposes (Morato et al., 2004). After their identification as true compound predicate, semantic properties of the light verbs which are called vector verbs are also examined.

Bhattacharyya et al. (2007) list the following ten Hindi light verbs to whom they called vector verbs:

i. *Daalnaa* 'put'
ii. *Lenna* 'take'
iii. *Denna* 'give'

iv. *Uthana* 'wake'
v. *Jannana* 'go'
vi. *Parnaa* 'lie'
vii. *Baithnaa* 'sit'
viii. *Maarna* 'kill'
ix. *Dhamaknaa* 'throb'
x. *Girnaa* 'fall'

These light verbs may also occur as polar or main verbs, but as light verbs their core semantic sense is bleached, and they acquire new semantic properties. Light verbs contribute to finality, definiteness, negative value, manner of the action and attitude of the speaker. As a result of their execution of the constituency tests, some automatic methods can be employed to extract Hindi compound predicate V_1+V_2 from a corpus (Bhattacharyya et al. 2007).

Regarding the insertion of complex predicate into electronic lexical resources like WordNet, two options are proposed (Bhattacharyya et al., 2007). One is to enter a noun first and link it with the compatible light verb, and second, N+V can be listed as a single entity with its unique syntactic and semantic features which are distinctive from its constituents. Syntactically, there is no problem especially in terms of the argument structure of the verb and its subject-verb agreement properties. However, it posed a problem that the semantic feature of the conjoined term is not compositional. For example, *chalaang maarnaa* means to 'dive'. Its light verb *maarnaa* means to beat or kill which is not entailed in N+V holistically (Bhattacharyya et al., 2007). Another example is 'to run/ *dorr lagaana*' which is not compositional of 'to put/*lagana*'.

2.5 Semantic Contribution of Nouns in Conjunct Predicate

In conjunct predicate constructions, the noun is a predicating element along with the light verb. The presence of two predicating elements representing a single meaning is a challenge for a linguistic theory that maps syntax and semantics (Vaidya et al., 2014). According to their analysis, nouns in conjunct predicate are a predicting element with the light verbs and both predicating elements combine to function as a single syntactic constituent. The semantic classes of nouns are identified by their compatibility with

light verbs which govern the case marking on subject arguments and assign thematic roles to them as well.

The lexicalized meaning of nouns is specific and includes all uses of nouns regardless of the situation (Levin & Rappaport, 1994). Nouns can be considered as a collection of attributes that are possessed by some entities to which those nouns refer. Any entity in the world may contain many attributes which are lexicalized by a noun. Two nouns may refer to the same entity with similar attributes, but their lexicalization may be expressed differently such as lorry and bus.

The nominal part of the N+V instance not only controls the argument structure of the clause, but also governs the case marking on the arguments.

Mohanan (1994) made use of the example 29 (a/b) and 30 (a/b) to elaborate the governing features of host noun in N+V instances:

29.

a. Raam-ne apnaa homwark kiyaa.
   Ram-M.SG.ERG self-G homework-M.SG.NOM do-M.SG.PERF
   'Ram did his homework.'

b. Niinaa-ne raam-ko kitab dii.
   Nina-F.SG.ERG Ram-M.SG.DAT book-F.SG.NOM give-F.SG.PERF
   'Nina gave ram a book.'

30.

a. Raam-ne mohan-par b^{h}arosaa kiyaa
   Ram-M.SG.ERG Mohan-M.SG.LOC reliance-M.SG.NOM do-M.SG.PERF
   'Ram relied on Mohan.'

b. Niinaa-ne kahaanii-par dyhaan diyaa.
   Nina-F.SG.ERG story-F.SG.LOC attention-M.SG.NOM give-M.SG.PERF
   'Nina paid attention to the story.'

In above examples 29 (a) and 30 (a), the subject arguments bear a semantic role of an agent, and there are things done i.e., 'homework' and 'reliance'. However, there is an

argument 'Mohan' in 30 (b) which is not licensed by 'Do/*Kar*'. This argument is licensed by the host noun 'Reliance/ *b*^{h}*arosaa*'.

Similarly, if examples 29 (b) and 30 (b) are studied, the locative case '*PER'* is not licensed by 'Give/ *Dia*', but it is the host noun 'Attention/ *dyhaan*' that governs the case marking of the argument. Thus, it is proved that the semantic properties of nouns are linked with thematic roles and cases which determine the distinct meaning of the event or state mentioned in the sentence (Mohanan, 1994).

At the same time, this proposition does not come with the null participation of light verb in N+V instance. Light verbs do have control over the semantic role of the arguments in the clause. Here I would quote the examples used by Mohanan (1994) to elaborate on the contribution of light verbs in N+V instances (see both the versions of Examples 31 and 32).

31.

a. Kamre-kii safaaii hui.
   Room-M.SG.GEN cleaning-F.SG.NOM happen/become-F.SG.PERF
   'The room got cleaned.'

b. Raam-ne kamre-kii safaaii kii.
   Ram-M.SG.ERG room-M.SG.GEN cleaning-F.SG.NOM do-F.SG.PERF
   'Ram cleaned the room.'

32.

a. Kamraa saaf huaa.
   Room-M.SG.NOM clean M.SG.NOM happen/become-M.SG.PERF
   'The room became clean.'

b. Raam-ne kamraa saaf kiyaa.
   Ram-M.SG.ERG room-M.SG.NOM clean do-M.SG.PERF
   'Ram cleaned the room.'

Light verb 'Do/*Kar*' can assign agentive role to the subject in examples 31

(b) and 32 (b), whereas the light verb in N+ V instance 'Become/*Hu*' is incapable of doing so.

Both nouns and verbs are open classes of words with a continuous flux of more words to the dictionary every year. In English language, there are fewer verbs as compared to nouns although there cannot be a sentence without a verb. There are 43,363 nouns and 14,190 verbs in the Collins English Dictionary. The Urdu language has almost 700 main verbs (Ahmed, 2010). Verbs are more polysemous than nouns. For instance, verbs have 2.11 semantic senses; whereas nouns have 1.74 (Fellbaum, 1990). The higher polysemous nature of verbs reflects that they are semantically more flexible than nouns. Semantic value of verb is related to the syntactic context means the noun arguments they occur with; similarly, the meaning of nouns is also linked to the verbs with which they are used. This notion is very close to the notions highlighted by both Fellbaum (1990) and Levin (1993). Similar thought has been fortified by Mohanan (1994) who says that given the meanings of words their syntactic positioning is very much predictable. The most frequently used verbs are the most polysemous and the nature of their polysemy is related to the nouns with which they are used. In 1990, Fellbaum suggested this phenomenon to have been incorporated in WordNet to reduce the semantic ambiguity by linking the verb synsets to the probable collocated noun synsets. The relationship between head and modifier in English noun compounds has been a focus of research in the field of semantics. This head-modifier relationship in an English noun compound is heavily dependent on whether the referent is a natural object or an artifact (Levin, Glass, & Jurafsky 2019). Two different types of noun compounds have been investigated: artifact-headed compound and natural-kind-headed compound. It was hypothesized that the modifier in an artifact-headed compound refers to essence associated with that artifact: whereas the modifiers in a natural compound correspond to inherent properties or the native habitat of that natural object.

Almost every language has some means to linguistically categorize the nouns. In South-East Asian languages, a numeral classifier is employed; whereas, in African and Indo-European languages, highly grammatical noun classes and genders are focused on. Noun classifying devices are morphemes which denote important features of the entity to which that noun refers (Aikhenvald, 2006). Noun categorization comes in different shapes: (i) Noun classes. (ii) noun

classifiers, (iii) numeral classifiers, (iv) classifiers in possessive construction, and (v) verbal classifier including two rare types (vi) locative and (vii) deictic classifier. Members within one class share a common semantic core which may differ in the morpho-syntactic conditions and preferred semantic features (Aikhenvald, 2006).

Classes of nouns present a deep understanding of how the whole world is categorized through nominal categories related to different parameters such as humanness, gender, animacy, form, shape, temporal, spatial, functional event, and natural essence. They tend to share common semantic essence but behave differently as far as morpho-syntactic contexts are concerned. Animacy, gender and humanness are parameters on which classes of grammatical agreement in some languages are based. Noun classes can be semantically transparent, but their categorization is based not only on meaning, but also on morphological and phonological criteria (Aikhenvald, 2006). These criteria are comprehended through noun agreement with the modifier or the verb in its syntactic context.

Cross-linguistic features of noun classes are summarized as follows (Aikhenvald, 2006):

i. Number of noun classes is limited and countable.
ii. Each noun belongs to at least one class; sometimes, one noun can be categorized in more than one class as well.
iii. Noun classes are usually categorized based on some semantic features such as animacy, gender, shape, size, and humanness. Animacy, here, refers to a semantic feature based on which is determined how alive or responsive the referent of the noun is.
iv. Some syntactic constituents must agree in gender with the noun in a sentence. These agreeing constituents may be adjectives, numbers, demonstrative, pronoun articles, adverbs, and elements within a predicate of the clause.

Semantics of the noun classes in the languages of the world revolves around the following parameters:

i. Gender (masculine and feminine): In many Afroasiatic languages, it is the most frequently referred semantic property based on which a noun class can be identified.
ii. Human or non-human: In some Dravidian language, this semantic parameter is used to determine noun classes.

iii. Rational and non-rational: It is a popular semantic parameter in Tamil and other Dravidian languages (Fedson, 1993).

iv. Animate and inanimate: It is considered as an important semantic feature for noun classification.

There is a term *neuter* which may refer to irrational, inanimate, or residue gender with no clear semantic feature.

Rules for determining the semantic classes of nouns are quite complex. For example, Dixon (2005) reported four noun classes in Dyirbal language:

Class I: Male human, non-human animates

Class II: Female human, water

Class III: Non-flesh food

Class IV: A residue class which covers everything else.

A typical gender system in case of Australian languages consists of four parameters such as masculine, feminine, vegetable and residual (Dixon, 2005). There are separate noun classes for insects and places in Audian and Bantu languages respectively (Corbett, 1991). In an important work, semantic rationale for assignment of semantic classes of nouns is presented for German language (Craig, 1986). Natural sex principle such as male and female adults of almost every specie is used for gender markers; whereas neuter is non-sex specific and juvenile entities. Cloth, precipitation, wind, and minerals are assigned masculine gender. Different types of knowledge have feminine gender. The neuter gender is assigned to solids and types of metal. This notion negates the proposition that no real semantic basis is used for gender assignment in different Indo-European languages. For inanimate and non-human animates, noun class assignment is even more opaque. The feminine class consists of female human animates, sun and mostly smaller songbirds. The rest of the animates come in masculine class; however, neuter comprises almost the rest of the inanimate.

These nouns are independent lexical items that have generic semantic features. In Australian languages, there are numerous noun classifiers (Aikhenvald, 2006). In Yidin Australian language, there are twenty noun classifiers which are of two kinds:

i. Inherent Nature: Human, fauna, flora, natural and artefacts

ii. Function: Edible flesh food, edible non-flesh food, habitable, drinkable, moveable, and purposeful noise

These morphemes which collocate with numeral, or quantifiers are frequently used in isolating languages of Southeast Asia, Korean, Turkish, Japanese, Fusional Dravidian, Indic and some agglutinating North Amazonian languages of South America. The semantics of the referring noun is the determining feature for a noun classifier. Specific semantic features include animacy, functional properties (i.e., object with a handle, etc.), arrangement (bunch, hierarchy, etc.) and physical properties such as dimensionality, shape, consistency, nature. In several New Guinea languages and Kana, a Cross-River language from Nigeria, there is no classifier for animates (Aikhenvald, 2006). When counted, they are classified by shape or by their function. In almost every language, use of numeral classifiers or quantifiers depends on the semantics of the noun which is determined by whether the noun is countable or not. For example, in the English language 'much' is used for non-countable nouns; whereas 'many' is used for countable nouns. Similarly, the use of quantifiers is also dependent on the semantic properties of the referent nouns. For instance, *five packs of pencils, two piles of hay, two schools of whale* and so on.

There are three kinds of possessive constructions:

(a) Relational Classifiers

(b) Possessed Classifiers

(c) Possessor

Cross-linguistically, the classificatory verbs mostly belong to the semantic category of handling, motion, and existence/location. Objects are characterized by their specific position and state; a tree grows whereas liquids flow. Classification of verb is different from the lexical selection of a verb in terms of physical properties or the position of a noun mostly as an object (Downing, 1977). Verbs make consistent paradigmatic differentiation in the choice of semantic properties for the nouns throughout the verbal lexicon. Classificatory verbs define a set of paradigmatic choices for the sets of verbs that depend on the physical properties of nouns. These verbs cannot be categorized as classificatory as the correlation between the choice of verb and the physical properties of the objects is not paradigmatic.

Locative preposition and postposition classify the noun in terms of its animacy or physical properties which include shape and form. Locative classifiers are found in Palikur, Arawak Brazilian language and South American Indian languages of the Carib family (Aikhenvald, 2000). Deictic classifier occurs on deictic pronouns within a nominal phrase where head is a noun with inherent properties and position in space.

The focus of formal semantic treatment on natural nouns and how they differ from artifact nouns has been discussed by Grimm and Levin (2017). Their work highlights the unique semantic properties of artifact nouns. The distinguishing feature between natural and artifact noun is of natural essence and of intended function respectively. For instance, noun *sun* is a natural object which can be defined by its natural essence; whereas, noun, *needle* is categorized by the intended function performed by it such as sewing, pricking, etc. It supports the proposal that names of an entity may be semantically transparent, but natural and artifact nouns are categorized by their essence and function respectively. Artifact nouns include as associated event which represents the artifact's intended use that involves modality or temporal components (Nichols, 2008). Furthermore, detailed account of countability for artifact nouns is also investigated. Idiosyncratic lexical behavior of compound names of artifact and natural nouns were analyzed to elicit that modifier in natural nouns talks about body related feature whereas in case of artifact noun, it is the function which modifies the head of the noun phrase (Levin et al., 2019). The countability features of artifact nouns include two interpretations: object and kind-level. They are called countable nouns when a single entity marks the minimal associated event. The minimality condition does not bound artifact noun such as 'furniture' to single entities. Hence, it leads to a non-countability grammatical category of those artifact nouns.

In a theory of lexical knowledge, one of the properties of verb is to combine with noun arguments to form sentences (Levin, 1985, 1989). These research studies imply that the semantic components of verbs and verb classes are responsible for their syntactic behavior. Means the type of argument a verb takes is linked with the semantic component of verb.

Some distinguishing features and a superordinate term are used to define and organize common nouns in WordNet (Miller et al., 1990). The superordinate semantic relation is a hierarchical semantic relation with hyponymy at the top as the most specific term and hypernym as the most general time at the lowest level. For example, 'Evangelist' is a hyponym of the general noun 'Human Being' which can go further general if we move upward in this hierarchy at the most general superordinate term 'Living Being'. In this way, all new nouns are also contained in already defined noun classes. Antonymy is another way to classify nouns, but it is not employed as the fundamental

classifying principle. Miller et al., (1990) suggest classifying nouns into twenty-five noun classes based on distinctive primitive semantic components mentioned in Table 2.1 as mentioned below.

Table 2.1

List of 25 Unique Beginners for Noun in WordNet

{act, action, activity}	{natural object}
{animal, fauna}	{natural phenomenon}
{artifact}	{person, human being}
{attribute, property}	{plant, flora}
{body, corpus}	{possession]
{cognition, knowledge]	{process}
{communication}	{quantity, amount}
{event, happening}	{relation}
{feeling, emotion}	{shape}
{food}	{state, condition}
{group, collection}	{substance}
{location, place}	{time}
{motive}	

Source: Miller et al.,1990

In WordNet, nouns are interconnected in a network of three kinds of semantic relations: hyponymy, meronymy and antonymy.

Theory for categorizing common nouns which extends named-entity classification used 26 semantic labels to classy common noun (Ciaramita & Johnson, 2003). This theory was preferred because it improves accuracy which helps in learning and classifying new nouns. It has got an ability to automatically place new words in the existing lexical hierarchy after identifying the syntactic and semantic properties of new words and consequently extend the synset hierarchies in different semantic relations, and a broader semantic class is determined automatically. They added another broad semantic class: supersense, to the inventory of already existing twenty-five noun primes of common nouns.

Mohanan (1994) highlights the abstract nature of nominal host in abundant N+V instances.

A related study about the Bangla language talks about the classes of nouns in a syntactic context where they are used in conjunct predicates (Ghosh, 2015). This type of classification of nouns can reflect the argument structure of verbs including the type of argument structure, case markers, and collocated postpositions. It helped in building the ontology of nouns which is more useful in different Natural Language Processing programs. It is a better way of classifying nouns as compared to a classification based on lexical semantics which is used in WordNet. Syntactic properties of nouns such as selection of arguments, case-markers on arguments not only help in classifying nouns but also aid to understand the features of conjunct predicate in Bangla. It is specific to Bangla language that argument can take different cases in the presence of the same single theta roles. Due to agglutinating properties of the most Indian languages, the morphosyntatctic properties of nouns are linked with their semantic properties. This unique way of noun classification involved the development of syntactic frames of conjunct predicate which are an integral part of building a knowledgebase for verbs of a language. The work provided broad-coverage subcategorization frames for conjunct predicate (N+V) in Bangla (Ghosh, 2015). In this study, the argument of the verb is distinguished from the nominal part of conjunct predicate (N+V). The nominal in the conjunct predicate carries information regarding the argument structure of the predicate. After checking the noun compatibility with frequently occurring Bangla verbs such as 'to do/*kora';* 'to become */hoa*';* 'to feel/*paoa*', and 'to offer/*deoa*' nouns are categorized according to the type and number of argument/s along with the possible case marking on these arguments.

Twenty-one syntactic frames of conjunct verbs with 'to do/*Kora*' have been found with the most frequently found Patient Class (Class 7) followed by the second frequent class Verbalizer (Class 5) (Ghosh, 2015):

i. Benefactive Class (Patient+ instrument)
ii. Associative Class (Agent Associative)
iii. Animate Locative Class (Agent Locative)
iv. Quality Noun Class (Agent Possessor)
v. Verbalizer
vi. Alocona Type (Agent Location)
vii. Patient Class (Agent Patient)
viii. Source Agreement Class (Agent Source)

ix. Locative Argument Class (Agent Location)
x. Locative Type with Postposition
xi. Genitive-Marked Argument Patient
xii. Sentential Complement Type
xiii. Patient +Genitive Agreement Type
xiv. Genitive Argument
xv. Only [-animate] Genitive Argument
xvi. Associative + Optional Locative Argument Class
xvii. Inanimate Patient + Benefactive Argument Class
xviii. Plural Benefactive +Patient Argument
xix. [+ Animate] Genitive Marked Patient
xx. Animate Genitive Marked Patient Argument
xxi. Patient + Locative

Second most common verb in Bangla N+V conjunct verbs is 'to give/ *Deoa*' which loses its core semantics and acts as a verbalizer except in 4th type as mentioned below:

i. Accusative Inanimate Argument with zero marker
ii. Genitive Argument with -r marker
iii. Locative Argument with -te/-e marker
iv. Recipient animate argument with -ke marker

Third type of Bangle N+V conjunct verb with 'to feel/*Paoa*' has got Experiencer subject and a noun argument; whereas Bangla Conjunct verb with 'to become/*Hona*' has got experiencer subject and emotional experiencer noun argument.

The above-mentioned three categories of Bangla N+V conjunct verbs have presented the basic tenets of cognitive linguistics which are applied in the domain of translation to come up with the classification of nouns. In this study, only two types of nouns are mentioned based on the number of arguments they take: nouns with one argument, and nouns with two arguments (Ghosh, 2015). Nouns with one argument are classified based on the semantic roles and the case marking they bear. Ghosh (2015) curtailed the number of these classes to 15:

i. Noun taking Inanimate Patient Argument 'refusal'
ii. Noun taking [-Animate] Genitive-marked Patient Argument 'damage'
iii. [+/-Animate] Genitive marked Patient Argument 'worship'
iv. Noun taking Sentential Complement Argument 'care'

v. Nouns taking Sentential Complement Argument ‘escape’
vi. Noun Taking sharing Agent Argument with Verb ‘bath’
vii. Emotional Experience noun taking one Experiencer Subject with Genitive marking ‘joy, sorrow’
viii. Noun Experiencer with Physical Realization ‘laugh, tear, vomiting, urine’
ix. Noun taking animate *ke*- Marked Benefactive Argument ‘help’
x. Noun taking Source Argument with Postposition ‘return’
xi. [-Animate] Genitive Argument ‘announcement’
xii. Attributive Nouns taking Genitive-marked Argument ‘praise’
xiii. Only an [-animate] Genitive Argument ‘inauguration’
xiv. Nouns taking Locative Argument ‘living’
xv. Nouns taking Associative argument with postposition ‘quarrel’

Nouns with two arguments are also classified according to the semantic roles and the case marking on those arguments:

i. One Patient and one Genitive Argument ‘forward’
ii. One associative and one Optional Locative Marked Instrument Argument ‘assistance’
iii. One Inanimate Patient and Benefactive Argument ‘dedication’
iv. One Plural Benefactive and one Patient Argument ‘assistance’

Ghosh (2015) envisions to build NounNet based on the above-mentioned hierarchical classification of nouns which is based on syntactic and semantic behavior of nouns in Bangla conjunct verbs (N+V). Another byproduct of this research is the syntactic frames of Bangla conjunct predicates (N+V).

Semantic structure (SEM STR) mentioned by Mohanan (1994) refers to meanings which can intervene with the syntactic or morphological linguistic realizations, and it is constructed out of semantic elaboration and entailment drawn from the universal inventory. It is not necessary that organization of elements at SEM STR is same to the event structure in the real world. It is due to the pure grammar internal level of structure of SEM STR. She did not intermingle the information of SEM STR related to complex predicate with the non-linguistic conceptual representation in terms of entailment.

As there is a link between attributes lexicalized by verbs and nouns, in the same way, it may be asserted that the attributes lexicalized by nouns also affect their morphosyntatctic behavior. Similar morphosyntatctic properties of nouns include the attributes of countability which encompasses number morphology, use of determiner and modification by number.

It is hard to characterize the precise difference between count nouns and the mass nouns. This distinction is based on the concept of individuation which is a grammatically relevant semantic component of meaning. Morphosyntactic features of nouns among other lexical items in the construction carry information about countability which is dependent on individuating attribute of the nominal element. So individuating attributes is a semantic notion that helps to classify many entities based on common grammatical behavior (Levin & Rappaport, 1994). For instance, count nouns allow plural morphemes (-s, -es, etc.) modification by cardinal quantifiers (one, two, etc.) and determiners showing plurality (many). Mass nouns, on the other hand, neither permit plural marking nor cardinal quantifiers nor determiners implicating plurality except on kind interpretation. So, in the nominal domain, as believed by Levin and Rappaport (1994), morphosyntactic countability can be utilized to identify individuatable features in the noun domain. They assert that liquids and homogenized substances are mass nouns which do not fall under individuated units. On the other hand, physical objects with clear physical boundaries are individuatable which include two semantic classes of nouns: natural and artifact (Soja et al., 1991). Some reasons to separate granules from aggregates such as bead and beans which are treated morphosyntactically as count nouns (Wierzbicka, 1988). It is asserted that individual units of aggregate are distinguishable and may be counted.

2.6 Semantic Contribution of Light Verbs in Conjunct Predicate

In Compound Predicates (V_1+ V_2), the first verb is referred to as *polar verb* whereas the second verb is called as *vector verb* or *light verb* (Ahmed, 2010; Ahmed & Butt, 2011; Butt, 2010; Kiani, 2013; Mushtaq, 2015; Schmidt, 1999). Vector verb is said to lose its original sense when combined with polar verb in a compound predicate. Vector verb, light verb, intensifying verb, compound auxiliary and explicator verb are the terms which are interchangeably used (Schmidt, 1999). Final verbs in N+V and Adj +V constructions are called light verbs and are usually the syntactic head of the

construction (Kiani, 2013). Another reason for calling it a 'light verb' is due to nouns because it bears the main predictive burden in an N+V construction (Mohanan, 1994). Different terms have been used for Urdu final verbs depending on the functions they perform. Light verb is the term which has been used by Butt (1995, 2001, & 2005) in her studies which is continued to be used in the same fashion in other studies as well (Maxwell, Brown & Lynn, 2009). The semantic contribution of polar verb is comparatively more than the vector verb which seems to be semantically bleached (Schmidt, 1999). Light verbs contribute to the completeness, suddenness of an action (Ahmed, 2010). Bowern (2008) is of the similar opinion that light verbs are an integral part of any complex predicate construction as in Hindi, Urdu, German and Roman. Light verbs are semantically incomplete or defective and can be categorized based on their syntactic context i.e., proverb. Light verb takes up morphological features in compound predicate. It inflects with tense, aspect, and agreement morphology when functioning as auxiliary.

*De /*Give, *lE /*take, *Aa /*come, *jA /*go, *DAl /*insert, *paR/*fall, *beTH/*sit, *uTH/*rise, *dE/* give, *rakH/*put, *ban/*get make, *lag/*touch/hit, *nikal/*come out, *Tahar/*stop and *cal/*move are some of the light verbs quoted in an important work on Hindi compound verb (Hook, 1972).

Despite the systematic synchronic relation between main verb and the light verb, their semantic and syntactic properties are not identical (Mohanan, 1994). Butt (1995) listed 13 light verbs for Urdu, and Bukhari (2009) highlighted 17 light verbs for Gojri.

Akhtar (2000) introduces 8 light verbs in Punjabi which carry aspectual information in V_1+V_2 construction. Furthermore, he claimed that information about volitionality can always be carried by Complex predicate in Urdu and Punjabi. Akhtar (2000) believes that volitionality in Urdu and Punjabi may be dependent on an extra inserted phrase i.e., *jaan buj kai-* (intentionally). Light verbs are not completely semantically void. This phenomenon can easily be understood by realizing the difference between 'take a bath' and 'give a bath'. These verbs are neither semantically full nor empty, but semantically bleached in some way.

Unlike alternations used by Levin (1993) to classify English verbs, semantic classes of Urdu main verbs have been drawn based on acceptability/unacceptability of these polysemous main verbs with the three most frequent light verbs i.e., 'Give/*Dia'*, 'Take/*Lia'* and 'Go/*Ja*'. These verb classes are used for sense disambiguation in the

case of polysemous Urdu main verbs. So, it can safely be asserted that compatibility of with light verbs played a decisive role in classifying Urdu main verbs.

Abbas and Khan (2009) use the term 'modal auxiliary' for helping verbs. Butt (2003) investigates the difference between light verbs and auxiliaries. According to her findings, light verbs form a separate syntactic. The syntactic properties of light verbs are different from that of auxiliaries and the main verbs. When light verbs are reduplicated, it distinguishes them from auxiliaries.

Auxiliaries are used in a sentence to support the main verb; they denote different features of tense, mood, voice, and aspect. In Urdu, when subject has a nominative case, tense and aspectual auxiliary agrees with the subject in terms of number, person, and gender. There are auxiliary verbs which contribute to the sense of completion of the action. Most specifically *chukna* along with other Urdu light verbs such as 'Go/*Jaa*'; 'Get/*Lena*'; 'Give/*Dia*' aid the completion of actions (Haq, 1906). The light verbs which contribute to the meaning of ability, permissibility and possibility are 'Can*/Sakna*'; 'Give/*Dia*' and 'Find/*Pana*'; whereas 'Do/*Kar*'; 'Go/*Jaa*' and 'Put/*Rekha*' communicate continuity of action and habitualness of the subject. Suddenness of action is communicated through the light verbs 'Get up/*Uthna*', 'Sit/*Bethna*'; 'Touch/*Lagaa*' and 'Take out/*Nikala*'**.**

Butt (1995) categorizes some common light verbs in Urdu as listed below in Table 2.2:

Table 2.2

Common Urdu Light Verbs

Based on (di)transitive (Ergative Subject)	Based on Intransitives (Nominative Subject)
le 'take'	*aa* 'come'
de 'give'	*jaa* 'go'
daal 'put'	*par*.'fall'
maar 'hit'	*mar* 'die'
nikaal 'pry out'	*nikal* 'emerge'
	cuk 'finish'
	baith 'sit'
	uth 'rise

(Source: Butt, 1995)

For her work on permissive and aspectual complex predicates in Urdu (Butt, 1995), she makes use of the following light: 'Take/*Le'*; 'Give/*De'*; 'Put/*Daal'*; 'Fall/*Par'*; 'Rise/*Uth'* and 'Go/*Jaa'* due to their broadly representative nature. She finds that 'Take/*Le'*, 'Give/*De'* and 'Put/*Daal'* are transitive light verbs which require Ergative case markers on the subject argument in the perfective form. In addition to this, she also points out that these light verbs carry semantic information of conscious choice and completion of an action. As she finds, 'Fall/*Par'* denotes no conscious control, and it requires a nominative case on the subject argument in the clause. Butt (1995) shares an eye-catching observation that same sets of light verbs (do, take, come, go, give, hit) are used for cross-linguistic studies related to complex predicates.

Butt (2003) deduces that unlike main verbs, light verbs are dependent on a distinct predicative part and do not contribute to their separate event. They give information about the type of event i.e., benefactive, unexpected, agentive, volitional, or accidental. In the case of compound predicate V_1+V_2, the order of two verbs is reversible. For example, 'Hit/*maar day*' and 'Hit/*day maar' which* do not bring any significant change in its semantic sense.

An effort to distinguish between auxiliaries and light verbs has been made by Butt (2010) who believes that light verbs form a distinct syntactic class which differs from auxiliaries and main verbs. Furthermore, light verbs are dependent on another predicative part which is not the case with main verbs as they may cover the complete meaning (Butt, 2010).

Raza (2011) finds that 'Do/*Kar*'; 'Become/*Ho*'; 'Give/*De*'; 'Take/*Lay'*; 'Come/*Aa*' and, 'Go/*Ja*' are the most frequently occurring light verbs in Urdu. Their nature of being the most frequent light verb affects the reason to include them in the present study on Urdu conjunct predicate.

Modals in Urdu give the semantic sense of ability, possibility, willingness, and obligation. 'Should/*Chahiye*' and 'Can/*Sak*' are the most frequently used modals in the Urdu language.

At the surface level, N+Modal may look like a complex predicate, but their syntactic and semantic features are remarkably different. For instance, firstly there are no compatibility restrictions applied to their productivity with other non-finite verbs, and, secondly, they always impart a clear modal meaning. Urdu reference dependency

structures, the modal is annotated separately using LFG, and kept separate from complex predicates (Ahmed et al., 2012).

Both light verbs and auxiliaries can serve as main verbs in different constructions. If only looked at the surface level, their distinction seems complex, but when analyzed syntactically, it becomes clear that auxiliaries do not contribute to the argument structure in the clause, and their contribution is limited to tense and aspect information. Auxiliaries may be added to the series and can become lengthy. In such cases, they contribute to the duration of the action. Only selective nouns are compatible with some light verbs. This compatibility depends on the semantic and syntactic orientation of both noun and light verb.

2.7 Compatibility between Noun and Light Verb in Conjunct Predicate

Combinatory restrictions between a noun host and the light in conjunct predicate are driven by the semantic features of nouns and light verbs according to the semantic and syntactic relevance proposition (Levin, 1993). Here this notion of compatibility can be predicted if we are familiar with the conventional noun classes compatible with light verbs. Newly added nouns to the lexicon can also be predicted based on their semantic relations with already existing hypernyms. To extract a complete knowledge of a probable combination restrictions can be yielded from an exhaustive inventory of N+V instances in a language.

Being conjunct predicate a complex verb and consists of two lexical entities i.e., noun and light verb semantic and syntactic information related to light verb also influence the mutual compatibility and controls the argument structure and case marking in the construction. So light verbs are also categorized based on their argument-taking ability (Butt, 1995). Transitivity of light verbs gives sufficient information regarding the prediction of the noun class it may form a combination, and other syntactic information of the clause. In the case of confirmation of an existing connection between semantic and syntactic information of Urdu nouns and light verbs in a conjunct predicate, it may validate the notion of Levin (1993) for the Urdu language as well.

2.8 Influence of Light verbs in a Conjunct Predicate Construction on the Subject Case-Marking

The case-marking system in Urdu is a phenomenon that is realized by three morphological processes: case clitics, inflected stem forms and postpositions (Mohanan, 1990). The nominative case, also known as null case, does not take an

overt case marker (Kachru, 1980). This direct case marked nominal always agrees in term of gender, number, and person with the verb regardless of its position in the sentence. In the absence of nominative case marked argument NPs, verb assumes default past perfective singular third person *'aa'* form. In the case of Hindi-Urdu, verb agrees with the left-most nominative case marked argument (Khan, 1989). Urdu case markers and their functions have been presented in Table 2.3 as mentioned below:

Table 2.3

Urdu Case Markers and their Functions

Case Marker	Case	Function
-□	Nominative	Subject/Object
-ne	Ergative	Subject/Agentive
-ko	Accusative	Object/Theme/Patient
-ko	Dative	Subject/*Indirect Object/Goal
-me/par	Locative	Subject/Oblique
-ka/ki/ke	Genitive	Subject/Object/Possessor
-se	Instrumental	Subject/Oblique

*DO = direct object, IO = indirect object

Function (Adapted from Mohanan, 1990 & Butt, 1993 with minor changes)

Butt (1995) claimed that Urdu is not a split ergative language even though it makes use of ergative case clitic in certain situations. Ergative languages group the subject of intransitive verbs with the objects of transitive verbs for various syntactic phenomenon (Dixon, 1991; Van Valin, 1990).

Both accusative and dative cases are marked by a homophonous marker '-ko'. According to Butt (1995)'s claim accusative '-ko' may be optional but dative '-ko' is obligatory. Kiani (2013) illustrated this phenomenon by giving the following examples 33 (a), (b), (c) and (d):

33.

a) *us-ne kursi-ko uthayaa*
   s/he.3.SG.ERG chair-F.SG-ACC lift.M.SG.PERF
   'S/he lifted the chair.'

b) *us-ne kursi utha.ii*

s/he.3.SG.ERG chair-F.SG-NOM lift-F.SG.PERF
'He lifted the chair.'

c) *ustaad-ne shagird-ko kitaab dii*
teacher-M/F.SG.ERG student-M/F.SG.DAT book-F.NOM give-F.SG.PERF
'The teacher gave student a book.'

d) ***ustaad-ne shagird kitaab dii***
teacher-M/F.SG.ERG student-M/F.SG.NOM book-F.NOM give-F.SG.PERF
'The teacher gave a book to student.'

In case of pronominal object, accusative takes an oblique form '*usse'*. In Urdu, the use of '*ko*' is very prolific. Apart from the marker of accusative and dative cases, '*ko'* is also used to mark specificity (Khan, 1987); furthermore, it denotes spatial and temporal adverbs (Ahmed, 2006).

Urdu pronominals are obligatorily marked for overt accusative case in the object position. The following Table 2.4 provides the nominative and accusative/ oblique forms of Urdu pronouns.

Table 2.4

Urdu Pronominal Forms in Nominative and Accusative /Oblique Case

Person	Nominative		Accusative / Oblique	
	Singular	Plural	Singular	Plural
1	*main* (I)	*həm* (We)	*Mujhee* *mujh ko/se* (Me)	*həmẽẽ* / *həm ko/se* (Us)
2	*tum*(You)	*tum*/*aap* (You)	*tumhẽẽ* (You)	*aap ko*/*se*) (You)
3	voh (He/She/It)	voh (They)	ussee (Him/Her/It)	unhẽẽ /un/ko/(Them)

Locative and instrumental cases are realized by '*me*/*per'* and '*se'* respectively. Genitive case is shown by different inflexions of 'k-' depending on the number and gender of possessor noun phrase.

In her study, Mohanan (1997) asserts that this four leveled information about predicate is also present at the sentence level and discourse level. The interconnectedness of case marking of the subject argument and agreement in Hindi N+V compound and N+V complex predicate is revealed with the help of different layers of syntactic information. This relation was specially explored in terms of levels of structures in which the connection between semantic and syntactic information is best contained and can be explained thereafter. It focuses on the ability of predictability of some parts on basis of information about other parts in the sentences. Exploration of that information based on which this connection is explored (Mohanan, 1997). The organization and connection between these levels of structure is permitted by the standard grammatical principles is also one of the main contributions toward exploring the information required for predictability between the case marking and agreement in Hindi complex predicate (Mohanan, 1997).

Mohanan (1997) put up a finding that a grammatical subject may take any case marking available in that language, and the choice of case marking is linked to the meaning. This proposition makes the phenomenon of case selection independent of the grammatical functions. It clearly highlights the interconnectedness between the syntactic principles of case marking with the semantic information.

The interpretation of conjunct predicate (N+V) at four levels of structures i.e., semantic structures, argument structure, grammatical function structure and grammatical category structure reveal some unconventional patterns of case marking and agreement (Mohanan, 1997).

2.9 Argument Structure of Conjunct Predicate

Argument structure carries information about the syntactic valency of a predicate and complex predicates as well. All predicates have arguments (Butt, 1995). For example, the predicate *build* requires at least two semantic arguments: *builder* and the thing being *built* which stood in a one-to-one connection with grammatical functions such as *Subject* and *Object* respectively (Butt, 1995). Argument structure of predicate was given a separate status which triggers the way to form a theory in upcoming studies related to the assignment of arguments to grammatical functions (subject, object). The

Lexical Mapping Theory (LMT) was mentioned in a few studies (Levin, 1986; Bresnan & Kanerva, 1989; Alsina & Mchombo, 1990; Bresnan & Moshi, 1990; Alsina, 1990; Bresnan & Zaenen, 1990). These studies mapped *Subject* argument to the semantic role *Agent*; whereas *Object* argument is mapped onto *Themes* and *Patients*. Predicates are categorized into different subclasses based on their valency which is their argumenthood. Intransitive, transitive and ditransitive verbs such as *cough*, *hug* and *give* are called monadic, dyadic and triadic respectively (Mohanan, 1994). Monadic verbs take only one argument which is subject. For example, 'he coughed' does not require any other argument to complete the meaning in the clause. Dyadic verbs require two arguments which are subject and object. For instance, the sentence 'mother hugged her child' carries a subject (mother) and an object (child) to give a complete meaning. Triadic verbs require three arguments which are subject, object and an indirect object. For example, the sentence 'mother gave a gift to her child' needs a subject (mother), object (gift) and an indirect object (child) to complete the meaning.
At the same time, it also reflects the vivid connection of semanticity between arguments (Grimshaw & Mester, 1988; Mohanan, 1994).

The nominal host of a conjunct predicate may agree with its light verb, and as a result of this agreement, this internal nominal will be declared as an argument (Mohanan, 1994). Another feature of an internal nominal host of a conjunct predicate as an argument is that it may undergo passivization. But at the same time, she also says that the agreement of internal nominal with the light verb with a conjunct predicate is a puzzling situation and it may pose a paradox owing to the duality present in the representation of conjunct predicate (Mohanan, 1994).

Vaidya et al., (2014) work on types of nominal predicate with light verbs 'Do/*Kar*' and 'Become/*Hu'* like Ahmed and Butt (2011). They present a lexicalized Feature-based Tree-Adjoining Grammar (TAG) analysis to deal with two predicating elements and to address the mapping between syntax and semantics (Vaidya et al., 2014). TAG analysis is an attempt to deal with the highly productive combination of nouns and light verbs which pose a great challenge for computational grammar. Syntactic-semantic interface nature resulted in two different approaches to deal with the argument structure of light verb constructions. First, one is a noun centric analysis of the light verb construction where noun draws all the arguments and light verb assigns theta mark to the arguments (Grimshaw & Mester, 1988; Kearns, 1988). The second approach favors argument sharing between the noun and the light verb (Butt, 1995; Ahmed et al., 2012).

The 'Noun-Centric Analysis' and 'Multi-word Analysis' can better be represented by examining the sentence '*Ali ne sabaq yaad kiya* /Ali memorized the lesson' as given below in Figures 2.2 and 2.3 respectively:

Figure 2.2

Noun-Centric Analysis

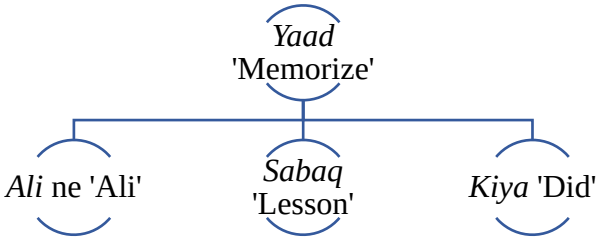
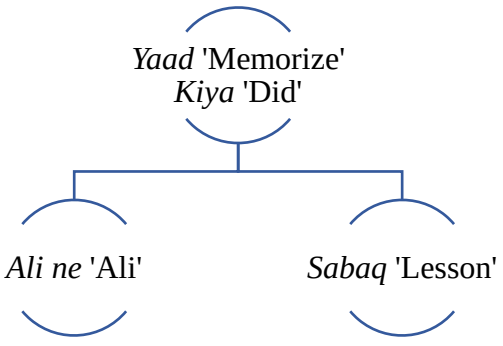


Figure 2.3

Multi-word Analysis



The second view proposes argument sharing between the noun and the light verb as they both contribute to the argument structure of the Light Verb Centric (LVC) which is referred to such analyses as verb-centric analyses (Butt, 1995; Ahmed et al., 2012). Tree-adjoining Grammar (TAG) is used to describe the syntax of natural languages in a formal tree-writing system by Bangalore and Joshi (2010). In Tree-adjoining Grammar (TAG), fragments of phrase structure tree are labeled with both terminal and non-terminal nodes.

2.10 Semantic Roles in Conjunct Predicate Constructions

Instead of looking at words as containers of meanings, they can be perceived as actors which play some roles (Yule, 2020). Same can happen to the light verbs and conjunct predicates in case of their identification as single constituent. Verbs may assign semantic roles to other noun phrases present in the sentences. Light verbs in conjunct predicates contribute to assigning semantic roles to subject arguments present in the sentence. Semantic roles are/or thematic roles also referred as 'theta roles' which is a commonly used term in the relevant literature.

Arguments bear semantic roles such as agent, goal, experiencer to their arguments. Prominently there are two schools of thought regarding the representation of semantic roles labels expressed at the argument structure of a predicate: Firstly, the most theories believe that semantic role information is carried at the argument structure; whereas, the second group of theories only believe in information related to number of arguments a predicate can take, the semantic relation and their comparative significance (Mohanan, 1994).

Identifying and extracting the semantic relationships between different lexical entities is crucial for natural language processing programs. In such tasks, main and light verbs are very important because they assign semantic roles to the arguments and other adjuncts in the sentence. So, verbs are studied to identify the arguments and their relation between predicate (Hwang et al.,2010). This specific task is called Semantic Role Labeling (SRL). Multiword Expressions (MWEs) usually pose a difficulty to this task. These multiword expressions include complex predicate which is a challenge for Semantic Role Labeling task.

Hwang et al. (2010) address the PropBank annotation of light verb constructions across languages such as English, Arabic, Chinese and Hindi. They take advantage of the already existing NomBank developed by Meyers et al., (2004) with corresponding semantic roles labels for noun predicates. Light verb constructions are a type of complex predicate which have been studied under different frameworks and from perspectives in different studies (Alsina, 1997; Butt, 1995; Mohanan, 1997).

34.

Larkay ne kahani yaad ki

Boy-M.SG.ERG story-F.SG.NOM memorized-F.SG do-F.SG.PST

'Boy memorized the story.'

In Example 34, one role is taken by the subject 'Boy/*larka'*. The noun which performs the action is technically known as the agent. Another role is taken by 'Story/*kahani*'. It is known as the theme when the direct object happens to be an inanimate entity. In the case of animate objects, the semantic role assigned to them is referred to as patient (Koeneman & Zeijlstra, 2017). Agent performs the action on theme/patient denoted by the verb. The patient/theme is affected by or undergoes the action. For instance, *'The police interrogated the culprit'.* Here, *the police* are the ones performing the activity of *interrogation* and hence referred to as *agent*. On the other hand, *the culprit* undergoes interrogation and is, therefore, said to be a *patient*. In majority cases, agents are human entities, but there is significant evidence when a natural force, creature, or machine plays the role of agent which causes action or affects something (Yule, 2020). For instance, in Urdu we can use Example 35 to explain this.

35.

Barish ne sara maza kirkara kar dia

Rain-F.SG.ERG all Fun-M.SG.NOM Spoil Do-M.SG.PST give-M.SG.PERF

'Rain spoiled all the fun.'

Here rain is a natural force which behaves as an agent. Similarly, a creature can also be an agent as shown in Example 36.

36.

Kutay nay larkay ka picha kia

Dog-M.SG.ERG boy-M.SG.GEN chase do-M.SG.PERF

'Dog chased a boy.'

If an agent makes use of another entity to accomplish a task, this overt entity assumes the role of an instrument (see Example 37 for 'Knife/*Churi'* as a noun having an instrumental semantic role).

37.

Qasaii nay churi say bakray ko zibah kia.

Butcher-M.SG.ERG knife-F.SG.INST goat-M.SG.ACC slaughter-M.SG.NOM do-M.SG.PST.

'Butcher slaughtered the goat with a knife.'

In this Example 37, 'Knife/C*huri'* assumes the semantic role of an instrument. When a noun phrase is used to designate a semantic role of an entity which feels, perceives, or states something, it assumes the semantic/theta role of an experiencer. In the following Example 38, 'Patient/*Mareez*' is not someone on which an action is performed but it experiences a feeling of wellness.

38.

Mareez ne behtar mehsoos kia.

Patient-M/F.SG.ERG better felt do M.SG.PST.

'Patient felt better.'

All the Examples (35-38) presented in this section, are used to highlight the phenomenon where two semantic entities i.e., a noun and a light verb combine to generate a syntactic category of conjunct predicate. A nominal element becomes the main predicational component, whereas light verb in such construction, assigns the case marking and the relevant semantic roles to the argument in the sentence.

Auxiliary and serial verbs are studied in Italian complex predicate to distinguish between their semantic roles with a conclusion that serial verb provides a semantic role to complex predicate whereas auxiliary does not contribute to the assignment of semantic roles (Rosen, 1997). She put forward a theory within the framework of Relational Grammar that serial verbs always occur before auxiliaries. This theory constraints that no complex predicate with an auxiliary can further be combined with a serial verb but only another auxiliary verb.

2.11 Tense/Aspect

Platts (1874) presents two categories of tenses with four types. These categories have been represented in Table 2.5.

Hindi has six tenses i.e., present, past, future, present prefect, past perfect and habitual past (Koul, 2008). Tense information comes from auxiliaries. Tense gives the time information of an event; whereas aspect is independent of time which defines duration, continuity, repetition and perfectiveness of the event (Naseer & Hussain, 2010).

Table 2.5

Categories of Tenses

x	
Imperfect (Represents incomplete action)	Perfect (Represents complete action)
1. Aorist	1. Past Absolute or indefinite
2. Perfect Imperfect	2. Proximate or Present Perfect
3. Past continuous or Imperfect	3. Remote or Past Perfect
4. Simple Future	4. Future Perfect or Past Potential

(*Source: Platts, 1874)*

The imperfective aspect in the Urdu language is realized by '*ta/ti-*' along with the present form of auxiliary *'hona'* which inflects for number as '*hai'* and '*hãĩn'* for singularity and plurality respectively. The morphological form of 't-' suffix agrees in number, person, and gender with the subject if the case is nominative. Representation of continuous aspect involves aspectual auxiliary '*raha'* which again agrees in number, gender, and person with the subject. Simple past in the Urdu language is marked by *aa*/*e*/*ii*. Past form of the verb which does not agree in number, gender and person with the subject which carries ergative case marker *'Ne'*. Future tense markers i.e., '*ga/gi/gay'* agrees with subject in person, number, and gender in presence of nominative case marker with subject (Mushtaq, 2015).

2.12 Different Theoretical Frameworks used to Study Different Complex Predicates

The phenomenon of complex predicates has been investigated by different researchers, who applied diverse theoretical frameworks based on their distinct research objectives.

Clause union is emerged out of Relational Grammar (RG) as reported by a few research studies (Aissen & Perlmutter, 1983; Davies & C. Rosen, 1988).

2.12.1 Lexical Functional Grammar

Lexical Functional Grammar advocates the syntactic phenomenon to be laid out in separated but interrelated levels of representation (Kaplan & Bresnan, 1995). It has got representation of Grammatical functions as F-Structure, and that of syntactic constituents as C-structure. It consists of three major modules: the lexicon, the syntax and the semantics. Firstly, Lexicon takes care of word formation with essential

morphological, phonological, semantic, and syntactic illustrations. Secondly, this lexicon enters syntax to construct sentences following the syntactical rules at the level of Functional (F) structure and constituent (C) Structure. Thirdly, semantics is represented at Semantic (S) structure which is studied in relation to the syntax for the probable connections. Grammatical functions are illustrated at F-Structure which are not same but related to Constituent (C) structure. A typical English sentence like 'An energetic player hit the ball' is made up of two pieces: noun phrase and a verb phrase. This verb phrase (VP) is itself made up of two entities: a verb and a noun phrase (NP). And the former NP is also analyzed into its parts: determiner, adjective and noun itself. When broken down to its terminal ends means the bottom structure, it results in the lexical entities out of which the sentence was constructed. F-structure considers a sentence as it consists of attributes which include the features of number, gender, and tense which are also interpreted as subject, predicate and object. Argument structure and semantic structure are also explained using Lexical Functional Grammar (LFG). Passivization is explained as 'lexical' according to LFG, and active -passive relation is explained as a relation between two types of verbs and not between different trees. Both active and passive verbs are entered in the lexicon. Lexical Functional Grammar is said to be an efficient model for both its in depth appeal to linguists and its efficiency for linguistic parsing and the sound formalism for computational linguists.

Alsina (1993) proposes that complex predicates are multi-headed and consist of more than one grammatical element. Based on these findings, it is claimed that complex predicates could be explored to their fuller extent following the theoretical framework of Lexical-Functional Grammar (LFG), an off-shoot of Generative Grammar. The argument structure of a complex predicate is a bit complex because two or more semantic heads combine to form a single syntactic entity (Butt, 1995). Lexical Functional Grammar (LFG) has been a frequently referred theoretical framework to illustrate the mismatch between syntactic and semantic information of complex predicate in terms of grammatical functions and constituency structure (Butt, 1995), but at the same time it has been insufficient to explain the semantically complex and syntactically discontinuous single nature of complex predicates. Butt (1995) distinguished between *permissive, instructive, and aspectual complex predicates* using the criteria of argument structure. Butt (1995) was successful in formulating a clear approach to distinguish between the serial verb and the complex predicate after conducting a detailed study on Permissive and Aspectual complex

predicate. She preferred Jackendoff 's (1992) theory of Conceptual Semantics for suggesting a detailed argument structure and talked about the limited scope of theory proposed by Alsina (1993) for the formation of Urdu complex predicate.

2.12.2 Semantic Componential Approach

Use of finite set of universal semantic-conceptual components to break down a lexical item into its smallest and simplistic semantic form can be a general explanation of this componential approach. To study verb semantics, it has been the most celebrated approach in the last quarter of the twentieth century. (Gruber, 1976; Jackendoff, 1985; Katz & Fodor, 1963; Lakoff, 1970). They advocated the possibility of semantically decomposing a word into its smallest semantic units in one way or the other way (Katz, 1971). These smallest units are termed as semantic primitives or atomic predicates. Famous study of how McCawley (1968) came up with the componential semantic analysis of verb '*Kill'* into "*cause to become not alive"* drew discussion (McCawley, 1968) and criticism by some upcoming studies of that time (Fodor, 1970; Shibatani, 1972).

A binary system is used to decompose the semantic primitives of a word, usually a noun. This binary system uses + and – to represent the existence and non-existence of certain semantic primitives related to a lexical item. Not all words are capable of being analyzed this way. For instance, it is difficult to decompose an abstract noun or a verb using binary system. Even though semantic componential approach has been declared as an inadequate theory of semantic analysis (Chomsky, 1972), some upcoming researchers of that time continued to suggest the dividing the verbs into some conceptual categories such as PLACE, MANNER, PATH, ACTION, STATE, and EVENT, etc. (Jackendoff, 1985, 1988, 1992).

Componential semantic analysis can best be explained in terms of entailment. It depicts the cooccurrence of the semantic components. However, verbs cannot be exhaustively broken down into their components using only verbs (Fellbaum, 1990, 2010; Miller et al., 1990). Lexical entailment is a one-way relation. Ways and methods to analyze nouns and verbs are different owing to their distinct syntactic categories.

2.12.3 Relational Semantics

Instead of assuming hypothetical finite semantic primes as in Componential Semantics, Relational Semantics considers semantics of lexical items as part of language users' mental lexicon, and the smallest units of analysis (Fellbaum, 1990). Though different, sometimes these two approaches complement each other. Probably, this is the reason that relational semantics is used to develop WordNet (Fellbaum, 1990). WordNet development prioritizes relational analysis and avoids semantic decomposition. In WordNet, the semantic relation among verbs relates and undergoes the phenomenon of entailment. For instance, CAUSE which is one of the important subcategory of generative semantics, is used to explain verb pairs in WordNet. In WordNet, CAUSE has a relation which links verb pairs such as *teach-learn* and *show-see.* This relation is capable of distinguishing between causative (transitive) and anticausative (intransitive) senses of some verbs such as BREAK, ROT and MOVE. There may be seen two super-predicates in English: Verbs of Change and Stative Verbs as quoted by Miller et al. (1990) because of the part of super-predicate CHANGE and STATE respectively as they refer to very basic concepts starting with topmost 'unique beginner'. The hyponyms of these superordinate general verbs are linked with each in the hierarchy using the semantic relations. As the concept of lexical inheritance, IS-A relation, is used to explain the semantic relation between nouns, NEG is used to organize adjectives; similarly, the lexical entailment is usually used to explain the semantic relation between verbs.

Verbs are looked as lexical items which need arguments to complete sentences (Levin, 1985, 1989a). These studies try to establish that the different syntactic behaviors of verbs are derived from its semantic component. Thus, try to prove that verb classes can also be made based on their semantic values. This notion of semantic values was also highlighted in an elaborate research study which supported the idea that children were able to decipher the difference between semantic-based verb classes based on their unique syntactic behavior (Pinker, 1989). Later, another study proved the intuition of children to sense syntactic-semantic constraints of action words to anticipate the meaning depending on the grammatical neighborhood of the sentence (Gleitman, 1990). The relation between the semantic structures of arguments of complex predicates and their syntactic structures has been also explored (Mohahan, 1994).

Predictability of syntactic distribution based on presumed meaning has got its traces in the Case Grammar (Fillmore, 1968), Generative Grammar (Lakoff, 1970, 1971, 1973), Government Binding Theory (Chomsky, 1982) and the Linking Mechanism.

As cited in Levin (1993), Bloomfield (1933) considers lexicon an appendix of grammar which is a list of irregularities. Lexicon bears an information regarding the idiosyncratic behavior of the lexical items; however, Levin (1993) added that the knowledge of a lexical item possessed by a language speaker implies that there is more to lexical knowledge than mere characteristic word-specific features. Levin (1993) elaborates on this phenomenon by presenting a case study of verbs which are argument taking lexical items. She highlights the ability of native language users to judge the possibility of occurrence of verb with different combination of arguments and adjuncts in different syntactic conditions. She elaborates the given proposition by presenting the case of English verb that alternation in the expression and arrangement of arguments is accompanied by change of meaning. In simple words, if the position of arguments is changed, it will result in its meaning also. Levin (1993) believes that the semantic classes of the lexical entities can be determined based on the syntactic context they occur in. Levin (1993) emphasized the arrangements of argument and their impact on the overall meaning of the clause which are termed as alternations. The concept of alternation and its different types can best be studied by the examples 39-42 quoted by Levin (1993):

Locative Alternations

39.

a. Sharon sprayed water on the plants.
b. Sharon sprayed the plants with water.

40.

a. The farmer loaded apples into the carts.
b. The farmer loaded the cart with apples.

41.

a. * Monica covered a blanket over the baby.
b. Monica covered the baby with a blanket.

42.

a. * Gina filled lemonade into the pitcher.

b. Gina filled the pitcher with lemonade.

The above-mentioned examples show that locative alternations result in semantic change; furthermore, not all verbs allow this alternation. For example, 40 (b) suggests that cart is full whereas this meaning is not present in 40(a).

Diathesis alternation is a one of the various *Transitivity Alternations* present in English language which involves a change in a verb's transitivity. There are many English verbs which can be used as both intransitive and transitive verb with a clear difference in the meaning of the clause e.g., see 43(a) example inchoative variant and 43(b) causative variant. This *causative/inchoative alternation* is not present for verb 'appear' see example 44(a) for this.

Diathesis Alternation

43.

a. The window broke.
b. The little boy broke the window.

44.

a. A rabbit appeared out of the magician's hat.
b. * The magician appeared a rabbit out of his hat.

Levin (1993) quoted an example of the word 'Gally' used by Hale and Keyser (1987) to reflect on the effect of syntactic context on the probable meaning of lexical item. It showed that people tried to decipher the meaning of a new word based on its syntactic arrangement and expression of arguments used in the clause. If the meaning of a verb is known to the people, they can predict its syntactic behavior. Or, if a new verb is presented in different syntactic contexts, the perceived meaning will be influenced by them.

The present research is an effort to explore the semantic-syntactic organization of the Urdu conjunct predicate. Levin's (1993) assumption that there is a link between semantics and syntactic context of the lexicon has been a great source of knowledge for drawing the semantic classes of nouns in conjunct predicates in the present research. The theory of semantic-syntactic connectedness is a guideline to identify the compatible semantic classes of nouns in the constituent structure of conjunct predicate (N+V) with the help of semantic-syntactic features of light verbs. Noun classes emerged from the study of conjunct predicates may be the sets of semantically related nouns sharing a

range of linguistic properties such as their compatibility with a selected number of light verbs which may further be interpreted in terms of the semantic roles of subject arguments licensed by light verbs. Surface syntactic functions of arguments and case markers are determined by light verbs in conjunct predicates. There are some examples of noun + light verb instances in the Urdu language where ergative case marker *'Ne'* does not supports an agentive role. In sentence, 'I borne/tolerated it/*main nay bardashat kia hay*' *Kia* assigns experiencer role to the subject which can be argued further based on the semantic value of related nouns in N+V instances. Tolerance/*Bardashat* is a Noun Cognition which entails the sense of cognitive activity in which the subject may not necessarily exercise the volitional agentive role.

Levin's (1993) proposition about the behavior of verb such as its argument structure and the licensed case marking correspond to the meaning it carries, can also be used to interpret the semantic classes of Urdu nouns in N+ V instances.

Three studies are presented to support the Essence Hypothesis related to the semantics of noun compounds (Levin et al. 2019). Firstly, a corpus of almost 1700 proven compounds in two domains: natural and artifact was analyzed. It was elicited that the compound from natural domain tends to refer to the essential properties; on the other hand, compounds with artifact as modifiers tend to evoke events. They anticipate the same tendencies in the production of the novel noun compounds. To study the semantic classes of nouns in Urdu conjunct predicate, Levin's (1993) theory of verb taxonomy has been followed as it talked about the correlation between semantic and syntactic contexts. Levin demonstrates it in her work (1993), lexical items with similar meaning tend to exhibit similar syntactic behavior. This theory provides linguistically motivated entries in lexicon for verbs or nouns which entail the information of meaning and syntactic expression.

2.12.4 Lexical Semantics

Lexical semantics deals with the nature of the meaning of lexical items and their mutual relations. It tends to answer how these meanings are learned and stored in human minds (Paradis, 2012). As explained by Structuralists, language is an independent intra-linguistic network of lexical relations. The Structuralists classify these lexical relations into two broader categories i.e. syntagmatic and paradigmatic (Cruse, 1986). Paradigmatic relations between words call for the possibility of them being a substitute

based on similar meanings. The words which can be substituted are called synonyms and hyponyms. For example, the synonyms *clear* and *transparent* can be substituted in a sentence: The stone is clear; The stone is transparent. *Finch* and *bird* are hypernym and hyponym which means there a hyponymy relation between them can also be substituted in a sentence: *Finch is injured; The bird is injured.*

Syntagmatic approach on the other hand refers to lexical relation between words present in a sentence. The collocation and c-occurrence patterns are explained in terms of company of a word which encompasses the contextual factors. The word with the same lexicographical representation may mean different depending on its syntactic context. Paradis (2012) used the following examples to elaborate on the syntagmatic relation between words and how the syntactic context determines the meaning of a lexical item: *I complained to the waiter because the sole was burnt; I complained to the waiter because the sole was burnt; The soles are made of rubber*. The word *sole* is an example of homonym.

There exist three different types of lexical relation (Paradis, 2012): (i) word with same form but different meaning: *homonyms and polysemy*; (ii) words with similar meanings but different forms: *synonyms and hyponyms*; (iii) Word with different forms and different meanings: *antonyms*.

Word sense disambiguation is disambiguating the sense of lexical items in a corpus using computational linguistic programs. The Natural Language Processing programs such as machine translation and speech processing have critical use of word sense disambiguation. The efficacy of different tasks such as text retrieval, document classification and document clustering are also improved using word sense disambiguation phenomenon. Due to the complex morphological system of Urdu, NLP research and word sense disambiguation research is in its nascent stage (Abid et al., 2018).

Different words with similar meanings are contained in a set which are called as synsets (Miller et al., 1990). WordNet is organized the words in relations. It represented semantic relation between synsets through pointers. There is a lexical relation between two synsets. Hyponymy/hypernymy is semantic relation between specific and general which is also represented as IS-A relation. It has a hierarchical semantic relation which originates from single superordinate (the general hypernym) and ends at subordinate (the specific hyponym). This semantic hierarchy which is represented with arrow between hypernyms and hyponyms is used in the formation of information retrieval

system. This information retrieval system is called as inheritance system (Miller, 1993; Miller et al., 1990).

Meronymy is a part-whole semantic relation which is also called HASA relation. In Wordnet, the meronymy relation is represented by pointer or parenthetical grouping between two synsets. These pointers build complex network of lexical items with meronymy relation (Miller et al., 1990).

Lexical entailment between two words *P* and *Q* demands a proposition that coexistence/interdependence of both *P* and *Q*. For example, *Snore* lexically entails *Sleep* (Fellbaum, 1990). It is a unilateral lexical relation; *P* entails *Q* does not necessitate that *Q* entails *P*.

2.13 Complex/Conjunct Predicates and English WordNet

WordNet is quite accurate, free of cost and an easily accessible online lexical resource which is used by computational linguists to develop different text based natural language programs. The famous *'Five Papers'* constitute the rudimentary works on English WordNet with the comprehensive explanation on its structure, how nouns, adjectives and verbs are represented and how it is designed to relate to other languages (Fellbaum, 1990; Gross & Miller, 1990; Miller et al., 1990; Tengi, 1998). It is different from traditional lexical resources developed by lexicographers which are based on an alphabetical order (*Princeton WordNet Online - Google Search*, 2022.). Efforts are made to simplify the tedious process of using alphabetical dictionaries by digitizing them and making computers read them as soon as a word is typed in it. Soon, it is discovered that computers can just be used as a quick turner. WordNet is the proposed plan for a more reliable combination of conventional print dictionaries and contemporary technology (Miller et al., 1990). As mentioned by Miller et al. (1990), Miller and Laird (1976) propose the term 'Psycholexicology' which incorporates the concept of using the phonological, syntactic, semantics and lexical components to work together. After the works of decades, the synergy of psychologists and linguists are successful at developing an online lexical database at Princeton University aligned to the psycholinguistic theories of human lexical memory. It is the popular reason to call it a Princeton WordNet (Miller et al., 1990; *Princeton WordNet Online - Google Search*, 2023). It is organized based on the word meanings and not on word forms. In this Princeton English WordNet, English nouns, verbs, and adjectives are categorized into sets based on semantic similarity, which are usually called as synsets.

As mentioned in Miller et al. (1990), the WordNet contains 95,600 lexical items which have both simple words (51,500) and collocations (44,100). These lexical forms are then categorized into 70,100 sets of synonyms known as synsets. A further dive into its statistics will reveal syntactic supremacy with the elicitation of a probe as a noun 79% times, adjectives 65% and verbs 43%. There has been a continuous influx of more lexical items to date as it is an online resource which is getting rich with the passage of time. This kind of syntactic categorization is further eased by organizing nouns in the hypernymy relations, verbs in the entailment relations and adjectives and adverbs in N-dimensional hyperspaces.

How concepts structures are mapped onto lexical items is done by addressing four forms of information such as phonetic form, argument structure, semantic features and grammatical features (Bierwisch & Schreuder, 1992; Higginbotham, 1985).

WordNet connects a word with a lexicalized concept along with its syntactic role within. This theory was also advocated in a few studies (Levin, 1985, 1989, 1993; Levin et al., 2019).

This proposition triggers at least three investigations: first, whether it is a single word or is an utterance based on a location of some words, second, what is the nature of lexicalized concept that the lexical entity expresses, and the third, what different syntactic categories the lexical entity can play.

2.14 WordNets in Different Languages

German WordNet is a remarkable merger of conceptual ontological information with lexical semantics with an integration of a computing tool to decipher the semantic-syntactic interaction (Hamp & Feldweg, 1997). GermaNet is different from the early version of English WordNet, as the former contains one or more syntactic frames for every verb sense and then encodes it with the lexical level. GermaNet made it possible to retrieve semantic information from larger corpora. Japanese WordNet employed the three methods for its extension: increasing the cover, linking it to the examples in corpora and other resources such as SUMO and GoiTaikei (Bond et al., 2009).

As this work is focused on mapping of nouns in Urdu Conjunct Predicate to English WordNet, following the matrix may be interesting and a highly relevant concept here (to retain its actual essence, I used the quote as it is mentioned in Miller et al.,1990):

"An entry in a cell of the matrix implies that the form in that column can be

used (in an appropriate context) to express the meaning in that row. Thus, entry $E_{1,1}$ implies that word form F_1 can be used to express word meaning M_1. If there are two entries in the same column, the word form is polysemous; if there are two entries in the same row, the two-word forms are synonyms (relative to a context)."

This matrix style shown in Figure 2.4 can be adapted to map Urdu Conjunct predicate to the semantic concepts in Urdu WordNet. Hence, as we proceed, more theories may come into consideration depending on the cross linguistic complexities.

WordNet, being a multidimensional lexical resource, distinguishes between the semantic relations and the lexical relations. It includes the emphasis on semantic relations without ignoring the lexical relations (Miller et al., 1990).

Figure 2.4

Concept of a Lexical Matrix

Illustrating the Concept of a Lexical Matrix:
F_1 and F_2 are synonyms; F_2 is polysemous

Word Meanings	Word Forms				
	F_1	F_2	F_3	. . .	F_n
M_1	$E_{1,1}$	$E_{1,2}$			
M_2		$E_{2,2}$			
M_3			$E_{3,3}$		
⋮				. . .	
M_m					$E_{m,n}$

Polysemy, synonymy, antonymy, hyponymy, and meronymy are the semantic relations which are listed among words in WordNet. However, the lexical relation has been limited to the inflectional morphological relation which were only incorporated in the interface of WordNet and not in its database as mentioned by Miller et al. in 1990. Incorporating inflectional morphological relation into WordNet proved to be a more arduous task for computer scientists than was imagined. Looking at the complexity of adding the morphological relations of English language, it can be anticipated that doing

the same for the Urdu language may pose even more technical challenges due to its combatively richer agglutinating nature.

My focus on Noun classes of Urdu conjunct predicate demands an in-depth study of nouns in WordNet (Miller et al., 1990). Hyponym which is a superordinate semantic relation generates a hierarchy of the inheritance system and provides a basis for organizing noun in English WordNet. Distinguishing features are considered to categorize nouns into twenty-five noun classes/primes. Attributes of noun e.g., Small, yellow, etc. are expressed by adjectives; whereas functions are in verbs.

Miller et al., (1990) use the metaphor of ornaments on a Christmas tree for nouns in WordNet to avoid the circularity lexicographer may experience while defining a word. And the semantic relation in which these nouns are woven is hypernymy which is also simply called as 'IS-A' relation. This superordinate relation stretches from general (hypernym) to the most specific term i.e., hyponym. A noun has usually a single superordinate, but it may have various hyponyms. As it is already mentioned above that values of attributes are given by adjectives. Nevertheless, nouns are called to be the arguments for attributes (Miller et al., 1990).

Predicate is claimed as the most important lexical and syntactic component of a language because no sentence can be complete without it as per the conventional definition of grammar. Predicate provides the relational and semantic framework for its sentence (Fellbaum, 1990; Levin, 1993).

The predicate-argument structure which is also known as subcategorization provides a link for a probable syntactic context of the sentence. The connection between nouns, their thematic roles and cases relate to distinct semanticity of event or states as expressed by sentence (Ahmed & Butt, 2011; Levin, 1993). As a result of this licensing condition, semantic properties of the noun class are distinguished which aid to complete the overall frame. Triangulation of semantic and syntactic information or let us say this semantic-syntactic interface is considered as the integral component of the subcategorization frame of a verb. It is also acclaimed that all this information is part of a language user's mental lexicon (Fellbaum, 1990; Levin, 1993). Richer polysemous nature of verb calls for the change in the meaning of a verb depending on its syntactic context i.e., depending on the kinds of noun arguments it occurs within the sentence. This less cohesive nature of verbs makes the semantic analysis of verbs a challenging task.

This concrete information regarding the connection between semantic and syntactic relations of nouns and verbs may serve as a great help for the present study which focuses on developing semantic classes of Nouns in Urdu conjunct predicate (N+V).

The most frequent English verbs such as *have and be* as mentioned by Fellbaum (1990) are the most polysemous as their semantic value heavily depends on the noun they occur with. One solution was proposed to overcome ambiguity in WordNet due to the polysemous nature of verbs that the verb synset could have information about the nouns compatible with these verbs (Fellbaum, 1990).

Simple verbs as well as phrasal verbs are part of English WordNet. All these verbs are collectively divided into 15 files based on their semantic values.

Except one, all 14 verb files represent distinct semantic domains (Fellbaum, 1990):

i. *Verbs of Bodily Functions and Care*

The verbs in this file relate to actions or events which are independent of the control of its argument. Means, the subject in the sentence has no agentive thematic role. Most of these verbs are intransitive, and they refer to involuntary actions such as *snort* and *wink* as explained by Fellbaum (1990). Some body verbs are transitive in nature, especially when action is performed on someone else or somebody part of the subject.

ii. *Verbs of Change*

This file is quite flexible and carries one of the largest number of verbs. Dowty (1986), as quoted by Fellbaum (1990), was a proponent of decompositional semantics and believed that all verbs are decomposable as Stative predicates. In addition to this, operator 'Become' is referred as a component of lexical make-up of verbs (Dowty, 1986). His claims that all verbs can be categorized as verbs of change under two conditions: one as intransitive when the Stative predicate is joined with 'Become' operator, and second, as transitive when there is an additional operator 'Do'. WordNet further subcategorizes the verbs of change into its superordinate verbs of change (Fellbaum, 1990). Some more verbs are derived from these verbs as their troponym (specific actions). Some change verbs are derived from noun by adding suffix *-ize*. This particle derivational is not common in the Urdu language, it is yielded by adding nouns to a light verb in form of a conjunct predicate. And, at the same time, some verbs of actions are formed by adding suffix

-ify or *-en* to the adjectives. Conjunct predicate where an adjective constitutes a predicate with an instance of adjective + light verb.

iii. Verbs of Communication

This file includes verbs of verbal (both speaking and writing) and nonverbal communication. The former category is rich in terms of its troponyms and derivation into binominals. These files also cater to animal noise and noise produced from inanimate sources.

iv. Verbs of Competition

This category of verb contains the semantic areas of sports, games, and warfare. It is characterized by the presence of composite troponyms (consist of two root words) and denominals.

v. Verbs of Consumption

Verb entries in this file intransitive which calls for unergative subjects. In other conditions, they may take cognate objects. Such as verb *drink* may take a cognate noun *drink* as an argument.

vi. Verbs of Contact

It is the largest verb file with mostly troponyms of few base contact verbs. Some of them entail the instrument arguments i.e., subject or object carries a semantic role of an instrument. Different types of arguments are entailed by some of the verb of contact: material, container, body part, etc.

vii. Verbs of Cognition

This file overlaps with communication verbs which means it may contain troponyms about mental activity as well as its articulation.

viii. Verbs of Creation

These verbs may be transitive when the direct object points to the creation. At the same time, it may refer to the change in the form of the material as an argument.

ix. Verbs of Emotion

Move as to create emotion, and 'Move' as to travel, these are two basic semantic senses into which this category of verb is partitioned. The first is called 'contained' motion, and second has a sense of locomotion (Pinker, 1989).

x. Verbs of Emotion or Psych

It is an interesting file of verbs which consists of two syntactically distinct classes based on the semantic behavior of their arguments: first, when subject is an animate Experiencer and object is the Source (*fear, miss, adore, love, despise*); second, when

subject is the Source and object is the animate Experiencer (*amuse, charm, encourage, anger*). These examples are used by Fellbaum (1990) to elaborate on the concept of verbs of emotions.

xi. *Stative Verbs*

Troponyms of verbs *being* and *having* fall in this file. Stative verb class is small and resembles adjectives as they usually have opposite terms.

xii. *Perception Verbs*

These verbs refer to a sense of perception with the help of five physical senses. These include several troponyms which specify the kind of perception. Arguments mostly take the semantic role of Experiencer or Source of perception.

xiii. *Verbs of Possession*

These are mostly driven from three basic semantic concepts: *have*, *give* and *take* (Fellbaum, 1990). Elaboration of these basic verbs is realized through troponyms.

xiv. *Verbs of Social Interaction*

Verbs in this file cater to the different forms of social life such as religion, politics, family, education, business, law, etc. Most verbs are denominals denoting the specialized semantic senses in different spheres of social life.

xv. *Weather Verbs*

It is the smallest verb file in which all verbs are intransitive. Means, they do not require any argument. And '*It'* is used as semantically empty expletive. Most of the weather verbs are homonyms of nouns such as *hail*, *rain*, *snow, thunder*, etc. (Fellbaum, 1990).

This division highlights the two major conceptual categories: Event and State which were primarily discussed by Jackendoff (1988). It is claimed that defining the semanticity of verbs within the paradigm of lexical semantics is quite arduous especially for WordNet which avoids the semantic decomposition in favor of relational semantics unlike other the formerly practiced approaches (Fellbaum, 1990).

Though WordNet circumvents the speakers' knowledge about the semantic and syntactic correlation of verbs, it still includes at least one sentence example for every listed verb in it. These example sentences hint the subcategorizing features of the verbs. So, WordNet can be used to study the semantic-syntactic constraints of verbs propagated by Levin (1985). So, the synsets with common sentence frames can be used to compare their semantic properties.

2.15 Mapping of Urdu Nouns onto English WordNet

In the early development of WordNet, insertion of N+V instances posed a challenge for computational linguists. A separate list of nouns is prepared for this purpose. For verb collocations, it was even more difficult. This challenge is also met with making the exception list and the several rounds of data parsing (Tengi, 1998). WordNets in different languages are being developed using the Princeton English WordNet as a gold standard. It involves the challenges of catering the cross linguistic differences. Especially when there is a condition like complex predicate and conjunct predicate V_1+V_2, Adj+ V and N+V constitute a verbal predicate and conjunct predicate, respectively.

Inadequacy of hypernymy linkage between English and Hindi WordNet is due to the sense granularity of the languages; therefore, its linkage method is suggested based on bilingual mapping (Singh et al., 2016). Bilingual mapping creates a linkage between two languages using translation and/or transliterations. It helps in capturing the uniform semanticity between languages. As a result, as advocated by Singh et al. (2016) it also enhanced the Hindi WordNet which as a result can be a valuable resource to develop multilingual applications such as machine translation and cross language information retrieval. Urdu WordNet is an attempt to facilitate natural language processing and computational linguistics programs such as Words Similarity, Word Sense Disambiguation, Information Retrieval Extraction and of course Machine Translation (Ahmed & Hautli, 2011).

Similar challenges are faced by Urdu WordNet which is in its nascent stage of development yet (Adeeba & Hussain, 2011; Saeed et al., 2019; Zafar et al., 2012). Initially, Urdu WordNet was developed following the rules of Hindi WordNet. The development of Urdu WordNet is an attempt to build a lexical resource to facilitate and augment natural language processing and computational linguistic programs such as Information Retrieval and Extraction, Word Sense Disambiguation, Word Similarity, and Machine Translation. Due to the time-consuming and costly way of construction from scratch, the Expansion Approach is preferred, and lexical information is extracted from Hindi based on similarities between Urdu and Hindi. Apart from the different writing systems, both the languages, Hindi and Urdu, share a good collection of words, their morphological rules, and their semantic values. Hindi data was to be transliterated into Urdu, so software was developed, but it could not ensure the precision of the task,

so it is manually cleaned. In addition to this, new unique words are added to Urdu WordNet and arranged into sets of synonyms (synsets). Sets of synsets aligned with the source language, Hindi, are needed to build a WordNet in a target language, Urdu in this case, which is not possible always. So, another approach, a Merge Approach, is explored (Zafar et al., 2012). This merge approach is based on building taxonomies in the Urdu language, making sysnets and defining semantic relations and finally mapping onto the senses found in the source WordNet, Princeton WordNet in this case. It involved the use of bilingual dictionaries along with the knowledge of native speakers of the Urdu language. Data organized this way is subjected to manual parsing afterwards as well. Construction of Urdu WordNet using the merge approach tried to address the issues encountered while using the former Expansion Approach (Adeeba & Hussain, 2011).

The Merge Approach required rich linguistic resources in Urdu along with the proficient and reliable team of native speakers of the Urdu language. Lexical resources related to Urdu conjunct predicates such as N+V inventory, semantic classes of noun host and a set of combinatory restrictions may augment Urdu WordNet.

Urdu when dealt with the perspective of Natural Language Possessing perspective like other languages of South Asia posed an issue of treatment of complex predicate (V_1+V_2) and conjunct predicates which are of two types (N+V) and (Adj+V). The approach of development of a Dependency Bank is followed to analyze Complex predicates and their types in Hindi and Urdu (Ahmed et al., 2012). Though Lexical Functional Grammar was the referred framework within which Complex predicates were examined, the presented guidelines to analyze all types of complex predicates in Hindi and Urdu used a different and independent framework. Moreover, the designs of the developed dependency banks are kept general to fit complex predicates found in different languages.

Large and productive instances of V1+V2, N+V and Adj +V in Urdu makes it difficult for an NLP to deal with them and needs an explicit framework to deal with them. The absence of an efficient lexical resource for their production and treatment is tried to overcome by a Reference Dependency Bank for complex predicates (Ahmed et al., 2012). This Reference Dependency bank also provides a framework for related lexical items such as modals and auxiliaries. Without this differentiation, Natural Language Processing (NLP) in Urdu may fail to yield accurate part of speech tag sets. Urdu

ParGram grammar is also another effort to describe all complex predicates in Urdu using LFG (Butt & King, 2007). On the other hand, the reference dependency bank for Urdu complex predicates is kept quite general to enable cross linguistic and theoretically independent analysis (Ahmed et al., 2012). This reference dependency grammar is considered as a 'Seed Bank' for a comprehensive study of Complex predicate across different languages of South Asia with a focus on Urdu and Hindi. In Urdu aspectual complex predicate, main verb is followed by an inflected light verb which carries an information about the kind of action or event and does not influence the argument to the overall construction. The number of arguments in the sentence are contributed by the main verb. Whereas the finite light verb influences the case marking of the subject and bares some combinatory restrictions between the main verb and itself (light verb). Second types of V_1+V_2, a permissive complex predicate, is composed of an infinitive main verb and a finite light verb where both the verbs contribute to the argument structure of the construction.

In Urdu Conjunct predicate, which is the focus of the current research work, light verb '*Kar'* provides mostly two arguments i.e., doer and the action performed. The nominal host in conjunct predicate licenses one argument in 'to memorize/*Yaad Karna*'. It happens only in the case of an incorporated noun host which does not agree with the light verb. In case of non-incorporation, noun in N+V instance will not contribute any argument in the clause. It can also be termed as an agreeing N+V conjunct predicate. This noun is also called as a grammatical object of the construction (Mohanan, 1994). Some of these agreeing nouns also allow modifiers which are identified as a different class of conjunct predicate (N+V) (Ahmed & Butt, 2011). Modals and auxiliaries are significantly different in their syntactic distribution from complex predicates. Both parts of the complex predicates contribute to the argument structure of the clause except for aspectual predicate in which arguments are only controlled by the main verb, but light verb do have an impact on case marking withing the paradigms of combinatory compatibilities between the main verb and the light verb. These combinatory restrictions are not exhibited by modal and auxiliaries which are compatible with almost all verbs (Ahmed et al., 2012).

Lack of an inventory of Urdu N+V and an efficient tool to recognize the true conjunct predicate is found to be a hinderance which widens the gap to make Urdu a digitally enabled language. With the identification of this problem, this research work aims at building an inventory of Urdu N+V instances with the exploration of semantic and

syntactic compatibility restrictions. To draw the semantic classes of compatible noun hosts in N+V instances, the semantic sense of these nouns is mapped onto the existing senses of nouns in English WordNet. This work may fill the knowledge gap by providing a useful Urdu linguistic resource necessary for Natural Language Processing (NLP) and computational linguistic programs and applications (Khalid et al., 2017).

2.16 Summary

The above section started with an introduction of taxonomy of complex predicates and its subtypes found in different languages including the Urdu language. It described the issue raised due to their intricate and complex linguistic nature which needs to be addressed to contribute to making Urdu a digitally enabled language. Both stem and the light verb have exhibited their semantic and syntactic contribution such as argument structure, case marking, semantic roles to the subjects in clause. Sufficient literature is reviewed on the taxonomy of nouns in different languages with a culmination of twenty-five noun primes suggested by Miller et al., (1990) onto them the noun classes of Urdu conjunct are mapped in the following chapter of results. With the discovery of compatible noun classes in Urdu conjunct predicate, it is hoped to draw a pattern for the combinatory conventions of Urdu nouns and light verbs. To develop a tool for the identification of true conjunct predicate, various tests employed to test complex predicate in other languages are explained in this chapter. A rationale for using 'Agreement' is considered as an efficient test for the identification of Urdu conjunct predicates. Detailed and explicit literature on Urdu case marking scheme not only aided the readability of glossary of example sentences in the thesis, but also contributed to the study of an interconnected notion of semantic roles. The section also explains the Lexical Functional Grammar which aided to resolve some structural issues involved in the interpretation of Urdu conjunct predicates. And Levin's (1993) proposition of semantic and syntactic relevance is elaborated with an aim to apply it to draw a connection between noun and the light verb in Urdu conjunct predicate (N+V). Towards the end of the chapter, Princeton English WordNet is presented as an electronic lexical tool and a standard using which WordNets in other languages including Urdu are constructed. Comparatively early stage of Urdu WordNet necessitates the mapping and insertion of conjunct predicates as they are productively used in the Urdu language.

Chapter 3: Research Methodology

This chapter aims to illustrate the research design and methodology to scale up the study of Urdu conjunct predicates by developing a comprehensive and sizeable N+V inventory and explore it to the level where some generalizations can be made about the semantic classes of Urdu nouns in N+V. This section also elaborates the analytical method through which the compatibility patterns between Urdu nouns and light verbs are studied. Due to the continuous additions of nouns to the lexicon as an open class category, Urdu N+V instances are highly productive, and drawing all possible N+V instances in the Urdu language is no less than an arduous task. Though the inventory developed in this research may not be exhaustive, still it may be rich lexical resource to study the semantic classes of Urdu nouns compatible with light verbs. Furthermore, it aims to provide patterns for semantic and syntactic connections necessary for a successful combination between noun and light verb as a conjunct predicate. Not all Urdu N+V instances may be claimed as true conjunct predicate, this complexity may pose a difficulty for their insertion in a linguistic resource as a single verbal constituent. To overcome this issue of identification of true Urdu conjunct predicate, a tool is required. Development of such recognition tool may pave the way for the insertion of Urdu conjunct predicate in electronic lexical resource such as WordNet. Filling this knowledge gap, the problems posed because of intricacy of conjunct predicates may be resolved to augment the digitally enabled status of the Urdu language. More Natural Language Programs, and computational linguistic application may be developed using this lexical resource and tool.

3.1 Research Design

A comprehensive research design is needed to accomplish this goal which can facilitate the achievement of the research objectives. Before finalizing the research design, it is very important to review the nature of research objectives and goals. Only an effective research design enables a researcher to answer the research questions. A clear research design encompasses comprehensive information related to the research objectives, overall nature of approach, rationale for the selection of corpus/participants, methods to collect data, and procedures to collect and organize data along with detailed mention of data analysis methods.

To elaborate on all these topics, let me start by reviewing my research objectives. The first research objective is to develop an inventory of Urdu N+V instances. The second research objective is to draw the semantic classes of Urdu nouns in conjunct predicates (N+V) for which semantic senses of all compatible nouns are mapped onto the semantic senses of nouns in English WordNet. Exploration of compatibility patterns and other combinatory restrictions between Urdu nouns and light verbs is the third research objective. The final and fourth research objective is to develop a tool for the identification of true Urdu conjunct predicate for which Agreement Test will be employed.

The first research goal of development of an Urdu N+V inventory is descriptive in nature. Second research question is of an exploratory nature where Urdu noun senses are explored in English WordNet and mapped. The third research inquiry is about the epistemological explanation and description of Levin's (1993) proven reality between semantic value and syntactic context of a lexical item in a clause. Development of a tool to identify true Urdu conjunct predicates is based on the methodology of employment of Agreement Test to check the noun incorporation by light verb in the clause.

The present work on Urdu conjunct predicate is based on the already existing ontological information. The theoretical background of noun taxonomy of twenty-five noun primitives (Miller et al., 1990) has been a guiding principle to explore the epistemological inquiry to find the similar noun semantic classifications in the Urdu language. In addition to this, the exploration of a correlation between the semanticity and the syntactic context is interpreted keeping in mind the proposition of Levin's (1993) taxonomical principles of English verbs.

The selection of reliable corpora, which is representative of all domains and contain regional diversity, was the first research step which is explained in detail in section 3.3 in this chapter. In case of scarcity of required data i.e., N+V instances it needs to be supplemented by a native speaker intuition to make the inventory more holistic. To do all this, two Urdu corpora i.e., Urdu Universal Dependency Treebank Corpus and Urdu WordNet Wordlist 1.0 were selected. In the former corpus, enough Urdu N+V instances were extracted, but to investigate the host noun compatibility with every light verb in the list, the data is fortified with the help of Native Urdu speaker intuition. For this purposes, native speakers Urdu were interviewed. The later corpus is a list of different

words parts of speech form which only nouns are extracted, and the process of native speaker intuition is applied to check their compatibility with light verbs.

After mapping all nouns in Urdu N+V instances, it is anticipated to see a pattern of compatible semantic classes of Urdu nouns. It not only helps in achieving the goal of drawing semantic classes of nouns in Urdu N+V instances, but also helps in understanding the combinatory restrictions between Urdu nouns and light verbs, which is the third research objective. It helps to understand the connection between semantic features and the syntactic context of a word. Levin's (1993) theory of semantic-syntactic interface correlation has been a guiding theory to execute this research.

To insert these Urdu N+V instances in Urdu WordNet is a challenge which is dependent on the identification of true conjunct predicate in Urdu. In case of their identification as single syntactic constituent, their insertion in Urdu WordNet aims to augment the digitally enabled status of the Urdu language.

Before all the above research steps are mentioned in a detailed manner in the subsequent sections, the complete theoretical framework has been described in a detailed manner.

3.2 Theoretical Framework

To study the complex predicate in the field of linguistics, many theoretical frameworks have been followed to yield varying research objectives. Lexical Functional Grammar, Government Binding Theory, Relational Grammar, and Relational Semantics are amongst the frequently referred paradigms.

The data-driven methodology is followed in an aid to Functional Approach where some generalizable instances are found with the approved notion of interconnectedness of semanticity and syntactic contexts (Atkins & Levin, 1995). There are some significant studies with theoretical contributions to semantic and syntax relevance (Beth et al., 1995; Beth & Rappaport, 1988; Pinker, 1989). The power of language lies in two principles: one is an irregular arbitrary sound-meaning connection and the other is a separate set of combinatory rules and system (Pinker, 1998). These principles make use of associative memory and symbol-manipulating rules. Pinker (1998) believes that linguistic information also has the capacity to combine abstract ideas. And it enables human beings to talk about 'all kinds of wild and wonderful ideas' (Pinker, 1998).

Here in the present work on the semantic classes of nouns in Urdu conjunct predicate, Levin's (1993) theory of verb taxonomy has been followed as it talks about the

correlation between semantic and syntactic contexts. As she demonstrates it in her work (1993), lexical items with similar meaning tend to exhibit similar syntactic behavior. This theory provides linguistically motivated entries for lexicon verbs or nouns which entails the information of meaning and syntactic expression. As cited in Levin (1993), Bloomfield (1933) referred that lexicon is really an appendix of grammar which is a list of irregularities. According to Bloomfield (1933), lexicon bears minimum information regarding the idiosyncratic behavior of the lexical item; however, Levin (1993) adds that the knowledge possessed by a language speaker regarding a lexical item implies that there is more to lexical knowledge than mere characteristic word-specific features. Levin (1993) elaborates on this phenomenon by presenting a case study of verbs which are argument taking lexical items. She highlights the ability of the native language user to judge the possibility of occurrence of verb with different combination of arguments and adjunct in different syntactic conditions. She elaborates the given proposition by presenting the case of English verb that alternations in the expression and arrangement of arguments is accompanied by change of meaning. Levin (1993) asserts that the semantic classes of the lexical entities can be determined based on the syntactic context they occur in. The present research would be an effort to explore the semantic-syntactic organization of the Urdu conjunct predicate. To draw upon the semantic classes of nouns in conjunct predicates (N+V), Levin's (1993) assumption that there is a link between semantics and syntactic context of the lexicon, has been a great source of knowledge. This theory has been a guideline to identify the compatible semantic classes of nouns in the constituent structure of conjunct predicate (N+V) with the help of syntactic features of different light verbs. Noun classes emerged from this study of conjunct predicates are semantically related nouns which share a range of linguistic properties such as their compatibility with some light verbs which are interpreted in terms of case marking argument structure and semantic roles thereafter. Tense and aspect also play an important role in this categorization. In Urdu conjunct predicate or simply let me say the instances of N+V, the surface syntactic forms of arguments and the case markers are determined by light verbs in conjunct predicates; furthermore, semanticity of nouns also affects the semantic roles taken by the subject argument in a sentence. Ergative case maker usually triggers an agentive role of the subject argument in the perfective aspect of tense i.e., past. There are examples of noun + light verb instances in the Urdu language where ergative case marker *'Ne'* does not assign

agentive role. In sentence, '*main ne bardashat kia hay*' it assigns experiencer role which may further be argued based on the semantic value of the related nouns in N+V instances.

According to Levin's (1993) proposition, behavior of verbs such as expression and interpretation of its arguments, corresponds to the meaning they carry. The correlation between their semantic orientation may be used to interpret the semantic classes of Urdu nouns in conjunct predicate (N+V). This research ventures to systemize and delimit the aspects of Urdu conjunct predicate.

Conjunct predicate, a quite productive sub-type of the complex predicate, has been under debate for its four levels of structure. i.e., Semantic Structure (SEM STR), Argument Structure (ARG STR), Grammatical Function Structure (GF STR), and Grammatical Category Structure (GC STR). Mohanan (1994) calls it an asymmetry dependency because the grammatical features and categories are predictable from the SEM STR and ARG STR; whereas the other way round is not proven possible yet (Butt, 1995; Mohanan, 1994).

The present study is an attempt to test the information shared by Mohanan (1994) for Urdu conjunct predicate, and it also inquires the validity of Levin's (1993) theory of connection between semanticity and the syntactic behavior of verbs.

Studying the semantic and syntactic patterns related to conjunct predicates is an arduous task because of the productive combination of its open class category of nouns hosted with the light verb.

A significant sized inventory of a conjunct predicate (N+V) was required based on which some general patterns of syntactic and semantic constraints and compatibility can be drawn. This information was necessary to train a program to recognize the conjunct predicate as a single grammatical constituent. Though the occurrence of conjunct predicate (N+V) is quite frequent in the Urdu language, and some work has also been done on it but not any lexical resource with the adequate information related to its four levels of structures i.e., SEM STR, ARG STR, GF STR, GC STR is available which can be used to train a program in Natural Language Processing.

Based on the information of earlier works on complex predicate (Butt, 1995; Mohanan, 1994), this work is an effort to elaborate the hidden ontology of host noun and the light verb in Urdu conjunct predicate.

Semantic classes of nouns are further elaborated keeping the noun classification used by Miller et al., (1990) and their compatibility is checked with different transitive or

intransitive light verbs. The elaborated information about the argument structure, semantic role and case marking is used to categorize different semantic classes of the conjunct predicate in Urdu.

The research approach resembles to that of positivists because of its quantitative quest of cooccurrence and finding the knowledge based on the existing single truth i.e., underlying logic which enables an Urdu language user to select and form conjunct predicate. And the same information is revealed in an explicit manner so that it can be used to improve the comparatively less-resourced status of the language.

3.3 Choice of Urdu Corpora

So, the first step to materialize the research plan of developing an inventory of Urdu N+V instances was to choose an annotated Urdu Corpus from which these instances can be extracted. Furthermore, this inventory may help to study the semantic classes of nouns in conjunct predicates (N+V), which is the second research objective. A good number of N+V instances collocations were required to draw some patterns. For this purpose, I chose Universal Dependency Urdu Treebank Corpus because it is annotated with different grammatical and categorial structures. Universal Dependency Treebank Urdu Corpus 'UD_URDU_UDTB' is a research product of Hindi/Urdu Tree Bank project. It is also placed at githib.com (Bhat et al., 2017) which can be used by researchers for research and academic purposes.

My rationale for using Universal Dependency Urdu Corpus is many folds which include an incorporation of the broader representation of Urdu dialect, its large size, free availability, and its multilayered annotated nature. It gives linguistic information related to all four levels of structures as it follows the Lexical Functional Grammar (LFG) framework which was also followed by Mohanan (1994) for the thorough understanding of Hindi verbs. It is annotated for morphological features such as gender, number, and tense of the base word. Grammatical functions i.e., subject, object, predicate etc. have been annotated. For every word in the sentences, its lemma form has also been described with its universal and regional part of speech. It followed the framework of Universal Dependency (also called as Stanford Dependency) for the consistent annotation of grammar which increases its usability and applicability for making linguistic applications in the field of Natural Language Processing.

Universal Dependency Treebank Urdu Corpus lists 5130 sentences with 138077 tokens altogether which include punctuation marks as well. A word which is considered as a

single token does not have space in between the characters. This corpus has 652 different types of tokens which include letters and punctuation, and to annotate them at the token level, 16 Universal Part of Speech (UPOS) are used. Universal dependency structures tend to maximize parallelism which allows language specific features to come in. It can be called universal taxonomy inclusive of language-specific features. Its design is principled around the dependency structures in which syntactic relations are annotated which are widely used in most of the Natural Language Processing (NLP) systems. Basic annotation units are words with different morphological properties and syntactic relations. Regarding its recoverability, it offers transparent mapping from input text to word segmentation. It also represents the basic orthographic tokens along with the syntactic words i.e., lemma, the basic word form is presented along with the actual syntactic word in use. Lemma represents the semantic content of the word, and Part of Speech (POS) represents the grammatical class. POS tagging divides the token entries into three categories: open class words, closed class words and others which include punctuation and symbols. Morphological features represent lexical and grammatical properties of the lemma the root word. These morphological features include three types: *Lexical*, *Inflectional Nominal*, and *Inflectional Verbal.*

Lexical features include PronType, NumType, Poss, Reflex, Foreign, and Abbr.

Inflectional Nominal include Gender, Animacy, Noun Class, Number, Case, Definite, and Degree.

Inflectional Verbal are Verb Form, Mood, Tense, Aspect, Voice, Evident, Polarity, Polite, Person, and Clausivity.

In the syntactic annotation layer, content words are connected by the syntactic relations. For every sentence, function words are attached to the content words they modify, whereas punctuation is attached to the head. Universal dependency presents taxonomy of 37 universal syntactic relations with three types of structures such as nominal, clauses and modifiers with the explicit information about the core arguments and other adjuncts with the presentation of language-specific subtypes. All this information is presented in the form of a non-projective tree. There is a two-level architecture. One is universal relations which allow cross-linguistic comparison. The other is subtypes to capture language-specific phenomena. While presenting the Universal Dependencies in an invited talk at Ancient Treebanks: Reports and Dreams (2019), Nivre shared that of the total languages used in Universal Dependencies, 45 % are Indo-European languages (De Marneffe et al., 2021).

The genre of this cross-language corpus if arranged from high frequency of occurrence to low, it ranges from news to fiction, non-fiction, Wikipedia, legal, spoken, blog, grammar, religious, social, web, reviews, medical, academic, poetry, email and learner essays.

Universal dependency corpus along with its CoNLL-U format is freely available for any research, whereas a formal written permission is required for any commercial use of the resource (More et al. 2018). By using this resource, I also agree to the further use of my data for any further advancement in the field of linguistics.

Theoretical framework used by Universal Dependency Treebanks is derived from Lexical Functional Grammar (LFG). The most interesting part of Universal dependencies is its intersection from syntax to semantics. It starts with a gradual addition of the C-structure, POS information which led to the determination of the head of the dependency structures (Dione, 2020). Interface of Urdu Dependency Treebank Corpus is presented in Figure 3.1 below:

Figure 3.1

Universal Dependency Urdu Treebank Corpus (De Marneffe et al., 2021)

ان	یہ	PRON	PRP	Case=Acc|Number=Plur|Person=3|PronType=Prs	29	obl	_	
پر	پر	ADP	PSP	AdpType=Post	16	case	_	ChunkId=NP5|ChunkType=child
سونٹزر	سونٹزر	PROPN	NNPC	Case=Nom|Gender=Masc|Number=Sing|Person=3	19	compound		
لینڈ	لینڈ	PROPN	NNP	Case=Acc|Gender=Masc|Number=Sing|Person=3	21	nmod	_	
کی	کا	ADP	PSP	AdpType=Post|Case=Nom|Gender=Fem|Number=Sing	19	case	_	
کمپنی	کمپنی	NOUN	NN	Case=Acc|Gender=Fem|Number=Sing|Person=3	27	iobj	_	
کو	کو	ADP	PSP	AdpType=Post	21	case	_	ChunkId=NP7|ChunkType=child
غیرقانونی	غیرقانونی	ADJ	JJ	Case=Acc	24	amod	_	ChunkId=NP8|ChunkType=child
طور	طور	NOUN	NN	Case=Acc|Gender=Masc|Number=Sing|Person=3	27	obl	_	
پر	پر	ADP	PSP	AdpType=Post	24	case	_	ChunkId=NP8|ChunkType=child
کنٹریکٹ	کنٹریکٹ	NOUN	NN	Case=Nom|Gender=Masc|Number=Sing|Person=3	27	obj	_	
دینے	دے	VERB	VM	Case=Acc|VerbForm=Inf	29	nmod	_	Vib=نا|Tam=nA|ChunkId=VG
کا	کا	ADP	PSP	AdpType=Post|Case=Nom|Gender=Masc|Number=Sing	27	mark	_	
الزام	الزام	NOUN	NN	Case=Nom|Gender=Masc|Number=Sing|Person=3	0	root	_	
ہے	ہے	AUX	VM	Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin|Voice=Act	29			

Patterns of noun classes with the collocated light verbs may lead to the development of a tool for the identification of conjunct predicates. Theoretical framework used for data collection i.e., Universal Dependencies makes it workable for natural language processing tool development for Urdu which is not fully resourced to date.

To increase the richness of data, and to make it inclusive of different Urdu dialects, I decided to use entries from Urdu WordNet 1.0 WordList (*Urdu WordNet 1.0 Wordlist*,

2013). This list was used as a language data with which Urdu WordNet was developed (Adeeba & Hussain, 2011). It consists of 5000 open ended words extracted from three multiple sources:

1. A corpus of 18 million words from online newspaper websites which have content of various domains such as culture, education, sports, politics, etc. (Ijaz & Hussain, 2007).
2. Center for Language Engineering (CLE) Urdu Digest Corpus from the diverse domains of news, literature, education, sports, health, humor, business, and international affairs which is augmented by an addition of 2000 words during the development of Urdu WordNet. This CLE Urdu Digest Corpus contains 100,000 written language data from both factual and fiction genres. All the text included in this corpus has been taken form the publications of a Pakistani Urdu magazine 'Urdu Digest' between 2003-2011. The language data in this corpus is divided into 348 Unicode Transformation Format (UTF-8 files arranged genre wise where every file consists of at least 300 words (Urooj et al., 2012). This linguistic resource is freely available for non-commercial research and pedagogical purposes.
3. Urdu Verb List which is extracted from Urdu Lughat (Urdu Online Dictionary). Released by Center for Language Engineering (CLE) Pakistan is also freely available for academic purposes (*Urdu Verb List*, accessed in June, 2022). This Verb List contains an exhaustive list of Urdu main verbs whereas there are only a few entries of complex predicates.

In this accumulated Urdu WordNet 1.0 Wordlist, inflectional forms of lexemes are not included. Some verbs have duplicate entries just to accommodate different acceptable spellings. Foreign words if they are part of Urdu Lughat were also included in it.

3.4 Data Collection Procedures

The ultimate motivating factor for using this resource is that 'Universal Dependencies Structure' is the tool which has become the first choice of natural language processing experts to develop useful tools in computational linguistics in the recent decade. I also used it to extract N+V collocations to inquire the semantic compatibility of nouns with the light verbs.

Initially, for my pilot study, I manually extracted N+V collocation by closely monitoring the sentences in UD_URDU_UTB. Multi-layered annotation provided an authentic revision to my intuition for picking it up a verb as a light verb. It also provided lemmas for different morphological forms of light verbs which further helped in their clear identification. In the initial step, I identified ten most frequently collocated light verbs: 'Do*/kar*', 'Become*/hu*', 'Is/*he*', 'Remain/*rah*', 'Put/*rakh*', 'Come/*aa*', 'Give/*dia*', 'Go/*ja*', 'Hit */laga*', and 'Take/*lia'*.

Then, I could see that manual extraction of N+V collocations to the extent for reaching some generalizable pattern was quite laborious; therefore, executed some improved method to pull all relevant instances from the corpus. In this matter, Lexico-Syntactic Pattern Extraction (LSPE) method is employed to pull all instances of N+V (Hearst, 1992, 1998). Firstly, the required searchable patterns are manually identified and then an algorithm is formulated to find the pattern in a corpus. Mining complex predicate from a parallel corpus has already been handled in a few research studies including Sinha (2009).

I am thankful to a reviewer during the preliminary presentation of the work that the project of Urdu conjunct predicate would be more representative of the local Pakistani dialect of Urdu if noun entries are included from a local Urdu corpus. So, keeping this view in mind, the list of verbs is extended by enriching Urdu nouns from Urdu WordNet Wordlist 1.0 (2013) which comprises 5000 content words. This lexicon is extracted from three sources including Ijaz and Hussain (2007), Urdu Verb List (https://www.cle.org.pk/software/ling_resources/urduverblist.htm), and CLE Urdu Digest Corpus 100K (https://www.cle.org.pk/clestore/urdudigestcorpus100k.htm). Then it is augmented by 2000 words during the development of Urdu WordNet (Zafar et al., 2012). The lexical entries in this list cover the disciplines from education, politics, literature, humor, sports, health, news, international affairs, finance, culture, sports, and business. Urdu Lughat has been an enriching source to fortify this corpus in terms of addition of verbs (http://oud.cle.org.pk/).

These 5000 words are manually sorted out for nouns only because the list includes not only nouns but verbs, adjectives, and adverbs as well. 126 unique nouns are selected, and it is tried that there is no duplication of nouns from the list of 154 nouns already crawled from Universal Dependency Urdu Treebank Corpus.

Gathered from two sources Urdu WordNet 1.0 Wordlist (2013) and Universal Dependency Urdu Treebank Corpus, I got a list of 280 unique nouns which were tested to check their semantic and syntactic compatibility with 10 light verbs: 'Do/*kar*', 'Become/*hu*', 'Is/*he*', 'Remain/*rah*', 'Put/*rakh*', 'Come/*aa*', 'Give/*dia*', 'Go/*ja*', 'Hit /*laga*', and 'Take/*lia'* .

For an upcoming future work in this domain, all naturally occurring light verbs in case of conjunct predicate may be included for discussion which could not be included in the present study to the time allowed to complete it and other research limitations.

For that purpose, I wish to include some Urdu data from Urdu Web (UrduWaC) available at Sketchengine (https://www.sketchengine.eu/corpora-and-languages/urdu-text-corpora/), but it is a non-tagged and a big sized corpus containing 53 million Urdu (Kilgarriff et al., 2010). Due to the insufficient time available for the present study, it could not be done within the timeframe available for this present study. As it is a non-tagged corpus, quite an adequate time was required to first tag the data and then extract the N+V instance from it. However, this venture can be left for some future advancement of the work.

Before the proper semantic and syntactic analysis of components of Urdu conjunct predicates i.e., Nouns and light verbs, their respective English translation is recorded from an Urdu to English dictionary (https://www.urduenglishdictionary.org/) keeping in mind the exact parts of speech (POS). Urdu nouns and verbs are also presented in English transliteration as all readers may not be expected to read Urdu writing system. English translation and transliteration of Urdu noun are provided in a table form using the MS excel sheets. English translation of the Urdu nouns is used to crawl the equivalent semantic sense from English WordNet (Fellbaum, 1998), and their WordNet ID is recorded in the parallel column.

Addressing the first research objective, an inventory of Urdu conjunct predicate was developed using these Universal Dependency Urdu Treebank Structures (Abdullah et al., 2021). An inventory of 1076 noun + light verbs instances was crawled from Universal Dependency treebanks which contained many duplicates. After the manual cleaning, data was reduced to 155 distinct nouns. These collocations include a good number of synonyms which are further stratified into a list of 15 unique noun classes (Abdullah et al., 2021).

In the later stage, this inventory crawled to come up with 280 unique N+V instances by adding handpicked data from Urdu WordNet 1.0 Wordlist . The choices of Urdu noun

and light verbs in the developed inventory is a good representative of all life domains as it is drawn from rich multidisciplinary Urdu corpora. By following this procedure of using two Urdu corpora, computationally pulled the required collocation from one Universal Dependency Urdu Treebank Corpus, aiding it by a manual handpicked data from Urdu WordNet 1.0 Wordlist, an inventory of 280 Urdu N+V could be developed. It addressed the first research goal of the thesis, and thereafter enabled to achieve the upcoming research objectives of drawing semantic classes of Urdu nouns in conjunct predicates and help fighting to understand all allowable semantic and syntactic constraints for the compatibility of Urdu nouns and light verbs.

In the next section, complete method of mapping of Urdu noun senses to the noun senses in English WordNet has been described. It is hoped to discover some pattern of semantic classes of Urdu nouns compatible with light verbs in N+V construction.

3.5 Data Analysis Procedures

For the pilot study, in an initial step, corpus has been surveyed to extract 45 nouns and their semantic and syntactic compatibility with ten most frequent light verbs mined through corpus, and simultaneously it is documented using an MS Excel sheet. The possibility of applying constituency tests such as Addition of Accusative Case Marker to the Noun, Movement, Wh-questioning, Coordination and Agreement was examined (Bhattacharyya et al., 2007; Carnie, 2012; Mohanan, 1994). The Agreement Test yielded consistent results, that is why the agreement test has been picked up as a tool to identify Urdu true conjunct predicates and a single syntactic constituency. The motivation behind applying the Agreement tests was to decide the single constituency of these conjunct predicates (N+V) so that a right decision can be made to insert them into Urdu WordNet which is an electronic lexical resource.

An Urdu noun list was developed along with its English translation and transliteration, and all the nouns are tested for their compatibility with ten light verbs one by one. Not relying only on extracted instances of Urdu N+V, each noun is tested for its compatibility with all ten light verbs one by one. For this purpose, the intuition of a native Urdu speaker is relied on. Acceptability/unacceptability of N+V instance by a native Urdu speaker is something which can easily be considered as an authentic source to lay down the data explicitly. Considering the intuition of native speakers for linguistic reliability has been advocated by many prominent researchers (Chomsky,

2002; Spenser, 1973). The notion of native speaker intuition was applied to analyze the data handpicked from Urdu WordNet Wordlist 1.0.

When a list of 280 unique nouns was prepared from two corpora, all these nouns are subjected to semantic and syntactic compatibility with these 10 light verbs: ('Do/*kar*', 'Become/*hu*', 'Is/*he*', 'Remain/*rah*', 'Put/*rakh*', 'Come/*aa*', 'Give/*dia*','Go/*ja*', 'Hit /*laga*', and 'Take/*lia'*).

This stage yielded 28100 example sentences to check the probable compatibility of Urdu nouns with light verbs. This procedure is followed to yield the research objective of finding the semantic and syntactic restrictions on the formation of conjunct predicate. MS Excel is used for effective tabulation and comparable recording of these compatibilities as depicted in Figure 3.2. These entries are studied to comprehend the semanticity linked to the host noun and the compatible light verbs.

Figure 3.2

Semantic and Syntactic Compatibility of Nouns with the Light Verbs

	Urdu Noun	English Noun	WN ID	Noun Class	Noun. Class code	Do/ Kar	Become/Hu	Is/ He	Remain/ rah	Put/ Rakh	Come/ aa	Give/ Dia	Go/ Ja	Hit/ Laga	Take/ Lee
Yaad/Memory	Yaad	Memory	{05943778}	<noun.cognition>	[09]	1	1	1	1	1	1	0	0	0	0
Intizaar/ Wait	Intizar	wait	{01065863}	<noun.act>	[04]	1	0	1	0	0	0	0	0	0	0
Mashwara/Advice	Mashwara	Advice	{06684229}	<noun.communication>	[10]	1	0	1	0	0	0	1	0	0	1
Mutalba/ Demand	Mutalba	Demand	{07205939}	<noun.communication>	[10]	1	1	1	0	0	0	0	0	0	0
Zikar/Mention	Zikar	Mention	{06778981}	<noun.communication>	[10]	1	1	1	0	0	0	0	0	0	1

Ilzaam/Accusation	Ilzaam	Accusation	{07248890}	<noun.communication>	[10]	0	1	1	0	0	0	1	0	1	1
Khasara/Loss	Khasara	Loss	{13530399}	<noun.process>	[22]	0	1	1	0	0	0	0	0	0	1
Izhaar/Expression	Izhaar	Expression	{06893299}	<noun.communication>	[10]	1	1	1	0	0	0	0	0	0	1
Qadam/Step	Qadam	Step	{00175261}	<noun.act>	[04]	0	0	1	0	1	0	1	0	0	0
Daaway/Claim	Daaway	Claim	{06742613}	<noun.communication>	[10]	1	1	1	0	0	0	0	0	0	0
Irtikaab/Confession	Irtikaab/	Confession	{01041339}	<noun.act>	[04]	1	1	0	0	0	0	0	0	0	0
Raqam/Document	Raqam	Document	{07230743}	<noun.communication>	[10]	1	1	1	0	0	0	1	0	1	1
Mouqa/Chance	Mouqa	Chance	{00804290}	<noun.act>	[04]	0	0	1	0	0	0	1	0	1	0

English transliteration is used to show Urdu noun in the first column which is followed by its English translation in the first column. For clarity, this transliteration and its English translation is listed in the next two columns. Just to keep the authenticity and reliability of the English translation in the same POS, an online Urdu to English dictionary is used (https://www.urduenglishdictionary.org/). It comes with an Urdu character keyboard which helps in typing an Urdu word on it (See Figure 3.3). Special attention was paid to picking the translation in the same POS i.e., noun.

To see the influence of semanticity of Urdu nouns on its syntactic behavior, the semantic classes of Urdu nouns were figured out by mapping them onto noun senses in English WordNet.

To keep the semantic sense of these two parallel words in two languages transparent, Princeton WordNet was used as a gold standard for selecting the right semantic sense ID and along with its semantic class (Fellbaum, 1998).

Figure 3.3

Urdu Keyboard Used in an Online Urdu-English Dictionary

This task was not simple because one lexical entry put into WordNet may present multiple senses across different parts of speech, and every sense has a unique WordNet ID.

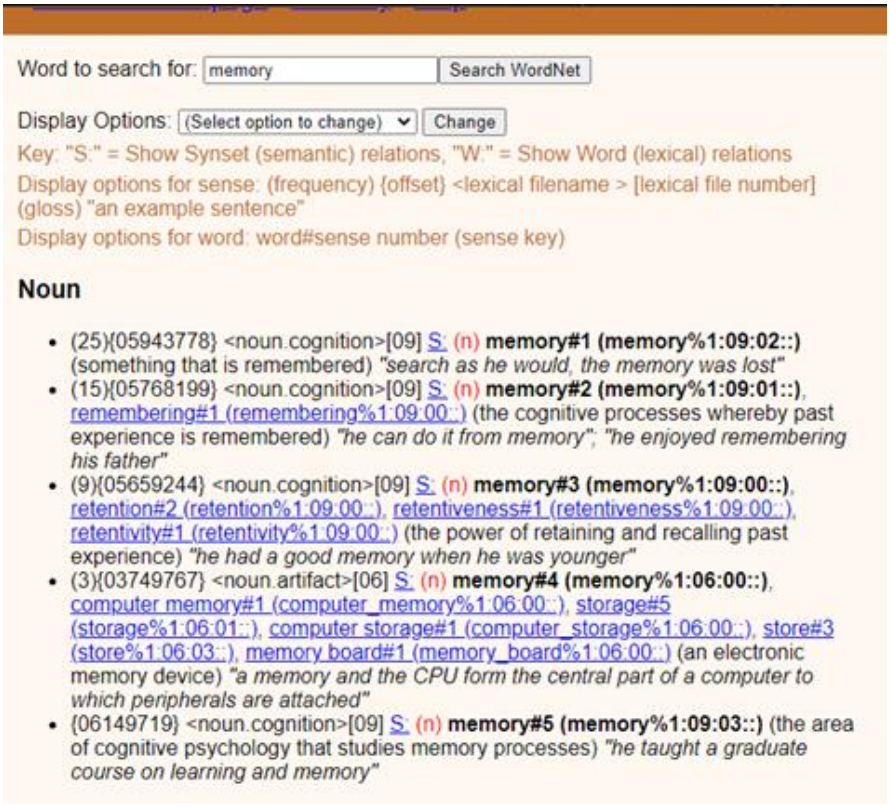
The procedure of picking up the right semantic sense was to go through all the listed senses and read the glossary to understand and pick only the intended as listed in the list of 280 Urdu nouns finalized for the work. For example, the English noun *'Memory'* presented five senses along with their unique WordNet IDs and noun classes:

1. (Noun Cognition) Something that is remembered
2. (Noun Cognition) A cognitive process whereby a past experience is remembered
3. (Noun Cognition) The power of retaining and recalling past experience
4. (Noun Artifact) An electronic memory device
5. (Noun Cognition) The area of cognitive psychology that studies memory processes

Complexity of the task may be understood by looking at the interface of the English WordNet mentioned below in Figure 3.4.

Figure 3.4

Multiple Senses of Noun 'Memory' in Princeton English WordNet (Fellbaum et al., 2010)



Information about the semantic class, here, played an important role in deciding the sense. One of the five senses of English noun *'Memory'* is noun artifact (something related to computer memory) which exactly is not a part of discussion at this point of study. So, it was eliminated at the first stage. Finalizing one sense among all noun cognition was even more tedious. The WordNet ID "{06149719} <noun.cognition>" was also not considered as the query is not about a branch of psychology. From the remaining four senses, again it was analyzed that which sense was being considered with respect to its holistic meaning because of compatibility of Urdu noun with the light verbs. Careful consideration of the possible case markers and tenses in an Urdu sentence enabled the choice of the following sense ID from WordNet. Its unique WordNet ID was documented in an MS Excel sheet to record the data for Urdu conjunct predicate.

"(25){05943778} <noun.cognition>[09] S: (n) **memory#1 (memory%1:09:02::)** (something that is remembered)' (*Princeton WordNet Online - Google Search*, 2023.)

The taxonomy of 25 noun primes was used as a referring theory to map the semantic classes of Urdu (Miller et al., 1990). Noun class 'Noun Cognition' is among the 25 noun classes as mentioned in the previous chapter of this study. The class code [9] is one of the forty-five lexicographer files into which synsets are organized during the development of WordNet https://wordnet.princeton.edu/documentation/lexnames5wn. These files are grouped logically based on the syntactic category. In mapping between equivalent Urdu and English nouns, my rationale for keeping the class code same as that of *Lexicographer Files* in WordNet, is to be consistent and to increase the usability of my data for any upcoming advanced work in my domain.

An example of noun *'End'* would be quite helpful to understand the intricacy of the process of the WordNet sense selection. This noun has got 15 semantic senses as a noun according to WordNet. As far as the semantic classes of *'End'* are concerned, among the 15 semantic senses, 2 are *Noun Act*, 2 are *Noun Location*, 1 is *Noun Time*, 1 is *Noun Event*, 2 are *Noun Cognition*, 1 is *Noun State*, 2 are *Noun Location*, and 2 are *Noun Act*. Among all these sense with different noun classes, the following sense and noun class was picked up and documented following the process of seeking suitability of concordance with the light verbs:

"(1){00789388} <noun.act>[04] S: (n) **end#11 (end%1:04:00::)** (the part you are expected to play)" (*Princeton WordNet Online - Google Search*, 2023.)

The process of sense selection gets lengthy and intricate with the polysemous nouns. It is a challenge when there is a tie between two senses with the same class. The word *'End'* has three different semantic senses with the same semantic class i.e., 'Noun Act' which is obviously sorted out keeping the semantic sense of Urdu noun and syntactic compatibilities with light verbs. There was only one sense and only one semantic class in case of 'Eradication/*Insidad*', so it was easily picked up.

The intended metaphoric entries of some lexical entries such as 'Poison/*Zehar*' which is used in terms of 'talking ill of others with the intention of harming them emotionally or physically' is not listed in WordNet.

Duplicity of synonyms of noun was avoided to allow more diversity in the list of nouns; however, noun 'Sacrifice/*Qurban*' was listed twice for its two distant semantic senses:

1. ‘Sacrifice/*Qurban*’ (Noun possession)
2. ‘Sacrifice/*Qurban*’ (Noun Act) (animal sacrificial)

I decided to use the same WordNet IDs for different senses and noun class Lexicographer File codes (as shown in Table 3.1) in my work. Consistency of uniform standard codes may enhance the prolific nature of my data to be used as a useful lexical resource in some Natural Language Processing (NLP) applications.

Table 3.1

Noun Classes and their Lexicographer File Codes

Noun Class Name	**Code**	**Noun Class Name**	**Code**
<noun.communication>	10	<noun.location>	15
<noun.Tops>	3	<noun.object>	17
<noun.act>	4	<noun.person>	18
<noun.animal>	5	<noun.phenomenon>	19
<noun.artifact>	6	<noun.plant>	20
<noun.attribute>	7	<noun.possession>	21
<noun.body >	8	<noun.process>	22
<noun.cognition>	9	<noun.quantity>	23
<noun.event>	11	<noun.shape>	25
<noun.feeling>	12	<noun.state>	26
<noun.food>	13	<noun.substance>	27
<noun.group>	14	<noun.time>	28

(Source: Princeton WordNet Online 2023)

Once the sense selection process of nouns in the list of 280 was completed, the next step was to segregate the nouns based on their noun classes. Sense selection was the first and the primary step which has already been discussed in detail in the previous heading. Selection of noun classes was tried to keep objective by following the descriptive approach instead of any subjective selection. It is just as if it emerged after comparing the sense of Urdu nouns with those of what already listed in the Princeton WordNet. The comparison and selection were not depended on their lexical representation but on their semantic senses. The complete semantic detail i.e., WordNet

Sense, WordNet ID, Noun Class and noun class codes is recorded in an Excel sheet. The next step was to segregate the entries depending on the noun classes. Consequently, the data is shifted to the separate sheets on that Excel file and renamed according to the names of the classes. First column gives the English transliteration of Urdu noun, and its English translation is listed in the second column followed by the WordNet ID in curly brackets, noun class is kept in angle brackets, and its code in square brackets in the subsequent columns in the MS Excel sheet.

3.6 Semantic Orientation of the Light Verbs

Next step was to elaborate the semantics details of the light verbs chosen for the study. So, the same method was repeated as practiced for nouns:

1. Look for the closed sense of verb in Princeton English WordNet, and record its ID
2. Choose the verb class
3. Write the WordNet lexicographer file ID used for the verb classes

Syntactic information about the light verb is also recorded i.e., transitivity which has already been used to categorize light verbs (Butt, 1995). Argument taking ability of verb is vital information about a verb which may significantly impact the overall semantic and syntactic behavior of the Urdu conjunct predicate. So, before diving into the semantic sense of light verbs, they are divided into two categories e.g., intransitive and transitive/ditransitive, and then their semantic information is completed according to the procedure mentioned above in this section.

3.7 Semantic Compatibility of Urdu Noun with Light Verbs

Now, each noun is tested to be used with a set of ten light verbs. Light verbs for this purpose are presented in parallel columns in English transliteration and translation as well. In case of a possibility of concordance (ability to occur together in N+V order) of noun with the respective light verb, '1' is inserted in the column as a code to depict 'Yes'. On the other hand, '0' is used to indicate 'No' compatibility.

A pattern emerged in a pilot study about the instances of noun compatibility with light verbs which further call for verification. Following questions were drafted to validate some hypothetical findings:

1. Are all nouns the most compatible with the transitive light verb *'*Do/*Kar*'?
2. Does a noun exhibit the same compatible behavior with both intransitive light verbs *'*Come/*Aa*' and 'Go/Ja'?

3. Does a noun behave in the same way with two ditransitive light verbs *'Give/Dia*' and *'Take/Lia*'?
4. What are characteristics of the noun which is the most prolific in terms of its compatibility with light verbs?
5. Which light verb is the most prolific in terms of its compatibility with nouns?

If the data is studied horizontally, the total sum reflects the productive nature of an individual noun. This means, how prolific some specific noun is in terms of its compatible concordance with different light verbs.

First, the collocated instances of N+V are found in the Universal Dependency Urdu Treebank Corpus. The corpus includes all the morphological inflections of light verbs which they take due to number, gender, singularity, and tense forms. Due to the scarcity of sufficient instances in corpus, intuition of native speaker of the Urdu language employed to confirm the probable compatibility of nouns with other light verbs which could not be found in corpus. The productivity of these nouns in terms of their compatibility with light verbs are calculated both horizontally and vertically. The vertical sum of a column can tell the prolific nature of a light verb in terms of its ability to collocate with different nouns; whereas the horizontal sum may reflect the ability of an Urdu noun to go with the chosen light verbs.

Subsections are made which are presented in the next chapter (result) with the semantic orientation of each noun class, their compatibility with the light verbs, and case marking, argument structure, and semantic roles.

3.8 Mapping of Urdu Conjunct Predicate onto English WordNet

Mapping between Urdu conjunct predicate and its English representative verbal lexical item was not a straightforward process for the following reasons:

1. The constituents of conjunct predicate i.e., Noun and light verb have their own semantic features which are modified when they joined together to represent a complex verb i.e., conjunct predicate here.
2. Sometimes the metaphoric interpretation causes diversion from the literal translation of the words while mapping between two languages.
3. The real marvel is to see the resultant verb class which is a product of a noun class and a verb class of light verb. It may be represented as follows:

"Semantic Class of Conjunct Predicate = Semantic Class of Noun **X** Semantic Class of Light verb"

So, the first step was to understand the meaning of Urdu nouns when they are compatible with different light verbs. The holistic semantic sense of conjunct predicate got changed depending on the attached light verb. Mapping of semantic sense of resultant Urdu conjunct predicate to English WordNet requires bilingual competency along with a reference from the authentic bilingual dictionary to understand the overall semantic class difference.

The process of mapping Urdu conjunct predicate to English WordNet can be summarized in simple steps as follows:

1. The resultant/holistic semantic sense of Urdu conjunct predicate is finalized using the same process as that of selection of the semantic senses of its constituents i.e., noun and light verb.
2. Verb sense ID is recorded in the data file along with its verb class and class ID.
3. The same is repeated with each noun class with different light verbs one by one.
4. In the first column, it is Urdu noun, in the second column it is the English translation of Urdu conjunct predicate (N+V). Here 'V' denotes the ten light verbs chosen for this study.
5. Next column constitutes the WordNet ID of the resultant verb senses in curly brackets
6. Verb class is mentioned in angle brackets and then the verb class code in square brackets constitute the last column in the table

The resultant verb classes can be filled in the main compatibility Excel sheet (as shown in Figure 3.2) to inquire the overall trend of outcome class of conjunct predicate in English language.

Not all semantic senses of Urdu N+V would be easy to map onto English WordNet. There may be some metaphoric transformation of the resultant conjunct predicate which may be different from the literal meaning of its components i.e., noun and light verb. For instance, it was very difficult to map an Urdu noun '*Out'* to Fan out.

Overt dative, accusative, genitive or ergative case marker on arguments blocks their agreement with the predicate. Urdu Copula verb *hay* seems neutral in all instances. Testing the pattern of agreement between the constituents of the conjunct predicate may

be helpful in deciding how many of the total recorded Urdu N+V instances are conjunct predicates.

The agreement was tested using elaborated instances of Urdu N+V. An interesting pattern emerged from the analysis of data that noun host in conjunct predicate does not agree with the light verb; whereas it only agrees with unmarked argument as explained in Examples 45-48.

45.

Larkay ko **sabaq** yaad **hua**

Boy-M.SG.DAT lesson-M.SG.NOM learn-F.SG became-M.SG.PERF

'Boy learnt the lesson.'

46.

Larkay nay **nazam** yaad **ki**

Boy-M.SG.ERG poem-F.SG.NOM learn-F.SG do-F.SG.PERF

'Boy memorized the lesson.'

47.

*Larkay nay maa ko **yaad kia***

Boy-M.SG.ERG mother-F.SG.ACC miss-F.SG do-M.SG.PERF

'Boy missed his mother.'

48.

*Larkay nay bhai ko **yaad kia***

Boy-M.SG.ERG brother-M.SG.ACC Miss-F.SG do-M.SG.PERF

'Boy missed his brother.'

Urdu sentences where there is no agreement between the stem noun and the light verb, these N+V instances can be considered as single syntactic unit with a holistic unique semantic feature. It may facilitate the insertion of such constituents in Urdu WordNet, an electronic lexical resource.

3.9 Summary

The present research was designed around achieving the research objectives step by step. This section aimed to provide detailed responses to a few questions. For instance, 'how the research goals are met?', 'What steps have been taken to materialize the goal

of development of Urdu N+V inventory?', 'Which Urdu corpora are used and why?', 'What data analysis methods are employed to navigate through mapping of Urdu nouns to English WordNet', 'Which set of hypotheses is followed to assess the combinatory restrictions between Urdu noun and light verbs?' and 'Why and what test have been approved efficient to identify true Urdu conjunct predicate?'. A proper rationale has been provided to justify the need to choose and adopt a particular corpus and method respectively. Achievement of the first objective of development of Urdu N+V inventory facilitated the accomplishment of the next objective of drawing semantic classes of Urdu nouns in N+V instances by mapping them onto the noun senses in English WordNet. Emerged semantic pattern of conjunctive noun classes helped in understanding the combinatory principles between noun and light verbs. Finally, a test was devised to identify true Urdu conjunct predicate which may pave the way for their insertion in the Urdu WordNet. It may help to achieve the bigger goal of improving the digital literacy status of Urdu.

Chapter 4: Results

The research philosophy and method are described in a detailed way in the previous section. It includes specific and explicit information regarding the choice of corpora and data analysis procedures employed to accomplish the research objectives. This chapter talks about what results have been retrieved from employing those methods. Results have been presented in the order of research objectives. The development of Urdu N+V inventory is the first finding of the present research. To move towards the second objective of drawing the semantic classes of nouns in Urdu N+V instances, this Urdu N+V inventory was used to check the compatibility of every listed noun with ten light verbs. Mapping of compatible Urdu nouns with light verbs to the semantic senses of nouns in English WordNet helped to discover the pattern of semantic classes of compatible Urdu nouns. Miller et al., (1990) taxonomy of 25 unique beginners of nouns in English WordNet is used as framework to map the semantic classes of Urdu nouns in N+V instances. The compatibility pattern and semantic mapping reveals that only 14 semantic classes of Urdu nouns mostly show compatibility with light verbs. This breakthrough discovery of Urdu nouns taxonomy in conjunct predicates (N+V) is termed as 'Urdu Conjunctive Noun Classes' and culminates the second research objective of the present research work.

The third objective is to understand the semantic and syntactic combinatory restrictions between nouns and light verbs in Urdu conjunct predicates (N+V). Data explicitly presented for the previous research objective i.e., semantic classes of Urdu nouns in conjunct predicate (N+V instances) helped in understanding the compatibility pattern between Urdu nouns and light verbs. Furthermore, Levin's (1993) theory of semantic and syntactic correlation and connection is followed to examine the semantic and syntactic correlating contributions of Urdu nouns and light verbs in a construction which include argument structure, case marking and semantic roles of the subject argument in the clause.

The fourth and finest objective is to come up with a tool to identify true Urdu conjunct predicates. The Agreement Test is applied to see the incorporated state of host noun in an N+V instance. In case of no agreement between the host noun and the light verb, this noun is called as an incorporated host noun and not an argument in the clause. Urdu N+V instances which are identified as true conjunct predicates are treated as single syntactic constituent and thus can be inserted in Urdu WordNet, an electronic lexical

resource. It aims to augment the efficacy of Urdu Natural Language Programs (NLP) and different computational applications.

4.1 Development of Urdu N+V Inventory

The development of an inventory of Urdu N+V instances is the first targeted goal for which an Urdu corpus i.e., Universal Dependency Urdu Treebank which is a product of Hindi/Urdu Treebank Project (Bhat et al., 2017) is used. As it is an annotated corpus, it facilitated the extraction of N+V instances though the annotation can be subjected to revision. The extracted list had to be manually cleaned as it also contained some adjectives and compound nouns without a light verb. This inventory of 1076 N+V instances listed the entries of Urdu nouns with different inflectional forms of light verbs.

Urdu N+V inventory shown in Table 4.1 may serve as useful and most-needed lexical resource for building Natural Language Programs in Urdu.

Identification of Urdu conjunct predicate in terms of noun incorporation is directly linked to the noun classes of nouns and verbs and their mutual semantic and syntactic restrictions. To study all this, development of an inventory of Urdu N+V was the first step even before the identification of true conjunct predicates. Once an inventory of a substantial number of Urdu N+V is developed, the complete semantic sense and classes of listed Urdu nouns and light verbs can be investigated and documented in a systematic way which can further be fortified with the information related to the probable argument structure, case marking on arguments and semantic roles of the arguments in the construction.

Table 4.1

Urdu Noun + Verb Inventory

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
1	قیمت لگائی	11	وضاحت -کرنی	21	واپس-آیا	31	ہلاک-کیا	41	نقصان-اٹھانا
2	یکسوئی کر	12	وضاحت -کی	22	واپس-چلا	32	ہلاک-ہو	42	نقصان-پہنچا
3	یکسوئی کی	13	وصول-کر	23	واپس-لے	33	ہلاک-ہوئے	43	نقصان-پہنچاتا
4	یقینی-بنائے	14	ورزش-کریں	24	واپس-ہو	34	ہراساں-کر	44	نقصان-پہنچتا
5	یقینی بنایا	15	واقف-کرایا	25	وابستہ-ہے	35	ہدایت-دی	45	نقصان-پہونچا
6	یقینی-ہے	16	واقف-کروایا	26	وابستہ-ہےں	36	ہدایات-دی	46	نقصان-ہوا
7	یقین-دلایا	17	واقف-ہےں	27	ہے-آنے	37	ہجرت کریں_گے	47	نقد-دئے
8	یرغمال بنا	18	واقع-ہوئی	28	ہے-پانی:	38	ہاتھ-جھاڑ	48	نعرے-لگا
9	یاد-رہے	19	واقع-ہے	29	ہے-کی	39	نوش-فرمایا	49	نظراندا ز-کیا
10	یاد-ہوگا	20	واضح-رہے	30	ہیں-کر	40	نوٹ-لیا	50	نظر-آ

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
51	یاد-ہونا	61	ہوتا-دیا	71	نظر-آئیں	81	نشاندہی-کی	91	نافرمانی-کی
52	وکالت-کی	62	ہموار-ہوگی	72	نظر-آتا	82	نشاندہی-ہوتی	92	نافذ-کر
53	وفات-پائی	63	ہمت-ہے	73	نظر-آتی	83	نذرآتش-کر	93	نارائنا-راجو
54	وعدہ-کر	64	ہلاکت-ہوئی	74	نظر-آیا	84	نذر-ہوا	94	ناانصافی-ہو
55	وعدہ-کیا	65	ہلاک-کر	75	نظام-آباد	85	ناواقف-تھے	95	میڈرڈ-ماسٹرس
56	واضح-کر	66	نوازا-ہے	76	نصیب-ہوتا	86	نامزد-کیا	96	موقف-کر
57	واضح-کیا	67	نمائندگی-کر	77	نشر-کی	87	ناگیشور-راؤ	97	موقع-دیا
58	واضح-ہو	68	نمائندگی-کرتے	78	نشانہ-بنا	88	ناکہ_بندی-کر	98	موصول-ہوئے
59	واپسی-ہوئ	69	نگرانی-کریں_گے	79	نشانہ-بناتے	89	ناکام-بنا	99	موڑ-لے
60	واپس-آ	70	نقل-کر	80	نشانہ-بنایا	90	ناکام-ہوئے	100	موزوں-سمجھا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
101	نافرمانی -کی	111	مورد- ٹہرایا	121	منقطع- ہو	131	منع-کیا	141	منصوبہ_بند ی-کرنی
102	نافذ-کر	112	موجود- تھی	122	منعقد-ہو	132	منظور ی-دی	142	منصوبہ-بنایا
103	نارائنا- راجو	113	موجود- تھے	123	منعقد- کرتے	133	منظور ی-دیدی	143	منصوبہ- رکھتے
104	ناانصاف یاں-ہو	114	موجود- ہیں	124	منعقد- کروائے	134	منظور ی-دے	144	منسوخ-کر
105	قید-کر	115	موجود- ہے	125	منعقد- کریں	135	منظور ی-لی	145	منسٹر-بنایا
106	موقف- کر	116	مہم-چل	126	منعقد- کیا	136	منظور- کر	146	منتقل-کر
107	موقع-دیا	117	مہم- چلائی	127	منعقد- ہوا	137	منظور- کی	147	منتقل-کیا
108	موصول -ہوئے	118	مہم- چھیڑ	128	منعقد- ہوگا	138	منظور- کیا	148	منتقل-ہو
109	موڑ-لے	119	مہارت- رکھتے	129	منعقد- ہوگی	139	منظم- کی	149	منتشر-کر
110	قیادت- کی	120	منقطع- رہی	130	منع- کرتی	140	منظم- کیا	150	منتخب-کرنا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
151	منتخب-ہو	161	ملتوی-کر	171	مکمل-ہوئی	181	مفلوج-رہی	191	معطل-کر
152	منتخب-ہوئے	162	ملاقاتیں-کریں	172	مقفل-کر	182	مفت-دی	192	معطل-ہے
153	منافع-ہوا	163	ملاقات-کر	173	مقرر-کی	183	مغائر-ہے	193	معذرت_خواہی-کی
154	مناسبت-رکھتے	164	ملاقات-کریں	174	مقرر-کیا	184	معین-الدین	194	معاون-ہو
155	ممکن-تھا	165	ملاقات-کرینگے	175	مقرر-ہے	185	معمہ-بنی	195	معاہدہ-کیا
156	ممکن-ہوگی	166	ملاقات-کی	176	مقبول-ہو	186	معلومات-دے	166	معائنہ-کیا
157	ممکن-ہے	167	ملاقات-ہو	177	مقبول-ہیں	187	معلوم-ہوئیں	197	مظاہرہ-کر
158	ملوث-پائی	168	ملاحظہ-فرمائیں	178	مقابلہ-کر	188	معلوم-ہوا	198	مظاہرہ-کرتے
159	ملوث-ہیں	169	مکمل-فرمایا	179	مقابلہ-کریں	189	معلوم-ہوتا	199	مظاہرہ-کیا
160	ملوث-ہےں	170	مکمل-ہو	180	مقابلہ-کیا	190	معلوم-ہونا	200	مطلع-کیا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
201	مطالبہ-کر	211	مشابہت-رکھتے	221	مذمت-کی	331	مخاطب-کریں	341	محروم-کر
202	مطالبہ-کرتی	212	مسلط-کریں	222	مذاکرات-کر	332	مخاطب-کیا	342	محروم-ہو
203	مطالبہ-کیا	213	مسدود-ہو	223	مدد-لی	333	محفوظ-بنائے	343	محروس-ہیں
204	مشورہ-دیا	214	مسترد-کر	224	مدد-ملی	334	محفوظ-رہیں_گے	344	محدود-رکھا
205	مشورہ-ملا	215	مستثنیٰ-رکھا	225	مخلوعہ-تھی	335	محفوظ-کر	345	محدود-رہنا
206	مشتمل-ہے	216	مسابقت-کر	226	مختص-کیا	336	محسوس-کر	346	محاصرہ-کر
207	مشاورت-کی	217	مزید-بتایا	227	مخالفت-تھی	337	محسوس-کیا	347	مجبور-کر
208	مشاہدہ-کر	218	مزید-کہا	228	مخالفت-کی	338	محسوس-ہوئی	348	مثال-ملتی
209	مشاہدہ-کرتے	219	مرتب-کیا	229	مخاطب-تھے	339	محسوس-ہوا	349	متنازعہ-ہوئی
210	مشابہت-رکھتا	220	مربوط-کر	330	مخاطب-کر	340	محروم-رہیں	350	متعین-کیا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
351	متعارف-کرائی	361	مبتلا-کرتا	371	گنا-کر	381	گذارش-کی	391	کمی-آنے
352	متروک-ہو	362	مبارکباد-دی	372	گمان-پہنچتا	382	گذارش-ہے	392	کم-دی
353	متجاوز-ہو	363	مانگ-کی	373	گلے-لگا	383	گامزن-رکھا	393	کلکشن-کرتا
354	متاثر-کر	364	مائل-کر	374	گلپوشی-کی	384	گامزن-کر	394	کشمکش-کر
355	متاثر-ہو	365	لہر-چلنی	375	گزارش-ہے	385	گامزن-ہو	395	کٹوتی-کی
356	متاثر-ہوتا	366	لطف_اندوز-ہو	376	گریز-کیا	386	گاڑیوں-ہوتی	396	کامیابی-دی
357	متاثر-ہوتی	367	گولیاں-چلائیں	377	گرفتار-کر	387	کوشش-کر	397	کام-لے
358	متاثر-ہوگی	368	گولی-مار	378	گرفتار-کیا	388	کوشش-کریں_گے	398	کاروائی-کی
359	متاثر-ہونا	369	گواہی-دینے	379	گردی-بڑھا	389	کوشش-کی	399	کارروائی-کر
360	مبتلا-بتائے	370	گھیراؤ-کیا	380	گذر-ہوا	390	کوشش-ہے	400	کارروائی-کی

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
401	کمی-آنے	411	قرار-دیے	421	قبضہ-ہو	431	فوت-ہو	441	فراموش-کرنا
402	کم-دی	412	قرار-دے	422	قابو-پایا	432	فوت-ہوئی	442	فرار-بتایا
403	کلکشن-کرتا	413	قتل-کر	423	قائم-کر	433	فوت-ہوئے	443	فرار-ہو
404	کشمکش-کر	414	قتل-ہوا	424	قائم-کی	434	فروغ-رہا	444	فٹ-ہوتا
405	قیادت-کی	415	قبول-کرتا	425	قائم-کیا	435	فروخت-کی	445	فاش-کر
406	قطعیت-دے	416	قبول-کرنے	426	قائم-ہو	436	فروخت-کیا	446	فائز-ہےں
407	قسمت_آزمائی-کی	417	قبول-کریں_گے	427	قائم-ہوگا	437	فروخت-کیے	447	فائرنگ-کی
408	قربت-رکھتا	418	قبول-کی	428	قائم-ہونے	438	فراہم-کر	448	فائدہ-ہو
409	قربان-جائیں	419	قبضہ-کر	429	قائم-ہیں	439	فراہم-کرے	449	فائدہ-ہوگا
410	قرار-دیا	420	قبضہ-کیا	430	فیصلہ کیا	440	فراہم کی	450	فائدہ-ہے

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
451	غور-کر	461	عطا-فرمائی	471	ظاہر-کیے	481	طلب-کیے	491	ضرب-لگےگی
452	غور-کیا	462	عطا-فرمایا	472	ظاہر-ہو	482	طاری-ہو	492	ضبط-کر
453	غائب-کر	463	عطا-کی	473	ظاہر-ہوا	483	طئے-کرے	493	ضائع-ہو
454	عہد-کیا	464	عرض-کیا	474	ظاہر-ہوتا	484	طئے-کیا	494	ضائع-ہوتی
455	علامت-بن	465	عالم-ہو	475	ظاہر-ہوتی	485	ضم-کر	495	صلاح-دی
456	عمل-آئی	466	عالم-ہے	476	طے-ہوئی	486	ضرور-ی-ہوتی	496	صدارت-کریں_گے
457	عمل-آیا	467	عائد-کی	477	طلب-کرنا	487	ضرور-ی-ہے	497	صدارت-کی
458	عمل-کر	468	عائد-کیا	478	طلب-کی	488	ضرور-ت-ہے	498	شناخت-رکھتا
459	عمل-ہو	469	ظاہر-کی	479	طلب-کیا	489	ضرور-دےگی	499	شناخت-کی
460	عطا-فرما	470	ظاہر-کیا	480	طلب-کیں	490	ضرب-پڑ	500	شمولیت-کرینگے

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
501	شمار-ہوتا	511	شکار-ہوتا	521	شروع-ہو	531	شامل-رہا	541	شادی-کر
502	شگاف-ہو	512	شریک-تھا	522	شروع-ہوئی	532	شامل-کر	542	شادی-کی
503	شکست-دی	513	شریک-تھے	523	شروع-ہوا	533	شامل-کیا	543	سیر-کیا
504	شکست-دے	514	شریک-ہو	524	شروع-ہوگی	534	شامل-کیے	544	سونیا-گاندھی
505	شکایت-رہی	515	شریک-ہوں	525	شرکت-کر	535	شامل-ہو	545	سوال-کر
506	شکایت-کی	516	شروعات-کی	526	شرکت-کریں	536	شامل-ہوتا	546	سوال-کرتے
507	شکایات-مل	517	شروع-کر	527	شرکت-کریں_گے	537	شامل-ہیں	547	سوال-کیا
508	شکار-بنے	518	شروع-کریں	528	شرکت-کی	538	شامل-ہے	548	سوار-ہوتے
509	شکار-ہو	519	شروع-کی	529	شامل-تھی	539	شامل-ہےں	549	سہولت-ہوگی
510	شکار-ہوئے	520	شروع-کیا	530	شامل-تھے	540	شادی-رچائیگی	550	سہارا-لیا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
551	سنسنی-پھیل	561	سپرد-کروایا	571	زور-ڈالا	581	ریکارڈ-کیا	591	روانہ-ہوتی
552	سنجیدہ-ہیں	562	سبق-سکھائیں_گے	572	زور-لگا	582	رویہ-رکھنا	592	روانہ-ہوگا
553	سمجھوتہ-کیا	563	سامنا-رہا	573	زندہ-ہے	583	روشناس-کرائی	593	روانہ-ہوں
554	سمجھ-لینا	564	سامنا-کر	574	زمین_بوس-ہو	584	روشناس-کرایا	594	رہا-کر
555	سماعت-کر	565	سامنا-کرنا	575	زخمی-ہو	585	روانہ-کر	595	رہا-کیا
556	سماعت-ہو	566	سامنا-ہو	576	زخمی-ہوئے	586	روانہ-کی	596	رسید-کیا
557	سفر-کر	567	سامنا-ہے	577	زخمی-ہوا	587	روانہ-کیا	597	رد-کر
558	سفارش-کی	568	سازش-رچائی	578	زخم-چھوڑے	588	روانہ-ہو	598	رخ-کیا
559	سڑکوں-ہٹا	569	سازش-رچی	579	ریمارکس-کیے	589	روانہ-ہوئی	599	رحمت-کند
560	سَرد-کر	570	زور-دیا	580	ریمارک-کیا	590	روانہ-ہوا	600	رجوع-کیا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
601	رجوع-ہوئی	611	دیا-پہنچ	621	دھمکی-دی	631	دستیاب-تھیں	641	درہم_برہم-ہو
602	رجوع-ہوا	612	دوری-ہے	622	دھکا-لگا	632	دستیاب-رہے	642	درکار-ہوگا
603	رجوع-ہوں	613	دورہ-کر	623	دھرنا-دیا	633	دستیاب-ہو	643	درخواست-کرتا
604	رپورٹ-دی	614	دورہ-کریں_گی	624	دلاسہ-دیا	634	دستیاب-ہوئی	644	درخواست-کرتے
605	رپورٹ-کرنا	615	دورہ-کیا	625	دفع-ہو	635	دستیاب-ہوا	645	درخواست-کی
606	ربط-کریں	616	دُور-ہوتی	626	دعویٰ-کیا	636	دستیاب-ہوگی	646	درخواست-ہے
607	راج-کیا	617	دور-دور	627	دعوی-کیا	637	دستخط-کئے	647	درج-کر
608	ذہن-بنا	618	دور-کرنا	628	دعوی-ہے	638	دستخط-کیے	648	درج-کرائی
609	ذکر-کیا	619	دوچار-کیا	629	دعوت-دی	639	دستاربند ی-ہوگی	649	درج-کروا
610	دیدار-کیا	620	دوچار-ہےں	630	دعاء-کی	640	دریافت-کیا	650	درج-کی

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
651	درج-کیا	661	خیال-رکھا	671	خطاب-کریں	681	خالی-کرا	691	حوالہ-دیا
652	درج-کیے	662	خیال-رکھتی	672	خطاب-کیا	682	خارج-کر	692	حملے-کیے
653	درج-ہیں	663	خیال-کیا	673	خریدار ی-کر	683	خارج-کیا	693	حملہ-کر
654	دباؤ-تھا	664	خودکش ی-کر	674	خرچ-ہو	684	خاتمہ-کر	694	حمایت-کی
655	داخل-کروائی	665	خواہش-کی	675	ختم-کر	685	حیرت-ہوئی	695	حلف-دلایا
656	داخل-کی	666	خواہاں-ہے	676	ختم-ہو	686	حوصلہ_شک نی-کی	696	حلف-لیا
657	داخل-ہو	667	خلل-پڑ	677	ختم-ہوا	687	حوصلہ_افزا ئی-کریں	697	حل-کر
658	داخل-ہوتا	668	خلل-پڑا	678	ختم-ہونا	688	حوالے-کر	698	حکومت-کرتے
659	دائر-کی	669	خطرہ-ہے	679	خبردار-کیا	689	حوالے-کی	699	حکم-دیا
660	خیرمقدم-کیا	670	خطاب-کر	680	خامیاں-ہونے	690	حوالے-کیے	700	حکم-دے

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
701	حق-پہنچتا	711	حاصل-رہےگی	721	حاصل-ہو	731	چمک-اُٹھیں	741	جمع-ہو
702	حصہ-لیا	712	حاصل-کئے	722	حاصل-ہوئی	732	چلانے-ہے	742	جمع-ہوئے
703	حصہ-لیتے	713	حاصل-کر	723	حاصل-ہوئے	733	چسپاں-کر	743	جڑے-ہوئے
704	حسد-ہے	714	حاصل-کرتا	724	حاصل-ہوا	734	چاندی-ہو	744	جذبہ-ملا
705	حذف-کیا	715	حاصل-کرتے	725	حاصل-ہوگا	735	جوش-آ	745	جدوجہد-کر
706	حدبندی-کی	716	حاصل-کرنے	726	حاصل-ہے	736	جواز-ہو	746	جدوجہد-کرنی
707	حد-ہوتی	717	حاصل-کریں	727	چیلینج-کیا	737	جواب-دیا	747	جتن-کرے
708	حامل-رہا	718	حاصل-کی	728	چیلنج-کر	738	جواب-دیں_گے	748	جاننا-چاہا
709	حاصل-تھی	719	حاصل-کیا	729	چیلنج-کیا	739	جواب-دے	749	جانچ-کی
710	حاصل-رہا	720	حاصل-کیں	730	چھاپا-مارا	740	جھڑپ-ہو	750	جانچ-ہوگی

Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V	
751	جانبر-ہو	761	جاری-ہے	771	ثبوت-ہے	781	توثیق-کرتا	791	تکمیل-کرنا
752	جاری-تھا	762	جائےگا-ہو	772	ثابت-ہوتا	782	توثیق-کی	792	تکمیل-کی
753	جاری-رکھیں	763	جائیزہ-لیا	773	ثابت-ہوگا	783	توثیق-ہو	793	تقویت-دے
754	جاری-رکھے	764	جائزہ-لیا	774	ثابت-ہوگی	784	تھی-کر	794	تقسیم-رہی
755	جاری-کر	765	جائزہ-لیں	775	تیقن-دیا	785	تنقید-کی	795	تقسیم-فرما
756	جاری-کرتی	766	جائزہ-لے	776	تیراکی-کی	786	تنبیہ-کی	796	تقسیم-کئے
757	جاری-کرنے	767	جائزہ-لےگا	777	توقع-تھی	787	تماشہ-دیکھےگی	797	تقسیم-کر
758	جاری-کی	768	ٹیلی_ک اسٹ-کرےگا	778	توقع-ہے	788	تلقین-کی	798	تقسیم-کیے
759	جاری-کیا	769	ٹیلی_ک اسٹ-ہوگا	779	توجہ-دلائی	789	تلاش-کر	799	تقرر-کیا
760	جاری-کیے	770	ٹائم-ہے	780	توجہ-دی	790	تلاش-کریں	800	تقرر-ہوگا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
801	تقاضہ-ہے	811	تعلق-رکھتے	821	تشکیل-دی	831	ترتیب-دی	841	تجربات-کیے
802	تفویض-کی	812	تعلق-ہو	822	تسلی-ہوئی	832	تخمینہ-لگایا	842	تبدیلی-کی
803	تفصیل-بتائی	813	تعظیم-ہوتی	823	ترمیم-کی	833	تخلیہ-کر	843	تبدیل-کر
804	تعینات-تھے	814	تعبیر-کیا	824	ترقی-دی	834	تحلیل-کر	844	تبدیل-کرتے
805	تعینات-کر	815	تعاون-کرنا	825	ترقی-ہو	835	تحقیقات-کر	845	تبدیل-کیا
806	تعین-کریں	816	تعاون-کریں_گی	826	ترس-آیا	836	تحقیقات-کی	846	تبدیل-ہو
807	تعین-کیا	817	تصور-کریں	827	تردید-کی	837	تحفظات-دیا	847	تباہ-کی
808	تعمیر-کیا	818	تصور-کیا	828	ترجیح-دی	838	تحریر-رہتی	848	تباہ-ہو
809	تعلیم-دلواتے	819	تصدیق-کر	829	ترجیح-دیا	839	تحریر-کیا	849	تبادلہ_خیال-کرینگے
810	تعلق-بتایا	820	تصدیق-کی	830	ترجیح-دیں	840	تجویز-رکھی	850	تبادلہ_خیال-کیا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
851	تبادلہ-کر	861	پیوست-کر	871	پیش-کریں_گی	881	پیدا-کی	891	پوری-ہوئیں
852	تبادلہ-کہا	862	پیشکش-کی	872	پیش-کریں_گے	882	پیدا-ہو	892	پورا-کرینگے
853	تبادلہ-کیا	863	پیش_قیاسی-کی	873	پیش-کرےگا	883	پیدا-ہوئی	893	پورا-ہوا
854	تانتا-بندھا	864	پیش-آ	874	پیش-کی	884	پیدا-ہوئے	894	پوچھ-کر
855	تانتا-لگا	865	پیش-آئی	875	پیش-کیا	885	پیدا-ہوا	895	پھنس-گئے
856	تازہ-کرتا	866	پیش-آئیں	876	پیش-کیے	886	پیدا-ہوتا	896	پھانسی-لے
857	تاخیر-کی	867	پیش-آئے	877	پیر-مار	887	پیچیدہ-ہوتا	897	پناہ-لی
858	تائید-کر	868	پیش-آیا	878	پیداوار-کر	888	پیٹ-کی	898	پل-پڑےگا
859	تائید-کی	869	پیش-کئے	879	پیدا-کر	889	پوسٹ_مارٹم-کیا	899	پسند-آئی
860	پےدا-کی	870	پیش-کر	880	پیدا-کرتا	890	پوری-ہوئی	900	پسند-کرتا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
901	پسند-کرتی	911	بیان-دیا	921	برقرار-رکھا	931	برا-مانیں	941	باری-آئی
902	پسند-کرے	912	بیان-کیا	922	برقرار-ہیں	932	برآمد-کیا	942	بارش-ہوئی
903	پریشان-ہو	913	بوچھار-کر	923	برقرار-ہے	933	برآمد-کیے	943	بات_چیت-کر
904	پرورش-فرماتا	914	بہم-پہنچائی	924	برطرف-کر	934	بحال-ہو	944	بات_چیت-کریں_گے
905	پرواز-بھری	915	بھگدڑ-مچ	925	برداشت-کر	935	بٹوارہ-کیا	945	بات_چیت-کی
906	پرسہ-دیا	916	بھاگ-گئے	926	برداشت-کرنا	936	بتنگڑ-بنایا	946	بات_چیت-ہوئی
907	پتہ-چلا	917	بنیاد-ڈالی	927	برداشت-کرنی	937	باور-کروایا	947	بات-کہی
908	پتہ-لگایا	918	بند-دیکھی	928	برداشت-کیا	938	باہر-ہےں	948	بات-کی
909	پارک-کیا	919	بند-کر	929	برخواست-کر	939	باقی-تھا	949	بائیکاٹ-کیا
910	پابندیاں-لگائی	920	بند-ہو	930	برخواست-ہوا	940	باقی-ہے	950	ایوارڈز-دیئے

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
951	اہمیت-دی	961	انحصار-کرتا	971	انتہاء-رہی	981	امید-تھی	991	اقدامات-کرنا
952	اہمیت-رکھتا	962	انحصار-کیا	972	انتقال-کر	982	امید-کرونگا	992	اقدامات-کیا
953	اہمیت-رکھتے	963	انجام-پایا	973	انتقال-ہو	983	امید-کی	993	اقدامات-کیے
954	اہتمام-کیا	964	انجام-دئیے	974	انتظامات-کئے	984	امکان-ہے	994	افتتاح-کیا
955	انکشاف-کیا	965	انجام-دی	975	انتظامات-کیے	985	الوداع-کیا	995	اعلان-کر
956	انکار-کر	966	انجام-دیا	976	انتظام-رہےگا	986	الزام-تھا	996	اعلان-کرتی
957	اندیشہ-ہے	967	انجام-دیتے	977	انتظام-کر	987	الزام-لگایا	997	اعلان-کیا
958	اندازہ-کیا	968	انجام-دیں	978	انتظار-کرنا	988	الزام-ہے	998	اعتراف-کیا
959	اندازہ-لگا	969	انجام-دیں_گے	979	انتباہ-دیا	989	اکٹھا-کریں	999	اعتراض-کیا
960	اندازہ-ہوتا	970	انجام-دے	980	امید-بندھ	990	اقدامات-کئے	1000	اعادہ-کیا

	Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V		Urdu N+V
1001	اظہار-کر	1011	اصرار-کیا	1021	استعمال-کر	1031	استدعا-کی	1041	اختیارات-چھین
1002	اظہار-کیا	1012	اشارہ-دیا	1022	استعمال-کرتے	1032	استثنیٰ-دیا	1042	اختیار-دیا
1003	اطلاعات-ملی	1013	اشارہ-ہے	1023	استعمال-کرنا	1033	استثنیٰ-دے	1043	اختیار-کر
1004	اطلاع-دی	1014	اسکور-کی	1024	استعمال-کرنے	1034	اڑان-بھرے	1044	اختیار-کرتے
1005	اطلاع-ہے	1015	اسکور-کیے	1025	استعمال-کریں	1035	ادعا-کیا	1045	اختیار-کریں
1006	اضافہ-کیا	1016	استقبال-کیا	1026	استعمال-کی	1036	ادا-کر	1046	اختیار-کی
1007	اضافہ-ہو	1017	استفسار-کیا	1027	استعمال-کیا	1037	ادا-کرے	1047	اختیار-کیا
1008	اضافہ-ہوا	1018	استفادہ-کر	1028	استعمال-کیے	1038	ادا-کی	1048	احساس-ہوگا
1009	اضافہ-ہوتا	1019	استفادہ-کریں	1029	استعمال-ہوتا	1039	ادا-کیا	1049	احساس-ہے
1010	اضافہ-ہوگا	1020	استعمالات-ہوتے	1030	استعمال-ہوتے	1040	اخذ-کیا	1050	احتراز-کر

	Urdu N+V		Urdu N+V		Urdu N+V
1051	احتجاج-کر	1061	اتفاق-کیا	1071	آغاز-ہوا
1052	احتجاج-کیا	1062	اپیل-کرتے	1072	آغاز-ہوگا
1053	احتجاج-ہو	1063	اپیل-کی	1073	آشکارا-ہوتے
1054	احاطہ-کیا	1064	ابتر-بنایا	1074	آزاد-کر
1055	احاطہ-ہوگا	1065	ابتداء-کی	1075	آپریشن-کیا
1056	اچھی-ہوئی	1066	آویزاں-کیے	1076	آؤٹ-کر
1057	اجازت-دی	1067	آغاز-کر		
1058	اجازت-دے	1068	آغاز-کرےگی		
1059	اثر-پڑ	1069	آغاز-کیا		
1060	اتلاف-ہو	1070	آغاز-ہو		

Holistic semantic and syntactic restrictions of Urdu noun and light verbs in conjunct predicate include the information about the semantic contribution and syntactic compatibility of its constituents i.e., nouns and light verbs having an influence to license the case markers which in turn tend to govern the semantic roles acquired by the subject argument.

To map the semantic classes of Urdu nouns, English WordNet has been used a gold standard as this lexical resource became the source for developing semantic tools in different languages. This mapping may pave the way for the augmentation of Urdu

lexical tools by inserting the semantic sense of Urdu conjunct predicate as one single lexical entity in case of noun incorporation which can otherwise is dealt separately.

Development of Urdu N+V inventory provided a sound threshold to document the semantic classes of nouns and light verbs. English WordNet is used as gold standard to map semantic senses and semantic classes of Urdu nouns. The referred semantic sense is handpicked with utmost care though chances of error may be not excluded, and further improvement of the work can be an ongoing process to obtain an authenticated and accurate outcome of some related linguistic goal e.g., information retrieval, machine translation, text to speech or speech to text, etc.

After the selection of Urdu noun and verb semantic classes, their mutual syntactic compatibility was assessed and then the probable effect of some compatible light verb on the case marking and semantic role of the subject argument is analyzed. The semantic and syntactic combinations carry information about the noun incorporation which is further elaborated in upcoming section of discussion.

4.2 Semantic Classes of Nouns and Light Verbs in Urdu Conjunct Predicates (N+V)

280 unique nouns are extracted from two corpora i.e., Universal Dependency Urdu Treebank (Bhat et al., 2017) and Urdu WordNet 1.0 Wordlist (2013).

4.2.1 Urdu Nouns

Their compatibility with light verbs is checked and documented followed by identifying the contributing semantic senses of Urdu nouns. These semantic senses of Urdu nouns are then mapped onto the noun senses found in Princeton WordNet. The selection of semantic sense and their mapping onto English WordNet has been described in the previous methodology chapter. Elaboration of noun semantic illustrations yielded categories among the noun classes. It is followed by checking their syntactic compatibility with different light verbs. Before going into the detail of syntactic illustration, noun class data is segregated according to the semantic classes of Urdu nouns.

Among the 25 noun classes listed, only the following 15 semantic classes are found in list of 280 nouns with their compositional percentage in Table 4.2. This taxonomy of semantic classes of nouns in Urdu conjunct predicates is termed as ‘Conjunctive Urdu Nouns’.

'Noun Act' is the most abundant conjunctive noun class with 105 entries which is 37.5 % of the total number of nouns in the list of 280 Urdu nouns. The second largest semantic conjunctive noun class found here is 'Noun Communication' with its 65 instances which comprises 23.3 % of the whole list.

Table 4.2

15 Unique Conjunctive Urdu Nouns

Noun Class	Frequency	Percentage
Noun Act	105	37.5
Noun Communication	65	23.3
Noun Attribute	11	3.9
Noun Time	5	1.7
Noun State	13	4.6
Noun Possession	5	1.7
Noun Relation	1	0.3
Noun Process	4	1.4
Noun Phenomenon	2	0.7
Noun Person	1	0.3
Noun Artifact	3	1
Noun Feeling	17	6
Noun Event	16	5.7
Noun Cognition	29	10
Noun Group	3	1
Total	280	100

'Noun Cognition' has 10 entries; whereas, 'Noun Event', 'Noun Feeling', and 'Noun State' have 16, 17 and 13 entries with not very varying frequencies i.e., 5.7%, 6%, and 5% respectively. Both 'Noun Person' and 'Noun Relation' have only one entry. 'Noun Possession', 'Noun Process', and 'Noun Artifact' have 5, 4, and 3 instances respectively.

Following are the 10 noun semantic classes which could not be found in this list of 280 nouns:

1. Noun body
2. Noun animal
3. Noun food
4. Noun location
5. Noun object
6. Noun group
7. Noun plant
8. Noun quantity
9. Noun shape
10. Noun substance

4.2.2 Light Verbs

Table 4.3 lists the Urdu light verbs chosen for the research.

Table 4.3

Urdu light verbs

1	2	3	4	*5*	6	7	8	9	10
Do/Kar	Becom e/*Hu*	Is/ *He*	Remain / *rah*	*Put/ Rakh*	Come/ *aa*	Give/ *Dia*	Go/ *Ja*	Hit/ *Laga*	Take/ *Lee*

These ten selected light verbs are categorized in terms of their argument taking ability i.e., transitivity. In addition to this, the semantic classes of intransitive and di/transitive Urdu light verbs are shown in Table 4.4(a) and 4.4(b) respectively. Same method is followed to identify the semantic classes of light verbs, and then map them onto the

semantic senses of verbs found in English WordNet as employed to do so for Urdu nouns (mentioned explicitly in the previous 'Methodology' chapter).

Table 4.4(a)

Mapping of Semantic Class and Argument Taking Ability of Urdu Light Verbs

Transitivity	Mapping to WordNet		
Intransitive	WordNet ID	WordNet Verb Class	Verb Class Code
Come/*Aa*	{02749089}	<verb.stative>	[42]
Go/ *Ja*	{02749089}	<verb.stative>	[42]
Is/ *He*	{02610777}	<verb.stative>	[42]
Become/*Hu*	{02730133}	<verb.stative>	[42]

In can be seen from Table 4.4 (a) that the semantic class of all intransitive light verbs is 'Verb Stative'.

Table 4.4(b)

Mapping of Semantic Class and Argument Taking Ability of Urdu Light Verbs

Light verbs			
Transitivity	Mapping to WordNet		
Di/Transitive	WordNet ID	WordNet Verb Class	Verb Class Code
Do/*Kar*	{01716563}	<verb.creation>	[36]
Remain/*Rah*	{00117793}	<verb.change>	[30]
Put/ *Rakh*	{01496967}	<verb.contact>	[35]
Hit/ *Laga*	{01407698}	<verb.contact>	[35]
Take/ *Lee*	{02272834}	<verb.possession>	[40]
Give/ *Dia*	{02321848}	<verb.possession>	[40]

Table 4.4 (b) shows that the semantic classes of di/transitive light verbs show a diverse band, and it contains 'Verb Creation', 'Verb Change', 'Verb Contact', and 'Verb Possession'.

'Do/*Kar*', 'Become/*Hu*', 'Remain/*Rah*', 'Put/*Rakh*', 'Come/*Aa*', 'Give/ *Dia*', 'Go/*Ja*', 'Hit*/Laga*', and 'Take/ *Lee*' can also behave as main verbs, but here they are acting as light verb due to their bleached semantic value and the feature of sharing predicational component with noun in form of conjunct predicate.

'Take/*lee'*, and 'Give/*Dia'* are ditransitive; whereas 'Hit/*Laga*', 'Put/ *Rakh*', 'Remain/ *Rah*', and 'Do/ *Kar*' mostly acquire transitive syntactic behavior.

All the four intransitive light verbs have verb stative class according to WordNet, and the respective code is [42]. Among the transitive verbs, two are Verb Possession [40], two are Verb Contact [35], one is Verb Change [30] and the one is Verb Creation [36]. The argument structure of the light verb yielded some conclusions regarding transitivity. Butt (1995) has explicitly declared 'Come', and 'Go' as intransitive light verbs. Verb 'Is' is referred to as copula and intransitive verb. Copula joins subject (noun) with its complement (adjective) shows that subject does not directly involve in doing an action; furthermore, there is no object in the sentence. Thus, they may be categorized as intransitive verbs.

4.3 Noun Act

This section contains elaborated data with English transliteration, English translation and the detailed semantic mapping of 15 unique Conjunctive Urdu Nouns onto English WordNet. Data related to each Urdu Conjunctive Noun is segregated and presented in different subheadings.

4.3.1 Noun Act: Semantic Orientation

Table 4.5 lists the 105 most abundantly found 'Urdu Conjunctive Noun Act' with the English transliteration in the first column. Their English translations are provided in the second column followed by the mapped WordNet IDs in curly {} brackets. The last two columns show the semantic class of Urdu noun and the noun class code in angle < > and box [ ] brackets respectively.

Table 4.5

Mapping of Urdu Conjunctive Noun Act

Urdu Noun	English Noun	WordNet ID	Noun Class	Noun Class Code
Out	Out	{00130347}	<noun.act>	[04]
Operation	Operation	{14032291}	<noun.act>	[04]
Ahtajaj	Protest	{07224885}	<noun.act>	[04]
Adaagi	Payment	{01122769}	<noun.act>	[04]
Iqdaam	Action	{00038116}	<noun.act>	[04]
Intizam	Arrangement	{00941444}	<noun.act>	[04]
Anjaam	Completion	{00558456}	<noun.act>	[04]
Boycott	Boycott	{00206979}	<noun.act>	[04]
Barkhawst	Adjournment	{00216905}	<noun.act>	[04]
Barqarar	Maintenance	{00268366}	<noun.act>	[04]
Band	Block	{00563435}	<noun.act>	[04]
Phansi	Execution	{01166269}	<noun.act>	[04]
Postmortem	Postmortem	{00142216}	<noun.act>	[04]
Tabah	Destruction	{00217881}	<noun.act>	[04]
Tajarba	Experiment	{00640799}	<noun.act>	[04]
Taraqi	Advancement	{00282894}	<noun.act>	[04]
Takmeel	Completion	{00558456}	<noun.act>	[04]
Jidojehad	Effort	{00787849}	<noun.act>	[04]
Hassil	Achievement	{00035910}	<noun.act>	[04]
Hamla	Attack	{00974725}	<noun.act>	[04]
Hosla	Encouragement	{01213509}	<noun.act>	[04]
Hoslashikni	Discouragement	{01078648}	<noun.act>	[04]
Khudkushi	Suicide	{00223352}	<noun.act>	[04]
Doura	Tour	{00311492}	<noun.act>	[04]
Dedar	Show	{00521313}	<noun.act>	[04]

Kaam	Work	{00576778}	<noun.act>	[04]
Firing	Firing	{00989024}	<noun.act>	[04]
Farokht	Sale	{01116934}	<noun.act>	[04]
Qaim	Establishment	{00237945}	<noun.act>	[04]
Qubza	Occupancy	{01056259}	<noun.act>	[04]
Qatal	Murder	{00221389}	<noun.act>	[04]
Qurban	Sacrifice	{00228462}	<noun.act>	[04]
Karwai	Action	{00955074}	<noun.act>	[04]
Katoti	Cutting	{00388465}	<noun.act>	[04]
Jamah	Collection	{01016008}	<noun.act>	[04]
Guraiz	Avoidance	{00204602}	<noun.act>	[04]
Galay	Hug	{00418872}	<noun.act>	[04]
Goli	Shoot	{00123774}	<noun.act>	[04]
Mustarad	Rejection	{00204191}	<noun.act>	[04]
Taruf	Introduction	{00238889}	<noun.act>	[04]
Mudad	Help	{01210099}	<noun.act>	[04]
Mustarad	Rejection	{00204191}	<noun.act>	[04]
Mushawarat	Consultation	{01266543}	<noun.act>	[04]
Muntaqil	Transfer	{00316812}	<noun.act>	[04]
Munsooba-bandi	Planning	{01146646}	<noun.act>	[04]
Munaqid	Organization	{00237945}	<noun.act>	[04]
Mutakhib	Selection	{00162063}	<noun.act>	[04]
Nainsafi	Injustice	{00745914}	<noun.act>	[04]
Nafarmadari	Disobedience	{01182197}	<noun.act>	[04]
Nazar-e-atish	Fire	{00989024}	<noun.act>	[04]
Vakalat	Advocacy	{01216661}	<noun.act>	[04]
Varzish	Exercise	{00625978}	<noun.act>	[04]
Hijrat	Migration	{01125416}	<noun.act>	[04]
Nighdasht	Care	{00656128}	<noun.act>	[04]
Nikaah	Marriage	{01039028}	<noun.act>	[04]

Numaish	Exhibition	{00523201}	<noun.act>	[04]
Nafiz	Enforcement	{01129340}	<noun.act>	[04]
Gorofiqar	Contemplation	{00880604}	<noun.act>	[04]
Ibadat	Worship	{01030597}	<noun.act>	[04]
Sulah	Reconciliation	{01208317}	<noun.act>	[04]
Sangsaar	Stoning	{01166884}	<noun.act>	[04]
Zakat	Charity	{01091559}	<noun.act>	[04]
Gunah	Sin	{00758433}	<noun.act>	[04]
Rishwat	Bribery	{00777759}	<noun.act>	[04]
Zaleel	Insult	{01227516}	<noun.act>	[04]
Khidmat	Service	{01212066}	<noun.act>	[04]
Kharach	Expenditure	{01124922}	<noun.act>	[04]
Jang	War	{00975181}	<noun.act>	[04]
Jaanch	Inquiry	{00637971}	<noun.act>	[04]
Transfer	Transfer	{00316812}	<noun.act>	[04]
Training	Training	{00895653}	<noun.act>	[04]
Taqseem	Division	{00386599}	<noun.act>	[04]
Tasweer-kashi	Portrayal	{00549839}	<noun.act>	[04]
Tushkeel	Arrangement	{00941444}	<noun.act>	[04]
Turbiat	Training	{00895653}	<noun.act>	[04]
Takhleeq	Creation	{00910190}	<noun.act>	[04]
Tijarat	Trade	{01093829}	<noun.act>	[04]
Tehreer	Writing	{00931533}	<noun.act>	[04]
Tehqeeq	Research	{00638164}	<noun.act>	[04]
Tajarba	Experiment	{00640799}	<noun.act>	[04]
Tabadala	Transfer	{00316812}	<noun.act>	[04]
Pooch-gach	Inquiry	{00637971}	<noun.act>	[04]
Publishing	Publishing	{01103863}	<noun.act>	[04]
Halaaq	Kill	{00219879}	<noun.act>	[04]
Baqawat	Rebellion	{01179817}	<noun.act>	[04]
Barbaad	Destruction	{00217881}	<noun.act>	[04]
Bachat	Saving	{00193462}	<noun.act>	[04]
Araam	Rest	{01066072}	<noun.act>	[04]

Azmaish	Test	{01008617}	<noun.act>	[04]
Inaam	Gift	{01088005}	<noun.act>	[04]
Ijazat	Permission	{01141515}	<noun.act>	[04]
Shru	Beginning	{00236302}	<noun.act>	[04]
Tashakar	Thanks	{01211287}	<noun.act>	[04]
Shamil	Inclusion	{00373938}	<noun.act>	[04]
Qabu	Control	{00832109}	<noun.act>	[04]
Kaarwai	Operation	{00410304}	<noun.act>	[04]
Istamaal	Use	{00948944}	<noun.act>	[04]
Koshish	Try	{00787849}	<noun.act>	[04]
Anjaam	End	{00789388}	<noun.act>	[04]
Mouqa	Chance	{00804290}	<noun.act>	[04]
Irtikaab/	Confession	{01041339}	<noun.act>	[04]
Qadam	Step	{00175261}	<noun.act>	[04]
Intizaar	Wait	{01065863}	<noun.act>	[04]
Himayat	Support	{01218392}	<noun.act>	[04]

4.3.2 Noun Act: Compatibility with the Light Verbs

Every Noun Act is checked for its compatibility with ten light verbs i.e., 'Do/*Kar',* 'Become/*Hu*', 'Be/*He',* 'Put/*Rakh',* 'Come/*Aa*', 'Give'/*Dia,* 'Go/*Ja*', 'Take/*Lena*', 'Remain/*Rah',* and 'Hit */Laga'.* The last column with the heading 'N+V', if examined horizontally, mentions the productivity of each Noun Act which means its total number of compatibilities with light verbs. Columns with the headings of unique light verb exhibit the state of compatibility of a light verb with a Noun Act, and at the end if calculated vertically it tells the productivity of a light verb with different Noun Acts as mentioned in this Table 4.6. '0' is entered in case of no compatibility and '1' is entered in case of compatibility between Noun Act and light verbs.

Table 4.6

Compatibility Pattern of Noun Act with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Out	Out	1	1	1	0	0	0	0	0	0	0	3
Operation	Operation	1	1	1	1	0	0	0	0	0	0	4
Ahtajaj	Protest	1	1	1	1	0	0	0	0	0	0	5
Adaagi	Payment	1	1	0	0	0	0	0	0	0	0	2
Iqdaam	Action	1	1	0	0	0	0	0	0	0	0	2
Intizam	Arrangement	1	1	1	1	0	0	0	0	0	0	4
Anjaam	Completion	1	1	1	1	0	0	0	0	0	0	5
Boycott	Boycott	1	1	1	1	0	0	0	0	0	0	4
Barkhawst	Adjournment	1	1	0	0	0	0	0	0	0	0	2
Barqarar	Maintenance	0	0	0	1	1	0	0	0	0	0	2
Band	Block	1	1	1	1	1	0	0	0	0	0	5
Phansi	Execution	0	1	1	0	0	0	1	0	0	1	4

Postmortem	Postmortem	1	1	1	0	0	0	0	0	0	0	3
Tabah	Destruction	1	1	1	0	0	0	0	0	0	0	3
Tajarba	Experiment	1	1	1	1	1	0	0	0	0	0	5
Taraqi	Advancement	1	1	1	0	0	0	1	0	0	0	4
Takmeel	Completion	1	1	1	0	0	0	0	0	0	0	3
Jidojehad	Effort	1	1	1	1	0	0	0	0	0	0	4
Hassil	Achievement	1	1	1	0	0	0	0	0	0	0	3
Hamla	Attack	1	1	1	0	0	0	0	0	0	0	3
Hosla	Encouragement	1	1	1	1	1	0	1	0	0	1	7
Hoslashikni	Discouragement	1	1	1	0	0	0	0	0	0	0	3
Khudkushi	Suicide	1	1	1	0	0	0	0	0	0	0	3
Doura	Tour	1	1	0	0	0	0	0	0	0	0	2
Dedar	Show	1	1	0	0	0	0	0	0	0	0	2
Kaam	Work	1	1	1	1	0	1	1	0	0	1	7
Firing	Firing	1	1	0	0	0	0	0	0	0	0	2
Farokht	Sale	1	1	1	0	0	0	0	0	0	0	3
Qaim	Establishment	1	1	1	1	1	0	0	0	0	0	5
Qubza	Occupancy	1	1	1	1	0	0	1	0	0	1	6

Qatal	Murder	1	1	1	0	0	0	0	0	0	0	3
Qurban	Sacrifice	1	1	0	0	0	0	1	0	0	1	4
Karwai	Action	1	1	1	0	0	0	0	0	0	0	3
Katoti	Cutting	1	1	1	0	0	0	0	0	0	0	3
Jamah	Collection	1	1	1	1	1	0	0	0	0	0	5
Guraiz	Avoidance	1	1	1	0	0	0	0	0	0	0	3
Galay	Hug	0	0	0	0	0	0	0	0	1	0	1
Goli	Shoot	0	0	0	0	0	0	0	0	1	0	1
Mustarad	Rejection	1	1	0	0	0	0	0	0	0	0	2
Taruf	Introduction	1	1	1	0	0	0	0	0	0	0	3
Mudad	Help	1	1	0	0	0	0	1	0	0	1	4
Mustarad	Rejection	1	1	0	0	0	0	0	0	0	0	2
Mushawarat	Consultation	1	1	0	0	0	0	1	0	0	1	4
Muntaqil	Transfer	1	1	0	0	0	0	0	0	0	0	2
Munsooba-bandi	Planning	1	1	0	0	0	0	0	0	0	0	2
Munaqid	Organization	1	1	0	0	0	0	0	0	0	0	2
Muntakhib	Selection	1	1	0	0	0	0	0	0	0	0	2
Nainsafi	Injustice	1	1	1	0	0	0	0	0	0	0	3

Nafarmadari	Disobedience	1	1	0	0	0	0	0	0	0	0	2
Nazar-e-atish	Fire	1	1	0	0	0	0	0	0	0	0	2
Vakalat	Advocacy	1	1	0	0	0	0	0	0	0	0	2
Varzish	Exercise	1	1	1	0	0	0	0	0	0	0	3
Hijrat	Migration	1	1	0	0	0	0	0	0	0	0	2
Nighdasht	Care	1	1	0	0	0	0	0	0	0	0	2
Nikaah	Marriage	1	1	1	1	0	0	0	0	0	0	4
Numaish	Exhibition	1	1	1	0	0	0	0	0	1	0	4
Nafiz	Enforcement	1	1	0	1	0	0	0	0	0	0	3
Gorofiqar	Contemplation	1	1	0	0	0	0	0	0	0	0	2
Ibadat	Worship	1	0	0	0	0	0	0	0	0	0	1
Sulah	Reconciliation	1	1	1	0	0	0	0	0	0	0	3
Sangsaar	Stoning	1	1	0	0	0	0	0	0	0	0	2
Zakat	Charity	0	0	0	0	0	0	1	0	0	1	2
Gunah	Sin	1	1	1	0	0	0	1	0	0	1	4
Rishwat	Bribery	0	0	0	0	0	0	1	0	0	1	2
Zaleel	Insult	1	1	1	0	0	0	0	0	0	0	3

Khidmat	Service	1	0	0	0	0	0	0	0	0	0	1
Kharach	Expenditure	1	1	0	0	0	0	0	0	0	0	2
Jang	War	1	1	1	1	0	0	0	0	1	0	5
Jaanch	Inquiry	1	1	0	1	0	0	0	0	0	0	3
Transfer	Transfer	1	1	0	0	0	0	0	0	0	0	2
Training	Training	1	1	0	0	0	0	0	0	0	0	2
Taqseem	Division	1	1	1	0	0	0	0	0	0	0	3
Taswer-kashi	Portrayal	1	0	0	0	0	0	0	0	0	0	1
Tushkeel	Arrangement	1	1	1	0	0	0	0	0	0	0	3
Turbiat	Training	1	1	1	0	0	0	0	0	0	0	3
Takhleeq	Creation	1	1	1	0	0	0	0	0	0	0	3
Tijarat	Trade	1	1	0	0	0	0	0	0	0	0	2
Tehreer	Writing	1	1	1	0	0	0	0	0	0	0	3
Tehqeeq	Research	1	1	0	0	0	0	0	0	0	0	2
Tajarba	Experiment	1	1	1	0	0	0	0	0	0	0	4
Tabadala	Transfer	1	1	0	0	0	0	0	0	0	0	2
Poochgach	Inquiry	1	1	0	0	0	0	0	0	0	0	2

Publishing/	Publishing	1	1	0	0	0	0	0	0	0	0	2
Halaaq	Kill	1	1	1	0	0	0	0	0	0	0	3
Baqawat	Rebellion	1	1	1	0	0	0	0	0	0	0	3
Barbaad	Destruction	1	1	1	1	0	0	0	0	0	0	4
Bachat	Saving	1	1	0	0	0	0	0	0	0	0	3
Araam	Rest	1	1	1	1	0	0	1	0	0	1	5
Azmaish	Test	0	1	1	0	0	0	1	0	0	1	4
Inaam	Gift	0	0	1	0	0	0	1	0	0	1	3
Ijazat	Permission	0	1	1	0	0	0	1	0	0	1	4
Shru	Beginning	1	1	1	0	0	0	0	0	0	0	3
Tashakkar	Thanks	1	1	0	0	0	0	0	0	0	0	2
Shamil	Inclusion	1	1	1	0	1	0	0	0	0	0	4
Qabu	Control	1	1	1	1	1	0	1	0	0	1	6
Kaarwai	Operation	1	1	1	0	0	0	0	0	0	0	3
Istamaal	Use	1	1	1	0	0	0	0	0	0	0	3
Koshish	Try	1	1	1	0	0	0	0	0	0	0	3
Anjaam	End	1	1	1	0	0	0	0	0	0	0	3

Mouqa	Chance	0	1	1	0	0	0	1	0	1	0	4
Irtikaab	Confession	1	1	0	0	0	0	0	0	0	0	2
Qadam	Step	0	0	1	0	1	0	0	0	0	0	2
Intizaar	Wait	1	0	1	0	0	0	0	0	0	0	2
Himayat	Support	1	0	0	0	0	0	0	0	0	0	1
		93	92	60	21	9	1	17	0	5	15	

For the clarity of the presented data, the columns of WordNet ID, and Class Code have been omitted in Table 4.6.

The data exhibits that the light verb 'Do/*Kar'* is the most compatible light verb with this noun class i.e., 94 compatible and collocated instances out of total nouns in this list of 105 which is 89 % probability of compatibility; the intransitive light verb 'Become/*Hua*' has nearly the same level of compatibility as the listed Noun Acts. 'Go/*Ja*' is proved to be the least compatible light verb with Noun Act in this list with only one instance found here.

If we look at the data horizontally in this Table 4.6, 'Courage/ *Hosla*' is the most prolific noun act which is compatible with 9 out of ten light verbs in the list. It is noted that the noun act which is the most compatible with the light verb 'Give/*Dia*' are also compatible with 'Take/*lee*' except for only two nouns i.e., 'Chance/*Mouqa*' and 'Advancement/*Taraqi*'. That is why I would refer to these light verbs as a set of ditransitive light verbs. Noun act also exhibits similar semantic and syntactic compatibility behavior with two intransitive light verbs 'Come/Aa' and 'Go/Ja'. Either Noun Act was compatible with both, if not it was not compatibility with both. It can be called as a pair/set of intransitive light verbs though some exceptions are always there as 'Work/*Kaam*' is compatible with 'Come/*Aa'* whereas it does not go with 'Go/*Ja'*(see Examples 49 and 50). '*Kaam Ana*' gives a reading of 'work' or 'help' as a verb.

49.

Larka logoon kay kaam aya

Boy-M.SG.NOM people-PL.GEN work-M.SG come-M.SG.PERF
'Boy helped people.'
50.
*Larka Logoon kay kaam gia
Boy- M.SG.NOM People- PL. GEN Work- M.SG go-M.SG.PERF
*'Boy went people work'.

This perceived pattern set up a set of inquiries to be examined for other noun classes as well emerged in this research work:

i. Are all nouns the most compatible with the transitive light verb 'Do/*Kar'*?
ii. Does a noun exhibit the same compatible behavior with both intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'*?
iii. Does a noun behave in the same way with two transitive light verbs 'Give/*Dia'* and 'Take/*Lee'*?
iv. What are characteristics of the noun which is the most prolific in terms of its compatibility with light verbs?

4.4 Noun Communication

Nouns categorized as Noun Communication refer to speaking or activities of different types of utterances.

4.4.1 Noun Communication: Semantic Orientation

This Table 4.7 lists the 65 unique Urdu Conjunctive Noun Communication. English transliteration of Urdu Noun Communication, English translation, mapped WordNet ID, noun class and noun class code are the constituent columns in this Table 4.7 to show the semantic orientation of Urdu conjunctive noun classes.

Table 4.7

Mapping of Urdu Conjunctive Noun Communication

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Mashwara	Advice	{06684229}	<noun.communication>	[10]
Mutalba	Demand	{07205939}	<noun.communication>	[10]
Zikar	Mention	{06778981}	<noun.communication>	[10]
Ilzaam	Accusation	{07248890}	<noun.communication>	[10]
Izhaar	Expression	{06893299}	<noun.communication>	[10]
Daaway	Claim	{06742613}	<noun.communication>	[10]
Raqam	Document	{07230743}	<noun.communication>	[10]
Hadayat/	Instruction	{06597067}	<noun.communication>	[10]
Sabaq	Lesson	{06685497}	<noun.communication>	[10]
Ahmiat	Emphasis	{07099567}	<noun.communication>	[10]
Larayai	Fight	{07199051}	<noun.communication>	[10]
Shahadat	Witness	{06747451}	<noun.communication>	[10]
Inkashaf	Disclosure	{07228055}	<noun.communication>	[10]
Baatcheet	Communication	{06262268}	<noun.communication>	[10]
Jawab	Answer	{07215187}	<noun.communication>	[10]
Appeal	Appeal	{07201488}	<noun.communication>	[10]
Ikhtilaaf	Disagreement	{07195447}	<noun.communication>	[10]
Ishtihaar	Advertisement	{07263469}	<noun.communication>	[10]
Itilaa	Information	{06646883}	<noun.communication>	[10]
Izhaar	Expression	{06893299}	<noun.communication>	[10]
Imtihan	Exam	{07211681}	<noun.communication>	[10]
Interview	Interview	{07210735}	<noun.communication>	[10]
Order	Order	{07183274}	<noun.communication>	[10]
Taqeed	Insistence	{07206437}	<noun.communication>	[10]
Peghaam	Message	{06263820}	<noun.communication>	[10]
Taeed	Agreement	{06783066}	<noun.communication>	[10]
Tuzleel	Insult	{06728162}	<noun.communication>	[10]

Taqaza	Demand	{07205939}	<noun.communication>	[10]
Tareef	Appreciation	{07243771}	<noun.communication>	[10]
Taqreer	Speech	{07253354}	<noun.communication>	[10]
Tafseer	Explanation	{06751030}	<noun.communication>	[10]
Tanqeed	Criticism	{06723485}	<noun.communication>	[10]
Tuhmat	Accusation	{07248890}	<noun.communication>	[10]
Telephone	Telephone	{06282943}	<noun.communication>	[10]
Khushamad	Flattery	{06708324}	<noun.communication>	[10]
Daad	Appreciation	{07243771}	<noun.communication>	[10]
Daryaft	Exploration	{00310889}	<noun.communication>	[10]
Daawat	Invitation	{07200808}	<noun.communication>	[10]
Shikayat	Complaint	{07223590}	<noun.communication>	[10]
Manzoor	Agreement	{06783066}	<noun.communication>	[10]
Nakal	Copy	{06517508}	<noun.communication>	[10]
Munsookh	Cancel	{06880268}	<noun.communication>	[10]
Maalomaat	Information	{06646883}	<noun.communication>	[10]
Muaafi	Apology	{06645870}	<noun.communication>	[10]
Muzamat	Condemnation	{06722631}	<noun.communication>	[10]
Mukhatib	Address	{06367922}	<noun.communication>	[10]
Mubarakbad	Greeting	{06642524}	<noun.communication>	[10]
Guzarish	Request	{06525357}	<noun.communication>	[10]
Kiadat	Guidance	{06663446}	<noun.communication>	[10]
Zahir	Expose	{07230228}	<noun.communication>	[10]
Shanakat	Identification	{06898133}	<noun.communication>	[10]
Sifarish	Reference	{06778981}	<noun.communication>	[10]
Samanah	Face	{06890628}	<noun.communication>	[10]
Record	Record	{06659969}	<noun.communication>	[10]
Report	Report	07233130}	<noun.communication>	[10]
Dhamki	Threat	{07240675}	<noun.communication>	[10]
Khermakdam	Welcome	{06643829}	<noun.communication>	[10]
Hukam	Order	{07183274}	<noun.communication>	[10]

Jawab	Answer	{06758700}	<noun.communication>	[10]
Telecast	Telecast	{06635102}	<noun.communication>	[10]
Tasdeeq	Attestation	{06749154}	<noun.communication>	[10]
Tarmeem	Amendment	{06736384}	<noun.communication>	[10]
Elaan	Announcement	{06738907}	<noun.communication>	[10]
Aitraaf	Acknowledgement	{06641368}	<noun.communication>	[10]
Appeal	Appeal	{07201488}	<noun.communication>	[10]

With 65 instances, noun communication is the second biggest Urdu Conjunctive noun class found here in the list of 280 Nouns.

4.4.2 Noun Communication: Compatibility with the Light Verbs

Noun communication as shown in Table 4.8 is also the most compatible with the transitive verb 'Do/*Kar*' with 53 instances of compatibility among a list of 65 nouns in this class. 'Become/*Hu*' is the second largest highly compatible light verb in this Urdu Conjunctive noun class with 51 instances.

Table 4.8

Compatibility Pattern of Noun Communication with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/ *aa*	Give /*Dia*	Go /*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Mashwara	Advice	1	0	1	0	0	0	1	0	0	1	4
Mutalba	Demand	1	1	1	0	0	0	0	0	0	0	3
Zikar	Mention	1	1	1	0	0	0	0	0	0	0	3
Ilzaam	Accusation	0	1	1	0	0	0	1	0	1	1	5

Izhaar	Expression	1	1	1	0	0	0	0	0	0	1	4
Daawa y	Claim	1	1	1	0	0	0	0	0	0	0	3
Raqam	Document	1	1	1	0	0	0	0	0	1	1	4
Haday at/	Instruction	1	1	1	0	0	0	1	0	0	1	5
Sabaq	Lesson	0	1	1	0	0	0	1	0	1	1	5
Ahmia t	Emphasis	1	1	1	0	0	0	1	0	0	1	5
Laraya i	Fight	1	1	1	0	0	0	0	0	0	1	4
Shaha dat	Witness	0	1	1	0	0	0	1	0	0	1	4
Inkash af	disclosure	1	1	1	0	0	0	0	0	0	0	3
Baatch eet	Communication	1	1	1	0	0	0	0	0	0	0	3
Jawab	Answer	0	1	1	0	0	1	1	1	0	1	6
Appea l	Appeal	1	1	1	0	0	0	0	0	0	0	3
Ikhtila af	Disagreement	1	1	1	1	1	0	0	0	0	0	5
Ishtiha ar	Advertisement	0	0	1	0	0	0	1	0	1	0	3
Itilaa	Information	1	1	1	0	0	1	1	1	0	0	6
Izhaar	Expression	1	1	1	0	0	0	0	0	0	0	3

Imtihan	Exam	0	0	0	0	0	0	1	0	0	1	2
Interview	Interview	1	0	0	0	0	0	1	0	0	1	3
order	order	1	1	1	0	0	1	1	1	0	1	7
Taqeed	insistence	1	1	0	0	0	0	0	0	0	0	2
Peghaam	message	0	0	1	0	0	1	1	1	0	1	5
Taeed	agreement	1	0	0	0	0	0	0	0	0	0	1
Tuzleel	insult	1	1	1	0	0	0	0	0	0	0	3
Taqaza	demand	1	0	1	0	0	0	0	0	0	0	2
Tareef	appreciation	1	1	1	0	0	0	0	0	0	0	3
Taqreer	speech	1	0	1	0	0	0	0	0	0	0	2
Tafseer	Explanation	1	0	1	0	0	0	0	0	0	0	2
Tanqeed	criticism	1	1	1	0	0	0	0	0	0	0	3
Tuhmat	accusation	1	0	0	0	0	0	0	0	1	0	2
Telephone	Telephone	1	0	1	0	1	1	0	1	0	0	5
Khushamad	flattery	1	0	0	0	0	0	0	0	0	0	1
Daad	appreciation	0	0	0	0	0	0	1	0	0	1	2
Daryaft	exploration	1	1	1	0	0	0	0	0	0	0	3

Daawat	invitation	1	0	1	0	1	0	1	0	0	0	4
Shikayat	complaint	1	1	1	0	0	0	0	0	1	0	4
Manzoor	agreement	1	1	1	0	0	0	0	0	0	0	3
Nakal	copy	1	1	1	0	0	0	0	0	0	0	3
Munsokh	Cancel	1	1	0	0	0	0	0	0	0	0	2
Maalomaat	information	1	1	1	0	0	0	1	0	0	1	5
Muaafi	apology	0	0	1	0	0	0	1	0	0	1	3
Muzamat	condemnation	1	1	0	0	0	0	0	0	0	0	2
Mukhatib	address	1	1	1	0	0	0	0	0	0	0	3
Mubarakbad	greeting	0	0	0	0	0	0	1	0	0	1	2
Guzarish	request	1	1	0	0	0	0	0	0	0	0	2
Kiadat	guidance	1	0	0	0	0	0	0	0	0	0	1
Zahir	expose	1	1	1	0	0	0	0	0	0	0	3
Shanakat	identification	1	1	1	0	0	0	1	0	0	1	5
Sifarish	reference	1	1	1	0	0	0	1	0	0	1	5
Samanah	face	1	1	1	0	0	0	0	0	0	0	3
Record	Record	1	1	1	0	1	0	0	0	0	0	4

Report	Report	1	1	1	0	0	0	1	0	0	1	5
Dhamki	threat	0	0	0	0	0	0	1	0	0	1	2
Khermakdam	welcome	1	1	1	0	0	0	0	0	0	0	3
Hukam	order	1	1	1	0	0	1	1	1	0	1	7
Jawab	answer	0	0	1	0	0	1	1	1	0	1	5
Telecast	Telecast	1	1	0	0	0	0	0	0	0	0	2
Tasdeeq	attestation	1	1	1	0	0	0	0	0	0	0	3
Tarmeem	amendment	1	1	1	0	0	0	0	0	0	0	3
Elaan	Announcement	1	1	1	0	0	0	0	0	0	0	3
Aitraaf	acknowledgement	1	1	1	0	0	0	0	0	0	0	3
Appeal	appeal	1	1	1	0	0	0	0	0	0	0	3
		53	46	51	1	4	7	24	7	6	24	222

All Noun Communication in the list behaved same with the intransitive light verbs 'Come/*Aa*' and 'Go/*Ja*'. For both 'aa' and 'ja,' past tense is used with the inflectional morphological forms 'ayi/aya' and 'gaya/gai'.

Light verbs 'Give/*Dia'* and 'Take/*Lena*' also exhibit almost the same compatibility with a noun with few exceptions like with the nouns such as 'Treat/*Dawat'*, 'information/*Itlaa'* and 'Advertisement/*Ishtihaar'*. Once it is compatible, it is compatible in past and present forms of tense (See the variants of Example 51)

51.

i) 'dawat dena'
Invitation-F.SG.NOM give-M.SG.INF

‘To invite’

ii) *‘dawat lena’

*Invitation- F.SG.NOM take- M.SG.INF

* Take an invitation

iii) Itlaa dena

Information-F.SG.NOM give- M.SG.INF

‘To inform’

iv) *Itlaa lena

Information-F.SG.NOM take- M.SG.INF

* To take Information

v) Ishtihaar dena

Advertisement-F.SG.NOM give-M.SG.INF

‘To advertise’

vi) *Ishtihaar lena

*Advertisement-F.SG.NOM Nom take-M.SG.INF

* ‘To take advertisement’

Amongst the most productive noun communications in this list two are ‘Answer/*Jawab*’ and ‘Information/*Itilaa*’ compatible with 6 light verbs. Whereas ‘Order/*Order*’ which is a borrowed noun got 7 compatibilities.

Transitive light verb ‘Do/*Kar’* and intransitive ‘Become/*Hu*’ are the most prolific light verbs with almost 80% compatibility with noun communication listed here.

4.5 Noun Cognition

Nouns listed as Noun Cognition refer to different mental activities.

4.5.1 Noun Cognition: Semantic Orientation

Urdu nouns mapped as Noun Cognition are lexical items which refer to cognitive activities. This Table 4.9 lists 29 unique entries of Urdu Conjunctive Noun Cognition. All unique semantic senses of nouns are mapped onto the semantic senses of nouns found in English WordNet. Their Unique WordNet IDs and Class Codes are also provided along with them for their further NLP use.

Table 4.9

Mapping of Urdu Conjunctive Noun Cognition

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Yaad	Memory	{05943778}	<noun.cognition>	[09]
Roshni	Light	{06220526}	<noun.cognition>	[09]
Faisala	Decision	{05846685}	<noun.cognition>	[09]
Aitimaad	Confidence	{05705328}	<noun.cognition>	[09]
Bardaasht	Tolerance	{06214141}	<noun.cognition>	[09]
Bhoolchook	Mistake	{05904330}	<noun.cognition>	[09]
Tajzia	Analysis	{05780664}	<noun.cognition>	[09]
Tasawur	Idealization	{05933040}	<noun.cognition>	[09]
Justajo	Quest	{05778388}	<noun.cognition>	[09]
Dihan	Attention	{05710222}	<noun.cognition>	[09]
Sazish	Plot	{05916276}	<noun.cognition>	[09]
Yaqeen	Trust	{05705941}	<noun.cognition>	[09]
Vaada	Promise	{05959600}	<noun.cognition>	[09]
Nazarandaz	Ignorance	{05997167}	<noun.cognition>	[09]
Munzoori	Acceptance	{06203462}	<noun.cognition>	[09]
Mushahida	Observation	{05711376}	<noun.cognition>	[09]
Mehsoos	Feeling	{05685184}	<noun.cognition>	[09]
Mussalat	Imposition	{05841695}	<noun.cognition>	[09]
Gour	Attention	{05710222}	<noun.cognition>	[09]
Mehsoos	Feeling	{05925333}	<noun.cognition>	[09]
Himayat	Favor	{06210352}	<noun.cognition>	[09]
Hissa	Part	{05937794}	<noun.cognition>	[09]
Tawaqa	Expectation	{05953807}	<noun.cognition>	[09]
Taleem	Education	{05993172}	<noun.cognition>	[09]
Tasawar	Imagination	{05633248}	<noun.cognition>	[09]
Tarjeeh	Preference	{06210079}	<noun.cognition>	[09]
Pasand	Preference	{06210079}	<noun.cognition>	[09]

Bardasht	Tolerance	{06214141}	<noun.cognition>	[09]
Andaaza	Estimate	{05811387}	<noun.cognition>	[09]

With 29 entries, Noun Cognition is the third biggest class found in Urdu N+V instances.

4.5.2 Noun Cognition: Compatibility with the Light Verbs

These 29 entries of unique Urdu Conjunctive Noun Cognition are checked for their compatibility with ten light verbs. The result of compatibilities is shown below in Table 4.10.

Where '0' indicates no probable combination and '1' shows the possibility of the occurrence of Noun Cognition with light verb.

Table 4.10

Compatibility Pattern of Noun Cognition with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *H*e	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Yaad	Memory	1	1	1	1	1	1	0	0	0	0	6
Roshni	Light	1	1	1	0	0	0	1	0	0	1	5
Faisala	Decision	1	1	1	0	0	0	1	0	0	1	6
Aitimaad	Confidence	1	0	1	0	1	0	1	0	0	0	4
Bardasht	Tolerance	1	1	1	0	0	0	0	0	0	0	3

Bhool chook	Mista ke	0	1	1	0	0	0	0	0	0	0	2
Tajzia	Analy sis	1	0	0	0	0	0	0	0	0	0	1
Tasaw ur	Idealiz ation	1	1	1	0	0	0	0	0	0	0	3
Justaj o	Quest	1	1	1	1	1	0	0	0	0	0	5
Dihan	Attent ion	1	1	0	0	1	0	1	0	0	0	4
Sazish	Plot	1	1	1	0	0	0	0	0	0	0	3
Yaqee n	Trust	1	1	1	1	1	1	0	0	0	0	6
Vaada	Promi se	1	1	1	1	0	0	0	0	0	1	6
Nazar andaz	Ignora nce	1	1	0	1	1	0	0	0	0	0	4
Munz oori	Accep tance	1	1	0	0	0	0	0	0	0	0	2
Mush ahida	Obser vation	1	1	1	0	0	0	0	0	0	0	3
Mehs oos	Feelin g	1	1	1	0	0	0	0	0	0	0	3
Mussa lat	Impos ition	1	1	0	0	0	0	0	0	0	0	2
Gour	Attent ion	1	0	0	0	0	0	0	0	0	0	2
Mehs oos	Feelin g	1	1	0	0	0	0	0	0	0	0	2
Hima yat	Favor	1	1	1	0	0	0	0	0	0	0	5
Hissa	Part	1	1	1	0	1	0	1	0	0	1	6

Tawaqa	Expectation	1	1	0	1	1	0	0	0	0	0	5
Taleem	Education	0	0	0	0	0	0	1	0	0	1	2
Tasawar	Imagination	1	1	1	1	1	0	0	0	0	0	6
Tarjeeh	Preference	0	0	0	0	0	0	1	0	0	0	1
Pasand	Preference	1	1	1	1	1	0	0	0	0	0	5
Bardasht	Tolerance	1	1	1	1	0	0	0	0	0	0	4
Andaaza	Estimate	1	1	0	1	1	0	0	0	0	0	4
		26	24	18	10	11	2	7	0	0	5	103

For noun cognition also ‘Do/*Kar’* is the most compatible transitive light verb with 26 Noun Cognitions. Intransitive ‘Become/*Hu’* is the next productive light verb compatible with 24 Noun Cognitions.

Noun Cognitions are the least compatible with intransitive light verbs ‘Come/*Aa’* and ‘Go/*Ja’*. Furthermore, ‘Memory/*Yaad’* and ‘Faith/*Yaqeen’* are only compatible with light verb ‘Come/*Aa’*.

The two transitive light verbs ‘Give/*Dia’* and ‘Take/*Lee’* also exhibited the different compatibility behavior with. ‘Give/*Dia’* showed 24 % compatibility with 7 noun cognitions whereas ‘Take/*Lee’* is 6 % compatible only with 2.

For Noun Cognitions in this list, 60% is the productivity level in terms of compatibility with 6 out of ten light verbs.

4.6 Noun Attribute

4.6.1 Noun Attribute: Semantic Orientation

Table 4.11 displays the entries of unique Urdu Conjunctive Noun Attribute. Table format is consistent with the English transliteration, English Translation, listing of mapped semantic senses, WordNet IDs and Class Code of Noun Attribute according to WordNet.

Table 4.11

Mapping of Urdu Conjunctive Noun Attribute

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Faraham	Availability	{04726275}	<noun.attribute>	[07]
Assar	Effect	{04682671}	<noun.attribute>	[07]
Mukhalfat	Disagreement	{04756982}	<noun.attribute>	[07]
Mushabahat	Resemblance	{04754685}	<noun.attribute>	[07]
Hosla	Courage	{04864314}	<noun.attribute>	[07]
Takabur	Arrogance	{04895107}	<noun.attribute>	[07]
Insaaf	Justice	{04857348}	<noun.attribute>	[07]
Samana	Face	{04687095}	<noun.attribute>	[07]
Rutba	Position	{05087073}	<noun.attribute>	[07]
Zehar	Poison	{05173924}	<noun.attribute>	[07]
Ahmiat	Importance	{05175788}	<noun.attribute>	[07]

4.6.2 Noun Attribute: Compatibility with the Light Verbs

Table 4.12 shows the complete combinatory probability of listed Noun Attributes with ten light verbs.

Table 4.12

Compatibility Pattern of Noun Attribute with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give /*Dia*	Go /*Ja*	Hit/ *Laga*	Take/ *Lee*	**N + V**

Faraham	Availability	1	1	1	1	0	0	0	0	0	0	4
Assar	Effect	1	1	1	1	1	0	0	0	0	1	7
Mukhalfat	Disagreement	1	1	0	0	0	0	0	0	0	0	2
Mushabahat	Resemblance	0	0	1	0	1	0	0	0	0	0	3
Hosla	Courage	1	1	0	0	1	0	1	0	0	0	5
Takabur	Arrogance	1	0	0	0	0	0	0	0	0	0	1
Insaaf	Justice	1	1	1	0	0	0	0	0	0	0	3
Samana	Face	1	1	1	0	0	0	0	0	0	0	3
Rutba	Position	0	0	0	0	1	0	1	0	0	0	3
Zehar	Poison	1	1	1	0	0	0	1	0	1	1	6
Ahmiat	Importance	0	1	1	1	1	0	1	0	0	0	5
		8	8	7	3	5	0	4	0	1	2	38

Both transitive light verb ‘Do/*Kar*’ and intransitive ‘Become/*Hu*’ showed 80 % compatibility with noun attribute in the list.

According to the data compiled, none of the Noun Attribute in the list is compatible with wo transitive light verbs “Give/*Dia*’ and ‘Take/*Lee*’.

The transitive light verb ‘Give/*Dia*’ showed 40 % compatibility whereas ‘Take/*Lee*’ exhibited only 20%.

‘Effect/*Assar*’ is an attribute noun which showed compatibility with 70 % of the light verbs included in the study. ‘Poison/*Zehar*’ is the next most amiable noun attribute with 60 % compatibility. It includes both literal and metaphoric meanings. *‘Zehar hona’* refers to ‘feeling very bad’ or ‘getting the feeling of extreme dislike’ which is its metaphoric interpretations.

4.7 Noun Artifact

Table 4.13 lists mapping of Urdu Noun Artifact to English WordNet. It is a small conjunctive Urdu noun class.

4.7.1 Noun Artifact: Semantic Orientation

Only three Urdu Conjunctive Noun Artifact are found in the inventory of Urdu N+V. Their semantic senses are figured out following the detailed procedure and the mapped onto the similar sense of nouns in English WordNet. WordNet IDs and Class Code of Noun Artifact can be found in the following Table 4.13.

Table 4.13

Mapping of Urdu Conjunctive Noun Artifact

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Daramad	Export	{03311152}	<noun.artifact>	[06]
Baramad	Import	{03569847}	<noun.artifact>	[06]
Goli	Bullet	{02919696}	<noun.artifact>	[06]

Noun Artifact is a small noun class found in the data of 280 Urdu N+V with only three entries.

4.7.2 Noun Artifact: Compatibility with the Light Verbs

Table 4.14

Compatibility Pattern of Noun Artifact with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/R*ah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Daramad	Export	1	1	1	0	0	0	0	0	0	0	3
Baramad	Import	1	1	1	0	0	0	0	0	0	0	3
Goli	Bullet	02	02	02	00	00	00	00	00	11	00	17

Noun Artifact (as shown in Table 4.14) continued to show the same trend of highest compatibility with transitive LV 'Do/*Kar'* and intransitive LV 'Become/*Hu*' with 66% compatibility with light verbs. None of the noun artifacts is compatible with the set of intransitive light verbs 'Come/*Aa*' and 'Go/*Ja*' or transitive 'Give/*Dia*' and 'Take/*Lee*'. Both noun artifact 'Export/*Daramad*' and 'Import/*Baramad*' showed the same compatibility behavior with the light verbs. Here, the antonym semantic relation between these two nouns cannot be ignored.

4.8 Noun Process

Procedural analysis of the senses of these nouns listed in Table 4.15 resulted in their mapping as Noun Process.

4.8.1 Noun Process: Semantic Orientation

Noun Process is also a small Urdu Conjunctive noun class like the previous Noun Artifact.

Table 4.15 files the Noun Process with their mapped semantic senses according to English WordNet.

Table 4.15

Mapping of Urdu Conjunctive Noun Process

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Khasara	Loss	{13530399}	<noun.process>	[22]
Tagdeed	Renewal	{13569128}	<noun.process>	[22]
janam	Birth	{13554089}	<noun.process>	[22]
Zaya	Wastage	{13595655}	<noun.process>	[22]

4.8.2 Noun Process: Compatibility with Light Verbs

Urdu nouns listed as Noun Process refer to some unique process going on because of ongoing activity. Such as Waste/*Zaya* as a result of Use/*Istemaal*, and Loss/*Khasara* as a result of some Business/*Karobaar*.

Table 4.16

Compatibility Pattern of Noun Process with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/R*ah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Khasara	Loss	0	1	1	0	0	0	0	0	0	0	2
Tagdeed	renewal	1	1	0	0	0	0	0	0	0	0	2
janam	birth	0	1	0	0	0	0	1	0	0	1	3
Zaya	wastage	1	1	1	0	0	0	0	0	0	0	3
		2	4	2	0	0	0	1	0	0	1	11

Noun process (as shown in Table 4.16) is the first in the list of noun classes which showed less compatibility with transitive light verb 'Do/*Kar'* as compared to 'Become/*Hu'*.

The intransitive light verb 'Become/*Hu'* is the most productive in terms of its co-occurrence with the entries in the list of noun process.

Both intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'* did not show any compatibility with any noun process found in the list. However, 'Birth/*Janam'* is the only noun process here which is compatible with both the transitive light verbs 'Give/*Dia'* and 'Take/*Lee'*.

4.9 Noun State

4.9.1 Noun State: Semantic Orientation

With 13 entries, this Noun State represents 4.6 % of Conjunctive Noun class taxonomy as shown in Table 4.17. The semantic senses of 13 distinct Noun States are mapped onto the matching semantic senses of nouns in English WordNet.

Table 4.17

Mapping of Urdu Conjunctive Noun State

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Inhasaar	Dependence	{14024833}	<noun.state>	[26]
Tayyari	Preparation	{14054590}	<noun.state>	[26]
Dakhil	Inclusion	{13959337}	<noun.state>	[26]
Kamyabi	Success	{14498478}	<noun.state>	[26]
Mehroomi	Deprivation	{14517010}	<noun.state>	[26]
Mustood	Blockage	{14531553}	<noun.state>	[26]
Mukhalafat	Disagreement	{14005842}	<noun.state>	[26]
Mukarar	Setting	{14536861}	<noun.state>	[26]
Rehmat	Blessing	{14497848}	<noun.state>	[26]
Tamasha	Scene	{14429707}	<noun.state>	[26]
Tahafuz	Security	{14562870}	<noun.state>	[26]
Itehaad	Union	{14441799}	<noun.state>	[26]
Zindagi	Life	{13984978}	<noun.state>	[26]

4.9.2 Noun State: Compatibility with the Light Verb

If seen horizontally in Table 4.18, maximum compatibility of Noun State with light verbs does not go beyond 4.

Table 4.18

Compatibility Pattern of Noun State with Light Verbs

Urdu Noun	English noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Inhasaar	Dependence	1	1	1	1	0	0	0	0	0	0	4
Tayyari	Preparation	1	1	1	0	0	0	0	0	0	0	3
Dakhil	Inclusion	1	1	1	1	0	0	0	0	0	0	4
Kamyabi	Success	0	1	1	0	0	0	1	0	0	0	3
Mehroom	Deprivation	1	1	0	1	1	0	0	0	0	0	4
Mustood	Blockage	1	1	0	0	0	0	0	0	0	0	2
Mukhalafat	Disagreement	1	1	0	0	0	0	0	0	0	0	2
Mukarar	Setting	1	1	1	1	0	0	0	0	0	0	4
Rehmat	Blessing	1	1	1	0	0	0	0	0	0	0	3
Tamasha	Scene	1	0	0	0	0	0	0	0	0	0	1
Tahafuz	Security	1	0	0	0	0	0	1	0	0	0	2
Itehaad	Union	0	1	1	0	0	0	0	0	0	0	2

Zinda gi	Life	0	0	0	0	0	0	1	0	0	0	3
		10	10	7	4	1	0	3	0	0	1	37

Noun State continued to show similar compatibility behavior like most of the noun classes selected for the study i.e., highest compatibility with the transitive light verb 'Do/*Kar'*. Neither of the two intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'* are compatible with any of the noun states mentioned in the list.

The compatible behavior of noun state is not same with two transitive light verbs 'Give/*Dia'* and 'Take/*Lee'*; 'Give/*Dia'* got two entries whereas 'Take/*Lee'* is here with only one instance with Noun State. Noun State showed compatibility with maximum 4 different light verbs.

4.10 Noun Feeling

4.10.1 Noun Feeling: Semantic Orientation

Nouns included in this class of Noun Feeling are abstract whose semantic senses are mapped to the related semantic senses of nouns in English WordNet.

Table 4.19

Mapping of Urdu Conjunctive Noun Feeling

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Sharam	Shame	{07521808}	<noun.feeling>	[12]
Zakham	Wound	{07512262}	<noun.feeling>	[12]
Uljhan	Irritation	{07533707}	<noun.feeling>	[12]
Umeed	Hope	{07526972}	<noun.feeling>	[12]
Arzoo	Wish	{07501369}	<noun.feeling>	[12]
Bhaichara	Brotherhood	{07515281}	<noun.feeling>	[12]
Taajub	Amazement	{07524811}	<noun.feeling>	[12]
Taqleef	Pain	{07509503}	<noun.feeling>	[12]
Touba	Attrition	{07550088}	<noun.feeling>	[12]

Herat	Amazement	{07524811}	<noun.feeling>	[12]
Rahat	Comfort	{07507656}	<noun.feeling>	[12]
Ishq	Love	{07558676}	<noun.feeling>	[12]
Ibadat	Worship	{07516659}	<noun.feeling>	[12]
Nafrat	Hate	{07561835}	<noun.feeling>	[12]
Muhabat	Love	{07558676}	<noun.feeling>	[12]
Mayoos	Disappointment	{07555990}	<noun.feeling>	[12]
Hassad	Envy	{07565182}	<noun.feeling>	[12]

4.10.2 Noun Feeling: Compatibility with the Light Verbs

Table 4.20

Compatibility Pattern of Noun Feeling with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Sharam	Shame	1	1	1	0	0	0	0	0	0	0	3
Zakham	Wound	1	1	1	1	0	0	1	0	1	0	6
Uljhan	Irritation	0	1	0	0	0	0	0	0	0	0	1
Umeed	Hope	1	1	1	0	1	0	0	0	0	0	4
Arzoo	wish	1	1	1	1	0	0	0	0	0	0	5
Bhaichara	brotherhood	0	0	1	0	0	0	0	0	0	0	1
Taajub	amazement	0	1	1	0	0	0	0	0	0	0	2

Taqleef	pain	1	1	1	1	0	0	1	0	0	0	5
Touba	Attrition	1	1	0	0	0	0	0	0	0	0	2
Herat	amazement	0	1	1	0	0	0	0	0	0	0	2
Rahat	comfort	0	1	0	0	0	0	0	0	0	0	1
Ishq	love	1	1	1	0	0	0	0	0	0	0	3
Ibadat	worship	1	0	0	0	0	0	0	0	0	0	1
Nafrat	hate	1	1	1	1	0	0	0	0	0	0	4
Muhabat	love	1	1	1	1	0	0	0	0	0	0	4
Mayoos	disappointment	1	1	1	1	0	0	0	0	0	0	5
Hassad	Envy	1	1	1	0	0	0	0	0	0	0	3
		12	15	13	6	1	0	2	0	1	0	50

Noun Feeling (as shown in Table 4.20) is the second in the list of noun classes which showed more compatibility with the intransitive light verb with ‘Become/*Hu*’ as compared to the transitive light verb ‘Do/*Kar*’.

Noun feeling exhibited the same behavior with both intransitive light verbs ‘Come/*Aa*’ and ‘Go/*Ja*’ with no instance of compatibility.

On the other hand, Noun Feeling showed more compatibility with ‘Give/*Dia*’ as compared to ‘Take/*Lee*’. ‘Wound/*Zakham*’ is a noun feeling in a metaphoric use which showed 6 instances of cooccurrences with different light verbs.

4.11 Noun Phenomenon

4.11.1 Noun Phenomenon: Semantic Orientation

Only one entry of Noun Phenomenon i.e., 'Rain/*Baarish'* is found in Urdu inventory N+V as shown in Table 4.21 given below.

Table 4.21

Mapping of Urdu Conjunctive Noun Phenomenon

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Barish	Rain	{11521799}	<noun.phenomenon>	[19]

4.11.2 Noun Phenomenon: Compatibility with the Light Verbs

Table 4.22

Compatibility Pattern of Noun Phenomenon with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give/ *Dia*	Go/ *Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Barish	Rain	1	1	1	1	0	1	0	0	0	0	5

Noun phenomenon has only one entry in the list of 280 nouns chosen for the study. It showed co-occurrence with five light verbs as depicted in Table 4.22.

4.12 Noun Time

4.12.1 Noun Time: Semantic Orientation

Referring to temporal features, Noun Time has got 5 unique entries in Urdu N+V inventory (see Table 4.23).

Table 4.23

Mapping of Urdu Conjunctive Noun Time

Urdu Noun	English Noun	WN ID	Noun Class	Noun Code	Class
Waqat	Time	{15295388}	<noun.time>	[28]	
Ikhtatam	Ending	{15291722}	<noun.time>	[28]	
Rukhsat	Leave	{15164090}	<noun.time>	[28]	
Multavi	Delay	{15297015}	<noun.time>	[28]	
Khatam	Ending	{15291722}	<noun.time>	[28]	

4.12.2 Noun Time: Compatibility with the Light Verbs

Table 4.24

Compatibility Pattern of Noun Time with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/ *rah*	Put/ Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/ *Laga*	Take/*Lee*	N + V
Waqat	Time	0	1	1	1	1	1	1	1	0	1	9
Ikhtatam	Ending	1	1	1	0	0	0	0	0	0	0	3
Rukhsat	Leave	1	1	1	0	0	0	1	0	0	1	5
Multavi	Delay	1	1	0	0	0	0	0	0	0	0	2
Khatam	Ending	1	1	1	0	0	0	0	0	0	0	3
		4	5	4	1	1	1	2	1	0	2	21

Noun Time (as represented in Table 4.24) showed almost the same behavior with the intransitive light verb 'Become/*Hu*' and the transitive verb 'Do/*Kar*' with 4 and 5 entries respectively. Only one Noun Time 'Time/*Waqat*' compatible with 'Come/*Aa*' is also compatible with 'Go/*Ja'* which marked the similar productive with this pair of intransitive light verbs.

'Time/*Waqat*', a noun time, showed 90% productivity with the chosen light verbs except the most otherwise prolific transitive light verb 'Do/*Kar*'.

4.13 Noun Event

4.13.1 Noun Event: Semantic Orientation

The nouns shown in Table 4.25 demonstrated some events whose semantic value is mapped after finding the nouns with the similar semantic value in English WordNet.

Table 4.25

Mapping of Urdu Conjunctive Noun Event

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Insidad	Eradication	{07345613}	<noun.event>	[11]
Nuqsaan	Loss	{07302177}	<noun.event>	[11]
Takkar	Collision	{07316161}	<noun.event>	[11]
Itifaaq	Coincidence	{07331599}	<noun.event>	[11]
Intiqal	Death	{07370091}	<noun.event>	[11]
Drama	Drama	{07305001}	<noun.event>	[11]
Junbish	Movement	{07324399}	<noun.event>	[11]
Mulaqat	Meeting	{07429522}	<noun.event>	[11]
Vafaat	Death	{07370091}	<noun.event>	[11]
Moohim	Movement	{07324399}	<noun.event>	[11]
Muntashir	Dispersion	{07346359}	<noun.event>	[11]
Mulaqat	Meeting	{07429522}	<noun.event>	[11]
Muqabala	Competition	{07470961}	<noun.event>	[11]
Gaib	Disappearance	{07350316}	<noun.event>	[11]

Shikist	Defeat	{07490504}	<noun.event>	[11]
Paida	Birth	{07334902}	<noun.event>	[11]

4.13.2 Noun Event: Compatibility with Light Verbs

Noun Event are quite prolific in terms of their compatibility with light verbs as shown in Table 4.26.

Table 4.26

Compatibility Pattern of Noun Event with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain /R*ah*	Put / Rakh	Come/*aa*	Give/ *Dia*	Go/ *Ja*	Hit / *Laga*	Take/ *Lee*	**N + V**
Insidad	Eradication	1	1	0	0	0	0	0	0	0	0	2
Nuqsan/	Loss	1	1	1	0	0	0	1	0	0	1	5
Takkar	Collision	1	1	0	0	0	0	1	1	1	1	6
Itifaaq	Coincidence	1	1	1	0	0	0	0	0	0	0	3
Intiqal	Death	1	1	0	0	0	0	0	0	0	0	2
Drama	drama	1	1	1	0	0	1	0	0	1	0	5
Junbish	movement	1	1	1	0	0	0	0	0	0	0	3
Mulaqat	meeting	1	1	1	0	0	0	0	0	0	0	3
Vafaat	death	0	1	0	0	0	0	0	0	0	0	1
Moohim	Movement	1	0	0	0	0	0	0	0	0	0	1

Muntashir	dispersion	1	1	1	1	0	0	0	0	0	0	4
Mulaqat	meeting	1	0	0	0	0	0	0	0	0	0	1
Muqabala	competition	1	1	1	0	0	0	0	0	0	0	3
Gaib	disappearance	1	1	1	1	1	0	0	0	0	0	5
Shikist	defeat	0	1	1	0	0	0	1	0	0	0	4
Paida	birth	1	1	0	0	0	0	0	0	0	0	2
		14	14	9	2	1	1	3	1	2	2	49

Noun Event exhibited same number of compatibility with transitive light verb 'Do/*Kar*' and transitive light verb 'Become/*Hu*' i.e., 86%.

Noun Event co-occurred with both intransitive light verbs 'Come/*Aa*' and 'Go/*Ja*' only once.

Except 'Defeat/*Shikist*', the two Noun Events showed similar compatibility with two transitive light verbs 'Give/*Dia'* and 'Take/*Lee'* with two entries each. 'Collision/*Takkar* 'is the most productive Noun Event compatible with 6 different light verbs.

4.14 Noun Group

4.14.1 Noun Group: Semantic Orientation

Semantic orientation of Noun Group is presented in Table 4.28 with their mapped WordNet IDs, Noun Class and Class Code.

Table 4.28

Mapping of Urdu Conjunctive Noun Group

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Raha	Free	{07964544}	<noun.group>	[14]
Partnership	Partnership	{08077530}	<noun.group>	[14]
Hajoom	Crowd	{08290764}	<noun.group>	[14]

4.14.2 Noun Group: Compatibility with the Light Verbs

Table 4.29

Compatibility Pattern of Noun Group with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become/*Hu*	Is/ *He*	Remain/R*ah*	Put/ Rakh	Come/*aa*	Give /*Dia*	Go /*Ja*	Hit/ *Laga*	Take/ *Lee*	N + V
Raha	Free	1	1	1	0	0	0	0	0	0	0	3
Partnership	Partnership	1	1	1	0	0	0	0	0	0	0	3
Hajoom	Crowd	1	1	1	0	0	0	0	0	0	0	3
		3	3	3	0	0	0	0	0	0	0	9

Noun Group is a small group of nouns in the data of 280 nouns. It showed equal productivity with transitive light verb 'Do/*Kar*' and the intransitive light verb 'Become/*Hu*'. Not a single entry in this list of noun group showed any compatibility with two intransitive light verbs: 'Come/*Aa*' and 'Go/*Ja*', and two ditransitive light verbs: 'Give/*Dia*' and 'Take/*Lee*'.

4.15 Noun Possession

5.14 4.15.1 Noun Possession: Semantic Orientation

The semantic orientation of Noun Possession is displayed in Table 4.29.

Table 4.29

Mapping of Urdu Conjunctive Noun Possession

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Tax	Tax	{13330021}	<noun.possession>	[21]

Zamanat	Bail	{13371991}	<noun.possession >	[21]
Mehfooz	Store	{13388243}	<noun.possession >	[21]
Qurbani	Sacrifice	{13349095}	<noun.possession >	[21]
Faida	Benefit	{13317922}	<noun.possession >	[21]

5.15 4.15.2 Noun Possession: Compatibility with Light Verbs

Noun possession also follows the same trends exhibited by previously mentioned Urdu Conjunctive noun classes in terms of their combinatory restrictions with light verbs as shown in Table 4.30 mentioned below.

Table 4.30

Compatibility Pattern of Noun Possession with Light Verbs

Urdu Noun	English Noun	Do/ Kar	Become /Hu	Is/ He	Remain/ Rah	Put / Rakh	Come /aa	Give/ Dia	Go/ Ja	Hit / Laga	Take/ Lee	N+V
Tax	Tax	1	1	1	0	0	1	1	1	1	1	6
Zamanat	bail	1	1	0	0	0	0	1	0	0	1	2
Mehfooz	store	1	1	1	1	1	0	0	0	0	0	5
Qurbani	sacrifice	1	1	0	0	0	0	1	0	0	1	2
Faida	benefit	1	1	1	0	0	0	1	0	0	1	6

These five instances of Noun Possession constitute only 1 % of the Urdu noun data (N=280). Both the transitive light verb ‘Do/*Kar*’ and the intransitive light verb ‘Become/*Hu*’ showed 80 % compatibility with noun possession. Noun Possession kept on following the previously captured pattern of exhibiting a similar productive nature with the set of two intransitive light verbs ‘Come/*Aa*’ and ‘Go/*Ja*’, and similar with two

ditransitive light verbs ‘Give/*Dia*’ and ‘Take/*Lee*’ with 1 and 4 entries respectively. A borrowed word from English ‘Tax/*Tax*’ is one of the most productive possession nouns with compatibility with 6 light verbs.

4.16 Noun Person

4.16.1 Noun Person: Semantic Orientation

Only one instance of Noun Person could be found in Urdu N+V instances as shown in Table 4.31. The lexical entity listed as Noun Person is someone who undergoes some sort of physical or mental misery or ill treatment by someone.

Table 4.31

Mapping of Urdu Conjunctive Noun Person

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Shikaar	Victim	{10771761}	<noun.person>	[18]

4.16.2 Noun Person: Compatibility with the Light Verbs

The only one Noun Person is not very productive in terms of its collocation with light verbs as depicted in Table 4.32.

Table 4.32

Compatibility Pattern of Noun Person with Light Verbs

Urdu Noun	English Noun	Do/Kar	Become/Hu	Is/He	Remain/Rah	Put/ Rakh	Come/aa	Give/Dia	Go/Ja	Hit/ Laga	Take/ Lee
Shikaar	Victim	1	1	1	0	0	0	0	0	0	0
		1	1	1	0	0	0	0	0	0	0

Only one instance of Noun Person makes it unique in the list of noun data. This Noun Person is only compatible with ‘Do/*Kar*’, ‘Become/*Hu*’, and ‘Is/*He*’.

4.17 Noun Relation

4.17.1 Noun Relation: Semantic Orientation

Like Noun Person, this Noun Relation has also only one entry in Urdu N+V inventory.

Table 4.33

Mapping of Urdu Conjunctive Noun Relation

Urdu Noun	English Noun	WN ID	Noun Class	Noun Class Code
Shareek	Inclusion	{13816246}	<noun.relation>	[24]

4.17.2 Noun Relation: Compatibility with the Light Verbs

Noun Relation though a very small group of Urdu Conjunctive Noun Classes, displayed the usual compatibility pattern as displayed by the previously mentioned noun classes (see Table 4.34).

Table 4.34

Compatibility Pattern of Noun Relation with Light Verbs

Urdu Noun	English Noun	Do/Kar	Become/*Hu*	Is/*He*	Remain/R*ah*	Put/Rakh	Come/*aa*	Give/*Dia*	Go/*Ja*	Hit/*Laga*	Take/*Lee*	N+V
Shareek	Inclusion	1	1	1	1	0	0	0	0	0	0	4
		1	1	1	1	0	0	0	0	0	0	4

Like Noun Person, only one instance of Noun Relation is found in the list of 280 Urdu nouns. It is neither compatible with 'Come/*Aa*' nor 'Go/*Ja*' which can be taken as similar behavior with this set of intransitive light verbs. Similarly, neither of the two transitive light verbs 'Give/*Dia*' and 'Take/*Lee*' are compatible with this noun relation.

4.18 Summary

This chapter provides a detailed account of findings based on the research objectives. The development of an extensive Urdu N+V inventory with 1076 instances marks the diverse combinations of Urdu nouns and light verbs and fulfils the task mentioned as first research objective. It includes the varied morphological inflections of light verbs with same nouns which makes it quite an elaborate inventory.

This inventory was crawled to draw a list of 280 unique nouns which is used to accomplish the second research objective of drawing the semantic classes of |Urdu nouns in conjunct predicates (N+V). The semantic orientation of theses nouns is elaborated, and their semantic features are mapped onto the nouns in English WordNet with the similar semantic values. It helped to identify their semantic classes and class codes of Urdu nouns in (N+V) conjunct predicates. Furthermore, the compatibility of these 280 unique nouns with ten light verbs presents detailed data to elicit significant semantic and syntactic combinatory restrictions between Urdu noun and light verbs in conjunct predicates. This detailed result helped to find the compatibility patterns of Urdu noun and light verbs.

The fourth research objective, development of a tool to identify true conjunct predicates in Urdu, is based on in-depth analysis of already found results in the present section of results. How that test is applied to different examples and in what conditions the host noun is incorporated is discussed in the next chapter (Discussion).

Chapter 5: Discussion

This chapter discusses the issues involved in the development of an inventory of Urdu conjunct predicates (N+V). It includes an investigation of the patterns emerged in the semantic classes of Urdu nouns which are compatible with different light verbs. This inquiry involves the exploration of possible sets of semantic and syntactic restrictions on the compatibility of Urdu nouns with light verbs. This section also discusses the issues involved in the identification of true conjunct predicates in Urdu.

The present research work is supported by the literature review related to conjunct predicate as a type of complex predicate. A survey of works related to conjunct predicate includes detailed introductory works on the history of research conducted on this topic. Along with a detailed account of complex predicates in different languages, it focuses on complex predicates in the Urdu language where two types of complex predicates i.e., Aspectual and Permissive (Butt, 1995). As this phenomenon of the complex predicate is also prevalent in other Indo-Aryan languages such as Hindi, Bangla, Marathi, Pushto, Balochi, Sindhi, etc. argument structure of complex(conjunct) predicate is also elaborated (Mohanan, 1994, 1997). The seminal works on complex predicates also include important information related to conjunct predicate as a subtype of complex predicate (Butt, 1995; Mohanan, 1994, 1997). Urdu complex predicates got the closer attention of the researchers due to their significance in computational work (Ahmed, 2010; Ahmed et al., 2012; Ahmed & Butt, 2011; Butt & Geuder, 2001; Butt & King, 2007; Butt & Ramchand, 2005).

To facilitate the Natural Language Processing Programs in the Urdu language, some Urdu corpora are used to develop lexical resources for conjunct predicates in the Urdu language (Abdullah et al., 2021; Ahmed et al., 2012; Bhat et al., 2017; De Marneffe et al., 2021; Saeed et al., 2019; Sinha, 2009). Development of an Urdu N+V inventory marks the fulfillment of first research product of the present research which can contribute to the industry by filling the previous gap of lack of lexical resource present in the Urdu language (Abdullah et al.,2021).

It is fully realized that not all N+V instances can be classified as conjunct predicate. Therefore, the phenomenon of identifying real conjunct predicate is also discussed in this work but could not be entertained to its fullest due to its alternative focus of development of N+V inventory, mapping of semantic class of Urdu nouns in N+V instances and semantic and syntactic compatibility patterns between the constituents of

conjunct predicates (Carnie, 2012; Bhattacharyya et al., 2007). Screening the real conjunct predicate through multiple constituency tests can be implemented in some upcoming advanced work using the Urdu N+V inventory as a resource. Developing an inventory of Urdu Noun+ Light Verb and exploring their semantic and syntactic compatibilities and restrictions aim at making digital linguistic resources for the Urdu language which can be used in different Natural Language Processing Programs.

Making Urdu a digitally enabled language has been an underlying goal of the present research work for which the mapping of Urdu conjunct predicate onto English WordNet is initiated (Chua & Kulathuramaiyer, 2004; Fellbaum, 1998; Meng et al., 2013). Somewhat similar process has already been initiated for different Indian languages (Bhattacharyya et al., 2006).

This section aims to analyze the findings presented in the previous chapter. The first finding of this research is an inventory of Urdu conjunct Predicate (N+V) developed from Universal Dependency Urdu Corpus. Discovery of semantic classes of nouns compatible with light verbs in Urdu conjunct predicate is the second breakthrough discovery of this research which yielded after the mapping of noun semantic senses onto the noun senses found in English WordNet. The third main research objective is to validate Levin's (1993) proposition of correlation between semanticity and syntactic context of a lexical item by examining the semantic and syntactic orientation of nouns and light verbs in Urdu N+V instances. Results of selective compatibility of Urdu nouns with different light verbs help in understanding their semantic and syntactic connections. The fourth and last objective of this research is to develop a tool for the identification of true conjunct predicates in the Urdu language. Not all instances of Urdu N+V may be a true conjunct predicate. Only an incorporated noun host will make N+V a single syntactic constituent. This incorporated status of Urdu nouns is examined using an Agreement Test. In case of no agreement between noun host and light verb, it cannot be established as an argument in the clauses but has to be an incorporated entity of predicate. Thus, such confirmed conjunct predicates can be inserted in Urdu WordNet in the form of a single entry. Inventory of Urdu N+V instances, semantic classes of Urdu nouns in N+V instances, the detailed account of semantic and syntactic combinatory restrictions between Urdu nouns and light verbs, and a tool to identify true Urdu conjunct predicates are language resources which may help to bridge the knowledge gap and help to solve the problem arising due to the complex semantic and

syntactic nature of Urdu conjunct predicates (N+V). The developed lexical resources, knowledge, and a devised tool may augment the Natural Language Processing programs and other Urdu computational applications.

I started with the manual crawling of the instances of Urdu N+V instances for my pilot study which grew laborious with the large size of corpora I selected for my study (Bhat et al., 2017; *Urdu WordNet 1.0 Wordlist*, 2013; Urooj et al., 2012). So, an automatic and computational method was employed to pull the instances of N+ V from former i.e., Universal Dependency Urdu Tree bank. At the same time, some Urdu nouns are manually selected from the list of words used to develop Urdu WordNet. The number of Urdu nouns in the two selected corpora (Bhat et al., 2017; *Urdu WordNet 1.0 Wordlist*, 2013) was quite large, the task was to come up with a list of unique nouns which collocate with the selected light verbs. Due to the scarcity of Urdu N+V instances, the intuition of native of the Urdu language is employed to find the probability of any possible collocations of nouns with the light verbs used in this study. At the same time, Nouns extracted from Urdu WordNet Wordlist are also subjected to native intuition of Urdu speakers to come up with possible compatibilities especially with the under studied light verbs.

5.1 Inventory of Urdu Conjunct Predicate (N+V)

The Urdu language has got this feature of complex predicates present in its different forms i.e., compound complex V_1+V_2, Conjunct Predicate (N+V)/Adj+V). All these forms present numerous questions related to their semantic and syntactic interpretations which are provided with different explanations in various studies (Alsina, 1997; Alsina, 1993; Butt, 1995; Mohanan, 1994, 1997).

Being conjunct predicate (N+V) the focus of the present study, it needed a detailed account of Urdu noun combinations with light verbs. Keeping in mind that N+V is a very productive instance as noun is an open class word category and many inflections are also available for light verbs, it remained a difficult task to achieve (Ahmed & Butt, 2011). Not enough data was available to anticipate the possible instances of Urdu N+V. Without this lexical resource, further research on Urdu conjunct predicate was not quite achievable. A good sizeable inventory was necessary to fill this knowledge gap and provide me with enough linguistic ground to find semantic classes of Urdu nouns compatible with light verbs, and to further investigate the semantic and syntactic combinatory restrictions between Urdu nouns and light verbs.

Developing an inventory of Urdu N+V instances could have been more laborious if an annotated Urdu corpus was not available. Selection of Universal Dependency Urdu Corpus resolved this issue (Bhat et al., 2017). Roughly 1090 instances of N+V are found in this corpus which are cleaned for any entry of Adj+V and some others which are not N+V. Same noun has repeated instances with different morphological inflectional forms of the same light verb. Such entries may be used in some future study to understand the behavior of light verbs, auxiliaries and modals. This list is subjected to manual cleaning for any probable entry of Adj +V and other non N+V entry.

This inventory is not limited to only ten light verbs i.e., 'Do/*Kar',* 'Become*/Hu*', 'Be/*He',* 'Put*/Rakh',* 'Come*/Aa*', 'Give'/*Dia,* 'Go/*Ja*', 'Take*/Lena*', 'Remain*/Rah',* and 'Hit */Laga'* selected to study the semantic classes of Urdu nouns in N+V instances, and semantic and syntactic combinatory restrictions between Urdu nouns and light verbs, but it includes all the verbs naturally occurring with Urdu nouns in the Universal Dependency Urdu Corpus. The characteristics of this lexical inventory improve its efficacy in a number of ways. Firstly, the inclusion of all naturally occurring light verbs enhances the scope of this Urdu N+V inventory for any future usage as a good lexical resource. Secondly, this linguistic resource may provide researchers with a model of different morphological forms of compatible light verbs. Thirdly, as taken from a naturally occurring data, this inventory provides a record of foreign words such as 'Minister', 'Telecast', 'Time', 'Award', 'Appeal', 'Out' and 'Report' which are borrowed from English language. Instances of nouns with light verbs provide data to study the use of foreign words as part of complex/conjunct predicates in the Urdu language. Furthermore, this inventory includes singular and plural forms as separate entries:

- Use /*Istemal* and Uses/*Istemalaat*
- Power/*Ikhtiyar* and Powers/*Ikhtiyaraat*

This feature provides a venue to study the behavior of Urdu nouns in singular and plural forms with light verbs and laid the basis of further exploration of such issues related to conjunct predicate.

As an initial step, such collocations are searched in Universal Dependency Urdu Treebank Corpus (Bhat et al., 2017). The initial product of this research is an inventory of Urdu N+V instances which presents instances of Urdu N+V (Abdullah et al., 2021). This inventory gives an overview of semantic and syntactic restrictions imposed on the

compatible collocations of Urdu nouns and light verbs. Not all Urdu nouns were compatible with the light verbs used in the study. This distinctive behavior of noun and light verbs further triggered the quest for an underlying pattern. Urdu nouns are extracted from the Urdu WordNet List which is already a collection of lexical entities mined from Urdu Digest Corpus (Urooj et al., 2012, 2014). Including this local Pakistani Urdu corpus was an attempt to encompass a dialectal variation of Urdu into the study of Urdu conjunct predicates.

5.2 Semantic Classes of Nouns in Urdu Conjunct Predicate (N+V)

An inventory mentioned in the previous section contains 1076 instances of Urdu N+V, but some of the Urdu nouns have redundant entries due to different reasons. Firstly, some nouns have multiple entries because of their singular and plural forms. Secondly, the same nouns are listed multiple times due to their collocations with light verbs with different inflectional morphological forms.

To find the semantic classes of nouns in the Urdu conjunct predicate (N+V), an adequate list of unique Urdu nouns compatible with different light verbs was required. To fulfill this requirement, unique nouns are extracted from the developed inventory of Urdu conjunct predicate (N+V) which is mentioned in the previous section (See Table 7). To further enrich this list, Urdu nouns are handpicked from Urdu WordNet Wordlist (2013). The addition of more nouns to the list not only enriches the noun list with unique entries but also brings dialectal diversity to the study. In total, it makes a list of 280 unique Urdu nouns which are further checked for their collocation with ten light verbs. Including all naturally occurring light verbs was beyond the scope of work, that is why the task is limited to only ten light verbs i.e., Do/*Kar'*, 'Become/*Hu*', 'Be/*He'*, 'Put/*Rakh'*, 'Come/*Aa*', 'Give'/*Dia*, 'Go/*Ja*', 'Take/*Lena*', 'Remain/*Rah'*, and 'Hit */Laga'*.

'Be/*He'* may also be an inflectional variation of the verb 'Become/*Hu'*, but it mostly serves as a copula verb.

To understand the semantic classes of Urdu nouns, their semantic senses are mapped to the senses of nouns found in English WordNet (Princeton University, 2022). Mapping of noun sense and selection of noun semantic class undergo a systematic and objective process. This process does not only require bilingual competence along with some reliable dictionary, but an Urdu native speaker intuitive consultancy is also employed. An online version of WordNet (http://wordnetweb.princeton.edu/perl/webwn) is used

to explore the noun senses it presents. It gives the semantic senses, WordNet ID, example sentence, semantic class and its class code.

Meaning of Urdu noun is navigated keeping in mind the resultant meaning because of its collocation with light verb. This compositional meaning is usually different from the individual meaning of Urdu noun and light verb. For instance, 'Gift */Inaam'* if searched in isolation will give the following three meaning as a noun in English WordNet:

1. something acquired without compensation
2. natural abilities or qualities
3. the act of giving/receiving

Only the third semantic sense is equivalent to 'to give a gift /*Inaam dena'.*

Its noun class is Noun Act. Therefore, it is picked and mapped to the Urdu noun *'Inaam'*. Similar process of detailed investigation is undertaken to sort out the mapping of noun senses and selecting their noun classes.

For this purpose, the 25 noun classes introduced in the English WordNet are used as a guideline. English translation of 280 Urdu nouns and then further navigation among the multiple semantic senses of that translation was used to further decide on their noun classes. A spectrum of Urdu conjunctive nouns was found where only 15 classes of noun were present with different representation ratios. The absence of 11 noun classes from the spectrum clearly implies their non-compatibility with light verbs in the Urdu language.

Before the detailed interpretation of the percentage of different noun classes found compatible with light verbs in the form of conjunct predicate, let me classify the 25 classes of Urdu nouns into two broad categories: Conjunctive and Non-conjunctive Urdu noun classes (See Table 5.1). Conjunctive Urdu noun classes refer to the semantic classes of Urdu nouns which are compatible with different light verbs to form conjunct predicate. On the other hand, non-conjunctive Urdu noun classes are semantic classes of nouns which do not form conjunct predicate due to their non-compatibility with light verbs. [-RF] refers to low referential force whereas, [+RF] denotes high referential force.

Table 5.1

Semantic Inventory for Noun Classes in Urdu Conjunct Predicate (N+V)

Conjunctive Urdu Noun Class [-RF]	Non-conjunctive Urdu Noun Class [+RF]
Noun Act	Noun Tops
Noun Communication	Noun Body
Noun Attribute	Noun Animal
Noun Time	Noun Food
Noun State	Noun Location
Noun Possession	Noun Object
Noun Relation	Noun Group
Noun Process	Noun Plant
Noun Phenomenon	Noun Quantity
Noun Person	Noun Shape
Noun Artifact	Noun Substance
Noun Feeling	
Noun Event	
Noun Group	
Noun Cognition	

The names of these categories in Table 5.1 are self-explanatory which is the initial motivation behind creating these categories. This categorization also signals that Urdu conjunctive noun classes refer to Urdu nouns which are semantically and syntactically compatible with light verbs. Noun Top is the unique beginner for nouns which is the most generic and semantically void class (Miller et al., 1990). Before any semantic

categorization, all lexical items are considered in Noun Tops. It is designated with vague abstraction which generates the further categorization of lexical items. If hyponymy hierarchy is generated between nouns, noun tops will be the most general and the topmost entity. No noun tops are found with any of light verbs in the corpus. Being the most general in the semantic hierarchy, Noun tops are void of any specific semantic feature and only refer to some semantic neutral concept, and its immediate hyponyms may be an 'Idea' or a 'Thing'.

This phenomenon may be interpreted in such a way that Light verbs which are already semantically bleached (Butt, 1995; Mohanan, 1994) may not form a conjunct predicate with a semantically void noun class. There are studies which even call light verbs 'relatively meaningless' (Brinton, 2008). Noun Tops may be a prototype of meaningless nouns.

So, it can safely be asserted that two semantically void constituents may not formulate a conjunct predicate. Technically, it may be declared as the first theory formulated in this research thesis. A similar investigation can be done to find the same phenomenon in compound predicate where a light verb is preceded by a main verb. Other non-conjunctive noun classes include concrete nouns including two classes denoting living things e.g., Noun Animal, Noun Plant. They neither collocated with the light verb as crawled in Urdu corpora nor seemed acceptable to native language users based on their intuition. On the other hand, all conjunctive noun classes are abstract except Noun Person (See Table 4.31) and Noun Artifact (See Table 4.13) with only 1 and 3 instances respectively. Only one Noun Person found 'Victim/*Shikaar'* refers to a state which can be undergone by a person or any living being. The 99.07 % of the conjunctive Urdu noun classes constitute abstract entities of the world; it highlights a distinctive abstract semantic trait of Urdu nouns compatible with light verbs. In other words, it may be asserted that Urdu nouns denoting physical entities are found to be less compatible with light verbs to form a conjunct predicate. To comment on the semantic classes of Urdu nouns in conjunct predicates, based on the data I analyzed it can easily be asserted that the dominant number of Urdu nouns compatible with light verbs in N+V instances are abstract nouns. Similar traits can be tested for validity of this notion in a cross-linguistic study.

If interpreted in terms of Referential Force (RF) introduced by Simone and Masini (2007), all conjunctive Urdu noun classes are low on their [-RF] because they do not

refer to countable, concrete, ostensible entities. The degree of strong semantic sense of noun is referred to as 'Nouniness' (Simone & Masini, 2007). Like light verbs, the term 'light noun' is put forward on the scale of Referential Force to understand the semantic contribution of nouns. It has become very easy to interpret the nature of these conjunctive noun classes which behave as Light nouns due to their 'low nouniness scale'. So, it would be easy to say by using a term coined by Lyons (1977) that these [-RF] conjunctive noun classes are not the 'first order' nouns. It was also mentioned that in syntactic contexts, these light nouns do not involve pure referential actions (Simone & Masini, 2007). When a light noun [-RF] comes with a light verb, it tends to incorporate in verb instead of performing pure referential nominal task. This debate can be handled in a different way. Here it is referred to understand the semanticity of Urdu nouns especially when they are divided into two categories: conjunctive and non-conjunctive depending on the focus of the work.

All non-conjunctive noun classes except Noun Tops have [+RF] which denotes their strong referential ability.

So, if I try to use this concept of referential ability of the nouns in Urdu conjunct predicates, it would be rephrased as follows: Conjunctive noun classes found compatible with light verb in Urdu conjunct predicates are low on their Referential Power and can be treated as Light nouns due to their state of not being first order nouns. This proposition supports the theories propagated by Lyons (1977) and Simone and Masini (2007).

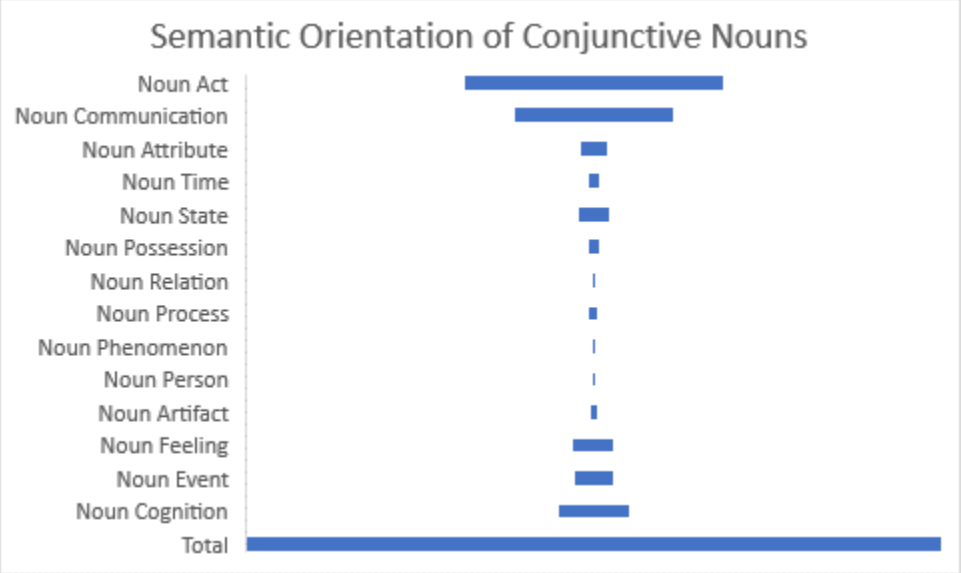
If we look at the percentage of each conjunctive noun class in the semantic inventory of Urdu conjunct predicate (See Figure 10), Noun Act is at the top of the list with 37.5 % representation. This higher representation of Noun Act may be interpreted in terms of its dominant semantic elements to denote action. Noun Communication is the second biggest class of conjunctive nouns which also conveys some semantic ability 'to communicate' as an action. The same phenomenon of referential power is termed as 'Intended Function' in a prominent work on artifact noun (Grimm, 2012; Grimm & Levin, 2017).

This finding relates to the second research objective of investigating the semantic classes of nouns in N+V based on their syntactic and semantic compatibility with light verbs. It also portrays the semantic composition spectrum of the conjunctive Urdu noun classes. Varying representation of conjunctive Urdu noun classes also reflects the semantic and syntactic makeup of Urdu conjunct predicate. For instance, NLP

applications can be provided with the data presented in Figure 5.1 about the probability of the semantic orientation of Urdu nouns in N+V collocations.

Figure 5.1

Semantic Orientation of Conjunctive Urdu Noun Classes



The data presented in Figure 5.1 clearly represent the compositional design of conjunctive Urdu noun classes.

The selection of semantic sense from the available listed semantic senses in WordNet was quite intricate, especially for the polysemous nouns. Here, the methodology of picking the semantic sense keeping in mind with the focus of its probable compatibility with the light verb played an important part in reaching the intended semantic sense which led to the selection of Urdu noun class. So, the categorization of conjunctive Urdu noun classes was not a straightforward selection process, it involved the intellectual and analytical contribution of native language user intuition. Use of native language user intuition cannot be undermined in theoretical linguistics based on its advocacy by pioneer linguists (Andow, 2015; Chomsky, 1966, 1992). Native speaker intuition is the decisive factor which helps in selection of a definite semantic sense of a noun with multiple semantic senses especially with the same semantic noun class. Following example may elaborate this intricate phenomenon of semantic sense selection:

The noun 'Fight/*Larayi*' has got five semantic senses in English WordNet (http://wordnetweb.princeton.edu/).

1. a hostile meeting of opposing military forces in the course of a war
2. the act of fighting; any contest or struggle
3. an aggressive willingness to compete
4. an intense verbal dispute
5. a boxing or wrestling match

The first two semantic senses belong to the same semantic class 'Noun Act' whereas the third and fourth are 'Noun Attribute' and 'Noun Communication' respectively. Fifth semantic sense is also Noun Act. The selection of matching noun class was not random but based on the native language user's intuition which led to the decision of picking the 4th semantic sense with semantic class 'Noun Communication' as the resultant mapped noun class.

Another polysemous Urdu noun 'Witness/*Shahadat*' has got five different semantic senses. First four of them belong to Noun Person and only the fifth one is Noun Communication:

1. someone who sees an event and reports what happened
2. a close observer; someone who looks at something (such as an exhibition of some kind
3. (law) a person who attests to the genuineness of a document or signature by adding their own signature
4. (law) a person who testifies under oath in a court of law
5. testimony by word or deed to your religious faith

Based on the composite meaning of Urdu N+V, the fifth semantic sense with semantic class 'Noun Communication' is selected and mapped.

The same motivation is employed to select the semantic sense and semantic class of all 280 nouns in the list. All this helped to construct the present spectrum of conjunctive noun classes (See Figure 5.1).

The noun class 'Noun Communication' presented an interesting case of Urdu noun 'Poison/*Zehar*'. When used with a light verb, 'Poison did/*Zehar kia'* it gave a distinct metaphoric meaning of 'spoiling and messing up' in Urdu which is distant from the literal meaning of its constituents i.e., noun and light verb. When mapped and translated into English, not only the literal meaning but its metaphoric meaning is also considered.

Urdu Noun 'Wound/*Zakham*' also presented an interesting case here. 'Wound give/*Zakham dena'* is used in metaphoric sense of 'inflicting mental pain' and its semantic class is 'Noun Feeling'; whereas 'Wound put/*Zakham lagna'* gives the literal meaning where it is 'Noun Act'. With some relevant examples, light verbs play an important part in associating meaning to a complex predicate and its subtypes i.e., conjunct Predicate (N/Adj +V). A study in future may be conducted on the metaphoric interpretation of conjunct predicate constructions. It may be another direction for conducting a future study to find interesting facts.

Urdu words borrowed from English language revealed some interesting information.

- Publishing
- Drama
- Training
- Transfer
- Tax
- Telephone
- Firing
- Record
- Report
- Telecast
- Postmortem
- Boycott
- Appeal
- Operation
- Out

Above mentioned Urdu nouns are some borrowed words which paved their way to the list of conjunctive Urdu nouns compatible with different light verbs. Most of them belong to 'Noun Act' when considered in collocation with a light verb in Urdu. At the same time, their semantic and syntactic compatibility was also high i.e., average compatibility was with 50 % of the light verbs in the list. Borrowing and then adaptation of these open class English words in Urdu interpreted their linguistic adaptation to the target language semanticity and syntactic context while maintaining their original

semantic orientation. All the above borrowed nouns except 'Tax' are compatible with transitive light verb 'Do/*Kar*' which result in the same transitive conjunct predicate which requires ergative case marker 'Ne' in perfective aspect to the subject in the sentence and contributes to its agentivity. Volitionality is an added semantic information which is attached to the willingness of the subject argument to perform the action with the conscious choice. Agentivity in Urdu is not bound to ergative case marker 'Ne' as 'Ne' is a result of perfective aspect and not of agentivity.

In addition to finding the borrowed words in Urdu conjunct predicates, mentioned below are the Urdu noun compounds found among the Urdu N+V instances:

- Brotherhood/ *bhai-chara*
- Mistake/*bhool-chook*
- Search/ *pooch-gach*
- Portrayal/ *tasweer-kashi*
- Attention/*qoro-fiqar*
- Fire/*nazar-e-atish*
- Planning/*munsooba-bandi*
- Flower/*gulpashi*

These Urdu compound nouns belong to two Conjunctive Urdu Noun Classes i.e., Noun Act and Noun Cognition. Furthermore, three types of Urdu compound nouns are found: noun + noun (*qoro-fiqar)*; noun + verb (*munsooba-bandi*); and noun followed by semantically void lexical item (*pooch-gach).* Some Urdu compound nouns have already some verbal components attached to them with the meaning of some action such as -*bandi*, *-pashi*, *-kashi*. These noun+verb Urdu compound nouns further cojoin with light verbs to form a conjunct predicate. These constructions can be represented by the expression: N_{n+v} +V. All the above compound nouns are compatible with light verb 'Do/*Kar*' which link (but not always) the Agentive semantic role of the subject argument with the licensing of ergative case marker 'Ne'. And all the resultant verbs (conjunct predicate) are transitive as well. This is so because the verb 'Do/*Kar'* is transitive and as a light verb, it does not lose this feature.

Urdu noun synonyms are intentionally avoided in the list of 280 Urdu N+V instances just to increase the uniqueness of this linguistic resource. When the semantic and syntactic behavior of nouns in semantic relation such as synonymy, antonymy and hyponymy is investigated, they tend to appear in similar syntactic contexts which

validates Levin's (1993) proposition of semantic and syntactic relatedness. This phenomenon can be elaborated by looking at the examples mentioned below:

- Advice/(*talqeen* and *mashwara)*
- Investigation/ (*pooch-gach* and *janch-partaal*)
- Movement/ (*junbish* and *harqat)*
- Amazement /(*taajub* and *herat*)
- Plot/ (*saaz baaz* and *sazish)*

Synonyms are kept in parenthesis. These synonyms exhibit nearly the similar semantic and syntactic compatibility with light verbs. It is also interpreted in the form of their similar case marking and semantic roles of subject arguments in the clause. Syntactic correlatedness in the case of verbs with similar meaning has already been explored by Levin (1993). Now Urdu synonyms nouns also exhibited the correaltedness between semantic values and their syntatctic behavior in a clause. This specific information can further be explored in detail for Urdu noun in different semantic relations which may also include antonymy, hyponymy, etc. The exhibition of similar syntactic behavior of nouns in some semantic relation lays the foundation for future work on the exploration of semantic and syntactic correlation of words in the semantic relations.

5.3 Semantic and Syntactic Connection between Nouns and Light Verbs in Urdu Conjunct Predicates (N+V)

Alongside commenting on the semantic orientation of compatible light verbs in Urdu conjunct predicate, it is very important to highlight significant syntactic features related to them such as transitivity.

Transitivity is the ability of a verb to project its argument/s to complete the proposition expressed by means of a sentence. 'How many arguments it will require to complete the semantic sense of the sentence' is a piece of information which is required to investigate the correlation between the semantic and syntactic context of a predicate (Levin, 1993).

When mapped to the English WordNet, all four Urdu 'intransitive' light verb found to be categorized in the 'Verb Stative' class (See Table 4.4a). These are verbs of being and having. On the other hand, the two ditransitive light verbs 'take' and 'give' belong to 'Verb Possession' (See Table 4.4b). The remaining 4 light verbs are transitive light verbs belonging to 3 different verb classes: 'Verb Contact', 'Verb Creation', and 'Verb

Change' (See Table 4.4b). Semantic classes of light verbs carry information which correlates with the semanticity of compatible conjunctive Urdu noun classes. As reflected in data, 'Noun Act' is the most compatible with transitive light verb 'Do/*Kar*'. Means there is something inherent in the argument structure of 'Do/*Kar*' which increases its compatibility with 'Noun Act'. Conjunct predicate carries the same argument structure as that of its cojoining constituent i.e., light verb. Nouns have referents but no argument. Mostly the verbal constituent of the conjunct predicate can influence argument structure though the nominal contribution to the argument structure is also discussed by Mohanan (1994) but it could not be found in case of Urdu N+V in the present study yet. 'Do/*Kar*' licenses ergative case marker 'Ne' is also linked to semanticity of Urdu noun 'Memory/*Yaad'* which increases the agentivity of the subject argument in the sentence. So, it can safely be propagated that the semantic and syntactic compatibility or combinatory restrictions between Urdu nouns and light verbs carry information about the overall semantic and syntactic context of the clause. There is enough semantic information mutually carried by noun and light verb when they cojoin to form a complex constituent i.e., conjunct predicate here.

5.4 Noun Act

This section corresponds to the semantic orientation, syntactic compatibility of Urdu Noun Acts mentioned in the previous chapter of results in Table 11 and 12 with light verbs, and their impact on the case marking and semantic role of the subject argument in the sentence.

5.4.1 Noun Act: Semantic Orientation

Lexical items listed in Noun Act may refer to names of some physical actions such as 'Destruction/*Tabah'*, 'Experiment/*Tajarba',* and 'Advancement/*Taraqi',* etc. which mostly involve physical actions. 'Encouragement/*Hosla*' and 'Work/*Kaam*' are Noun Acts which are highly compatible with seven out of ten light verbs included in the study except 'Hit/*Laga'*, 'Go/*Ja'* and 'Come/*Aa'*. Noun Act is the most compatible noun class with Urdu transitive light verb 'Do/*Kar*' which carries a semantic feature of executing some sort of action. So, the semantic properties of both noun and light verb complement each other in this formation of N_{act} +LV_{kar} conjunct predicate. Compatibility of Noun Act is also high with an intransitive light verb 'Become/*Hu*'. However, in the reading of sentences with an intransitive light verb 'Become/*Hu*', the

element of volitionality is not present. 'Become/*Hu'* is a stative verb and volitionality is associated with stative verbs.

5.4.2 Noun Act: Compatibility with the Light Verbs

Similar semantic and syntactic compatibility behavior of ditransitive light verbs 'Give/*Dia*' and 'Take/*lee*' is seen with Noun Act along with few exceptions. For instance, the Urdu nouns 'Chance/*Mouqa*' and 'Advancement/*Taraqi*' are compatible with 'Give/*Dia*' and not with 'Take/*Lee*'.

52.

- Mouqa Dena
  To give a chance
- *Mouqa lena
  Chance take
- *Taraqi Dena*
  To promote
- **Taraqi Lena*
  Promotion/advancement take

Mouqa lena or *Taraqi lena* does not seem appropriate or acceptable to a native Urdu speaker.

Compatibility of noun with 'Give/*Dia*' reflects the semantic feature involved in increased matter of conscious choice. So, the light verbs 'Do/*Kar*' and 'Give/*Dia*' reflect agentive semantic role of the subject in the sentence.

A set of intransitive light verbs, 'Come/*Aa'* and 'Go/*Ja*', exhibited the same uniform semantic and syntactic behavior with only one exception found in the list of Urdu Noun Act i.e., 'Work/*Kaam*'.

This pattern of intransitive light verbs, 'Come/*Aa*' and 'Go/*Ja*', and ditransitive light verbs 'Give/*Dia*' and 'Take/*Lee*', is taken as a pilot finding which is further tested for other conjunctive Urdu noun classes included in the study.

So, this study proceeded with the notion of testing the following hypotheses:

i. Conjunctive noun classes are more compatible with transitive light verb i.e., 'Do/*Kar'*.

ii. Compatibility with transitive light verb 'Do/*Kar*' adds to the agentivity of the subject and consequently increases the volitionality of action in the perfective form of sentence mostly with animate subjects.

iii. 'Become/*Hu*' is the intransitive light verb compatible with conjunctive noun classes though it does not involve any volitionality of the action.

iv. 'Come/*Aa*' and 'Go/*Ja*' constitute a pair of intransitive light verbs with nearly uniform semantic and syntactic compatibility with conjunctive nouns with few exceptions.

v. 'Give/*Dia*' and 'Take/*Lee*' form a pair of ditransitive light verb which show similar semantic and syntactic compatibilities with conjunctive Urdu noun classes with a few exceptions. And those few exceptions are due to the strong semantics sense of the conjunctive noun. The semantic sense of some conjunctive nouns calls for increased agentive mode of the action characterized with the element of volitionality which correlates their compatibility with only 'Give/*Dia*'.

5.4.3 Noun Act: Case Marking

Noun Act when comes with a light verb correlates with assigning three case markers to the subject argument in the sentence: ergative, dative and nominative. Conjunctive Urdu Noun Act class when collocated with four light verbs ('Do/*Kar*', 'Put/ *Rakh*', 'Give/*Dia*' and 'Take/*Lee*'), it licenses 'ergative case marker' of the subject. It reflects the agentivity of the subject which can also directly be interpreted as the presence of volitionality semantic element found in the sentence. It shows the willingness of the subject to execute the action. The correlation marker with the conscious choice has also been discussed in a number of related studies on argument in the complex predicate (Butt, 1995; Mohanan, 1994; Pandharipande, 1990). Butt (1995), however, raised the point of not-so-conscious choice with every instance of ergative case marker. This instance can be viewed in the following example 53 and 54:

53.

Sardi ne fasal tabah kar di

Winter-F.SG.ERG crop.F.SG.NOM destroy-NOM do-F.SG.PERF

'Winter spoiled the crop.'

54.

Baarish ne maza kirkara kar dia

Rain-F.SG.ERG joy.M.SG.NOM Spoil.M.SG.NOM do-F.SG.PERF
'Rain spoiled the fun.'

So, it can be asserted that volitionality correlates with ergative case markers on animate subjects.

Example 54 has been sourced from Kachru (1988) to elaborate the phenomenon of ergativity on inanimate subjective and their link with the feature of volitionlaity.

Nominative case marked the subject argument in N_{act} +V with all three types of light verbs: intransitive, transitive and ditransitive. Two transitive light verbs, 'Remain/ *Rah*', and 'Hit/*Laga*', license dative case markers on subjects in the sentence.

Subject arguments of N_{act} +$V_{intrans}$ in Urdu are marked with nominative case.

Subject arguments of N_{act} +V_{trans} overtly marked with ergative when sentence is realized in perfective aspect. Transitive light verbs like 'Become/*hona'* and 'Put/*lagna'* license overt dative case *'Ko'* on subject argument.

Accusative 'Ko' appears on Direct Objects (DO) under conditions of specific references. The Indirect Object (IO) of N_{act} +$V_{ditrans}$ ditransitive light verb is marked with dative 'Ko'.

5.4.4 Noun Act: Argument Structure

There is almost an equal probability of Urdu conjunct predicate productivity in terms of collocation of Noun Act with transitive and intransitive light verb as there are approximately equal number of instances of N+V recorded both with transitive and intransitive light verbs i.e., 'Do/*Kar'* and 'Become/*Hua'* respectively. 155 instances of N_{act} +$V_{intrans}$ were found; whereas, 163 instances were of N_{act} +V_{trans}. The presence of ergative and dative case correlates with the argument structure of transitive and ditransitive light verbs respectively. The four intransitive light verbs such as 'Come/*Aa*', 'Go/*Ja*', 'Become/*Hu*', and 'Is/ *He*' are usually marked with nominative case.

The light verbs, 'Do/*Kar*', 'Put/ *Rakh*', 'Give/ *Dia*', and 'Take/ *Lee*' license ergative case marker 'Ne' on the subject argument in the clause. And the rest of the transitive light verbs in the list such as 'Remain/ *Rah*', and 'Hit/ *Laga*' allow dative case '*Ko*'.

The following set of hypotheses is formed which can be tested for other conjunctive noun classes:

Here N_{act} denotes a Noun Act which is a conjunctive Urdu noun class, $V_{intrans}$ is an intransitive light verb , and V_{trans} is a transitive light verb.

i. There are nearly equal instances of N_{act} +$V_{intrans}$ and N_{act} +V_{trans} 155 and 163 respectively

ii. N_{act}+$V_{intrans}$ are expected to mark the subject argument with nominative case marker

iii. N_{act} +V_{trans} have two categories: first one ('Do/*Kar*', 'Put/ *Rakh*', 'Give/ *Dia*', and 'Take/ *Lee*') has ergative case marker 'Ne' when the verb is realized in perfective aspect, and the second category ('Remain/ *Rah*', and 'Hit/ *Laga*') is marked with the dative case marker '*Ko*' in Urdu conjunct predicate.

5.4.5 Noun Act: Semantic Roles

Two subcategories of N_{act}+V_{trans} have been discovered (see Table 5.2) based on the case markers they allow and semantic roles they assign to the subject argument: Agentive Ergative Nominative (AEN) and Experiencer Dative (ED). Dative can also be associated with recipient/beneficiary role. The one with the ergative case marker gives a reading of agentive subject argument in the sentence; the other with dative case marker gives a reading of an experiencer subject in the sentence. The nominative also correlates with light verbs to assign an agentive role to the subject when the sentence is in an imperfective aspect.

Table 5.2

Noun Act and Transitive Light Verbs

	Urdu Conjunct Predicate with Di/Transitive Light Verbs	
	Category A AEN	Category B ED
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF) Nominative (-PERF)	Dative
N_{act}+V_{trans}	N_{act} + ('Do/*Kar*') N_{act} + ('Put/ *Rakh*')	N_{act} + ('Remain/ R*ah*') N_{act} + ('Hit/ *Laga*')
N_{act}+ $V_{ditrans}$	N_{act} + ('Give/ *Dia*') N_{act} + ('Take/ *Lee*')	

5.5 Noun Communication

This section refers to the semantic orientation, syntactic compatibility with light verbs, case marking, argument structure and Semantic roles of Noun Communications presented in the previous chapter of results in Table 4.7 and 4.8.

5.5.1 Noun Communication: Semantic Orientation

The word choices picked as Noun Communication mostly represent actions which fulfil some communicative purposes. Noun Communication 'Order/*Order*' which is a borrowed word from English got the highest number of compatibilities with the seven light verbs in the list (see Table 4.8). If we observe the behavior of different light verbs with Noun Communication, the transitive light verb 'Do/*Kar'* is the most compatible with 53 instances of N+V. Similarly, collocation with intransitive light verb 'Become/*Hu'* and copula 'Is/ *He'* is also high with 46 and 51 instances, respectively. As in Noun Act, similar semantic information can be retrieved here in case of Noun Communication with light verb 'Become/*Hu'* which does not contain a feature of conscious choice.

5.5.2 Noun Communication: Compatibility with Light Verbs

The semantic and syntactic behavior of Noun Communication with the light verbs and other information such as formulaic syntactic behavior with the sets of different transitive and intransitive light verbs.

To increase the cohesion and cohesiveness of the work, findings are interpreted as given bellow:

i. Noun Communication has the highest compatibility instances with transitive light verb 'Do/*Kar'*.
ii. Collocation of Noun Communication with transitive light verb 'Do/*Kar'* adds to the semantic element of conscious choice of subject argument to the reading of the sentence.
iii. Amongst the intransitive light verbs, 'Become/hu' is the most productive stative verb; however, it does not add to the volitionality of the action.
iv. A pair of intransitive light verbs, 'Come/*Aa'* and 'Go/*Ja'*, constitutes the same semantic and syntactic behavior with seven Noun Communication lexical items without any exception.

v. 'Give/*Dia'* and 'Take/*Lee'* which form a pair of ditransitive light verbs also exhibited similar compatibility with 24 Noun Communication in the list. As it is also observed in Noun Act, an agentive semantic role with a feature of a conscious choice is found in the reading of such N+V collocations.

5.5.3 Noun Communication: Case Marking

Noun Communication has a semantic value when collocates with light verb assigns ergative, dative, genitive, locative and nominative case markers on its arguments. Noun Communication when collocated with transitive light verb 'Do/*Kar'* can mark the subject argument with both ergative and nominative cases with perfective and imperfective aspect respectively as shown in the examples 55 and 56 below:

55.

Kisaan ne mutalba kia
Farmer-M.SG.ERG Demand-M.SG.NOM Do-M.SG.PERF
'The farmer demanded.'

56.

Kisaan mutalba krta hay
Farmer-M.SG.NOM demand-M.Sg.NOM Do-M.SG.IMPERF is.SG.COP
'Farmer demands.'

Ditransitive light verbs 'Put/*Rakh'*, 'Give/*Dia'* and 'Take/*Lee'* when collocates with Noun Communication mark subject with the ergative case in the sentence (see Examples 57 and 58).

57.

Kisaan ne barish ka record rakha
Farmer-M.SG.ERG rain-F.SG.GEN record-M.SG.NOM put-M.SG.PERF
'Farmer kept the record of rain.'

58.

Ustaad ne parhai ka mashwara dia
Teacher-M/F.SG.ERG Study-F.SG.GEN advice-M.SG.NOM give.M.SG.PERF
'Teacher advised to study.'

59.

Tareekhdaan ne vaqia raqam kia

Historian-M/F.SG.ERG incident-M.SG.NOM document-M.G.NOM do-M.SG.PERF

'Historian documented the incident.'

A set of intransitive light verbs, 'Come/*Aa'* and 'Go/*Ja'*, and 'Become/*Hu'* collocating with Noun Communication license the nominative case marker.

Two transitive light verbs, 'Remain/*Rah'*, and 'Hit/*Laga'*, have low collocation probability with Noun Communication with only 1 and 5 Noun Communications respectively. In these collocations, light verbs license ergative, locative, and genitive I.e *'Ne', 'Per', 'Se'* and *'Ki'* case marker respectively (see Examples 60-66).

60.

Main ne aap se Ikhtilaaf Rakha

I-M/F.SG.ERG you-PL.LOC disagreement-M.SG.NOM put-M.SG.PERF

' I disagreed with you.'

61.

Mazdoor per Ilzaam laga

Laborer-M/F.SG.LOC accusation-M.SG.NOM put-M.SG.PERF

' Laborer got accused.'

62.

Ghar per Raqam lagi

House-M.SG-LOC money-F.SG.NOM put-F.SG.PERF

'Money was spent on house.'

63.

Devaar per Ishtahaar laga

Wall-F.SG.LOC advertisement-M.SG.NOM put-M.SG.PERF

'Advertisement is being posted on wall.'

64.

Aadmi per Tuhmat lagi

Man-M.SG.LOC accusation-F.SG.NOM put-F.SG.PERF

'Man got accused.'

66.

Bachay ki Shikayat lagi

Child-M/F.SG.GEN complaint-F.SG.NOM put-F.SG.PERF

'Child got complained.'

5.5.4 Noun Communication: Argument Structure

Similar productive instances of N_{com} +V in terms of collocation of noun communication are found with 111 and 112 collocations with intransitive and di/transitive light verbs respectively. The transitive light verbs collocation with Noun Communication marked the presence of ergative, dative, and genitive case markers on the subject argument in the sentences. Collocation of Noun Communication with intransitive light verb allows nominative case markers to the subject noun phrase in the construction.

Following are the findings for Noun Communication:

i. There are equal instances of $N_{com}+V_{intrans}$ and $N_{com}+V_{trans}$.

ii. $N_{com}+V_{intrans}$ mark the subject with nominative case marker.

iii. $N_{com}+V_{trans}$ have two categories: In category A, two transitive light verbs, 'Remain/ *Rah'*, and 'Hit/ *Laga'* have low compatibility with light verbs which license genitive case marker '*Ka'*. One instance of $N_{com}+V_{trans}$ '*Tuhmat Lagna'* licenses locative case marker '*Per'* as well. and other category B of transitive light verbs 'Do/*Kar'*, 'Put/ *Rakh'*, 'Give/ *Dia'*, and 'Take/ *Lee'* assign ergative case on subject with perfective aspect realization of the sentence; whereas in case of imperfective aspect, nominative case marks the subject argument.

5.5.5 Noun Communication: Semantic Roles

$N_{com}+V_{trans}$ allows Ergative '*Ne'*, Genitive '*Ka'* and Locative case markers '*Per'* on subject noun phrases. The one with ergative gives a reading of agentive role for subject in the sentence; the other one with genitive and Locative case marker gives a reading of an experiencer with subject in the sentence.

Table 5.3

Noun Communication and Transitive Light Verbs

	Urdu Conjunct Predicate with Transitive Light Verbs	
	Category A	Category B
	AEN	EGL
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF)	Genitive +Locative

	Nominative (-PERF)	
N_{com}+V_{trans}	N_{com} + (‘Do/*Kar’*)	N_{com}+ (‘Remain/ *Rah’*)
	N_{com} + (‘Put/ *Rakh’*)	N_{com} + (‘Hit/ *Laga’*)
	N_{com} + (‘Give/ *Dia’*)	
N_{com}+$V_{ditrans}$	N_{com} + (‘Take/ *Lee’*)	

5.6 Noun Cognition

This section contains information related to the semantic orientation, syntactic compatibility with light verbs, case marking, argument structure and semantic roles of subject arguments related to Noun Cognition listed in the previous chapter in Table 4.9 and 4.10.

5.6.1 Noun Cognition: Semantic Orientation

Noun Cognition refers to mental activity which is endured by an agent at the cognitive level. This semantic class of noun is quite prolific in terms of collocation with light verbs as many Noun Cognitions show 6 instances of N+V. Like previous two conjunctive noun classes i.e., Noun Act and Noun Communication, Noun Cognition showed the highest compatibility with the transitive light verb ‘Do/*Kar’.* It can be interpreted as a high probability of collocation due to the semanticity of Noun Cognition and the syntactic context of transitive light verb. It may also be interpreted as linked with an agentive thematic role of the subject. Following the previous trends of conjunctive noun classes, Noun Cognition also shows high compatibility with the intransitive light verbs ‘Become/*Hu’.* High compatibility with ‘Become/*Hu’* is void of any conscious choice, and it gives the experiencer role to the subject.

5.6.2 Noun Cognition: Compatibility with Light Verbs

Keeping in mind the prevalent semantic and syntactic behavior of already discussed conjunctive noun classes i.e., Noun Act and Noun Communication, it can be helpful in formulating a set of semantic and syntactic combinatory restrictions for Urdu conjunct predicate.

i. High compatibility of Noun Cognition with transitive light verb 'Do/*Kar'* is also recorded which is in accordance with the pattern of formerly discussed conjunctive classes.

ii. Even in the case of this Noun Cognition which denotes a mental activity, compatibility with 'Do/*Kar'* reflected the meaning of conscious choice in performing the cognitive action.

iii. The same high collocation of Noun Cognition with intransitive light verb 'Become/*Hu'* without an element of conscious choice is observed; however, it adds to the experiencer semantic role of the subject in the clause.

iv. Noun Cognition does not behave similarly in case of compatibility with a pair of intransitive light verbs 'Come/aa' and 'Go/Ja'. There is no compatibility found between Noun Cognition and intransitive light verb 'Go/Ja'; whereas only two instances of Noun Cognition are found with 'Come/aa'.

v. Noun Cognition also behaved almost dissimilar in terms of compatibility with a set of ditransitive light verbs i.e., 'Give/*Dia'* and 'Take/*Lee'* except only two common entries listed below in different variants of example 67.

67.

- Roshni Dena
  To give light
- Roshni Lena
  To take light
- Faisala Dena
  To give a decision
- Faisala Lena
  To take a decision

5.6.3 Noun Cognition: Case Marking

Noun Cognition has semanticity which allows ergative, dative, genitive and nominative case on different arguments in the clause i.e., subject, direct object and indirect object, in the sentence (see examples 68-70).

68.

Bachay ne maan ko yaad kia
Child-M/F.SG.ERG mother-F.SG.ACC mis-.M.SG.NOM do-M.SG.PERF
'Child missed the mother.'

69.

Bachay ko maan ki yaad ayi
Child-M/F.SG.DAT mother-F.SG.ACC mis-M.SG.NOM come-F.SG.PERF
'Child recalled his/her mother.'

70.

Bacha sabaq yaad karta he
Child-M/F.SG.NOM lesson-M.SG.NOM do-M.SG.IMP
'Child memorized the lesson.'

Noun Cognition with transitive light verbs i.e., 'Do/*Kar'*, 'Put/*Rakh'*, and ditransitive light verbs i.e., 'Give/*Dia'* and 'Take/*Lee'* are marked with ergative and nominative case markers in case of perfective and imperfective aspects respectively. On the other hand, the intransitive light verbs, Become/*Hu'* and 'Come/*Aa'* allow dative case marker. The copula 'Is/*He'* marks the syntactic context with genitive and accusative cases. Here with Noun Cognition, a transitive light verb 'Hit/ *Laga'* and an intransitive light verb 'Go/ *Ja'* did not show any collocation with Noun Cognition (See Table 4.10).

5.6.4 Noun Cognition: Argument Structure

A previously perceived pattern of similar number conjunctive Urdu noun class instances with both intransitive and transitive light verbs slightly deviate in case of Noun Cognition. There are 44 instances of $N_{cog}+V_{intrans}$; whereas there are 59 instances of $N_{cog}+V_{trans.}$ No instance of Noun Cognition is found with light verbs 'Go/ *Ja'* and 'Hit/ *Laga'*.

Findings are interpreted in a schematic way as follows:

i. Pattern of equal instances of $N_{cog}+V_{intrans}$ and $N_{cog}+V_{trans}$ is slightly deviated
ii. $N_{cog}+V_{intrans}$ are marked with dative case marker on the subject argument
iii. In category A of transitive light verbs, there are 10 instances of 'Remain/*Rah'*; on the other hand, no instance of Noun Cognition is found with 'Hit/ *Laga'*. However, the category B of transitive light verbs i.e., 'Do/*Kar'*, 'Put/ *Rakh'*, and ditransitive 'Give/ *Dia'*, and 'Take/ *Lee'* permit ergative and nominative case markers on the subject with perfective and imperfective aspect in the clause.

5.6.5 Noun Cognition: Semantic Roles

N_{cog}+ V_{trans} allows Ergative case '*Ne'* on subject which can be interpreted as the agentive semantic/thematic role of the subject (See Table 5.4).

Table 5.4

Noun Cognition and Transitive light Verbs

	Urdu Conjunct Predicate with Transitive Light Verbs	
	Category A AE	Category B ED
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF) Nominative (-PERF)	Dative
N_{cog}+V_{trans}	N_{cog}+ ('Do/*Kar'*) N_{cog} + ('Put/ *Rakh'*)	N_{cog}+ ('Remain/*Rah'*) N_{cog}+ ('Hit/ *Laga'*)
N_{cog}+$V_{ditrans}$	N_{cog} + ('Give/ *Dia'*) N_{cog} + ('Take/ *Lee'*)	

5.7 Noun Attribute

This segment gives correlating information such as semantic orientation, syntactic compatibility between Noun Attribute and light verbs, case marking, argument structure and semantic roles influenced by N_{att} +V listed in the previous chapter (see Tables 4.12 and 4.13.

5.7.1 Noun Attribute: Semantic Orientation

A list of Noun Attribute (see Table 4.12) has the semantic feature related to propositions about abstract concepts. These Noun Attributes refer to some ideas and concepts in the minds of people about the entities of the world. This conjunctive Urdu noun class, Noun Attribute, constitutes only 4% of the total conjunctive nouns found in the list of 280. 'Effect/*Assar'* and 'Poison/*Zehar'* are the two most productive Noun Attributes which are compatible with 6 and 7 light verbs respectively. 'Poison/*Zehar'* has a metaphoric reading as well in Urdu which could not be found in WordNet. Here, it would be interesting to notice that 'Poison/*Zehar'* when collocated with 'Kar/*Do'*,

'Become/*Hu'* and copula 'Is/ *He'* gives a metaphoric reading. However, when its instances are found with ditransitive light verbs 'Give/ *Dia'* and 'Take/ *Lee'* it carries literal meaning. So, I would suggest 'Poison/*Zehar' be* treated as Noun Substance in this case. And Noun Substance is a class for which I did not find any collocation of N+V as Urdu conjunct predicate. It forfeits the notion of not finding Urdu conjunct predicate with nouns referring to tangible things.

Entries listed in the list of Noun Attribute show equal collocation with transitive light verb 'Kar/*Do'* and intransitive light verb 'Become/*Hu'*. When comes with 'Kar/*Do'* assigns agentive semantic role to the subject along with the conscious choice; whereas, the 'Become/*Hu'* gives an experiencer reading to the subject.

5.7.2 Noun Attribute: Compatibility with Light Verbs

i. Following the previously found trends in the present study, Noun Attribute is also the most productive conjunctive Urdu noun class with the transitive light verb 'Do/*Kar'*

ii. High productivity with this transitive light verb exhibits the conscious choice of the subject undertaking the action contributing to some Noun Attribute

iii. Noun Attribute also reflects high collocation with the intransitive light verb 'Become/*Hu'*, but here the semantic feature of conscious effort cannot be seen in the reading of the subject argument in the sentence

iv. As far as the similar semantic and syntactic behavior with a set of ditransitive light verbs i.e., 'Give/*Dia'* and 'Take/*Lee'*, it cannot be seen except for the entries with 'Poison/*Zehar*' which is taken in literal meanings of Noun Substance. In that case, it can be expelled from the list of Noun Attribute.

v. With a set of intransitive light verbs, 'Come/*Aa'* and 'Go/*Ja'*, Noun Attribute showed similar 'no entry' of collocation.

5.7.3 Noun Attribute: Case Marking

Noun Attribute carries a semanticity which permits ergative, dative, and genitive case markers on the noun phrases in the clause in case of its compatibility with different light verbs (see Examples 71-73).

71.

Nurse ne paani faraham kia

Nurse-M/F.SG.ERG water-M.SG.NOM provide-M.SG do-M.SG.PERF

'Nurse provided the water.'

72.

Mareez ko assar hua

Patient-M/F.SG.DAT effect-M.SG Become-M.SG.PERF

'Patient got affected/improved.'

73.

Ummedwaar ki Mukhalfat hui

Candidate-M/F.SG.GEN Opposition-F.SG.NOM become-F.SG.PERF

'Candidate got opposed.'

A detailed syntactic study of Urdu conjunct predicate (N_{att} +V) shows that it allows two case makers on the subject argument when collocated with an intransitive light verb 'Become/*Hu'*. With the transitive light verb 'Remain/*Rah'*, Noun Attribute allows genitive; whereas another transitive light verb 'Hit/*Laga'* it draws dative case marker. Rest of the transitive light verbs in the list i.e., 'Do/*Kar'*, 'Put/*Rakh'*, and ditransitive light verbs i.e., 'Give/*Dia'*, and 'Take/*Lee'*, allow ergative case marker with perfective aspect in the sentence (See Example 74).

74.

Maan ne bachay ko hosla dia

Mother-F.SG.ERG child-M.SG.ACC encouragement-M.SG.NOM give-M.SG.PERF

' Mother encouraged the child.'

With imperfective aspect in the construction, 'Do/*Kar'*, 'Put/ *Rakh',* 'Give/ *Dia'*, and 'Take/ *Lee'* assign nominative case to the subject argument (See Example 75).

75.

Badsha insaaf karta he

King-M.SG.NOM justice-M.SG.NOM do-M.SG.IMP

'King (always) does justice.'

A set of intransitive light verbs i.e., 'Come/*Aa'* and 'Go/*Ja'*, did not show any collocation with Noun Attribute so their corresponding case marking could not be decided yet.

5.7.4 Noun Attribute: Argument Structure

In terms of slightly different number of instances with transitive and intransitive light verb, Noun Attribute exhibited a deviant behavior like that of Noun Cognition. There are 23 instances of $N_{att}+V_{tran}$; on the other hand, 15 instances are of $N_{att}+V_{intrans.}$ The difference was created specifically due to no instance of Noun Attribute with 'Come/*Aa'* and 'Go/*Ja'* which are intransitive light verbs.

i. A pattern of slightly different number of compatibilities of $N_{att}+V_{tran}$ and $N_{att}+V_{intrans}$ has been found

ii. $N_{att}+V_{intrans}$ are marked with dative case e.g., 'Become/*Hu'* also allows dative case with the subject.

iii. Two categories have been found among transitive light verbs in terms of their case marking: category (A) i.e., 'Do/*Kar'*, 'Put/ *Rakh'*, 'Give/ *Dia'*, and 'Take/ *Lee'*, allows ergative case markers with perfective aspect in the clause; whereas, with imperfective clause, it assigns nominative case marker to the subject. Category B i.e., 'Remain/ *Rah'* and 'Hit/ *Laga'*, allows genitive and dative respectively.

5.7.5 Noun Attribute: Semantic Roles

A pattern of thematic roles with two categories of conjunct predicate ($N_{att}+V_{trans}$) with transitive light verbs correlates with the case markers on the subject argument. Like Noun Cognition, two categories of conjunct predicate $N_{att}+V_{trans}$ based on different case markers and semantic roles of the subject arguments (See Table 5.5).

The subject argument takes an agentive role when marked with ergative and nominative cases in the presence of perfective and imperfective aspects in the sentences respectively. The subject takes a semantic role of experiencer when marked with the dative case.

Table 5.5

Noun Attribute and Transitive Light Verbs

Urdu Conjunct Predicate with Transitive Light Verbs	
Category A	Category B

	AEN	ED
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF)	Dative
	Nominative (-PERF)	
$N_{att}+V_{trans}$	N_{att}+ (‘Do/*Kar’*)	N_{att}+ (‘Hit/ *Laga’*)
	N_{att}+ (‘Put/ *Rakh’*)	N_{att}+ (‘Remain/ *Rah’*)
$N_{att}+V_{ditrans}$	N_{att}+(‘Give/ *Dia’*)	
	N_{att}+ (‘Take/ *Lee’*)	

5.8 Noun Artifact

This section aims to comment on semantic orientation, syntactic compatibility pattern, case marking, argument structure and pattern of thematic roles with N_{art}+V instances in the Urdu language mentioned in chapter 4 (See Table 4.13 and 4.14).

5.8.1 Noun Artifact: Semantic Orientation

Noun Artifact is a small conjunctive noun class with a limited (1%) representation in the list of 280 conjunctive nouns (See Table 4.13). Noun Artifact refers to the tangible and concrete objects mostly made by human beings. Low representation of Noun Artifact in the spectrum of conjunctive Urdu noun classes can also be interpreted as the scarce probability of such nouns forming conjunct predicate in Urdu as majority of conjunctive nouns refer to action, communicative act or some abstract proposition. Its compatibility with different light verbs is also limited. For instance, the Noun Artifact ‘Bullet/*Goli’* is only compatible with one light verb ‘Hit/ *Laga’*. This reading does not give the agentive reading to the argument in the subject position but only of an experiencer semantic role. Other two Noun Artifact entries show 100% with one transitive light verb ‘Do/*Kar’* and two intransitive light verbs i.e., Become/*Do’* and copula ‘Is/ *He’* (See Table 4.14).

5.8.2 Noun Artifact: Compatibility with Light Verbs

Noun Artifact exhibited similar productivity pattern with light verbs as that of Noun Act and Noun Communication. There is almost the same pattern of N_{art}+V instances of compatibilities with transitive and intransitive light verbs.

Despite an overall very low productivity with light verbs, Noun Artifact still shows the same pattern of high compatibility with the transitive light verb ‘Do/*Kar’*.

i. High compatibility pattern of N_{art} with transitive light verb 'Do/*Kar'* is prevailed
ii. With the transitive light verb 'Do/*Kar*' the semantic element of conscious choice is also persistent. Means 'Export/*Daramad'* and 'Import/*Baramad'* which are in an antonymy semantic relation with each other exhibit the similar combinatory scheme with light verb
iii. Noun Artifact also shows the same high productivity with intransitive light verb 'Become/*Hu'* and copula 'Is/*He'*, but the semantic factor of conscious choice cannot be found here on subject argument
iv. Noun Artifact exhibits non-complacent semantic and syntactic behavior with intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'*
v. No productivity of Noun Artifact is found with ditransitive pair of light verbs 'Give/*Dia'* and 'Take/*Lee'*

5.8.3 Noun Artifact: Case Marking

The semantic sense of Noun Artifact when collocates with light verbs collaborates to assign ergative, dative, genitive and nominative case markers to the subject argument in the clause. When it occurs with two intransitive light verbs, 'Become/*Hu'*, and 'Is/*He'*, Noun Artifact allows genitive case marking (See Examples 77-80).

77.

Aam ki baramad hui
Mango-M.SG.GEN import-F.SG.NOM become-F.SG.PERF
'Mango is imported.'

78.

Aadmi ne aam beramad kiay
Man-M.SG.ERG mango-M.PL.NOM Import-F.SG.NOM do-M.PL.PERF
'Man imported mangoes.'

79.

Aadmi ko goli lagi
Man-M.SG.DAT bullet-F.SG.NOM put-F.SG.PERF
' Man got shot.'

80.

Aadmi aam baramad kerta hai

Man-M.SG.NOM mango-M.PL.NOM Import-F.SG.NOM do-M.SG.IMP is-SG

'Man imports mangoes.'

Noun Artifacts take different case markers with different transitive light verbs.

Different syntactic context is found when its collocation is found with two different transitive light verbs. For instance, with 'Do/*Kar'* it draws ergative case; whereas, with 'Hit/ *Laga'* it allows dative. Here, it can be noticed that the choice of case marker depends on the semanticity of conjunctive noun.

5.8.4 Noun Artifact: Argument Structure

Noun Artifact, a small conjunctive Urdu noun class, exhibits an almost similar nature of productivity with both transitive and intransitive light verbs. This pattern coincides with previously mentioned conjunctive Urdu noun classes. There were 4 instances of $N_{art}+V_{intrans}$; similarly, there are 3 instances of $N_{art}+V_{trans.}$

i. A theme of equal number of collocations was persistent both in case of $N_{art}+V_{intrans}$ and $N_{art}+V_{trans}$

ii. $N_{art}+V_{intrans}$ allows genitive case markers to the subject argument

iii. Same as earlier, two categories of Urdu conjunct predicate ($N_{art}+V_{trans}$) with the transitive light verbs were found based on case markers they permit. Category A consists of 'Do/*Kar'* which as usual allows ergative case; on the other hand, the category B transitive light verb 'Hit/ *Laga'* permits dative case (See Table 47 in the next section).

5.8.5 Noun Artifact: Semantic Roles

Category A with $N_{art}+V_{trans}$ is consistent with the ergative case '*Ne'* which can be interpreted as the agentive semantic role of the subject.

The different syntactic behavior of $N_{art}+V_{trans}$ within the same argument structure of Urdu conjunct predicate can be viewed for difference in case marking and its consequent impact on semantic role attributed to the subject is given below:

Table 5.6

Noun Artifact and Transitive Light Verbs

	Urdu Conjunct Predicate with Transitive Light Verbs	
	Category A AE	Category B ED
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF) Nominative (-PERF)	Dative
N_{art} +V_{trans}	N_{art}+ (‘Do/*Kar’*)	N_{art}+ (Hit/ *Laga*)

5.9 Noun Process

This section elaborates on semantic orientation, syntactical combination with light verb, case marking, argument structure, and semantic role of subject argument in case of Urdu conjunct predicate with Noun Process listed in chapter 4 (See Tables 4.15 and 4.16).

5.9.1 Noun Process: Semantic Orientation

The lexical choices listed here as Noun Process; they refer to some process in action. This Noun Process has got a limited representation in the list of conjunctive nouns here, so it can be interpreted as less productive conjunctive noun. Noun Process has got very thin representation on a spectrum of Urdu conjunctive noun classes which reflects their low productivity with light verb forming an Urdu conjunct predicate. Noun Person is the most compatible with intransitive light verb ‘Become/*Hu’* as compared to the transitive light verb ‘Do/*Kar’*. It is the first instance when an Urdu conjunctive noun class is more compatible with an intransitive light verb as compared to a transitive. This very information may also be interpreted as an absence of a conscious choice on part of subject argument. The ditransitive light verbs ‘Give/*Dia’* and ‘Take/*Lee’* which otherwise assign agentive role to the subject, are not doing the same here in case of (‘birth/ *Janam*’ + ‘Give/*Dia’*) and (‘birth/ *Janam*’ + ‘Take/*Lee’*). Both ‘to give birth’

and 'to be born' do not bear the sense of conscious choice of subject as an agent. It can be asserted that the absence of volitionality feature is due to the semanticity of Noun Process (See Examples 81 and 82).

81.

Maan ne do bachoon ko janam dia
Mother-F.SG.ERG Two Children-PL.ACC birth gave-M.SG.PERF
'Mother gave birth to two children'.

82.

Shazia ne Lahore main Janam lia
Shazia-F.SG.ERG Lahore-M.SG.LOC Birth took-M.SG.PERF
'Shazia was born in Lahore'.

5.9.2 Noun Process: Compatibility with Light Verbs

Compatibility of Noun Process with the light verbs has got a similarity and a slight deviation at the same time. Let us talk about the deviation first:

i. Noun Process is only compatible with one transitive light verb 'Do/*Kar'* unlike formerly discussed conjunctive noun classes.
ii. Noun Process is the most productive with intransitive light verb 'Become/*Hu'*.
iii. No instance was found regarding the collocation of Noun Process with a set of intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'*. This compatibility behavior is consistent with that of Noun Artifact.

In terms of some consistency of data, description is listed as below:

iv. Noun Process behaved as formerly discussed conjunctive nouns in terms of uniform compatibility with a set of ditransitive light verb: 'Give/*Dia'* and 'Take/*Lee'*.

5.9.3 Noun Process: Case Marking

The phenomenon of influence of noun semanticity continues to reveal in terms of syntactic context of Noun Process. Noun Process when comes with transitive light verb 'Do/*Kar'*, permits ergative case marker (See Example 83 and 84).

83.

Larkay ne paani zaya kia
Boy-M.SG.ERG water-M.SG.NOM waste-M.SG.NOM do-M.SG.PERF

'Boy wasted water.'

84.

Qabeelay ne muhaeeday ki tajdeed ki
Tribe-M.SG.ERG agreement-M.SG.GEN renewal-F.SG.NOM do-F.SG.PERF
'Tribe renewed the agreement.'

On the other hand, in case of its compatibility with 'Become/*Hu'* it allows dative and nominative (See Example 85 and 86).

85.

Dukaandar ko khasara hua
Shopkeeper-M.SG.DAT loss-M.SG.NOM become-M.SG.PERF
'Shopkeeper got a loss.'

86.

Anaaj zaya hua
Wheat-M.SG.NOM waste-M.SG.NOM become-M.SG.PERF
'Wheat was wasted.'

Only one instance of Noun Process 'Birth/*janam'* is compatible with a set of ditransitive light verbs 'Give/*Dia'* and 'Take/*Lee'* which mark the subject argument with ergative case.

5.9.4 Noun Process: Argument Structure

Noun Process has a small representation in the spectrum of conjunctive Urdu noun classes with only 1% instances. Productivity of Noun Process with light verbs is also not very high. It showed only 30 % compatibility with light verbs included in the study. Among the 10 instances of Noun Process productivity with light verbs, there are only 4 collocations with transitive, and 6 with intransitive light verbs. So, intransitive light verbs show slightly increased productivity with Noun Process which may also be interpreted as an absence of an agentive role of the subject. The agentive semantic role of subject argument is less seen in case of Noun Process. It is assumed as a process which happens because of some other action e.g., 'Loss/*kasara'*

i. For Noun Orocess, the equilibrium of equal instances of $N_{proc}+V_{intrans}$ and $N_{proc}+V_{trans}$ is slightly disturbed

ii. $N_{proc}+V_{trans}$ licenses ergative case marker on the subject argument in the presence of direct object in the clause

iii. $N_{proc}+V_{ditrans}$ exhibits similar compatibility of behavior with 'Give/*Dia'* and 'Take/*Lee'* with presence of direct object, indirect object and a subject. These compatible ditransitive light verbs allow ergative case markers on subject argument.

5.9.5 Noun Process: Semantic Roles

$N_{proc}+V_{trans}$ is consistent with the Ergative case '*Ne'* which assigns agentive semantic role of the subject (See Table 5.7).

Table 5.7

Noun Process and Transitive Light Verbs

Urdu Conjunct Predicate with Transitive Light Verbs	
	AE
Semantic Roles	Agentive
Case Marker	Ergative
N_{proc} +V_{trans}	N_{proc}+ ('Do/*Kar')*
N_{proc} +$V_{ditrans}$	(N_{proc}+ Give/ *Dia*, Take/ *Lee*)

5.10 Noun State

Entries listed in Table 4.17 and 4.18 of chapter 4 refer to some sort of state such as 'Deprivation/*Mehroomi'*, Blessing/*Rehmat*, etc. This section discusses the semantic orientation, compatibility pattern of Noun State with light verbs, argument structure, case marking and semantic roles of subject argument influenced by N_{state} +V in the sentence.

5.10.1 Noun State: Semantic Orientation

The semanticity of Noun State refers to the abstract state of entities. This conjunctive noun class has got a low representation in the whole spectrum of conjunctive nouns with 4% N_{state} +V instances. The prominent abstract semantic feature of Noun State is consistent with the semanticity of other conjunctive noun classes. 'Disagreement/*Mukhalfat'*, 'Security/*Tahafuz'* and 'Union/*Itehad'* are some examples

of Noun State which carry semantic feature of some cognitive state of mind which can only be felt.

5.10.2 Noun State: Compatibility with Light Verbs

i. Noun State maintains the equilibrium of light verb compatibility between a transitive and intransitive light verb i.e. 'Do/*Kar'* and 'Become/*Hu'*

ii. In accordance with the previous pattern of conjunctive noun classes, compatibility with 'Do/*Kar'* reflects the conscious choice of an agentive subject present in the sentence. With noun 'Success/*Kamyabi'*, it needed another noun 'Obtainment/*Hasil'* in the serial to complete its acceptability for native language user i.e., '*Kamyabi Hasil Karna'*/ to get success'. The noun '*Hasil'* is comparatively higher in level of formality than '*Lena'*. *Hasil Karna* if paraphrased in Urdu may simply obtain '*Lena'. Hasil Karna* is used with some nouns which contain semantic features of formality.

iii. High rate of productivity with intransitive light verb 'Become/*Hu'* is observed but it is void of volitionality feature of subject argument in the clause.

iv. Like Noun Artifact, Noun State is also non-complacent with a pair of intransitive light verbs i.e., 'Come/*Aa'* and 'Go/*Ja'*. At the same time, it can be noted that Noun State exhibits uniform compatibility pattern with this set of intransitive light verbs.

v. Noun State is slightly more productive with 'Give/*Dia'* as compared to 'Take/*Lee'* which may reflect the prominent agentive role influenced by Noun State.

5.10.3 Noun State: Case Marking

Noun State expresses quite diverse case marking. Transitive 'Do/*Kar'* permits the ergative case on subject with perfective aspect in the clause. This trait is prevalent to the previously discussed conjunctive noun classes (See Example 87). In case of imperfective aspect, it allows nominative case (See Example 88).

87.

Fouj ne jawanoon per inhasaar kia
Army-F.SG.ERG soldiers-M.SG.LOC reliance-M.SG.NOM do-M.SG.PERF
'Army depended on soldiers.'

88.

Fouj jawanoon per inhasaar karti he
Army-F.SG.NOM soldiers-M.SG.LOC reliance-M.SG.NOM do-F.SG.IMP is-SG
'Army depends on soldiers.'

The intransitive light verb 'Become/*Hu'* allows dative case on subject argument with perfective aspect on the clause (See Examples 89-90).

89.

Aurat Ko kamyabi Hui
Women-F.SG.DAT success-F.SG.NOM become-F.SG.PERF
'Woman has success.'

90.

Logo ko Siasatdaan se mukhalfat hui
People-PL.DAT Politician-M.SG.LOC disagreement-F.SG.NOM become-F.SG.PERF
'People disagreed with the politician.'

'Is/ *He'* allows nominative on subject argument case marker on the subject argument in the clause (See Example 91).

91.

Bacha school main dakhil he
Child-M.SG.NOM school-M.SG.LOC admit-M.SG.NOM is-SG
'Child is admitted to school.'

The ditransitive light verb 'Give/ *Dia'* allows ergative case marking on subject argument (See Examples 92-93).

92.

Allah ne mareez ko zindagi dee
God-M.SG.ERG patient-M.SG.ACC life-F.SG.NOM give-F.SG.PERF
'God gave life to patient.'

93.

Police ne mujrim ko tahafuz dia
Police-F.SG.ERG culprit-M.SG.ACC protection-F.SG.NOM give-M.SG.PERF
'Police protected the culprit.'

Whereas 'Hit/ *Laga'* does not show any collocative instance with any of the Noun States mentioned in the list.

5.10.4 Noun State: Argument Structure

i. N_{state}+V_{trans} and N_{state}+$V_{intrans}$ show different number of instances with 20 and 17 entries respectively
ii. N_{state}+$V_{intrans}$ allows the dative case marking on subject argument
iii. Compatible transitive and ditransitive lights allow the standard pattern of ergative case marking which in turn activates the agentive role of the subject argument in the clause
iv. No instance of Noun State was found with a pair of intransitive light verbs i.e., 'Come/*Aa'* and 'Go/*Ja'*

5.10.5 Noun State: Semantic Roles

Table 5.8

Noun State and Transitive Light Verbs

Urdu Conjunct Predicate with Transitive Light Verbs	
	AEN
Semantic Roles	Agentive
Case Marker	Ergative (+PERF)
	Nominative (-PERF)
N_{state} +V_{trans}	N_{state} +Do/*Kar*
N_{state} +$V_{ditrans}$	N_{state} + Give/*Dia*

5.11 Noun Feeling

Urdu lexical entities listed in Tables 4.19 and 4.20 in chapter 4 fall in the category of Conjunctive Noun Class 'Noun Feeling'. The following section introduces semantic orientation of Noun Feeling, their syntactic compatibility with light verbs, argument

structure of N_{feel} +V, case marking and semantic roles of subject argument in the clause with N_{feel} +V.

5.11.1 Noun Feeling: Semantic Orientation

The class Noun Feeling refers to the range of emotions and feelings. The abstract entity classified as Noun Feeling such as 'Wound/*Zakham'* does not refer to a physical wound but an emotional abusive feeling. Same is the case of 'Worship/*Ibadat'* where it gives the sense of reverence and not the physical acts of worship. Other examples of Noun Feeling such as 'Hope/*Umeed'*, 'Love/*Muhabat'*, 'Hatred/*Nafrat'* etc., are abstract entities related to cognitive perception.

5.11.2 Noun Feeling: Compatibility with Light Verbs

i. Most Noun feelings are compatible with the transitive light verb 'Do/Kar' such as *Sharam karna*/to feel shame, *Arzoo karna*/to wish and *Umeed karna*/to hope etc. On the other hand, instead of '**Rahat Karna'*, a native Urdu language user will prefer '*Rahat Hasil karna'* (to get comfort). The lexical entity 'Obtainment/*Hasil'* is an added noun to obtain the formal meaning.

ii. Noun Feeling exhibits prolific behavior with intransitive light verb with 'Become/*Hu'* like Noun Process.

iii. Noun Feeling exhibited non-compliance with both intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'*

iv. The Noun Feeling lexical items which are compatible with 'Give/*Dia'* are not collocated with sister ditransitive light verb 'Take/*Lia'*

v. The only instance of noun feeling with 'Hit/ *Laga'* is 'Wound/*Zakham'* where it gives only a metaphoric reading due to the abstract semanticity of this Noun Feeling

5.11.3 Noun Feeling: Case Marking

Noun Feeling collocates with intransitive light verbs 'Become/*Hu',* and it permits dative case marker on subject phrase (See Example 94).

94.

Larkay ko muhabat hui

Boy-M.SG.DAT love-F.SG.NOM become-F.SG.PERF

'Boy fell in love.'

The transitive light verb 'Do/*Kar'* is productive in terms of its compatibility with Noun Feeling, and it allows an ergative case marker subject argument with a perfective aspect in the clause. This ergative case marker brings along the agentive role of subject with the conscious choice provided that the subject is an animate entity (See Example 95).

95.
Larkay ne muhabat ki
Boy-M.SG.ERG love.F.SG.NOM do-F.SG.PERF
'Boy loved (somebody).'

Ditransitive light verbs 'Put/ *Rakh'* and 'Give/ *Dia'* permits ergative case marker on subject (See Example 96 –98).

96.
Maan ne umeed rakhi
Mother-F.SG.ERG hope-F.SG.NOM put-F.SG.PERF
'Mother hoped.'

97.
Larkay ne dost ko zakham dia
Boy-M.SG.ERG dost-M.SG.ACC wound-M.SG.NOM give-M.SG.PERF
'Boy inflicted pain to (his) friend.'

98.
Larkay ne dost ko takleef di
Boy-M.SG.ERG dost-M.SG.ACC pain-M.SG.NOM give-F.SG.PERF
'Boy inflicted pain to (his) friend.'

'Hit/ *Laga'* a ditransitive light verb licenses dative case marker to subject argument 99 (See Example 99).

Larkay ko zakham laga
Boy-M.SG.DAT wound-M.SG.NOM hit-M.SG.PERF
'Boy got hurt (emotionally).'

5.11.4 Noun Feeling: Argument Structure

i. Two patterns of $N_{feel}+V_{intrans}$ incidents are seen. The overall productivity of Noun Feeling with intransitive light verb is greater than as compared to $N_{feel}+V_{trans}$

ii. Subject argument in $N_{feel}+V_{intrans}$ have dative case marking on subject argument

iii. Subject argument in $N_{feel}+V_{trans}$ show diverse case marking which includes ergative, nominative. $N_{feel}+V_{trans}$ permits ergative with perfective aspect. With imperfective aspects, the subject takes nominative case.

iv. $N_{feel}+V_{ditrans}$ instances with 'Put/*Rakh'* and 'Give/ *Dia'* allow three arguments to the clause: subject, direct object and indirect object. Such instances allow ergative case on the subject.

The phenomenon of two categories regarding the case marking on arguments prevailed in case of $N_{feel}+V_{trans.}$ Category A would be consistent with an ergative case marker on its arguments with perfective aspect and nominative with imperfective aspect. Category B reflects a combination of two case markers i.e., nominative and dative (See Table 5.9).

5.11.5 Noun Feeling: Semantic Roles

The semantic roles of the arguments under the combination of $N_{feel}+V_{trans}$ are reflected as follows with reference to their different case marking scheme.

Table 5.9

Noun Feeling and Di/transitive Light Verbs

	Urdu Conjunct Predicate with Di/Transitive Light Verbs	
	Category A AEN	Category B DE
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF) Nominative (-PERF)	Dative
$N_{feel}+V_{trans}$	$N_{feel}+$ ('Do/*Kar'*)	$N_{feel}+$ ('Hit/ *Laga'*)
Case Marker $N_{feel}+V_{ditrans}$		Ergative $N_{feel}+$ ('Put/ *Rakh',* 'Give/ *Dia')*

5.12 Noun Phenomenon

In the list of 280 unique Urdu nouns, there is only one instance of Noun Phenomenon found as mentioned in Table 4.21 and 4.22 in chapter 4.

The following section comments on the semantic orientation, combinatory restrictions with light verbs, argument structure in case of N+V, case marking and semantic role of subject in the presence of N_{phen} +V.

5.12.1 Noun Phenomenon: Semantic Orientation

Noun Phenomenon is the smallest conjunctive Urdu noun class found with only one instance. This only instance 'Rain/*Barish'* refers to natural phenomenon. When collocated with transitive LV 'Do/*Kar'* it allows ergative which in turn supports an agentive argument, but this argument is usually 'Nature', or some supernatural entity usually resides in the concept of 'God' depending on the religion of the people who are using this expression. The semantic element of volitionality is inherent in the use of supernatural power as a subject.

5.12.2 Noun Phenomenon: Compatibility with Light Verbs

Noun Phenomenon showed compatibility with one transitive light verb 'Do/Kar'.

Agent in the case of this noun is usually a supernatural power 'God' or 'Nature'.

'Rain/*Barish'* is also compatible with intransitive light verbs 'Become/*Hu',* but it does not require an agent as a subject (See Example 100). This Noun Phenomenon is also compatible with 'Come/*aa'*.

100.

Roz barish hoti hai

Daily-Adv+time rain-F.SG.NOM become-F.SG.PERF is-SG

'It rains daily.'

5.12.3 Noun Phenomenon: Case Marking

In the presence of a transitive light verb i.e., 'Do/*Kar'*, it allows ergative and nominative case marker to the subject argument with perfective and imperfective aspect respectively (See Example 101- 103).

101.

Allah ne barish ki

God-M.SG.ERG rain-F.SG.NOM do-F.SG.PERF
'God rained.'

102.

Allah barish karta hai
God-M.SG.ERG rain-F.SG.NOM do-M.SG.IMP is-SG
'God rains'.

103.

Allah barishain karta hai
God-M.SG.ERG rain-F.PL.NOM do-M.SG.IMP is-SG
'God rains'.

Whereas in case of compatibility with intransitive light verbs i.e., 'Become/*Hu',* and *'*Come/*Aa',* it licenses nominative case (See Examples 104 and 105).

104.

Aaj barish hui
Today-Adv+time rain-F.SG.NOM become-F.SG.PERF
'It rained today.'

105.

Barish Ayi
rain-F.SG.NOM come-F.SG.PERF
'It rained.'

5.12.4 Noun Phenomenon: Argument Structure

i. The Noun Phenomenon 'Rain/*Barish'* is found in both combinations: N_{phen}+V_{trans} and N_{time} +$V_{intrans}$. With intransitive, Noun Phenomenon requires only one argument i.e., subject; whereas, with transitive light verb, it requires two arguments i.e., subject and object.

ii. In fact, both above-mentioned categories show two case marking designs on their arguments. For N_{phen}+$V_{intrans,}$ nominative case marking has been observed on the subject in the clause.

5.12.5 Noun Phenomenon: Semantic Roles

The Noun Phenomenon when compatible with the transitive light 'Do/*Kar'* assigns an agentive role to the subject argument which is usually a supernatural entity. With perfective aspect, it assigns an ergative case and agentive role to the subject.

When the clause is realized in imperfective aspect, it also assigns agentive role to the subject but marked with nominative case (See Table 5.10).

Table 5.10

Noun Phenomenon and Di/transitive Light Verbs

	Category A AEN
Semantic Roles	Agentive
Case Marker	Ergative (+PERF) Nominative (+PERF)
N_{phen} +V_{trans}	N_{phen} + ('Do/*Kar'*)

5.13 Noun Time

The lexical entities listed as Noun Time refer to temporal information related to the beginning, ending or duration of some activity. It is a small conjunct noun class with five entries listed in Table 4.23 and 4.24 of chapter 4.

The following subheadings discuss the correlation between semantic orientation, syntactic combinatory restrictions, argument structure, case marking and semantic roles in case of N_{time}+V instance in the construction.

5.13.1 Noun Time: Semantic Orientation

Noun Time refers to temporal information related to some events or some actions e.g., Time/*Waqat*, Ending/*Ikhtatam*, Delay/*Multavi.* These examples refer to different expressions of time which are abstract entities.

5.13.2 Noun Time: Compatibility with the Light Verbs

i. Noun Time is one of the conjunctive noun classes which has an equal number of compatibility instances with both transitive and intransitive LVs. With transitive light verb, it requires two arguments i.e., subject and direct object to complete the clause.

ii. Noun Time with intransitive set of LVs showed nominative case on only one argument present in the sentence i.e., subject.

iii. Transitive LV 'Do/*Kar'* and intransitive LV 'Become/*Hu'* are highly productive with Noun Time; at the same time, it assigns agentive and experiencer role to the subject.

5.13.3 Noun Time: Case Marking

Equilibrium of Noun Time with transitive and intransitive LVs resulted in a formulaic scheme of case marking. As it is also previously shown by Noun Feeling also, Noun Time when collocates with intransitive light verbs i.e., 'Become/*Hu'*, 'Come/*Aa'* and 'Go/*Ja',* it creates two different case marking patterns. It allows nominative case marker to the subject with 'Become/*Hu'* (See Examples 106 –112).

106.

Musaafir rukhsat hua
Passenger-M.SG.NOM leave-F.SG.NOM become-M.SG.PERF
'Passenger left.'

107.

Train rukhsat hui
Train-F.SG.NOM leave-F.SG.NOM become-F.SG.PERF
'Train left.'

108.

Mehfil berkhaast hui
Meeting-F.SG.NOM ending-F.SG.NOM become-F.SG.PERF
'Meeting ended.'

Intransitive light verbs i.e., 'Come/*Aa'* and 'Go/*Ja'* permit nominative case markers.

109.

Waqt aata hai
Time-M.SG.NOM come-M.SG.IMPis-SG
'Time comes.'

110.

Waqt jaata hai
Time-M.SG.NOM go-M.SG.IMP is-SG
'Time goes.'

Transitive light verb 'Do/*Kar'* allows ergative case on subject argument .
When compatible with ditransitive light verbs 'Give/*Dia'*, and 'Take/*Lia',* Noun Time allows ergative case on subject argument in clause (See Examples 111 and112).

111.

Maan ne bachay ko waqt dia
Mother-F.SG.ERG child-M.SG.ACC time-M.SG.NOM give-M.SG.PERF
'Mother gave time to child.'

112.

Talib-ilm ne ustaad ka waqt lia
Student-M/F.SG.ERG teacher-M/F.SG.GEN time-M.SG.NOM take-M.SG.PERF
'Student took time of teacher.'

5.13.4 Noun Time: Argument Structure

i. $N_{time}+V_{trans}$ require two arguments i.e., subject and object. There is only one argument required i.e., subject in case of $N_{time}+V_{intrans}$

ii. $N_{time}+V_{ditrans}$ instances require three arguments in the clause to complete the semantic sense of the clause i.e., subject, direct object and indirect object.

iii. Instances of $N_{time}+V_{trans}$ and $N_{time}+V_{ditrans}$ showed two subcategories. 'Do/*Kar'* as usual showed an ergative case marker on its arguments in perfective aspect. In imperfective aspect, instance with 'Do/*Kar'* assigns nominative case to the subject. On the other hand, category B has genitive case marker (See Table 5.10)

5.13.5 Noun Time: Semantic Roles

The pattern of two subcategories among transitive and ditransitive light verbs has emerged, and the resultant semantic role remained consistent as majority of conjunctive noun classes with minor deviations.

Table 5.11

Noun Time and Di/transitive Light Verbs

	Urdu Conjunct Predicate with Di/Transitive Light Verbs
	Category A
	AEN
Semantic Roles	Agentive
Case Marker	Ergative (+PERF)
	Nominative (-PERF)
$N_{time}+V_{trans}$	N_{time}+ (‘Do/*Kar’*)
$N_{time}+V_{ditrans}$	N_{time}+ (‘Give/*Dia’*, and ‘Take/ *Lee’*)

5.14 Noun Event

The class Noun Event refers to the concepts of some happening such as Meeting/ *Mulaqaat*, Movement/ *Junbish*, Dispersion/*Muntashir* etc. The following subheading try to correlate information related to semantic orientation, combinatory restrictions, argument structure, case marking and thematic roles in N_{eve} +V instances. The list of Noun Event is mentioned in Table 4.25 and 4.26 of section in chapter 4.

5.14.1 Noun Event: Semantic Orientation

This conjunctive noun class ‘Noun Event’ refers to the different events happening in a course of time. In case of some events, it is possible to physically monitor the existence of the event such as ‘Birth/*Paidayish’*, ‘Meeting/*Mulaqat’*, ‘Collision/*Takkar’*; on the other hand, some senses of these events are just a construction in mind i.e., ‘Defeat/*Shikist’*, ‘Movement/*Moohim’*, and ‘Loss/*Nuqsaan’*, etc.

5.14.2 Noun Event: Compatibility with Light Verbs

i. Compatibility patterns are almost the same in case of Noun Event with transitive and intransitive light verbs i.e., 'Do/*Kar'* and 'Become/*Hu'* respectively.

ii. Though showed uniform compatibility with a set of intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'*.

iii. A set of ditransitive light verbs i.e., 'Give/ *Dia'*, and 'Take/ *Lia'* which show a similar pattern of productive pattern with most conjunctive noun classes, did not show very eclectic compatibility theme except one noun 'Defeat/*Shikist'*. 'Defeat/*Shikist'* is only 'Given/*Di'* and not 'Taken/*Li'*. Agentive role of subject is contributed due to the semanticity of the noun 'Defeat/*Shikist'*.

iv. The Noun Event 'Collision/*Takkar'* is compatible with two ditransitive light verbs 'Give/ *Dia'*, and 'Take/ *Lia'*. Both the instances have the semanticity of metaphoric interpretation, and they did not give literal meaning of 'Collision/*Takkar'*.

5.14.3 Noun Event: Case Marking

Case marking induced by N_{eve} +V displayed a diverse range.

With transitive light verb 'Do/*Kar'* Noun Event assigns ergative case to the subject (See Examples 113-114).

113.

Barish ne fasal ka nuqsaan kia
Rain-F.SG.ERG Crop-F.SG.GEN loss-M.SG.NOM do-M.SG.PERF
'Rain damaged the crop.'

114.

Sadar ne wazir se mulaqaat ki
President-M.SG.ERG minister-M.SG.INS meeting-F.SG.NOM do-F.SG.PERF
'President met the minister.'

With 'Remain/ *Rah'* , Noun Event subject takes genitive case (See Examples 115-116).

115.

Aadmi ka dimaagh muntashir raha

Man-M.SG.GEN brain-M.SG.NOM dispersion-M.SG.NOM remain-M.SG.PERF

'Man' mind remained dispersed.'

116.

Bus ko gaari ki takkar lagi

Bus-F.SG.DAT car-F.SG.GEN collision-F.SG.NOM put.F.SG.PERF

'Bus got collided by car.'

Metaphoric interpretation of Noun Event with two ditransitive light verbs 'Give/ *Dia'* and Take/*Lee'* assigns ergative case to the subject (See Examples 117 and 118).

117.

Larkay ne afsar se takkar le

Boy-M.SG.ERG officer-M.SG.LOC collision-F.SG.NOM take-F.SG.PERF

'Boy messed up with officer.'

118.

Larkay ne afsar ko takkar de

Boy-M.SG.ERG officer-M.SG.ACC collision-F.SG.NOM give-F.SG.PERF

'Boy challenged the officer.'

Wheras the semanticity of 'to collide/*Takkar lagna'* carries literal meaning. This instance of N_{eve} +Hit/*Lagna* assigns dative case to the subject (See Example 119).

119.

Cycle ko gaari ki takkar lagi

Cycle-F.SG.DAT car-F.SG.GEN collision-F.SG.NOM put-F.SG.PERF

'Cycle collided by car.'

Ditransitive light verb 'Put/ *Rakh'* when collocated with Noun Event assigns ergative case to the subject (See Example 20).

120.

Mujrim ne bachay ko Gaib rakha

Culprit-M.SG.ERG child-M.SG.ACC disappearance-M.SG.NOM remain-M.SG.PERF
'Culprit grounded the child.'

Noun Event with intransitive light verb 'Become/*Hu'* permits dative case to the subject. And for 'Come/*Aa'* it is locative (See Example 121).

121.

Television per drama aya
Television-M.SG.LOC drama-M.SG.NOM come-M.SG.PERF
'There was a drama on television.'

5.14.4 Noun Event: Argument Structure

i. Instances with $N_{eve}+V_{intrans}$ get one argument i.e., subject.

ii. $N_{eve}+V_{trans}$ two arguments i.e., subject and object. There was an equal number of $N_{eve}+V_{intrans}$ and $N_{eve}+V_{trans}$ instances found in data.

iii. Single argument in $N_{eve}+V_{intrans}$ 'Become/Hua' displays dative case marking on subject argument (See Example 122).

122.

Larkay ko nuqsaan hua
Boy-M.SG.DAT loss-M.SG.NOM become-M.SG.PERF
'Boy had a loss.'

iv. Within the setting of $N_{eve}+V_{trans}$ as also observed earlier in case of formerly discussed conjunctive nouns, it showed the existence of two categories in terms of syntactic behavior of its arguments. Category A persistently allows ergative case marking on its subject arguments when there is perfective aspect in the sentence. In case of imperfective aspect, subject takes nominative case with $N_{eve}+V_{trans}$ instances. And in category B, there is a range of genitive, nominative and dative case markers on arguments.

5.14.5 Noun Event: Semantic Roles

The semantic roles displayed by arguments in the syntactic context of N_{eve}+V_{trans} showed a diversity owing to the diverse case marking induced in this setting. The detail is seen in the Table 5.12 given below:

Table 5.12

Noun Event and Di/transitive Light Verb

	Urdu Noun Event with Di/Transitive Light Verbs	
	Category A	Category B
	AEN	ED
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF)	Dative
	Nominative (-PERF)	
N_{eve}+V_{trans}	N_{eve}+'Do/*Kar'*,'Put/ *Rakh'*	N_{eve}+Remain/*Rah* , 'Hit/lagi
N_{eve}+$V_{ditrans}$	N_{eve}+'Take/*Lee'*,'Give/ Dia	

5.15 Noun Group

Nouns mapped as Noun Group refer to the concept of some sort of grouping of people. The list of Noun Group is mentioned in Table 4.27 and 4.28 of chapter 4. The sub-sections below talk about the connection between semantic features, syntactic compatibility, argument structure, case marking and semantic roles of subjects in case of N_{grp} +V in the construction.

5.15.1 Noun Group: Semantic Orientation

Noun Group is a small conjunctive Urdu noun class with only 3 instances in the whole list of 280 conjunctive nouns (See Table 4.27). When the semanticity of nouns in this class is observed, it is deduced that the three entries refer to the group of abstract nouns. Their meanings are based on some abstract semantic concepts i.e., 'Free/ *Raha'*, 'Partnership/*Partnership'*. Only one noun group 'Crowd/*Hajoom'* can be seen as a physical group of people gathered somewhere.

5.15.2 Noun Group: Compatibility with Light Verbs

i. All instances of Noun Group are compatible with 'Do/*Kar'* and 'Become/*Hu'*.
ii. No other light verb showed any compatibility with Noun Group

5.15.3 Noun Group: Case Marking

Noun Group when collocates with light verbs influences the pattern of case marking on subject argument.

With transitive light verb 'Do/*Kar'*, it allows ergative on argument. Similar behavior is exhibited by all conjunctive Urdu nouns with 'Do/*Kar'*. Noun Group is consistent with previously explored conjunctive noun classes (See Examples 123 and 124 given below).

123.

Police ne qedion ko raha kia
Police-F.SG.ERG prisoners-PL.ACC free-M.F.NOM do-M.SG.PERF
'Police freed the prisoners.'

124.

Company A ne Company B se partnership kia
CompanyA –F.SG.ERG Company B-F.SG.INS partnership-F.SG.NOM do-F.SG.PERF
'Company A did the partnership with company B.'

With intransitive light verb, 'Become/*Hu'*, it allows nominative and genitive case marking on subject argument (See Examples 125 and 126).

125.

Qedi raha huay
Prisoners-PL.NOM free-M.SG.NOM become-PL.PERF
'Prisoners got free.'

126.

companioon ki partnership hui
Companies-Pl.GEN partnership-F.SG.NOM become-F.SG.PERF
'Companies got partnership.'

5.15.4 Noun Group: Argument Structure

i. $N_{grp}+V_{intrans}$ displayed the double instances as compared to $N_{grp}+V_{trans.}$

ii. Arguments in Noun Group syntactic contexts do not show quite variant case marking. For instance, the only productive transitive light verb 'Do/*Kar'* permits an ergative case marker on subject; whereas the intransitive light verb occurred with Noun Group induces nominative case markers on the subject argument.

5.15.5 Noun Group: Semantic Roles

As previous conjunctive Urdu noun classes, the argument with ergative case marker in $N_{grp}+V_{trans}$ 'Do/*Kar'* gives agentive role to the subject. However, no two categories could be observed in $N_{grp}+V_{trans}$ setting due to no more combination with other transitive light verb other than 'Do/*Kar'.*

5.16 Noun Possession

These nouns refer to some sort of material possession which is taken or given in some way. The entries categorized as Noun Possession and their compatibility with light verbs are listed in Tables 4.29 and 4.30 in chapter 4.

5.16.1 Noun Possession: Semantic Orientation

Noun possession has only 4 instances in the list of 280, so it is one of the small conjunctive noun classes (See Table 4.30). The nouns refer to the abstract concept of possession which is either kept, given, or taken.

5.16.2 Noun Possession: Compatibility with Light Verbs

i. Noun Possession 'Bail/*Zamanat'* like some other conjunctive Urdu nouns (Success/*Kamyabi*) requires 'Obtainment/ *Hasil'* + 'Do/*Kar'* to achieve the appropriate meaning. Here the semanticity of 'Obtainment/ *Hasil'* can be further investigated to state its semanticity explicitly. For now, it can be stated

that it contributes to the level of formality involved in the action. The semantic element of conscious choice is also present here in this reading.

ii. Noun Possession showed slightly different compatibility patterns with different light verbs. For instances 'Become/*Hu'* is 100% compatible with the Noun Possession, and 'Do/*Kar'* also shows high compatibility productivity with Noun Possession.

iii. A set of ditransitive light verbs, 'Give/*Dia'* and 'Take/*Lia'*, shows similar pattern of compatibility with Noun Possession

iv. Two intransitive light verbs, 'Come/*Aa'* and 'Go/*Ja'*, are also synchronous in terms of schematic productive pattern with this conjunctive noun.

The above two points encourage a formulaic pattern of light verb compatibility combinations with conjunctive nouns.

5.16.3 Noun Possession: Case Marking

Noun Possession and transitive light verb 'Do/*Kar'* permits ergative on subject argument with perfective aspect in the sentence (See Example 127).

For imperfective aspect in the clause, subject takes nominative case when a Noun Possession is followed by transitive light verb 'Do/*Kar'* (See Example 128).

127.

Kisaan ne anaaj mehfooz kia

Farmer-M.SG.ERG wheat-M.SG.NOM store-M.SG.NOM do-M.SG.PERF

'Farmer stored the wheat.'

128.

Kisaan anaaj mehfooz karta he

Farmer-M.SG.NOM wheat-M.SG.NOM store-M.SG.NOM do-M.SG.IMP is-SG

'Farmer stores wheat.'

Noun Possession when compatible with ditransitive light verb 'Give/*Dia'* allows ergative case marker on subject argument with perfective aspect in construction (See Examples 129-130).

129.

Aadmi ne larkay ki zamanat di
Man-F.SG.ERG boy-M.SG.GEN bail-F.SG.NOM give-F.SG.PERF
'Man bailed the boy.'

130.

Larki ne qurbani de
Girl-F.SG.ERG sacrifice-F.SG.NOM give-F.SG.PERF
'Girl sacrificed.'

With imperfective aspect clause, subject argument takes nominative case with Noun Possession and ditransitive light verb (See Example 131).

131.

Larki qurbani deti he
Girl-F.SG.NOM sacrifice-F.SG.NOM give-F.SG.IMP is-SG
'Girl sacrifices.'

Another ditransitive light verb 'Put/*Rakha'* also allows ergative case on subject with perfective aspect in sentence (See Example 132).

132.

Larke ne saaman mehfozz rakha
Boy-M.SG.ERG luggage-M.SG.NOM safe-M.SG.NOM put-M.SG.PERF
'Boy kept the luggage safe.'

Intransitive light verb 'Become/*Hua'* licenses dative case on subject with both perfective and imperfective aspect (See Examples 133 and 134).

133.

Larkay ko fiada hua
Boy-M.SG.DAT benefit-M.SG.NOM become-M.SG.PERF
'Boy got benefited.'

134.

Larkay ko fiada hota he

Boy-M.SG.DAT benefit-M.SG.NOM become-M.SG.IMP is-SG

'Boy gets benefit.'

Whereas the light verb 'Give/*Dia'* when compatible with Noun possession along with perfective aspect in the clause takes ergative case marker on the subject (See Example 135).

135.

Larki ne aadmi ki zamanat di

Girl-F.SG.ERG man-M.SG.GEN bail-F.SG.NOM give-F.SG.PERF

'Girl bailed for man.'

5.16.4 Noun Possession: Argument Structure

$N_{pos}+V_{intrans}$ allows one argument i.e., subject .

$N_{pos}+V_{trans}$ needs two arguments to complete the construction i.e., subject and an object. Noun Possession is compatible with 9 intransitive and 11 transitive light verbs (See Table 36). The notion of two categories regarding the case marking and semantic role of the subject with the transitive light verb is also prevalent here in case of $N_{pos}+V_{di/trans}$. Category A arguments have 'Do/*Kar'* as usual in addition to 'Put/*Rakh'*, 'Give/*Dia'*, and 'Take/*Lia'*. In category B, subject argument has Locative case marker in the sentence with 'Hit/*Laga'* as a light verb (See Example 136).

136.

Aamdi per dus Hazar tax laga

Man-M.SG.LOC Ten Thousand Tax-Nom put-M.SG.PERF

'Man got ten thousand tax this year.'

5.16.5 Noun Possession: Semantic Roles

As there are two categories of arguments in the case of $N_{pos}+V_{di/rans,}$ it has a direct effect on the semantic roles subject arguments bear (See Table 5.13).

Table 5.13

Noun Possession and Di/transitive Light Verb

	Urdu Conjunct Predicate with Di/Transitive Light Verbs	
	Category A AEN	Category B NAG
Semantic Roles	Agentive	Experiencer
Case Marker	Ergative (+PERF) Nominative (-PERF)	Locative
$N_{pos}+V_{trans}$	N_{pos}+ (‘Do/*Kar’)*	N_{pos}+(‘Hit/*Laga’*)
$N_{pos}+V_{ditrans}$	N_{pos}+(‘Put/*Rakh’*, ‘Give/*Di’*, ‘Take/ *Lee’*)	

5.17 Noun Person

Among the list of 280 unique nouns, there is only one instance mentioned in Table 4.31 and 4.32 that could be mapped as Noun Person i.e., ‘Victim/*Shikaar’.*

5.17.1 Noun Person: Semantic Orientation

Noun Person has only one instance recorded here. The lexical entity ‘Victim/*Shikaar’* may refer to an animal or human being depending on the contextual discourse. Here, it is picked as a Noun Person as it denotes a human being who is being victimized.

5.17.2 Noun Person: Compatibility with Light Verbs

Regarding the semantic and syntactic compatibility with light verbs, Noun Person resembles a lot with the Noun Group (See Tables 4.31 and 4.32). It is only compatible with three light verbs i.e., one transitive and two intransitives. There seems to be a connection between a smaller number of entries and their less productivity in terms of compatibility with different light verbs. Conjunctive Urdu noun classes with greater instances tend to be more productive with light verbs as evident from the previous findings.

5.17.3 Noun Person: Case Marking

Arguments with Noun Person in the syntactic context transitive and intransitive light verb tend to get ergative and nominative case markers respectively (See Examples 137-139).

137.

Company ne logon ka shikar kia
Company-F.SG.ERG people-PL.GEN victim-M.SG.NOM do-M.SG.PERF
'Company victimized people.'

138.

log company ka shikar huay
People-PL.NOM company-F.SG.GEN victim-M.SG.NOM Become-M.PL.PERF
'People got trapped by company.'

139.

log company ka shikar hain
People-PL.NOM company-F.SG.GEN victim-M.SG.NOM are-PL
'People are trapped by the company

5.17.4 Noun Person: Argument Structure

i. $N_{per}+V_{intrans}$ requires a subject argument to complete the sentence. Intransitive light verb gives an experiencer role to the subject.

ii. $N_{per}+V_{trans}$ needs two arguments i.e., subject, direct object. Transitive light verb assigns an agentive role to the subject.

5.17.5 Noun Person: Semantic Roles

$N_{per}+V_{trans}$ gives an agentive reading to the subject which is correlated to the ergative case licensed by the semanticity of noun and light verb in the presence of perfective aspect.

With imperfective aspect, $N_{per}+V_{trans}$ assigns a nominative case to the subject (See Table 5.14).

Table 5.14

Noun Person and Di/Transitive Light Verb

	Urdu Conjunct Predicate with Transitive Light Verbs
	Category A
	AEN
Semantic Roles	Agentive
Case Marker	Ergative (+PERF)
	Nominative (-PERF)
$N_{per}+V_{trans}$	$N_{pos}+$ ('Do/*Kar'*)

5.18 Noun Relation

Only one noun has been mapped as Noun Relation among unique list of Urdu nouns i.e., 'Inclusion/*Shirkat'* listed in Table 4.33 and 4.34 of chapter 4.

5.18.1 Noun Relation: Semantic Orientation

Like Noun Person, and Noun Phenomenon, this conjunctive Urdu Noun Relation has also only one representative instance in the list of conjunctive nouns (See Table 39). The lexical entry as Noun Relation refers to a concept of involvement/inclusion in some physical or emotional activity.

5.18.2 Noun Relation: Compatibility with the Light Verbs

If we arrange small conjunctive Urdu noun classes in terms of their compatibility with light verbs in ascending order, then it will be Noun Person, Noun Relation and Noun Phenomenon with 3, 4, and 5 compatibilities.

Noun Relation is equally compatibility with transitive 'Do/*kar'* and intransitive light verb 'Become/*hu'* (See Table 4.34).

5.18.3 Noun Relation: Case Marking

The semanticity of Noun Relation and compatibility with transitive light verb 'Do/*Kar'* supports ergative and nominative case marking on its subject argument in case of perfective and imperfective aspects respectively (See Examples 140 and 141).

140.

Adil ne Ijlass main shirkat ki

Adil-M.SG.ERG Meeting-M.SG.LOC Inclusion-F.SG.NOM do-F.SG.PERF

'Adil attended the meeting.'

141.

Adil Ijlass main shirkat karta he

Adil-M.SG.NOM Meeting-M.SG.LOC Inclusion-F.SG.NOM do-F.SG.IMP he-SG

'Adil attends the meeting.'

Intransitive light verb 'Become/*hu'* assigns nominative case to the subject (See Example 142).

142.

Adil Ijlass main Shareek Hua

Adil-M.SG.NOM Meeting-M.SG.LOC Inclusion-F.SG.NOM become-M.SG.PERF

'Adil attended the meeting.'

5.18.4 Noun Relation: Argument Structure

i. Both N_{rel} +$V_{intrans}$ 'Become/*Hu'*, and N_{rel}+V_{trans} 'Do/*Kar'* exhibited compatibility. N_{rel} +$V_{intrans}$ requires one argument and N_{rel}+V_{trans} needs two arguments to complete the sentence.

ii. One instance of N_{rel}+V_{trans} permits with ergative on subject argument.

5.18.5 Noun Relation: Semantic Roles

Noun Relation when compatible with transitive light verb 'Do/Kar' assigns ergative case to the subject argument with perfective aspect in the sentence. Whereas imperfective aspect in the clause with the condition N_{rel}+V_{trans} allows nominative case to the subject argument (See Table 5.15). This condition is persistently constant with other conjunctive Urdu noun classes as well.

Table 5.15

Noun Relation and Di/Transitive Light Verb

	Urdu Conjunct Predicate with Transitive Light Verbs
	Category A
	AEN
Semantic Roles	Agentive
Case Marker	Ergative (+PERF)
	Nominative (-PERF)
$N_{rel}+V_{trans}$	N_{rel}+ ('Do/*Kar'*)

5.19 Insertion of Urdu Conjunct Predicate in Urdu WordNets

For inserting the Urdu conjunct predicate in Urdu WordNet, two provisions may be put forward: Firstly, both host noun and light verb in a N+V instance be mapped to the relevant semantic senses separately; secondly, N+V instance as a unit with a distinct semantic structure may be mapped onto relevant predicate semantic sense. The similar proposition has also been suggested in the case of different types of complex predicates in Indian languages (Bhattacharyya et al., 2007). It has been elaborated that the holistic meaning of N+V instance may not be compositional of the semantic sense of its components. And it may also pose problems in realization of its argument structure which resultantly may not carry usual functional information of the lexicon. Mapping of host noun and light verb separately may not yield the resultant semantic sense across language programs such as text-to-text and speech-to text application, etc. For different information retrieval applications, it may also decrease the precision of the application. Therefore, mapping the whole sense of N+V instance onto the same semantic sense of a predicate present verb may be a better and more efficient solution.

5.19.1 Categorization of Urdu Conjunct Predicate

Urdu conjunct predicates N+V are categorized based on semantic classes of host noun. As a result of the present study, fifteen (15) categories of Urdu conjunct predicates have emerged (see Table 5.16).

Table 5.16

Categories of Urdu Conjunct Predicates

Urdu Noun Class	Category of Urdu Conjunct Predicate	Urdu Noun Class	Category of Urdu Conjunct Predicate
Noun Act	N_{act}+V	Noun Time	N_{time} +V
Noun Communication	N_{com}+V	Noun Event	N_{eve} +V
Noun Cognition	N_{cog}+V	Noun Group	N_{grp} +V
Noun Attribute	N_{att} +V	Noun Possession	N_{pos} +V
Noun Artifact	N_{art} +V	Noun Person	N_{per} +V
Noun Process	N_{proc} +V	Noun Relation	N_{rel} +V
Noun State	N_{state} +V		
Noun Feeling	N_{feel} +V		
Noun Phenomenon	N_{phen} +V		

If the clear distinction of lexical and syntactic complex predicates based on the argument structure of the light verb (Williams, 1997) is followed for Urdu Conjunct predicates (N+V) instances, then N+V instances with transitive light verb will be *Syntactic* conjunct predicates. On the other hand, N+V instances with intransitive light verbs will be *Lexical* conjunct predicates. Bhattacharyya et al., (2007) proposition of lexical and syntactic compound verbs is different from it. They categorized **V1inf-e +lagnaa** and the **V1inf +paRnaa** constructions in Hindi as Syntactic Compound Verbs (SCpdVs) because of the modal auxiliary nature of V2. Therefore, Bhattacharyya et al., (2007) suggested dealing syntactic compound verbs (SCpdVs) in the syntax. For V1stem+V2, another category of Lexical Compound Verbs (LCpdVs) is suggested

because of the unpredictable nature of V2. Lexical compound verbs function as a single complex of semantic and syntactic properties. It is proposed to include these lexical compound verbs in the lexical knowledge base so that their automatic extraction may be made possible (Bhattacharyya et al., 2007).

In the present work, it is clearly seen that the semantic classes of nouns in Urdu conjunct predicates have clear preferences for different light verbs. In the previous sections of this chapter, it has been tried to explicitly talk about the impact of syntactic context of Urdu N+V conjunct predicates of their syntactic context in terms of the compatibility of host noun with different light verbs on the case marking and semantic roles of the subject argument in the clause. Most of the conjunctive noun classes have equal compatible instances with both transitive and intransitive light verbs. It may be interpreted that same conjunctive Urdu noun class has got the ability to form both Lexical and Syntactic conjunct predicates if followed the classification advocated by William (1997).

5.19.2 Connection between the Argument Structure, Semantic Sense and Case Marking

The present investigation of the semantic and syntactic properties of Urdu nouns and light verbs in N+V instances validates Levin's (1993) proposition that the expression and interpretation of an argument in a clause are heavily influenced by meaning of both host noun and the light verb. Here, the accumulative behavior of Urdu N+V instance is used to draw semantic classes of Urdu nouns. It resulted in 15 conjunctive noun semantic classes which are drawn from the semantic classes of nouns mentioned in English WordNet (Miller et al., 1990). The elaboration of conjunctive Urdu noun classes based on the argument behavior and compatibility with different light verbs may pave the way for a *Theory of Lexical Knowledge* about Urdu nouns.

The present interpretation of Urdu nouns is heavily colored by Levin (1993) theory of a connection between the semantics of predicate and its syntactic contexts. The meaning of Urdu nouns has influenced their compatibility with different light verbs; furthermore, the tense and the case marking of the subject argument have a clear impact on the semantic roles of the subject arguments in the clause.

The term *Elaborated Argument Structure* used for interconnectedness of aspectual light verbs and case marking of the holistic reading of complex predicate is also seen for Urdu conjunct predicate (Butt, 1995). It addresses the issues related to complex

argument structure but simple functional structure. As different conjunctive semantic classes of the host noun in N+V instances reflected a pattern of compatibility with transitive and intransitive light verbs, it reflected on the argument structure of Urdu N+V instance as a conjunct predicate which in turn is also related to the case marking and the semantic roles of its subject argument. N+ $V_{transitive}$ assigns agentive role to the subject argument; whereas, the same semantic noun class with intransitive light verb N+ $V_{intransitive}$ gives experiencer role to the subject argument

5.20 Identification of True Urdu Conjunct Predicate

Many of the listed complex predicates (Butt, 1995 & 2019; Davison, 2005; Kachru, 2006; Mohanan, 1994; Montaut, 2016) lack a clear criteria of categorization as conjunct predicate. The most common light verbs that are listed in the literature which merge with the appropriate nouns or adjectives in Urdu conjunct predicates are: 'Do/*Kar',* 'Become/*Hu*', 'Be/*He',* 'Put/*Rakh',* 'Come/*Aa*', 'Give'*Dia,* 'Go/ *Ja*', 'Take/*Lena*', 'Remain/*Rah',* and 'Hit /*Laga*'.

However, the semantic class of nouns that merge with these light verbs are so varied in their essence that it requires an insightful interpretation even by experts to determine which two elements constitute a conjunct predicate and which ones do not.

A common and prolific transitive light verb in terms of its compatibility with most semantic classes of Urdu nouns is 'Do/*Kar'.* However, a lot of examples of conjunct predicate formed with 'Do/*Kar'* cited in literature do not strictly fall under the category of a conjunct predicate. A simple heuristic H* proposed by Bhattacharyya et al. (2007) is not sufficient to identify lexical compound verbs (LCpdVs) which says that a Polar verb which is followed by a vector from a selected list of verbs can form a complex predicate. So, it cannot be simply said that a conjunctive Urdu noun class when cojoined with one of the preselected light verbs forms a true Urdu conjunct predicate. The following sentences (See Examples 143 and 144) picked up from exhaustive research on Urdu conjunct predicates indicate a common problematic diagnosis of this category of Urdu conjunct predicate.

143.

bacco – ne **kām** **kiyā**

children-PL.ERG work.M.SG.NOM do-M.SG.PERF
'The children finished the job/ assignment.'

144.

baccõ – ne mujh - se **muqāblā** **kiyā**
children-PL.ERG me - ABL competition-M.SG.NOM do-M.SG.PERF
'The children competed with me.' (Fatima, 2021, P.225)

In both the above examples 143 and 144, *kam kerna* (lit. to do work) and *muqabla kerna* (to compete) are exemplified as conjunct predicate where 'Work/*Kam'* and 'Competition/ *Muqabla'* are the respective nouns that are claimed to have merged with the light verb 'Do/*Kar'* to from the conjunct predicate.

In fact, in neither case the so called 'conjunct predicate' happens to be true conjunct predicate. The verb 'Do/*Kar'* is a transitive verb which carries two arguments in the above examples, an NP in the subject position and an NP in the direct object (DO) position. This is visible in the above-cited examples (143 and 144) by the agreement the verb shows with the DO. It is a well-known fact that the verb in Urdu agrees in terms of number, person and gender with its leftmost overtly unmarked argument (Khan, 1989. Gair & Wali, 1989). When no unmarked argument is available in the sentence, the verb assumes the default agreement which form is 3rd person, masculine, singular. Below are some examples (See Examples 145-147) to illustrate the verb-agreement phenomenon in Urdu.

145.

bacce kam **karte** **hain**
Children-M.PL work-M.SG.NOM do.M.PL.IMP be.PL
'The children work.'

146.

bacciyo – ne **kam** **kiya**
Children-F.PL.ERG work-M.SG.NOM did-M.SG.PERF
'The girls did the work/ The girls worked.'

147.

bacciyo – ne cuRiyoo – ko mezoo pe **rakkha**

Children-F.PL.ERG bangles.F.PL.ACC/Spec tables.F.PL.LOC on put-M.SG.PERF

'The girls placed the bangles on the tables.'

The verb in 145 agrees in number and gender with the leftmost, overtly unmarked argument noun phrase in subject position. In 146, the leftmost unmarked noun phrase argument is in Direct Object position with which the verb agrees in number and gender. Since all the arguments in 147 are overtly marked by some case marker, the verb assumes the default agreement of 3rd person masculine singular form.

The verb-agreement phenomenon as evidenced in the above examples supports the fact that not every noun adjacent to a light verb will constitute a conjunct predicate. Given the fact that verbs in Urdu agree only with their arguments, it is intuitively awkward to assume that a verb can agree with its own internal constituent part. This would be tantamount to claim that the "*es*" (/z/) morpheme in *does* agrees with the "*do*" part of the verb and has nothing to do with, say, 3rd person singular subject of a simple sentence with which English verbs usually agree.

The verb 'Do/*Kar'* as a light verb in Urdu selects only conjunctive semantic classes of noun (See Table 4.2) that merge with it to form an Urdu conjunct predicate. Unlike the noun 'Work/*Kam'* in 145 and 146 above which triggers verb agreement, the light verb *kerna* in examples 148, 149 and 150 below does not agree with its conjunct component, *yaad*, in terms of the ϕ-features.

148.

Bacco ne **sabaq** yaad kiya

Children-M.PL.ERG **lesson-M.SG.NOM** memory-F.SG.NOM **did.M.SG.PERF**

'The children memorized the lesson.'

149.

Bacco ne nəzmee yaad kii

Children-M.PL.ERG poems-F.PL.NOM memory-F.SG.NOM do-F.PL.PERF

'The children memorized the poems.'

150.

Bacco ne jawab yaad- kiye

Children-M.PL.ERG answers-M.PL.NOM memory-F.SG.NOM do-M.PL.PERF

'The children memorized the answers.'

In the three examples (148, 149, 150) above, the argument in the subject position is overtly marked with the ergative case marker *ne* which blocks the verbal agreement with it and hence the verb agrees with the argument in the direct object position in terms of number and gender. The noun *yaad*-F.SG. (memory), however, does not trigger any agreement with the verb 'Do/*Kar'* which assumes *masculine, singular* form in 148, *feminine, plural* form in 149 and *masculine, plural* form in 150 respectively. This proves that the noun 'Memory/Y*aad'* is not an argument of the light verb 'Do/*Kar'* and instead is an integral part of the conjunct predicate 'Memorize/ *Yaad-kerna'*. 'Be remembered/*yaad-hona'* and 'Remember/Y*aad-anaa'* do form conjunct predicate.

5.21 Significance of Research Contribution

The verb agreement phenomenon, therefore, becomes a salient test to decide which nominal elements can merge with a verb to constitute a conjunct verb and which ones are, in fact, an argument of the same 'light verb' and create a mere illusion of a conjunct predicate. The verb-agreement test, however, can be problematic with some light verbs such as *'*Come*/Aa*' or *'*Become*/Hu*' when they form a conjunct predicate with the nouns of the type 'Memory/*Yaad'*. While none of the so called "conjunct verbs" listed and repeated below in example 151 under *Natural Event* are conjunct predicate (Fatima et al., 2021).

151.

Natural Event (Fạtima, 2021 et al., P.228)

a. *rāt honā* 'to be night'

b. *barsāt honā* 'to be rainy'

c. *garmī honā* ‘to be hot’

d. *ţhanḍ honā* ‘to be cold’

The above data (See Example 151) cannot be treated as instances of conjunct predicate as the nominal part of each of the verbal complex agrees in number and gender with the verbal element of the serial verb. This is made evident in the sentences below where *a-2* and *a-3* are unacceptable instances of the “conjunct verb” *rat hona* because the verb fails to agree with the noun *rāt* either in gender or in number. Yet, example 152 *a-1* and *a-4* are acceptable because the verb agrees with the preceding noun in terms of number and gender.

152.

a-1.	ra*t*	*huii*	(The night fell)
	night.F.SG.	became.F.SG.	
a-2.	**rat*	*hua*	(The night fell)
	night.F.SG.	became.M.SG.	
a-3.	**rat*	*huii*	(The nights fell)
	night.F.SG.	became.F.Pl.	
a-4.	*ratee*	*huii*	(The nights fell)
	nights.F.Pl.	became.F.Pl.	

It’s obvious that if the verb *‘*Become*/Hu*’ inflects with changing ϕ-features of the noun ‘Night/ *rat’*, the said noun cannot be part of the claimed conjunct predicate *rat hona* ‘to be night’. It has to be an argument, in fact a subject of the intransitive verb in the above sentences and external to the serial verb complex, in order for the verb to agree with it. As mentioned earlier, a genuine conjunct predicate in Urdu is formulated with conjunctive semantic classes of noun such as ‘Memory/Y*aad’* or ‘Sight/*Nazar’*, involving psychological feature of mental perception, that can merge with light verb *‘*Become*/Hu*’, *‘*Come*/Aa*’ and ‘Do/*Kar’*. Like examples (148-149) cited above of sentences with conjunct predicate ‘to Memorize/*Yaad-kerna'*, ‘be remembered/Y*aad-hona’* is easy to establish as a conjunct predicate by applying the verb-agreement test (see examples 153 and 154).

153.

Bacco ko sabaq yaad hua
Children-M.PL.DAT lesson-M.SG.NOM memory-F.SG.NOM became-M.SG.PERF
'The children memorized the lesson.'

154.

Bacco ko kahaniya yaad *huii*
Children-M.PL.DAT stories-F.PL.NOM memory-F.SG.NOM became-F.PL.PERF
'The children memorized the stories.'

In both these sentences, the light verb 'Become*/Hu*' agrees in number and gender with the unmarked direct object 'Lesson/*Sabaq'* and 'Stories/*Kahaniya'* respectively but not with the noun 'Memory/*Yaad'* making it a constituent part of the conjunct predicate 'be remembered/Y*aad hona'*.

The conjunct predicate '(remember/recall/miss someone or something)/*Yaad-anaa'*, however, is not easy to determine by applying verb-agreement rule. In sentences such as 155 – 157 below, the light verb appears to agree with the unmarked direct object in terms of number and gender.

155.

us ko ghar yaad ayaa
him DAT home-M.SG.NOM remembrance-F.SG.NOM came-M.SG.PERF
'He remembered/ missed (his) home.'

156.

us ko bite din yaad ayee
Him-DAT gone-by days-M.PL.NOM remembrance-F.SG.NOM came-M.PL.PERF
'He longed for the days gone by.'

157.

us ko puranii batee yaad ayii
Him-DAT old chitchat-F.PL.NOM remembrance-F.SG.NOM came-F.PL.PERF
'He recalled the olden (days) chitchat.'

In none of the above sentences (155-157) the light verb *'*Come*/Aa*' agrees with its constituent nominal host 'Memory/*Yaad'* in terms of the relevant ϕ-features. The subject argument in the above sentences is marked by *ko* (the dative case marker) which blocks the verb agreement with it leaving the unmarked direct object to take over the verb-agreement. This proves that at least in these examples 'Memory/*Yaad'* is not an argument of the verb *'*Come*/Aa*' and therefore it genuinely forms the nominal component of the conjunct predicate '*Yaad-anaa'*. However, in sentences where the light verb *'*Come*/Aa*' seems to agree in number and gender with the noun 'Memory/*Yaad'*, it becomes difficult to classify '*Yaad-anaa'* as a conjunct predicate. The contrast in the following sentences (See Examples 158 and 159) clearly brings out the problem.

158.

us ko ghar yaad ayaa
him -DAT home-M.SG.NOM memory-F.SG.NOM came-M.SG.PERF
'He remembered/ missed (his) home.'

159.

us ko ghar ki yaad ayii
him -DAT home.M.SG GEN memory.F.SG came.F.SG.PERF
'He remembered/ missed (his) home.'

Both the sentences are almost synonymous in meaning and yet the verb *'*Come*/Aa*' agrees with the NP 'Remembrance of home/ *Ghar ki yaad'* headed by the noun 'Memory/*Yaad'* in terms of number and gender. Thus, *yaad-anaa* in the first sentence, 158, functions as a conjunct predicate, but in the second sentence 159, *yaad-anaa* doesn't appear to be the conjunct predicate as the light verb *'*Come*/Aa*' agrees with the noun 'Memory/*Yaad'* making it an argument of the verb.

There are two possible ways to resolve this apparent complexity. The first one is a little cumbersome which assumes that *yaad-anaa* is a conjunct predicate which is present in the underlying structure in the following form (See Example 160).

160.

? us ko[ghar ki yaad] yaad ayii
him DAT home.M.SG GEN memory-F.SG memory came-F.SG.PERF
'He remembered/ missed (his) home.'

Because the noun 'Memory/*Yaad'* is copied in close adjacency to each other, a language specific rule may be postulated for Urdu which would delete the copy of 'Memory/*Yaad'* or copy of any noun or adjective that is part of a conjunct predicate and occurs adjacent to a phonologically identical noun or adjective which is part of the argument of the light verb.

161.

us ko [ghar ki yaad] $_{\text{arg NP}}{}^{1}$ ~~yaad~~ – ayii
'He remembered/ missed (his) home.'
(1: arg NP = argument NP)

This is not a novel language specific rule to be postulated for Urdu. Similar hypotheses have been put forward while explaining the phenomenon of *re-merging* in the Minimalist Framework of Syntax (Carnie, 2012; Koeneman & Zeijlstra, 2017).

The other rather simplistic explanation is to assume that the appearance of '*Yaad-kerna'* is not an instance of the conjunct predicate. Instead, the sentence carries a transitive verb *'*Come/*Aa*' whose two arguments appear in the subject and Direct Object (DO) position and the noun 'Memory/*Yaad'* is part of the direct object NP and not a constituent of conjunct predicate (see example 161).

To summarize, verb-agreement phenomenon in Urdu, and probably many other South Asian languages, can be used as a test to identify the noun that apparently looks to be the part of a conjunct predicate but, in fact, is an argument of the light verb.

5.22 Summary

This chapter is a detailed commentary on the need to develop Urdu N+V inventory, unique features of this developed lexical resource and its probable link with Natural Language Processing (NLP) industry. It highlighted the limitations of not having only unique Urdu nouns with a repetition of same noun in both singular and plural forms

along with different morphological forms of light verbs because of data extraction from natural language corpus (Universal Dependency Urdu Treebank Corpus here). Urdu N+V inventory repeats many instances of Urdu N+V which present diverse probabilities of Urdu nouns combinations with light verbs. This inventory is not limited to only ten light verbs selected for the present work. In Universal Dependency Urdu Treebank Corpus when a noun is linked with a verb with a dependency label 'compound' are chosen as instances of Urdu conjunct predicate (N+V). In this process, all naturally occurring light verbs linked with host nouns and labeled as compound are extracted. For the next research objective of drawing and mapping of semantic classes of Urdu nouns in conjunct predicate, a list of 280 unique Urdu N+V instances are drawn after adding Urdu data from Urdu WordNet Wordlist 1.0. The process of selection and then mapping of semantic senses of Urdu nouns to English WordNet was tried to make it systemic and less subjective by following the intuition of native speaker of the Urdu language. Detailed and laborious manual procedure of checking the compatibility of each Urdu noun with all ten light verbs yielded a large language data which is presented in the 4th chapter. After the comprehensive compilation of language data, there emerged a pattern of finding only 14 semantic classes of Urdu nouns which were compatible with some of the light verbs chosen for the study. Most of the Urdu nouns which were compatible with light verbs were abstract and could only be mapped onto 14 semantic classes. It helped to emerge an inventory of 'conjunctive Urdu noun classes' which are low [-RF] on their 'nouniness' and are not first order nouns'. Conjunctive Urdu nouns are low on their referential power and thus may be called 'light nouns'.

Later in chapter these 15 conjunctive Urdu nouns are separated from the list of 280 nouns, and they are dealt with detail regarding their important features such as semantic orientation, compatibility with light verbs, argument structure, case marking and semantic roles.

Most conjunctive Urdu nouns exhibited similar traits which include the licensing of ergative case and agentive role of the subject argument with the perfective aspect realization of the clause. In case of imperfective aspect, subject argument takes nominative case and experiencer semantic role. In the spectrum of 15 conjunctive Urdu noun classes, a very small number of Urdu nouns refer to tangible objects of the world with the majority referring to abstract entities. The element of volitionality seems linked with things: perfective aspect, semanticity of host noun, transitivity of light verbs, case

marking and semantic role of subject argument. With imperfective aspect in the clause, most conjunctive Urdu noun classes when compatible with light verbs, license nominative case on the subject arguments along with their agentive role. A tendency of equal compatibility instances of Urdu nouns with both transitive 'Do/*Kar'* and intransitive light verb 'Become/*Hu'* is seen. A set of intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'* exhibits similar pattern of compatibility within the same conjunctive Urdu noun. A similar compatibility pattern is shown by a set of two ditransitive light verbs 'Give/*Dia'* and 'Take/*Lee'* with very few exceptions which are listed in an elaborate manner. Semantic classes of Urdu nouns when compatible with di/transitive light verbs unveiled a correlating pattern of case marking, sematic roles and aspect which are also mentioned separately in tabulated form. *Hasil Karna* is a frequently found instance of N+V which is usually added to an Urdu noun such as 'Success/*Kamyabi'.* Hasil is an Urdu noun which may be semantically paraphrased to 'Get/*Pana* or 'Take/*lena'* but it adds to formality of the clause. So, it would be more appropriate if it says *Kamyabi hasil karna* instead of *Kamyabi lena* (to get success). It is an instance of N_{N+V}+V which can be studied in some future study.

This section sums up the retrieved categorization of Urdu conjunct predicates with a discussion of their probable mechanism of insertion in Urdu WordNet and its possible impact on Urdu computational programs. In all this process of the insertion of Urdu conjunct predicates to Urdu WordNet, their true identification is very important which is discussed in a detailed manner using examples after applying agreement test. This tool of using agreement test proved to be an effective way to decide on true nature of Urdu conjunct predicates which can easily be mapped and inserted in the lexical tool such as Urdu WordNet. Finally, it can be asserted that the present study is an elaborate account of Urdu conjunct predicates which bridges the knowledge gap and thus helps to solve the problem of lack of useful lexical Urdu resources needed to augment the low resourced status of the Urdu language.

Chapter 6: Conclusion

In the present work, the knowledge gap related to Urdu conjunct predicate (N+V) such as an inventory of Urdu N+V, semantic classes of Urdu nouns in conjunct predicate, combinatory restrictions between host noun and light verb, and a tool to test true Urdu conjunct predicates is attempted to bridge. The Urdu language with a speech community of more than 280 million people and having an institutional status is widely used for different domains of life such as education, print and digital media, judiciary, and administration. Despite a large collection of print and digitized literature, and the spread of functionalities, Urdu continued to be less resourced in terms of digital lexical resources to be used in Natural Language Processing (NLP). The study of conjunct predicate in Urdu, and the ways Urdu nouns are combined with light verbs to make a single constituent has been a bottleneck for NLP experts. Complex predicate is widely used in Urdu where a vector verb is cojoined with light

verb, or a noun host comes with a light verb to produce compound and conjunct predicates respectively. Conjunct predicate (N+V), a type of complex predicate, instances are so productive that to date not an exhaustive list of Urdu conjunct predicate has been developed. There was a knowledge gap of ontological information about the combinatory restrictions between Urdu nouns and light verbs. For this, a good sizeable inventory of Urdu conjunct predicate is required. It may help in anticipating the Urdu N+V instances, identifying the semantic classes of Urdu nouns and mapping them onto English WordNet. Such information may not only enable a computational application to extract Urdu conjunct predicates from a corpus, but also contribute to the precise interpretation of the incorporated noun host with light verb as a conjunct predicate. The mutual compatibility pattern of Urdu nouns and light verbs helps in understanding the combinatory restrictions between them. The theory related to the connection between the semantics and syntactic context of a clause helps to draw patterns which can be used as linguistic information required for computational linguistics. Furthermore, not all instances of N+V can be recognized as true conjunct predicate because all noun hosts are not incorporated into verbal predicate. And in case of non-incorporation, these nouns will be arguments of predicate. It required a tool which can identify true Urdu conjunct predicate. As a result, only the true conjunct predicate can be inserted in a lexical resource such as WordNet as a single constituent.

6.1 Research Outcomes

The development of an inventory of Urdu N+V instances was a much-needed linguistic resource for Urdu which may provide a basis of semantic and syntactic combinatory restrictions. Emerged themes of semantic classes of host noun and their compatibility with light verbs explicitly describe the semantic-syntactic orientation of Urdu Conjunct predicate N+V. Application of agreement test may also highlight the incorporated nature of host noun which would be an argument otherwise in case of non-compliance to the agreement test.

The first contribution of the present work is to develop an Urdu N+V inventory. For this purpose, the first step Universal Dependency Urdu Treebank Corpus was chosen due to its multilayer annotated nature (Bhat et al. 2017). This Urdu N+V inventory included compound nouns whose concatenating morphological process can be discussed in some future study (Levin et al., 2019). List of these compound Urdu nouns has also been given in the 4th chapter (page 203). Compound nouns found in the extracted noun list belong to three categories morphologically: N+N, N+V; and N+ S_o (semantically void lexical item) e.g., 'Attention/*Qoro-fiqar'*, 'Planning/*Munsooba-bandi'* and 'Search/*Pooch-gach'.* Varied and, but not limited to Noun Act and Noun Cognition, varied conjunctive Urdu noun classes got representation in these compound nouns. Surprisingly, all compound nouns are compatible with transitive light verb 'Do/*Kar'* which joined with these compounds to allow ergative case marker. Resultantly, it assigns an agentive role to the overtly marked subject argument along with the volitional element marked with the conscious choice to do so. This Urdu N+V inventory also listed some borrowed Urdu nouns which otherwise in source language (English) also behave as verb, but they require a light verb in the Urdu language to give the verbal meaning. 90 % of the borrowed nouns found in the list belong to noun class 'Noun Act' which correlated with their compatibility with transitive light verb 'Do/*Kar'* and then consequently license the ergative case marker clearly found with the element of volitionality.

Then, to introduce a dialectal variation, Urdu N+V instances are also mined from Urdu WordNet Wordlist 1.0 (Wordlist, 2013). Extracted 280 Urdu nouns from these two Urdu corpora are checked for their compatibility with ten light verbs: 'Do/*Kar',* 'Become/*Hu*', 'Be/ *He',* 'Put/*Rakh',* 'Come/*Aa*', 'Give'*Dia,* 'Go/ *Ja*', 'Take/*Lena*', 'Remain/*Rah',* and 'Hit /*Laga*'. A list of Urdu nouns found in combination with light

verbs are mapped onto English WordNet looking at their semantic senses and then their respective noun classes are documented according to the noun taxonomy introduced by Miller et al., (1990). The resultant noun class composition yielded after this investigation reflected the emerging pattern of conjunctive Urdu noun classes (See Table 41). It clearly indicated the semantic classes of Urdu nouns which show compatibility with light verbs. Conjunctive Urdu noun classes are semantic classes of nouns which refer to mostly abstract entities rather than tangible objects found in the world. The names of these emerged conjunctive Urdu noun classes are quite reflective of their referred concepts: Noun Act, Noun Communication, Noun Attribute, Noun Time, Noun State, Noun Possession, Noun Relation, Noun Process, Noun Phenomenon, Noun Person, Noun Artifact, Noun Feeling, Noun Event and Noun Cognition. It is found that noun classes referring to tangible objects are less likely to occur with light verbs. In case of occurrence, their instances are really limited. I found two Conjunctive Urdu noun classes, Noun Artifact and Noun Person, which refer to concrete and tangible objects but in holistic way. For example, only two Noun Artifacts have been found which are in antonym semantic relation with each other i.e., 'Export/*Daramad'* and 'Import/*Baramad'*. And 'Victim/*Shikar'* is the only Noun Person in the list of conjunctive Urdu noun classes which refer to an entity which has been affected in some way. In my perception, 'Victim/*Shikar'*, is not necessarily a person but it can also be any living entity, plant or animal etc. which has been victimized, but in WordNet there has been only one noun class i.e., Noun Person which has been mentioned for lexical item 'Victim'. Hence, the noun classes in WordNet can also undergo a revision, and missing senses or classes can be inserted. This phenomenon of noun class specificity in case of host nouns in conjunct predicate can be tested for its cross linguistic validity. For example, the found proposition of having predominately abstract nouns as a host in conjunct predicates can also be tested in other languages such as Burushaki, Punjabi, etc.

Another reason to emphasize the abstract nature of conjunctive Urdu noun classes in N+V instances is the metaphoric conceptual representation of 'Poison/*Zehar*' which does not denote the physical substance 'poison' but refers to 'something hard/unpleasantnt to bear'. Similarly, *Zakham dena* means to inflict pain and not the literal meaning of 'Wound/*Zakham*'. Now, with the detailed examination of 280 unique nouns, it may be clearly asserted that conjunctive Urdu noun classes mostly refer to the

abstract entities without concrete physical realization. The metaphoric representative concept of noun host in Urdu conjunct predicate can make the basis of an advance work which can then be inserted in Urdu WordNet as mostly these semantic senses are missing.

Some words with similar semantic sense such as 'Advice/*talqeen* and *mashwara*' and 'Movement/ *jumnbish* and *harqat*' when tested for their syntactic contexts show similar constructions and compatibility pattern with light verbs. This phenomenon highlights the matching syntactic realization of Urdu nouns in different semantic relations such as synonyms. It fortifies the Levin (1993) proposition of connectedness between meaning and the probable sentence constructions. Further investigation of the proposition for large number of synonyms and antonyms of Urdu nouns in N+V instances can help to amass the entries in conjunctive Urdu noun classes discovered in the present research work.

For this study, Urdu nouns in semantic relations such as hyponymy, antonymy and synonymy are concerned, are intentionally avoided in the list of N+V instances just as an effort to encompass varied and unique semantic senses (Levin, 1989). Semantic class of an Urdu noun influences the argument structure, case marking and semantic role of the subject argument. This relatedness can be investigated for nouns with similar semantic sense. These synsets say synonyms are pilot tested to have similar syntactic realization to validate the proposition of semantic and syntactic relatedness as advocated for English verbs (Levin, 1985, 1989, 1993). This validation can further be investigated for words in other semantic relations such as hyponymy and in some future study.

Urdu nouns included in the present are common nouns and due to the open class of noun category, continuous influx of more lexical items is expected. If the hypothesis related to the similar syntactic context of synonyms is confirmed because of above-mentioned proposed future study, then it is likely to put the proper nouns and new added nouns to the hierarchical hyponymic relationship as their hyponyms. Noun class within an inheritance system of hypernymy mostly remain the same (Miller et al., 1990). This is the way more nouns can be added to enrich the entries in the lexical resource of conjunctive Urdu noun classes. It paves the way for more connected future studies and may contribute to augment process of the digitization of the Urdu language.

Intransitive light verbs i.e., 'Come/*Aa'*, 'Go/ *Ja'*, 'Is/ *He'* and 'Become/*Hu'* in Urdu N+V instances when mapped onto English WordNet revealed their semantic verb class

'Verb Stative'. On the other hand, the transitive light verbs i.e., 'Do/*Kar'*, 'Remain/ *Rah'*, and 'Hit/*Laga'*, are found on the spectrum of Verb Creation, Verb Change, and Verb Contact. 'Take/*Lee'*, 'Put/ *Rakh'* and 'Give/*Dia'* are three ditransitive light verbs which are mapped as Verb Possession.

Whenever a conjunctive Urdu noun class cojoins with the transitive light verb 'Do/*Kar'*, it permits an ergative case marker along with the assigning the agentive role to the subject argument in the presence of perfective aspect in the clause. Another prominent semantic information carried by this collocation of N+V is the conscious choice of agent to perform an action. Compatibility of an intransitive light verb 'Become/*Hu'* though equal in numbers for most conjunct Urdu noun classes, but it is devoid of assigning the agentive role to the subject argument and therefore the element of volitionality is also absent. For large number of conjunctive noun classes, intransitive light verbs 'Come/*Aa'* and 'Go/*Ja'* showed similar compatibility pattern with light verbs. Similarly, most of the conjunctive Urdu noun classes if compatible with 'Give/*Dia*', tend to exhibit parallel compatibility with 'Take/*Lee*'. The semantic sense of noun and the argument structure of light verb together influence the case marking and the semantic role of the subject argument in the construction. When investigated, noun class with transitive light verbs forms two categories of Urdu conjunct predicate: one with ergative case marker (A) and other with dative/genitive/locative case marker on subject argument (B). Former category (A) assigns an agentive role to the subject argument whereas it is experiencer in case of later. Category (A), 'Give/*Dia'* and 'Take/*Lee',* 'Do/*Kar'*, and 'Put/ *Rakh'* seem to allow ergative or nominative case marker which then contribute to assign agentive role to the subject argument along with perfective and imperfective aspects in the clauses respectively. Category (B) of transitive light verbs, 'Remain/ *Rah'* and 'Hit/*Laga'* allow dative, genitive and locative case markers to the subject argument whose semantic role is not agentive.

Hence the ergative case is linked to the agentive role of subject role because of specific semantic noun class along with effect of selective compatible transitive light verb i.e., category (A) transitive light verbs. Whereas category (B) of transitive light verbs is void of the capability to assign agentive role to the subject argument. So, it can be asserted that not all transitive light verbs found in Urdu N+V instance assigns agentive role to the subject argument along with the overt ergative case marker. The semantic verb classes of category (A) light verbs include 'Verb Possession', Verb Creation', and

'Verb Contact'. Whereas the semantic verb classes of transitive light verb of (B) transitive light verbs are 'Verb Contact' and 'Verb Change'.

Addressing the third research objective, this research study elicited the range of possible syntactic constraints on the combinatory possibilities of noun and light verb in Urdu N+V instances. The discovered theory of semantic and syntactic combinatory restrictions between the compatible host Urdu nouns in N+V instance with transitive or intransitive light verbs elicits the relatedness between the semantic value and syntactic orientation of nouns and light verbs. It proved the validity of Levin's (1993) proposition about the connectedness between the semantic value of a lexical item and its syntactic context for the Urdu language as well. Levin (1993) presented the taxonomy of English verbs on the theory of interrelation between the semantic values and the syntactic context. A similar theoretical framework is employed to draw the semantic classes of Urdu nouns in Urdu conjunct predicate (N+V).

It helped to understand the combinatory restrictions of different semantic classes of nouns to collocate with light verbs. Ranging from licensing of case marking and its effect on the semantic role of the subject argument, everything seemed connected to the semantic senses of nouns and the argument structure of light verbs.

Now addressing the fourth research objective of the study, the identification of true conjunct predicate in Urdu, I need to cast an intense and holistic overview of everything which has been discovered up till now in the study.

Predicate agrees to the leftmost unmarked argument. In case of no unmarked argument present in the construction, it assumes the default agreement with 3rd person singular masculine. If an unmarked noun agrees to the predicate, it indicates that it is an argument and it cannot be called as an incorporated noun into the light verb. Not every noun which is adjacent to light verb cannot be a nominal host of conjunct predicate (N+V). Urdu Noun host in true conjunct predicate does not trigger any agreement with light verb. Hence, noun host in conjunct predicate cannot be an argument in the construction.

6.2 Research Limitation

Identification of true conjunct predicate through agreement test is not a very simple phenomenon. Constructions with similar noun host and light verb may pose two agreement patterns: (1) no agreement between noun host and light verb in the presence

of leftmost unmarked direct object as an agreeing argument, (2) agreement between noun host and light verb in case of absence of any other unmarked argument leaving noun host as the leftmost unmarked argument. In fact, in the case of point (2) mentioned above, the agreement is between Noun Phrase (NP) with host noun as its head and the light verb. This situation presented a dual but contrastive postulation of the situation. One that resolves the issue through *re-merging* in the Minimalist Framework of Syntax and declared it as a 'conjunct predicate' where a similar copied noun adjacent to a noun argument is assumed to be deleted. Others simply reject this N+V instance as a true conjunct predicate.

To a large extent, it resolved the issue of identification of true conjunct predicate and paved the way for their insertion into WordNet as single lexical item having a distinct semantic sense. Results of application of agreement test reflect that not all Urdu noun classes could pass the criteria of agreement test. Majority of the nouns which passed the agreement belong to the conjunctive noun classes. This proposition is strong enough and may nullify the claims of apparent merger of Urdu noun with light verbs in a related study (Fatima, 2021).

Addressing to the devised research objective of developing an inventory of Noun + Light Verb instances from Urdu corpora, a paper has been published (Abdullah et al., 2021). This paper contains an inventory of Urdu N+V instances extracted from Urdu Universal Dependency Treebank corpus. These Urdu N+V instances may be declared as conjunct predicate after they qualify the agreement test and other criteria defined for the identification of Urdu Conjunct predicate. This paper is getting attention by researchers who want to work on the semantic frames of light verbs in Urdu conjunct predicate (Jehangir & Azher, 2022).

Development of Urdu N+V inventory was a milestone in achieving the next research objective of investigating the semantic classes of noun in Urdu N+V conjunct predicate. Research revealed the clear taxonomy of 15 conjunctive Urdu noun semantic classes obtained by mapping the semantic senses onto WordNet. Predominantly, conjunctive Urdu noun classes refer to abstract entities which are also likely to pass the agreement test used for the identification of true Urdu conjunct predicates.

While investigating the semantic classes of nouns, the range of possible syntactic constraints on the combinatory possibilities of noun and light verb in Urdu N+V

instances is explicitly documented with clear information about connection between case marking, argument structure, aspect and semantic roles.

A tool for the identification of true Urdu conjunct predicate was a research aim achievement of which was heavily depended on already realized objectives of the research. Wading through the inventory of Urdu N+V, conjunctive noun semantic classes along with their combinatory restrictions and connection between semantic and syntactic context, agreement test was another parameter to conclude the recognition of true Urdu conjunct predicate. Here the incorporated nature of noun host and its adjacency to the light verb came out as an integral proposition related to the recognition of conjunct predicate phenomenon.

Conducting a Linguistics and Multilingual Studies Research seminar at School of Humanities, Nanyang Technological University, Singapore on September 30, 2022 provided me an opportunity to present my findings to a very learned audience. Their feedback and questions gave me a chance to improve my work.

In the seminar, Prof. Randy LaPolla inquired about the distinction between default Masculine Singular (M. SG.) agreement and an overt Masculine Singular (M. SG.). It helped me to draft a clear distinction. When there is no unmarked argument, it facilitates default M. SG. agreement. Overt M. SG. agreement occurs to the leftmost unmarked argument with M. SG phi features. 'Default agreement M. SG' is indifferent to the phi features of arguments (number, gender, tense) due to the presence of overt case markers which block their agreement from the predicate. Whereas the unmarked leftmost argument tends to show its agreement with the predicate which clearly depicts its phi features. It is not necessarily M. SG. as it depends on the features of unmarked arguments. So, it can be concluded that *case-marking* and *position* of arguments define the phenomenon of no agreement, agreement, and default agreement. This is how we can trace the difference between default M. SG. and Overt M. SG. Agreement.

Prof. K. K. Luke gave me an opportunity to explicitly talk about the exact number of Urdu N+V instances I explored to obtain answers to my research questions. I had a list of 280 Urdu nouns extracted from two Urdu corpora which were tested for their compatibility with 10 light verbs, thus making it a data of 2800 test Urdu N+V im/possibilities. In addition to this, he indicated the possibility of interesting phenomenon of testing compatibility between Noun and Light verb in other languages. Resultantly, the research provides a criterion for the insertion of true conjunct predicate as single constituent in Urdu digital resource like WordNet. It may augment the status

of the Urdu language as a digitally enabled language. Development of N+V inventory, semantic classes of Urdu noun host, semantic-syntactic combinatory restrictions, and a tool to identify true conjunct predicate will help computational linguistics to improve the efficiency of Natural Language Processing programs such as Word Sense Disambiguation, Information Retrieval, Machine Translation, Speech to Text, and Text to Speech, etc. for the Urdu language.

6.3 Recommendation for Future Studies

The present study was an effort to map the constituents of Urdu N+V onto WordNet. To further advance the study, the semantic classes of resultant true conjunct predicate (N+V) can be investigated and mapped onto WordNet to boost the cross linguistic interlinked semantic lexical resource cross linguistically. Being an open class category, insertion of more nouns in the Urdu language extends the scope of the future work in this area. The research could have been more fruitful and comprehensive if Urdu nouns in semantic relations such as synonyms, antonyms and hyponyms are included in the study to test the connection between their similar syntactic constructions. Noun being an open class word undergoes a regular addition to the language lexicon. Newly added nouns which may be hyponyms of some existing hypernyms, can be added to the hierarchical inheritance system of hyponym. It will also provide a chance to investigate the uniform semantic class of words at different hierarchical in a hyponym relation. It could have been increased the size of Urdu N+V conjunct predicate inventory, but due to time constraints it could not be done in this study.

References

Abdullah, F., Ahmed, T., & Anjum, U. (2021). Urdu conjunct predicates (N+ V) inventory from Urdu Universal Dependency Corpus. *Corprorum: Journal of Corpus Linguistics*, *4*(1), 15–29. https://journals.au.edu.pk/ojscrc/index.php/crc/article/view/265/214

Abid, M., Habib, A., Ashraf, J., & Shahid, A. (2018). Urdu word sense disambiguation using machine learning approach. *Cluster Computing*, *21*(1), 515–522. https://doi.org/10.1007/s10586-017-0918-0

Ackerman, F., & LeSourd, P. (1997). Toward a lexical representation of phrasal predicates. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 67–106). Cambridge University Press.

Adeeba, F., & Hussain, S. (2011). Experiences in building Urdu WordNet. *Proceedings of the 9th Workshop on Asian Language Resources*, 31–35. https://aclanthology.org/W11-3406.pdf

Ahmed, T. (2010). The interaction of light verbs and verb classes of Urdu. *Proceedings of Verb 2010*, 10–15. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=99d4d8eeac2919a2572833cb863936b2363eae79#page=21

Ahmed, T., & Butt, M. (2011). Discovering semantic classes for Urdu NV complex predicates. *Proceedings of the Ninth International Conference on Computational Semantics (IWCS 2011)*. https://aclanthology.org/W11-0132.pdf

Ahmed, T., Butt, M., Hautli, A., & Sulger, S. (2012). A reference dependency bank for analyzing complex predicates. *LREC*. Eighth conference on international language resources and Evaluation, Istanbul, Turkey. https://kops.uni-konstanz.de/entities/publication/bd29c87c-137c-4c69-ad87-b278f8f7c52b

Ahmed, T., & Hautli, A. (2011). A first approach towards an Urdu WordNet. *Proceeding of Conference on Language and Technology 2010*, 1–14. https://kops.uni-konstanz.de/entities/publication/4f5a2e5a-cbcf-41ed-902d-73366bf7a320

Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices: A typology of noun categorization devices*. OUP Oxford.

Aikhenvald, A. Y. (2006). Classifiers and noun classes: Semantics. In K. Brown (Ed.), *Encyclopedia of languages and linguistics* (Vol. 1, pp. 463–471). Elsevier. http://www1.elsevier.com/homepage/sal/ell2/

Aissen, J. L., & Perlmutter, D. M. (1983). 10 Clause reduction in Spanish. In D. M. Perlmutter (Ed.), *Studies in relational grammar* (Vol. 2, pp. 1–30). The University of Chicago Press.

Akhtar, R. N. (2000). *Aspectual complex predicates in Punjabi.* [PhD Thesis]. University of Essex.

Alsina, A. (1993). *Predicate composition: A theory of syntactic function alternations.* (Vol. 1–I and II). Stanford University.

Alsina, A. (1997). A theory of complex predicates: Evidence from causatives in Bantu and Romance. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 203–246). Cambridge University Press.

Alsina, A., & Mchombo, S. (1990). The syntax of applicatives in Chichewa: Problems for a theta theoretic asymmetry. *Natural Language & Linguistic Theory*, *8*(4), 493–506.

Amberber, M., Baker, B., & Harvey, M. (2010). *Complex predicates: Cross-linguistic perspectives on event structure*. Cambridge University Press.

Andow, J. (2015). How "intuition" exploded. *Metaphilosophy*, *46*(2), 189–212. https://doi.org/10.1111/meta.12127

Atkins, B. T. S., & Levin, B. (1995). Building on a corpus: A linguistic and lexicographical look at some near-synonyms. *International Journal of Lexicography, 8*(2), 85–114. https://doi.org/10.1093/ijl/8.2.85

Baker, L. D., & McCallum, A. K. (1998). Distributional clustering of words for text classification. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 96–103. https://dl.acm.org/doi/pdf/10.1145/290941.290970

Baker, M. C. (1997). Complex predicates and agreement in polysynthetic languages. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 247–288). Cambridge University Press.

Bashir, E. (1993). Causal chains and compound verbs. In M. Verma (Ed.), *Complex predicates in South Asian languages*. Manohar Publishers.

Beckwith, R., Fellbaum, C., Gross, D., & Miller, G. A. (2021). WordNet: A lexical database organized on psycholinguistic principles. In U. Zernik (Ed.), *Lexical acquisition: Exploiting on-line resources to build a lexicon* (pp. 211–232). Psychology Press.

Beth, L., Hovav, M. R., & Keyser, S. J. (1995). *Unaccusativity: At the syntax-lexical semantics interface* (Vol. 26). MIT press.

Beth, L., & Rappaport, M. (1988). Nonevent-er nominals: A probe into argument structure. *Linguistics, 26*(6), 1067–1084. https://doi.org/10.1515/ling.1988.26.6.1067

Bhat, R. A., Bhatt, R., Farudi, A., Klassen, P., Narasimhan, B., Palmer, M., Rambow, O., Sharma, D. M., Vaidya, A., Ramagurumurthy Vishnu, S., & Xia, F. (2017). The Hindi/Urdu Treebank Project. In N. Ide & J. Pustejovsky (Eds.), *Handbook of*

linguistic annotation (pp. 659–697). Springer Netherlands. https://doi.org/10.1007/978-94-024-0881-2_24

Bhatt, R. (2005). Long distance agreement in Hindi-Urdu. *Natural Language & Linguistic Theory*, *23*(4), 757–807.

Bhatt, R. (2008). *NV compounds and Agreement*. http://people.umass.edu/bhatt/papers/eflu-aug18.pdf

Bhattacharyya, P., Chakrabarti, D., & Sarma, V. M. (2006). Complex predicates in Indian languages and WordNets. *Language Resources and Evaluation*, *40*, 331–355. https://doi.org/10.1007/s10579-007-9032-x

Bond, F., Isahara, H., Fujita, S., Uchimoto, K., Kuribayashi, T., & Kanzaki, K. (2009). Enhancing the Japanese WordNet. *Proceedings of the 7th Workshop on Asian Language Resources*, 1–8. https://aclanthology.org/W09-3401.pdf

Bowern, C. (2008). The diachrony of complex predicates. *Diachronica*, *25*(2), 161–185. https://doi.org/10.1075/dia.25.2.03bow

Bril, I. (2007). Nexus and juncture types of complex predicates in Oceanic languages: Functions and semantics. *Language and Linguistics*, *8*(1), 267–310. https://shs.hal.science/halshs-00281187/

Brinton, L. J. (2008). "Where grammar and lexis meet": Composite predicates in English. In E. Seoane, L. Couso, M. José, & T. Fanego (Eds.), *Theoretical and empirical issues in grammaticalization* (Vol. 77, pp. 33–53). John Benjamins. https://www.jbe-platform.com/content/books/9789027290212-tsl.77.04bri

Butt, M. (1993a). Conscious choice and some light verbs in Urdu. In M. Verma (Ed.), *Complex predicates in South Asian languages* (Vols. 31–46, p. 45). Manohar Publishers.

Butt, M. (1993b). Hindi-Urdu infinitives as NPs. In Y. Kachru (Ed.), *South Asian language review: Special issue on studies in Hindi-Urdu* (Vol. 3, pp. 51–72). South Asian Language Association of Australia (SALAA)

Butt, M. (1995). *The structure of complex predicates in Urdu*. Center for the Study of Language (CSLI). https://ojs.ub.uni-konstanz.de/jsal/dissertations/diss-butt.pdf

Butt, M. (1998). Constraining argument merger through aspect. In E. Hinrichs, A. Kathol, & T. Nakazawa (Eds.), *Complex predicates in nonderivational syntax* (Vol. 30, pp. 73–113). The Academic Press. https://doi.org/10.1163/9780585492223_004

Butt, M. (2010). The light verb jungle: Still hacking away. In M. Amberber, B. Baker, & M. Harvey (Eds.), *Complex predicates: Cross-linguistic perpectives on event structure.* (pp. 48–78). Cambridge University Press.

Butt, M. (2019). Complex predicates and multidimensionality in grammar. *Linguistic Issues in Language Technology*, *17*(4), 1–14. https://journals.colorado.edu/index.php/lilt/article/view/1425

Butt, M., Bögel, T., Hautli, A., Sulger, S., & Ahmed, T. (2012). Identifying Urdu complex predication via bigram extraction. *International Conference on Computational Linguistics*, 409–424. https://kops.uni-konstanz.de/server/api/core/bitstreams/2e710dca-4097-4872-884e-1f1e08f9b946/content

Butt, M., & Geuder, W. (2001). On the (semi) lexical status of light verbs. In N. Corver & H. Riemsdijk (Eds.), *Semilexical categories: On the content of function words and the function of content words* (pp. 323–370). De Gruyter Mouton. https://doi.org/10.1515/9783110874006

Butt, M., & King, T. H. (2007). Urdu in a parallel grammar development environment. *Language Resources and Evaluation*, *41*(2), 191–207. https://link.springer.com/article/10.1007/s10579-007-9042-8

Butt, M., & Ramchand, G. (2001). Complex aspectual structure in Hindi/Urdu. In M. Liakata, B. Jensen, & D. Maillat (Eds.), *Oxford University Working Papers, in Linguistics, Philology & Phonetics* (Vol. 6, pp. 1–30). University of Hawaii at Manoa. https://www.ling-phil.ox.ac.uk/files/owp2001.pdf#page=7

Carnie, A. (2012). *Syntax: A generative introduction* (Vol. 18). John Wiley & Sons.

Cattel, R. (1984). Composite predicates in English. *Syntax and Semantics*, *17*, 5–12. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=12073919

Chomsky, N. (1966). Explanatory models in linguistics. In E. Negal, P. Suppes, & A. Tarski (Eds.), *Studies in logic and foundations of mathematics* (Vol. 44, pp. 528–550). Elsevier. https://www.sciencedirect.com/science/article/pii/S0049237X09706172

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. MIT Press.

Chomsky, N. (1992). Explaining language use. *Philosophical Topics*, *20*(1), 205–231. https://www.jstor.org/stable/43154643

Chomsky, N. (2002). *On nature and language*. Cambridge University Press.

Chomsky, N. (2013). *Studies on semantics in generative grammar* (Vol. 107). Walter de Gruyter.

Chua, S., & Kulathuramaiyer, N. (2004). Semantic feature selection using WordNet. *IEEE/WIC/ACM International Conference on Web Intelligence (WI'04)*, 166–172. https://doi.org/10.1109/WI.2004.10115

Ciaramita, M., & Johnson, M. (2003). Supersense tagging of unknown nouns in WordNet. *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 168–175. https://aclanthology.org/W03-1022

Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press. https://books.google.com/books?hl=en&lr=&id=xDSBaet2uSsC&oi=fnd&pg=PR11&dq=lexical+semantics&ots=9D7bmfMQAg&sig=SAH8PwJVj9Jc_hUXzMgopgfo9jM

Davies, W., & Rosen, C. (1988). Unions as multi-predicate clauses. *Language*, *64*, 52–88. https://doi.org/10.2307/414785

De Marneffe, M.-C., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal dependencies. *Computational Linguistics*, *47*(2), 255–308. https://doi.org/10.1162/coli_a_00402

Dione, C. M. B. (2020). From LFG To UD: A Combined Approach. *Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020)*, 57–66. https://aclanthology.org/2020.udw-1.7

Downing, P. (1977). On the creation and use of English compound nouns. *Language*, *53*(4), 810–842. https://doi.org/10.2307/412913

Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: Semantics or pragmatics? *Linguistics and Philosophy*, *9*(1), 37–61. https://www.jstor.org/stable/25001231

Dukes, M. (2000). The projection of arguments: Lexical and compositional factors. *Language*, *76*(2), 476–477. https://doi.org/10.1353/lan.2000.0069

Durie, M. (1997). Grammatical structures in verb serialization. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 289–354). Cambridge University Press.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2023). *Ethnologue: Languages of world*. Ethnologue (Essentials). https://www.ethnologue.com/

Ehsan, T., & Butt, M. (2020). Dependency parsing for Urdu: Resources, conversions and learning. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, 5202–5207. https://aclanthology.org/2020.lrec-1.640/

Evans, N. (1997). Role or cast? Noun incorporation and complex predicates in Mayali. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 397–430). Cambridge University Press.

Fatima, G., Nawab, R. M. A., Khan, M. S., & Saeed, A. (2021). Developing a cross-lingual semantic word similarity corpus for English–Urdu language pair. *Transactions on Asian and Low-Resource Language Information Processing*, *21*(2), 1–16. https://dl.acm.org/doi/abs/10.1145/3472618

Fedson, V. J. (1993). Complex verb-verb predicates in Tamil. In M. Verma (Ed.), *Complex predicates in South Asian languages*. Manohar Publishers and Distributors.

Fellbaum, C. (1990). English verbs as a semantic net. *International Journal of Lexicography*, *3*(4), 278–301. https://doi.org/10.1093/ijl/3.4.278

Fellbaum, C. (2010). WordNet. In R. Poli, M. Healy, & A. Kameas (Eds.), *Theory and applications of ontology: Computer applications*. Springer,Dordrecht. 10.1007/978-90-481-8847-5_10

Fillmore, C. J. (1968). Lexical entries for verbs. *Foundations of Language*, *4*(4), 373–393. https://www.jstor.org/stable/25000345

Fodor, J. A. (1970). Three reasons for not deriving “kill” from “cause to die”. *Linguistic Inquiry*, *1*(4), 429–438. https://www.jstor.org/stable/4177587

Foley, W. A. (1997). Polysynthesis and complex verb formation: The case of applicatives in Yimas. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 355–395). Center for the Study of Language (CSLI).

Foley, W. A., & Van Valin Jr, R. D. (2009). *Functional syntax and universal grammar*. Cambridge University Press. http://hdl.handle.net/11858/00-001M-0000-0013-442C-F

Folli, R., Harley, H., & Karimi, S. (2005). Determinants of event type in Persian complex predicates. *Lingua*, *115*(10), 1365–1401. https://doi.org/10.1016/j.lingua.2004.06.002

Gair, J. W., & Wali, K. (1989). Hindi agreement as anaphor. *Linguistics*, *27*(1). https://doi.org/10.1515/ling.1989.27.1.45

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, *1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Godard, D., & Samvelian, P. (2021). Complex predicates. In S. Müller, A. Abeillé, R. D. Borsley, & J.-P. Koenig (Eds.), *Head-driven phrase structure grammar* (p. 419). Language Science Press.

Grimshaw, J., & Mester, A. (1988). Light verbs and θ-marking. *Linguistic Inquiry*, *19*(2), 205–232.

Gross, D., & Miller, K. J. (1990). Adjectives in WordNet. *International Journal of Lexicography*, *3*(4), 265–277. https://doi.org/10.1093/ijl/3.4.265

Gruber, J. (1976). *Lexical structures in syntax and semantics* (Vol. 25). North-Holland Publishing Company.

Hale, K., & Keyser, S. J. (1997). On the complex nature of simple predicators. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (Vol. 64, pp. 29–65). Center for the Study of Language (CSLI).

Hamp, B., & Feldweg, H. (1997). Germanet-a lexical-semantic net for German. *ACL/EACL-97 Workshop Proceedings*, 9–15. https://aclanthology.org/W97-0802.pdf

Hautli-Janisz, A. (2014). *Urdu/Hindi motion verbs and their implementation in a lexical resource* [PhD Thesis]. University of Konstanz. https://kops.uni-konstanz.de/handle/123456789/28820

Hearst, M. A. (1992). Automatic acquisition of hyponyms from large text corpora. *The 14th International Conference on Computational Linguistics*, *2*, 539–545. https://aclanthology.org/C92-2082.pdf

Hearst, M. A. (1998). Automated discovery of WordNet relations. In C. Fellbaum (Ed.), *WordNet: An electronic lexical database* (Vol. 2, pp. 131–150). The MIT Press.

Higginbotham, J. (1985). On semantics. *Linguistic Inquiry*, *16*(4), 547–593. https://www.jstor.org/stable/4178457

Hook, P. E. (1972). *The compound verb in Hindi*. University of Pennsylvania.

Hook, P. E. (1978). The Hindi compound verb: What it is and what it does. In K. S. Singh (Ed.), *Readings in Hindi-Urdu linguistics* (pp. 129–154). National Publishing House.

Jackendoff, R. (1985). *Semantics and cognition* (Vol. 8). MIT press.

Jackendoff, R. (1988). Conceptual semantics. In U. Eco (Ed.), *Meaning and mental representations* (Vol. 1, pp. 81–97). Indiana University Press.

Jackendoff, R. (1992). *Semantic structures* (Vol. 18). MIT Press.

Jehangir, H., & Azher, M. (2022). Semantic Frames of the Urdu conjunct verb lagnā: A corpus-based study. *Corporum: Journal of Corpus Linguistics*, *5*(1), 1–23. https://eprints.lancs.ac.uk/id/eprint/174494

Jespersen, O. (1949). *A modern English grammar* (Vol. 4). Routledge. https://www.taylorfrancis.com/books/mono/10.4324/9780203715956/modern-english-grammar-historical-principles-otto-jespersen

Kaplan, R., & Bresnan, J. (1995). Lexical-Functional Grammar: A formal system for grammatical representation. In M. Dalrymple, R. M. Kaplan, J. T. Maxwell, & A. Zaenen (Eds.), *The mental representation of grammatical relations* (pp. 173–281). Center for the Study of Language (CSLI).

Katz, J. J. (1971). Semantic theory. In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantic theory. Semantics: An interdisciplinary reader in philosophy, linguistics and psychology* (pp. 297–307). Cambridge University Press.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39*(2), 170–210. https://doi.org/10.2307/411200

Khalid, K., Afzal, H., Moqaddas, F., Iltaf, N., Sheri, A. M., & Nawaz, R. (2017). Extension of semantic based Urdu linguistic resources using natural language processing. *15th Intl Conf on Dependable, Autonomic and Secure Computing*, 1322–1325. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.214

Kiani, Z. H. (2013). *The Syntax of complex predicates in Urdu* [MPhil Thesis, University of AJ & K, Muzaffarabad]. https://d1wqtxts1xzle7.cloudfront.net/30294154/The_Syntax_of_Complex_Predicates_in_Urdu-

Kiparsky, P. (1997). Remarks on denominal verbs. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (Vol. 64, pp. 473–499). Center for the Study of Language (CSLI).

Koeneman, O., & Zeijlstra, H. (2017). *Introducing syntax*. Cambridge University Press.

Lakoff, G. (1970). *Irregularity in syntax*. Holt Rinohart and Winston.

Lakoff, G. (1971). On generative semantics. In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology*

(Vol. 232, pp. 144–153). Cambridge University Press. https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.1973.tb49487.x

Lakoff, G. (1973). Lexicography and generative grammar: Hedges and meaning criteria. *Annals of the New York Academy of Sciences*, *211*, 144–153.

Levin, B. (1985). Lexical semantics in review. In B. Levin (Ed.), *Lexical project working papers* (pp. 1–62). MIT Press. https://cir.nii.ac.jp/crid/1571980073975166208

Levin, B. (1989). *Towards a lexical organization of English verbs*. University of Chicago Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago Press.

Levin, B., Glass, L., & Jurafsky, D. (2019). Systematicity in the semantics of noun compounds: The role of artifacts vs. natural kinds. *Linguistics*, *57*(3), 429–471. https://doi.org/10.1515/ling-2019-0013

Levin, B., & Rappaport, M. (1994). A Preliminary Analysis of Causative Verbs in English. *Lingua*, *92*, 35–77. https://doi.org/10.1016/0024-3841(94)90337-9

Manetta, E. (2019). Verb-phrase ellipsis and complex predicates in Hindi-Urdu. *Natural Language & Linguistic Theory*, *37*(3), 915–953.

McCawley, J. D. (1968). Lexical insertion in a transformational grammar without deep structure. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *4*(1), 71–80.

Meng, L., Huang, R., & Gu, J. (2013). A Review of Semantic Similarity Measures in WordNet. *International Journal of Hybrid Information Technology*, *6*(1), 1–12.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography*, *3*(4), 235–244. https://doi.org/10.1093/ijl/3.4.235

Mohanan, T. (1994). *Argument structure in Hindi*. Center for the Study of Language (CSLI).

Mohanan, T. (1997). Multidimensionality of representation: NV complex predicates in Hindi. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 431–471). Center for the Study of Language (CSLI).

Morato, J., Marzal, M. A., Lloréns, J., & Moreiro, J. (2004). WordNet applications. *Proceedings of Second International WordNet Conference, GWC 2004*, 20–23. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8e4b1c9e0141ed001057fa56824a513526cbf5a2

More, A., Çetinoğlu, Ö., Çöltekin, Ç., Habash, N., Sagot, B., Seddah, D., Taji, D., & Tsarfaty, R. (2018). Conll-ul: Universal morphological lattices for universal dependency parsing. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*. LREC 2018. https://aclanthology.org/L18-1608.pdf

Mushtaq, Z. (2015). *Grammatical analysis of Urdu auxiliary verbs* [PhD Thesis, University of Management and Technology Lahore]. https://www.cle.org.pk/Publication/theses/2017/Grammatical%20Analysis%20of%20Urdu%20Auxiliary%20Verbs.pdf

Naseer, A., & Hussain, S. (2010). Tense and Aspect in the Context of Urdu. *International Journal of Electrical & Computer Sciences*, *10*(05), 47–54.

Nichols, L. (2008). Lexical semantic constraints on noun roots and noun borrowability. *Studies in Language*, *32*(3), 683–700. https://doi.org/10.1075/sl.32.3.10nic

Pandharipande, R. (1990). Serial verb construction in Marathi. In A. Zwicky & B. Joseph (Eds.) *When verbs collide* (pp. 178–199). Ohio State University. Department of

Linguistics. https://kb.osu.edu/bitstream/handle/1811/82132/WPL_39_December_1990_178.pdf

Paradis, C. (2012). Lexical semantics. In *The encyclopedia of applied linguistics*. Wiley-Blackwell. https://portal.research.lu.se/files/5751127/4857965.pdf

Pinker, S. (1989). Language acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 359–400). MIT Press. http://20130924-29402-1lmcu79-libre-libre.pdf

Pinker, S. (1998). Words and rules. *Lingua*, *106*(1–4), 219–242. https://www.sciencedirect.com/science/article/pii/S0024384198000357

https://books.google.com/books?hl=en&lr=&id=xDSBaet2uSsC&oi=fnd&pg=PR11&dq=lexical+semantics&ots=9D7bmfMQAg&sig=SAH8PwJVj9Jc_hUXzMgopgfo9jM

Platts, J. T. (1874). *Grammar of the Hindustani or Urdu Language*. Wm. H. Allen & Company. https://books.google.com.pk/booksPlatts+(1874)+Categories+of+Tense

Princeton University. (2022, August 7). WordNet. https://wordnet.princeton.edu/

Reference corpus for term extraction | Sketch Engine. (2015, March 18). https://www.sketchengine.eu/corpora-and-languages/reference-corpus/

Rosen, C. (1997). Auxiliation and serialization: On discerning the difference. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 175–202). Center for the Study of Language (CSLI).

Saeed, A., Nawab, R. M. A., Stevenson, M., & Rayson, P. (2019). A sense annotated corpus for all-words Urdu word sense disambiguation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, *18*(4), 1–14. https://doi.org/10.1145/3314940

Schmidt, R. L. (1999). *Urdu, an essential grammar*. Routledge. https://doi.org/10.4324/9780203979280

Sharma, D. M. (2010). On the Role of NLP in Linguistics. *Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the Common Ground*, 18–21. https://aclanthology.org/W10-2103.pdf

Shibatani, M. (1972). Three reasons for not deriving 'kill' from 'cause to die' in Japanese. In J. P. Kimbell (Ed.), *Syntax and semantics* (Vol. 1, pp. 125–137). Seminar Press. https://brill.com/display/book/edcoll/9789004372986/BP000010.xml

Simone, R., & Masini, F. (2007). Support nouns and verbal features: A case study from Italian. In A. Grezka & F. Martin-Berthet (Eds.), *Verbes et classes sémantiques, Special issue of Verbum* (pp. 143–172). https://www.researchgate.net/profile/Francesca-Masini/publication/228772571_Support_Nouns_and_Verbal_Features_a_case_study

Singh, M., Shukla, R., Saraswati, J., Kashyap, L., Kanojia, D., & Bhattacharyya, P. (2016). Mapping it differently: A solution to the linking challenges. *Proceedings of the 8th Global WordNet Conference (GWC)*, *8*, 411–418. https://aclanthology.org/2016.gwc-1.57

Sinha, R. M. K. (2009). Mining complex predicates in Hindi using a parallel Hindi-English corpus. *Proceedings of the Workshop on Multiword Expressions: Identification, Interpretation, Disambiguation and Applications (MWE 2009)*, 40–46. https://aclanthology.org/W09-2906.pdf

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition, 38*(2), 179–211. https://doi.org/10.1016/0010-0277(91)90051-5

Tengi, R. I. (1998). Design and implementation of the WordNet lexical database and searching software. In C. Fellbaum (Ed.), *WordNet: An electronic lexical database* (pp. 105–128). The MIT Press.

https://books.google.com.pk/books?hl=en&lr=&id=Rehu8OOzMIMC&oi=fnd&pg=PA105&dq=Design+and+implementation+of+the+WordNet+lexical+database+and+searching+software&ots=IsnePhUVlb&sig

UniversalDependencies/UD_Urdu-UDTB. (n.d.). GitHub. Retrieved July 29, 2020, from https://github.com/UniversalDependencies/UD_Urdu-UDTB

Urdu to English dictionary—Google Search. (n.d.). Retrieved July 24, 2022, from https://www.urduenglishdictionary.org/

Urdu Verb List. (2011). https://www.cle.org.pk/software/ling_resources/urduverblist.htm

Urdu WordNet 1.0 Wordlist. (2013). Center for Language Engineering, UET, Lahore. https://www.cle.org.pk/software/ling_resources/UrduWordNetWordlist.htm

Urooj, S., Hussain, S., Adeeba, F., Jabeen, F., & Parveen, R. (2012). CLE Urdu digest corpus. *Proceeding of Conference on Language & Technology*, 47–53. https://www.cle.org.pk/clt12/

Urooj, S., Shams, S., Hussain, S., & Adeeba, F. (2014). Sense tagged CLE Urdu digest corpus. *Proceeding of Conference on Language & Technology (CLT2014)*. Conference on Language and Technology, Lahore. https://cle.org.pk/Publication/papers/2014/Sense%20Tagged%20CLE%20Urdu%20Digest%20Corpus.pdf

Vaidya, A., Rambow, O., & Palmer, M. (2014). Light verb constructions with ‘do’and ‘be’in Hindi. *Proceedings of Workshop on Lexical and Grammatical Resources for Language Processing*, 127–136. https://aclanthology.org/W14-5816.pdf

Van Valin Jr, R. D. (1990). Semantic parameters of split intransitivity. *Language*, *66*, 221–260. https://doi.org/10.2307/414886

Van Valin Jr, R. D., & LaPolla, Randy J. (1997). *Syntax: Structure, meaning, and function*. Cambridge University Press.

Wierzbicka, A. (1988). *The Semantics of grammar*. John Benjamins Publishing.

Williams, E. (1997). Lexical and syntactic complex predicates. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex predicates* (pp. 13–28). Center for the Study of Language and Information.

WordNet | A Lexical Database for English. (2022). WordNet. https://wordnet.princeton.edu/

Wunderlich, D. (1997). Cause and the structure of verbs. *Linguistic Inquiry*, *28*(1), 27–68. https://www.jstor.org/stable/4178964

Yule, G. (2020). *The study of language* (8th ed.). Cambridge University Press.

Zafar, A., Mahmood, A., Abdullah, F., Zahid, S., & Mustafa, A. (2012). Developing Urdu WordNet using the merge approach. *Proceedings of Conference on Language and Technology (CLT12)*, 55–59. https://www.cle.org.pk/Publication/papers/2012/8.pdf